



# JOURNAL OF INFORMATION TECHNOLOGIES

BİLİŞİM TEKNOLOJİLERİ DERGİSİ

Volume / Cilt 15 Number / Sayı 4 Year / Yıl 2022 Month / Ay October / Ekim





GAZİ ÜNİVERSİTESİ (GAZİ UNIVERSITY)  
BİLİŞİM ENSTİTÜSÜ (INSTITUTE OF INFORMATICS)

**BİLİŞİM TEKNOLOJİLERİ DERGİSİ**  
(JOURNAL OF INFORMATION TECHNOLOGIES)

ISSN: 1307-9697 e-ISSN: 2147-0715



Cilt (Volume): 15

Sayı (Issue): 4

Ekim (October) 2022

Sahibi (Owner)  
**Dr. Musa YILDIZ**  
Rektör (Rector)

**Genel Yayın Yönetmeni & Baş Editör**  
(General Publication Director & Editor in Chief)  
**Dr. Aslıhan TÜFEKÇİ**  
Bilişim Enstitüsü Müdürü  
Director of Institute of Informatics

**Yardımcı Editör**  
(Associate Editor)  
**Dr. Öner BARUT**  
Bilişim Enstitüsü Müdür Yardımcısı  
Vice Director of Institute of Informatics

**Yardımcı Editör**  
(Associate Editor)  
**Dr. Sinan TOKLU**  
Bilişim Enstitüsü Müdür Yardımcısı  
Vice Director of Institute of Informatics

**Editörler (Editors)**

❖ <b>Dr. Ahmet KARAARSLAN</b>	Ankara Yıldırım Beyazıt Üniversitesi (Ankara Yıldırım Beyazıt University)
❖ <b>Dr. Hüseyin POLAT</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. İbrahim KÖK</b>	Pamukkale Üniversitesi (Pamukkale University)
❖ <b>Dr. Mahir DURSUN</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Mehmet ŞİMŞEK</b>	Milli Savunma Üniversitesi (National Defence University)
❖ <b>Dr. M. Hanefi CALP</b>	Ankara Hacı Bayram Veli Üniversitesi (Ankara Hacı Bayram Veli University)
❖ <b>Dr. Mohammad Abdus SALAM</b>	Southern University and A&M Collage
❖ <b>Dr. Oktay YILDIZ</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Olgun DEĞİRMENÇİ</b>	TOBB ETÜ (TOBB Economics and Technology University)
❖ <b>Dr. Recep BENZER</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Serdar KULA</b>	Gazi Üniversitesi (Gazi University)

**Yayın Danışma Kurulu (Editorial Advisory Board)**

<b>Dr. Ahmet COŞAR</b> Turkish Aeronautical Association University, Turkey	<b>Dr. Aslanbek NAZİEV</b> Ryazan State University, Russia	<b>Dr. Bogdan PATRUT</b> Alexandru Ioan Cuza University of Iasi, Romania
<b>Dr. Deepak GUPTA</b> Maharaja Agrasen Institute of Technology, India	<b>Dr. Jafar A. ALZUBİ</b> Al-Balqa Applied University, Jordan	<b>Dr. Jolanta SABAITYTĖ</b> Vilnius Gediminas Technical University, Lithuania
<b>Dr. İlya LEVİN</b> Tel Aviv University, Israel	<b>Dr. Pınar KARAGÖZ</b> Middle East Technical University, Turkey	<b>Dr. Ufuk ÇAĞLAYAN</b> Yaşar University, Turkey

<b>Dr. Veysi İŞLER</b> <i>Hasan Kalyoncu University, Turkey</i>	<b>Dr. Victor Hugo Costa DE ALBUQUERQUE</b> <i>Universidade de Fortaleza, Brazil</i>	<b>Dr. Vijender Kumar SOLANKİ</b> <i>CMR Institute of Technology, India</i>
<b>Dr. Ebrahim KHOSRAVI</b> <i>Clayton State University, United States</i>		

<b>Dil Editörü</b> <i>(Language Editor)</i> <b>Dr. Çağla Gizem AKKAŞ</b> <i>Bilişim Enstitüsü</i> <i>Institute of Informatics</i>
---

<b>Teknik Koordinatör</b> <i>(Technical Coordinator)</i> <b>Muhammed Ali KOŞAN</b> <i>Bilişim Enstitüsü</i> <i>Institute of Informatics</i>
---

<b>Teknik Sorumlu</b> <i>(Technical Assistant)</i> <b>Merve ÇELEBİ</b> <i>Bilişim Enstitüsü</i> <i>Institute of Informatics</i>
---

<b>Teknik Sorumlu</b> <i>(Technical Assistant)</i> <b>Tuana İRKEY</b> <i>Bilişim Enstitüsü</i> <i>Institute of Informatics</i>
--

<b>Sekreterlik</b> <i>(Secretary)</i> <b>Bilge AKYÜZ ÜNLÜ</b> <i>Bilişim Enstitüsü</i> <i>Institute of Informatics</i>
--

<b>Bilişim Teknolojileri Dergisi uluslararası hakemli bir dergidir.</b> Journal of Information Technologies is an international refereed journal.
<b>Yazışma Adresi (Contact Address)</b> Tunus Cad. No: 35 Kavaklıdere Çankaya/ANKARA Telefon / Telephone: 0312 202 38 01 Faks / Fax: 0312 212 79 29
<b>Çevrimiçi Değerlendirme Sistemi (Online Evaluation System)</b> <a href="http://dergipark.gov.tr/gazibtd">http://dergipark.gov.tr/gazibtd</a> <b>E-posta (e-mail):</b> btd@gazi.edu.tr
<b>Bilişim Teknolojileri Dergisi 3 ayda bir (Ocak, Nisan, Temmuz, Ekim) yayınlanmaktadır.</b> Journal of Information Technologies is published every 3 months (January, April, July, October).

# Derin Öğrenme Modellerinin Doğruluk, Süre ve Boyut Temelli Ödünleşme Değerlendirmesi

*Araştırma Makalesi/Research Article*

 İsmail ARI,  Mustafa Barış ÇAMLI

Bilgisayar Mühendisliği, Özyeğin Üniversitesi, İstanbul, Türkiye

[ismail.ari@ozyegin.edu.tr](mailto:ismail.ari@ozyegin.edu.tr), [mustafa.camli@ozu.edu.tr](mailto:mustafa.camli@ozu.edu.tr)

(Geliş/Received:22.08.2021; Kabul/Accepted:01.08.2022)

DOI: 10.17671/gazibtd.976711

**Özet**— Makine öğrenmesi ve özellikle derin öğrenme modellerinin gerçek-zamanlı saha uygulamalarında operasyona alınması için üç ana kriterin aynı anda optimizasyonu gerekmektedir. Bunlar modelin tahmin doğruluğu, eğitim-test süreleri ile dosya boyutu olup ilgili çalışmalarda sadece iki kriter (örnek: doğruluk-süre) beraber göz önüne alınmıştır. Ancak, modellerin tahmin doğruluğunu artırmak için oluşturulan derin sinir ağlarının (DSA) eğitim süresi ve boyutunu artırdığı, boyutunu küçültmek için yapılan çalışmaların ise doğruluğunu düşürdüğü gözlemlenmiştir. Bu üç kriter arasında bir ödünleşme yapılması gerekmektedir.

Farklı optimizasyon tekniklerinin modelin performansına etkisini göstermek için, bu makalede DSA araştırma alanında sıklıkla kullanılan ResNet50, ResNet101, VGG16, VGG19 ve EfficientNet ön-eğitilmiş modellerini CIFAR10, CIFAR100 görsel veri kümeleriyle test ettik. Google Colab Pro ve Tensorflow sistemi üzerinde yaptığımız başarımların elde edilen önemli sonuçların arasında ağırlık quantizasyonun çok-boyutlu optimizasyonunda şu ana kadarki en başarılı teknik olduğu, ağırlık kümeleme ve transfer öğrenimi tekniklerinin ise ancak 2-boyutta fayda sağladıkları söylenebilir. Çalışmamızda ayrıca, literatürde ilk defa DSA'lar için bir operasyonel skor ve modelden-modele katman aktarımı metodunu tasarlayıp, sınadık. Oluşturulan çerçevenin, yeni geliştirilen DSA modellerinin operasyona sokulmadan önce çok-boyutlu değerlendirilebilmeleri için bir referans teşkil etmesi umuyoruz.

**Anahtar Kelimeler**— operasyonel makine öğrenmesi, derin öğrenme, derin sinir ağları, ön-eğitilmiş modeller, ResNet, VGG, CIFAR, doğruluk, eğitim süresi.

## Tradeoff Assessment of Deep Learning Models based on Accuracy, Time and Size

**Abstract**— Machine Learning and especially deep learning models need to be optimized over three main criteria concurrently, to be operationalized in real-time field applications. These criteria are model's accuracy, training-testing times and file size. Related work only considers two criteria (e.g. accuracy-time) together. However, it is observed that deep neural networks (DNN) designed to improve model accuracy can increase training time and size, while efforts to reduce model size can lead to lower accuracy. A trade-off needs to be made among these three criteria.

In this paper, to demonstrate the effects of different optimization techniques on model performance, we tested ResNet50, ResNet101, VGG16, VGG19, EfficientNet pre-trained models with CIFAR10, CIFAR100 image datasets, which are commonly utilized in the DNN research field. Important performance results obtained over Google Colab Pro and TensorFlow system show that weight quantization is the most successful technique so far in multi-dimensional optimization, while weight clustering and transfer learning techniques remain useful in 2-dimensions. In addition, we designed and tested a new DNN operational score and model-to-model layer transfer method for the first time in literature. We hope that our framework will constitute a multi-dimensional evaluation reference for DNN models before they are operationalized.

**Keywords**— operational machine learning, deep learning, deep neural networks, pre-trained models, ResNet, CIFAR, VGG, accuracy, training time.

## 1. INTRODUCTION

Machine Learning (ML) models can be broadly categorized as supervised or unsupervised: supervised models are trained and tested over labelled data, whereas unsupervised models depend on purely statistical properties and patterns in the datasets. Classifier models fall under supervised category, whereas clustering algorithms are unsupervised. In this paper, we focus on Deep Neural Network (DNN)-based classifier models, which have a high potential for use in real-time applications including autonomous vehicles [1], industrial quality control & maintenance [2], automated financial trading [3], IT security [4], Natural Language Processing (NLP) [5] and many more [6].

In ML classifiers, three performance criteria that need to be optimized simultaneously to achieve the highest operational efficiency are model accuracy, training-testing times, and file size. Figure 1 illustrates our framework for *Operational Machine Learning (OpML)*, where the hypothetical optimal point for all three criteria would be the  $\langle A:0, T:0, S:0 \rangle$  point, which represents the highest accuracy ratio obtained (*1.0-accuracy*) with minimal training-testing times (*0.0 sec*) and smallest model size (*0 MB*). For example, a *random* prediction that requires no training or state-keeping, would meet the time and size criteria, but have very low accuracy. As illustrated in Figure 1 with dashed arrows, trying to improve model accuracy usually results in increased model complexity followed by more training and testing times. In addition, trying to reduce the model size (e.g. by weight clustering or quantization) reduces accuracy, if not handled carefully. Finally, to improve model testing times researchers use more data, which increases the corpus size as well as model size, while also carrying the risk of model overfitting. When designing new models and/or tuning pre-trained models, the operational goal should be to stay within the inner triangles in Figure 1, while pushing the limits on either one of these criteria.

We can combine 3-dimensions into a simple, but efficient operational score (*OpScore*) as shown in Equation 1:

$$OpScore = -\log_{10}(1/(1 - Acc)) + \log_{10} Time + \log_{10} Size(MB) \quad (1)$$

where the optimal score is close to zero and between [0,5] for most practical cases (see Sec.4). Note that, increasing accuracy in Equ. 1 decreases the score, whereas increasing model time and size values increase the score, denoting increasing operational costs. The operational score can be used separately for training and testing times, a weighted average or total of both depending on the operational scenario. We will discuss model training, testing (or inference) and even pre-training (tuning & pruning) times in this paper to understand their implications on real-time field applications. While *OpScore* is directly applicable to different DNN models, a value-based comparison is done over the same dataset for fairness.

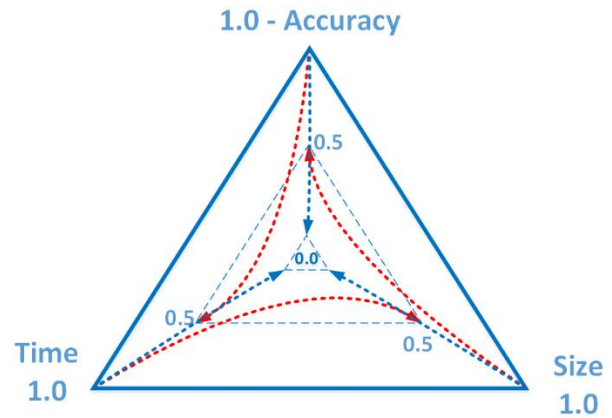


Figure 1. Dimensions of model performance criteria to be optimized concurrently are model accuracy, time and size. We use this scheme as a reference to compare novel DNN models and optimization techniques in terms of their operational efficiency.

Figure 2 illustrates the general structure and components of a DNN classifier model. As an example, the image of an animal such as a cat or a dog is passed through the DNN to generate a correct classification “cat” in text format. The *depth* of a DNN model is measured by the number of hidden layers. As DNN models get deeper they contain hundreds of convolutional layers as well as activation, pooling, regularization, attention, normalization, and dense layers that carry thousands of functions and millions of weights. These facts are valid for Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and their hybrids called R-CNN.

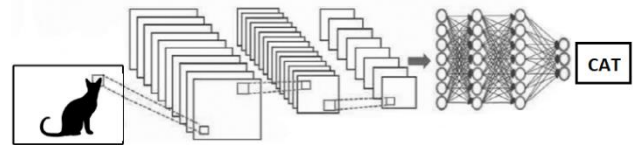


Figure 2. DNN, CNN or RNN models have layered structures containing millions of weights, which account for 100s of millions of floating-point operations over multiple epochs of training and testing.

After training, the models are serialized and saved in storage systems using special Hierarchical Data Formats (HDF) [7] or Open Neural Network Exchange (ONNX) [8] file formats. ONNX is becoming the de facto standard for easily converting and sharing pre-trained models. These models can be uploaded to public model repositories called “model zoo” [9] for online sharing. They are later downloaded locally, possibly in edge or Internet of Things (IoT) devices and operationalized by deserializing the model file into the memory. Next, every input (e.g. image) gets turned into a matrix of float-type variables by the input layer and passed through model layers to deliver correct classification results at the output layer. Convolutional layers generate feature maps from the input. Pooling layers (average or max pooling) down sample feature maps to provide translational invariance. In between these two,

there are activation functions (ReLU, Sigmoid, *etc.*) that are used to introduce nonlinearity. It takes millions of sample images to train accurate models. DNN models can take days to train over traditional Central Processing Units (CPU). Fortunately, massively parallel Graphical Processing Unit (GPU) and novel Tensor Processing Units (TPU) [10, 11] come to rescue to reduce these model training times to hours and minutes.

One potential caveat in model training is over-fitting, where the model strictly and only learns the training data. To avoid over-fitting, techniques such as data cross-validation and activation functions are used. Models may need to be retrained with new data in the field for continuous improvement, yet mobile edge devices [12] are usually not as powerful as the central servers or cloud resources. Therefore, field success depends on optimization of model training time as well as its accuracy and size, concurrently. Model testing does not require weight updates and therefore it is faster than model training. However, since inference time effects field performance it must be measured and optimized carefully.

Contributions of our paper are as follows:

- We develop a 3D framework and a score for evaluation and ranking of DNN model to be operationalized in the field,
- We present a detailed multi-dimensional evaluation of commonly used DNN models including ResNet, VGG and EfficientNet with different datasets,
- We calculate an *OpScore* and rank models using the same unified metric,
- We evaluate the performance effects of modern CPU accelerator chips, including GPU and TPU over DNN training and testing times,
- We quantify the cost of model optimization techniques such as weight clustering,
- We report on sensitivity of model accuracy over increasing number of dense & hidden layers “unfrozen” to be retrained,
- We test a new method that we call “model-to-model layer transfer”.

The rest of the paper is organized as follows. Section 2 gives the background on DNN models and their Key Performance Indicators (KPI) including model accuracy, training and testing times as well as size reduction techniques. Section 3 describes our methodology including experimental setup, datasets and the concept of transfer learning. Section 4 presents the detailed performance assessment of several DNN models and optimization techniques including model quantization and clustering using proposed tradeoff assessment scheme. Most of the contribution listed above are realized in this section. Section 5 discusses related work on model optimizations in comparison to our proposed framework and findings. Section 6 concludes the paper also describing future work.

## 2. BACKGROUND

In this section, we first give a quick background on KPI for OpML (model accuracy, time, size). Next, we describe properties of reference DNN models used in this research.

### 2.1. Model Accuracy

When a set of  $N$  inputs are fed into a classifier, the four possible outcomes of the classification are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), thus  $N=(TP+FP+TN+FN)$ . Accuracy (Eq.2) is simply measured as the ratio of “truly classified” instances over  $N$ . Model’s accuracy, recall, precision and F1-measure formulas are given below by Equations 2-5, respectively.

$$\text{Accuracy} = (TP+TN)/N \quad (2)$$

$$\text{Recall} = TP/(TP+FN) \quad (3)$$

$$\text{Precision} = TP/(TP+FP) \quad (4)$$

$$\text{F1-score} = 2/(1/\text{Precision}+1/\text{Recall}) \quad (5)$$

These are commonly used for model accuracy measures but have small semantic differences and implications. In cases where the items to catch are extremely rare (*e.g.*, a medical test to detect a rare cancer type) precision and recall should also be employed. Recall (Eq.3) focuses on model’s ability to correctly detect (*i.e.*, “not miss” or make “False Negative”) items that should be placed in the selected class, while Precision (Eq.4) denotes model’s ability “not to misplace or confuse” (FP) items belonging to other sets into the selected class. F1-score (Eq.5) is the harmonic average of recall and precision values, giving a unified view of both success metrics.

### 2.2. Model Training and Testing Time

NN are difficult to deploy on IoT and embedded systems with limited resources since they are both compute and memory intensive. Most of the recent R&D efforts in OpML focus on reducing model training and testing (*a.k.a.* inference, serving) times. CPU-GPU manufacturers also produce ML-accelerator chips such as Tensor Processing Units (TPU) to speed up floating point operations in DNN. In cases where a model contains mixed-precision values (float16, float32, float64) together, chip-level support for mixed-precision operations [13] is also crucial for high performance. Otherwise, every arithmetic-logic operation also requires a value conversion (among float16-32) having a diverse performance effect.

### 2.3. Model Size Reduction Techniques

Modern pretrained DNN model files generally range between 10 Megabytes (MB) to 1 Gigabytes (GB) in size. We compared two techniques for model size reduction in this paper, namely quantization [14, 15] and weight sharing (*a.k.a* weight clustering) [16]. Quantization is a technique

to reduce model size by reducing the number of bytes used for each NN weight value. For example, if we use a float16 value (2-bytes) instead of float32 (4-bytes) the model size should be approximately reduced by half. However, since the weights are highly tuned to make accurate predictions and reducing weight precision could also result in loss of accuracy.

In weight sharing technique [16], all the weights in a model are clustered using K-means algorithm (or alike) and centroid values of each cluster are used as a single, shared value for the weights that belong to that cluster. Thus, the number of DNN weights to maintain is reduced from millions to thousands. Potentially, this also results in caching of weight values in L1-L2-L3 caches, which could speed up model training and testing times. However, like quantization, replacing thousands of values with their shared value also results in loss of model accuracy. We test this hypothesis and present comparative results in Sec. 4.

#### 2.4. Models: ResNet50-101, VGG16-19, EfficientNet

Most image classification DNN models including ResNet50, ResNet101 and VGG were originally trained over the ImageNet dataset [17], but they can be retrained and used with other datasets. Table 1 shows their original and quantized sizes as well as layer counts. We can generally observe the 2x reduction in size due to float32 to float16 type change over all models. ResNet50 consists of 176 layers, which are inside 50 blocks (hence the name ResNet50) each consisting of a sequential group of <Conv-2D, Batch-Normalization, ReLU> layers, whereas ResNet101 consists of 346 layers inside 101 such carefully-engineered blocks. ResNet50 original model file is 103 MB whereas ResNet101 model is 180 MB, which is approximately proportional to the number of layers. Since we used the CIFAR datasets, which were different from the ImageNet dataset that was originally used to pretrain ResNet50-101 models, we had to retrain and test them. This would be the case with any operational ML scenario when pretrained models are employed in field applications. Note that the architecture of models (blocks and layers) are not affected by weight quantization techniques.

VGG16 was the winner of ImageNet classification competition in 2014 [18]. While it doesn't have as many layers as ResNet, it has very wide convolutional layers that carry a total of 138 Million parameters that reflect its original model size of 528 MB. VGG19 model is very similar in structure to VGG16 except 3 more convolutional layers in the middle, which increases VGG19 size to 549MB for a slightly better accuracy. When we downloaded models into Keras [19], we omitted the top-layers of the pretrained models, therefore Table 1 reports smaller model sizes without the top layers, obtained before transfer learning.

EfficientNet [20,21] studies model scaling issues and identifies that "carefully balancing network depth, width, and resolution can lead to better performance". They

Table 1: Originally used and quantized model sizes (MB).

Model Name	Original Size (MB)	Quantized Size (MB)	Number of Layers
ResNet50	103	48	176
ResNet101	180	86	346
VGG16	59	29	16
VGG19	80	40	19
EfficientNet-B0	16	8	238

proposed a method that uniformly scales all dimensions using a compound coefficient using which they designed a range of EfficientNet from B0→B7. They demonstrated effectiveness of this method up to 6x in speed and 8x in size over ResNet50-101-152. Table 1 shows the modest size of B0 compared to ResNets and VGG. We'll compare their time-accuracy performances in Section 4.

### 3. METHODOLOGY

This section describes our methodology for tradeoff assessment of different DNN models over different benchmark datasets. Compared techniques include quantization, weight sharing (*clustering*), transfer learning using top (dense + hidden) layer unfreezing (Top-1, Top-10), and direct model-to-model layer-based reuse of weights (*e.g.*, from ResNet101 to ResNet50). We trained each model for 10 Epochs, which allowed for accuracy convergence.

#### 3.1. Experimental Systems and Software

We used Google Colab Pro cloud system [10], TensorFlow [11] Keras [19] and Scikit-learn [22] Python libraries to calculate our model accuracies, training times and sizes. Google Colab Pro provides Intel Xeon® CPU @ 2.30GHz 16-core 32GB DRAM machines with different GPU generations: NVIDIA Tesla Pascal (P100) PCIE-16GB and Volta (V100) SXM2-16GB. A major difference between these two GPU's is that the newer V100 supports mixed-precision operations, while P100 does not. We will discuss the effects of system differences in performance.

#### 3.2. Datasets: CIFAR10 and CIFAR100

Canadian Institute For Advanced Research (CIFAR) has published labelled image datasets that have become the de-facto benchmark for image classification models. CIFAR10 and CIFAR100 include 60,000 32x32 bit color images, belonging to 10 different classes and 100 different classes, respectively. Image contents exclusively show animals (bird, cat, dog, horse, *etc.*), vehicles (automobiles, planes, ships, trucks), flowers, vegetables, trees and people. CIFAR10 has 6000 images from each of 10 class types, whereas CIFAR100 has 600 images from each of 100 types. In both datasets, 50,000 images are used for training and 10,000 for testing purposes. These datasets are identically and independently distributed (IID) [23], basically having no label and quantity skews. Since we compare models over a centralized system architecture in this paper, having non-IID datasets would be equivalent to

leaving some classes out of the question and render our results incomparable with respect to other related work in this domain. Yet, we included a sensitivity analysis having different splits (Table 3: 50-50%, 70-30%, 90-10%) for training and testing CIFAR datasets. In another study [24], we currently tested models with non-IID datasets with respect to their impact on distributed, federated learning scenarios.

### 3.3. Unfreezing Model Layers for Transfer Learning

Transfer learning allows a model, pretrained with a former dataset, to be unfrozen and retrained with the new dataset at hand. Top layers are the ones closer to the model output, which are also the fully-connected, dense layers. Since the input layers are the same for all models (ingesting 32x32 bit images), the output layer and the hidden layers are of utmost importance in improving model performance. Therefore, we start by retraining models' weights from Top-to-Bottom, where Top denotes the model output. We used the term "unfreeze" to denote the number of layers to be retrained. For example, when we "unfreeze" Top-1 (dense) layer the dense layer is retrained, but hidden layers are not affected, whereas when we unfreeze Top-10 layers, we update and adapt 1 dense layer + 9 hidden layers. VGG16-VGG19 models we used had 2 fully-connected (dense) layers [25] and 13-16 hidden layers, respectively. For fairness, we unfroze and retrained the same number of dense & hidden layers in each model type.

After downloading selected pretrained models into Keras for transfer learning, we initialized models' top layers with Keras' default Glorot (Xavier) uniform initializer. This initializer is known to fix the exploding & vanishing gradient problems associated with random initialization [26]. We can regard our new model-to-model layer transfer strategy as a coarse-granularity initialization technique where the top-layers of successor models (e.g. ResNet101 to ResNet50) are transferred with the presumptions that (a) there is an evolutionary relation between the two model generations which guarantees input-output compatibility (2) the successor model has higher accuracy and the predecessor model has smaller size. We report the results in Section 4.

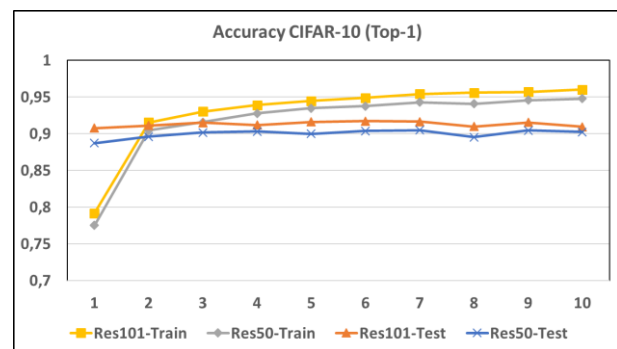
## 4. RESULTS AND DISCUSSION

In this section, we first present comparisons of the baseline ResNet50-ResNet101 models over selected datasets. We start by comparing their accuracies and continue with training-testing times over two different GPU systems. Then, we compare different model optimization techniques in terms of their impact on model accuracy, time, and size. We fix both training and testing epochs at 10 for all experiments for reference. While training models for 10 epochs provides weight updates and accuracy convergence, testing models for 10 epochs serves a statistical validation for accuracy and inference timing measurements.

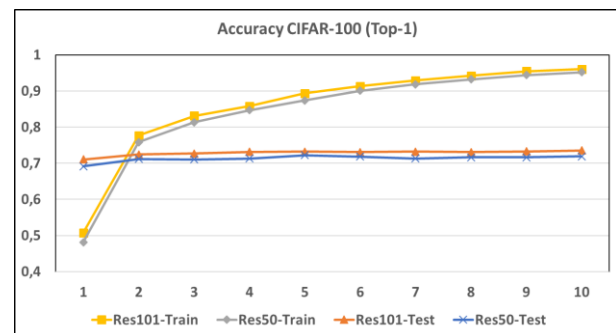
### 4.1. Analysis and Comparison of Model Accuracies

Figure 3 shows the comparison of classification accuracies for ResNet50-101 models over CIFAR10 and CIFAR100 datasets while retraining Top-1 (Dense) and Top-10 (1 Dense + 9 Hidden) layers. First, we observe in Figure 3a-3b that retraining of even the last (Top-1) layer in both ResNet50 and ResNet101 results in a significant improvement of accuracy. Since original models were trained using ImageNet, this case shows the immediate benefits of transfer learning for both CIFAR10 (from ~0.78-to-0.96) and CIFAR100 (from ~0.5-to-0.95).

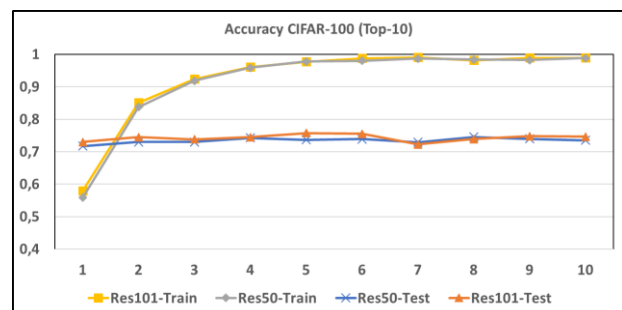
Second, training accuracies are always higher than testing accuracies and the difference between them denotes the amount of overfitting (since we have IID data). Third, ResNet101 only has *slightly higher* training and testing



(a) Accuracy with CIFAR10 dataset Top-1 retraining.



(b) Accuracy with CIFAR100 dataset Top-1 retraining.



(c) Accuracy with CIFAR100 dataset Top-10 retraining.

Figure 3. Comparison of accuracies for Resnet50 and Resnet101 models with (a) CIFAR10 Top-1, (b) CIFAR100 Top-1, and (c) CIFAR100 Top-10 layers retraining.



Table 2: ResNet50-101 accuracy values over CIFAR10-100 datasets using Recall, Precision and F1 measures.

Model	Dataset	Recall	Precision	F1
CIFAR10	Res50-Top10	0.92	0.92	0.92
	Res101-Top10	0.99	0.99	0.99
CIFAR100	Res50-Top10	0.72	0.76	0.74
	Res101-Top10	0.73	0.77	0.75

accuracy compared to ResNet50. However, the dataset's image complexity has a bigger role in the final result, *i.e.* CIFAR100 being more complex than CIFAR10 leads to lower accuracies (0.7 vs. 0.9). Otherwise, ResNet50-101 models have very close validation accuracies in both datasets. This proves that additional efforts and complexity added to create ResNet101 model *did not contribute significantly to higher accuracy* while it caused 60-80% increase in model training time and sizes. ResNet101 takes 1.6-1.8x the training time of ResNet50 (see Table3: P100: 470sec vs. 840sec, V100: 280sec vs. 470sec) and increases model size by ~1.74x (see: Table1: 103MB → 180MB). We therefore find that operationalizing models in the field to train & test them over different datasets can be as important as offline feature and model engineering.

In Figure 3b-3c, a comparison of models over retraining of Top-1 vs. Top-10 layers with CIFAR100 shows faster convergence in Top-10 layers (*e.g.*, 0.9 point reached at 3 epochs in Top-10 vs. 5 epochs in Top-1) and a slightly higher accuracy result with Top-10 at the end (Top-1: 0.95 vs. Top-10:0.98). Note that we refer to unfrozen layers and not the probably of a class prediction being in Top-1 vs. Top-5 lists, since Top-5 result in an unacceptably higher accuracies for field applications.

In Table 2, we briefly report different accuracy measures including Recall, Precision and F1 for ResNet50-ResNet101 models over different datasets. Since these benchmark datasets were IID for all class types, all of the measures gave similar accuracy results: ~0.9 for CIFAR10 and ~0.7 for CIFAR100. There are no label or quantity skews to cause reductions in precision or recall. We tested these issues in our recent work on distributed federated learning [24] where non-IID datasets [27,28] can have more impact on federated model performance [29].

While not shown here for brevity, we also tested a Random Model (*i.e.* naive reference) with CIFAR100 dataset for cross-examination and correctness control and we obtained accuracy: 0.01 (1/100), recall: 0.01, precision: 0.01, Area Under Curve (AUC): 0.50 as expected.

#### 4.2. Model Training-Testing Times and GPU Effects

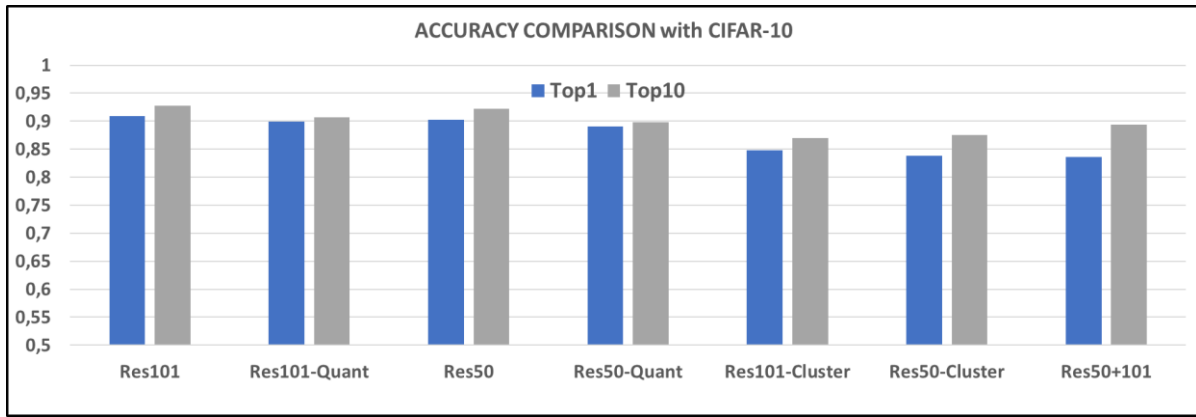
In this part, we compare training and testing times of the baseline ResNet50-ResNet101 models over CIFAR100 dataset and by unfreezing Top10 layers. As a part of our measurement and sensitivity analysis we run the experiments over two different GPUs (P100-V100) and change the ratio of training-testing datasets (50-50%, 70-30%, 90-10%) to validate per mini-batch processing times. We used TensorFlow Keras library and its *evaluate()* function. The *batch\_size* was 128 items for both training and testing.

Table 3 shows that the training time for ResNet50 was 26.8 sec/30,000 images (~0.9 ms/image) and for ResNet101 42.7 sec/30,000 images (~1.4 ms/image). For example, if a real-time application sets the inference time limit for the DNN model to be 1ms, ResNet50 would be eligible for field deployment whereas ResNet101 would not qualify. Depending on the ratio of increase in training data the training times increase and the testing times decrease, but the total time remains consistent around ~53 seconds for ResNet50 and ~84 second for ResNet101 over P100 GPU. We observe that both training and testing times of Resnet101 are 1.6-1.8x slower than ResNet50 for all data sizes (*e.g.* Res101/Res50 testing for 90-10%:  $8.2/4.9=1.67$ ), which is due to the model size and complexity.

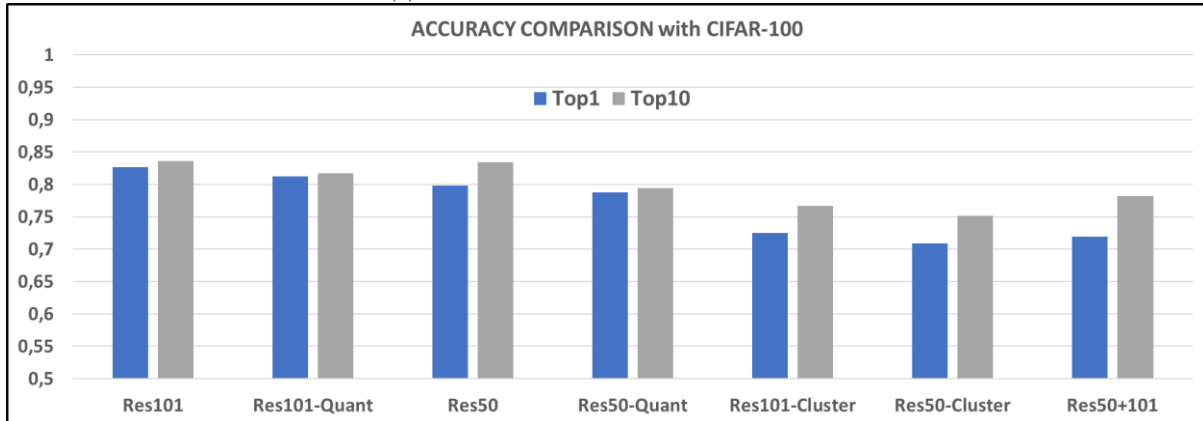
Training and testing times improved by ~45% for both models when we used Volta (V100) GPU, since it has ~45% more GPU cores (+TPU cores), higher memory bandwidth as well as memory capacity (32GB vs 16GB). For example, Res50 50-50% was 14.5 sec vs. 26.8 sec. In an offline analysis, we also compared Adam or SGD optimizers and observed no performance impact on training time. The total time was similarly ~29 sec. for ResNet50 and ~47 sec. for ResNet101 with V100.

Table 3: Effects of different GPU accelerators (P100,V100) on model training-testing time for different data ratios (50-50%, 70-30%, 90-10%) using CIFAR100 dataset and Top-10 layers for retraining.

GPU	Ratio (%)	Train Time (sec)		Test Time (sec)		Total Time (sec)	
		Res50	Res101	Res50	Res101	Res50	Res101
P100 (Pascal)	50-50	26.8	42.7	27.8	40.5	64.6	83.2
	70-30	37.4	59.9	14.7	24.4	52.1	84.3
	90-10	48.1	76.9	4.9	8.2	53.0	85.1
V100 (Volta)	50-50	14.5	23.7	15.0	22.7	29.5	46.4
	70-30	20.3	33.1	8.1	20.6	28.4	53.7
	90-10	25.9	42.5	2.7	5.2	28.6	47.7



(a) Accuracies of models with CIFAR10 dataset.



(b) Accuracies of models with CIFAR100 dataset.

Figure 4. Accuracy comparison of different pre-trained DNN models and their optimized (quantized, clustered, transfer learning, model-to-model layer transfer) versions after retraining of Top-1 and Top-10 layers.

4.3. Comparison of Different Model Optimization Techniques on Accuracy

Figure 4 shows the accuracy results for the basic Resnet50-Resnet101 models and their optimized (quantized, weight-clustered, transferred) versions over CIFAR10 (Fig.4a) and CIFAR100 (Fig.4b) datasets, respectively and in descending order of their accuracy achievements. We report the accuracies obtained after 10 epochs of retraining.

First, we observed that unfreezing Top-10 Layers resulted in higher accuracy compared to unfreezing only Top-1 layer. For Res50-cluster and Resnet101-cluster models, although we used a relatively large number of clusters  $K=8192$  to obtain high-precision weights, the accuracy drop over the original model was significant (~6%) compared to the quantization optimization. Finally, transferring top layers directly from ResNet101 to the ResNet50 model (Res50+101 model-2-model) resulted in the lowest accuracy before retraining. However, this strategy was still more promising than the Res50-Cluster model after retraining with both CIFAR10 and CIFAR100. This technique requires more investigation by careful selection of the transferred blocks.

Also, accuracies of all unmodified (Res50-Res101) and optimized (quantized, clustered, transferred) models are

higher for CIFAR10 (~0.85-0.92), then CIFAR100 (~0.65-0.75). Surprisingly, quantized versions of Res50-Res101 from float32-to-float16 resulted in a very small decrease (<2%) in the accuracy, although we gained a significant size reduction (see Sec 4.5).

Based on this subtle, but important difference we decided to conduct another sensitivity analysis of model depth over Resnet and VGG. Figure 5 shows that ResNet benefited more as we unfroze more layers, whereas VGG did not.

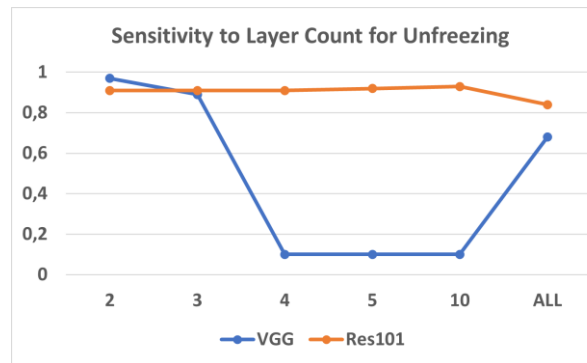


Figure 5: Sensitivity of Resnet101 and VGG model testing accuracy over different count of Top layers unfrozen and retrained for 10 epochs.

Table 4: Training and testing times with different NVIDIA Tesla GPUs (P100-V100).

GPU	Dataset-Layers	Resnet50	Res50-Quant	Res50-Cluster	Resnet101	Res101-Quant	Res101-Cluster	Res50+101
P100	CIF10-Top1	472	630	471	842	1062	847	838
	CIF100-Top1	476	642	476	841	1048	861	841
	CIF10-Top10	528	671	540	894	1108	892	896
	CIF100-Top10	540	691	524	895	1121	898	902
V100	CIF10-Top1	286	226	286	474	329	470	477
	CIF100-Top1	286	218	285	493	325	470	477
	CIF10-Top10	318	238	323	509	343	501	510
	CIF100-Top10	319	238	319	509	343	507	514

Including dense (Top 1-3) classifier layers to transfer learning is generally a good idea but moving further into the hidden layers does not always guarantee accuracy gains. This may be due to the fact that model architectures are originally designed and fit for their initial datasets, which is ImageNet dataset in this case.

#### 4.4. Comparisons of Training-Testing Times

Table 4 shows the total times for 10 training + 10 testing epochs with 50,000 training and 10,000 testing images (ratio ~83-17%) of original ResNet50-101 models and their optimized (quantized, clustered, transfer learned, and model-2-model layer transfer) versions over P100 and V100 GPUs. There are two main findings within these results: (1) clustering does not change the train-test times, but there is a preparation time cost for these models, (2) quantization is highly-sensitive to GPU's mixed-precision support at the hardware level: if there is support (V100) the resulting times are 25% faster, but if not (P100) the train-test times can be 30% slower. Resnet50+101 model-to-model layer transfer has similar time performance to weight clustering.

#### 4.5. 3D Comparisons Including Size Reduction

In Table 1 we showed that quantization can reduce model sizes by 50% (or 2x). Note that the accuracies presented in Figure 4 were comparable within 2% of original models. Quantization with mixed-precision support leads to the best results in multi-dimensional assessment. However, edge nodes currently don't have GPUs with mixed-precision support. These high-end GPUs are powerful, but expensive and energy-consuming devices. So, their operational use at the edge for mass scale is currently not feasible.

Figure 6 shows the model inference time (X-axis) vs. model accuracy (Y-axis) and bigger bubble sizes refer to the bigger model sizes. High accuracy, small inference time and smallest model sizes are preferable. From Fig. 6 we can see that EfficientNet-B0 Top-1-10 quant models have the highest accuracy, smallest inference time and size. In comparison, VGG19 models are at the outer rim of the chart with lower accuracy, higher inference times and relatively larger bubble sizes. Finally, Resnet101 models have relatively larger sizes and inference times compared to EfficientNet-B0 and Resnet50 models (respectively, their quant versions) although their accuracies are also

high. For each model, we observe that quantization increases the inference times while it makes the model sizes smaller. While Fig 6. is illustrative, it cannot help us make a final decision in model selection. Therefore, we use a new operational score to combine conflicting dimensions and rank our selected models. Table 5 shows the model rankings based on OpScore. We added a fictitious MIN and MAX model in the table, which picks the best and worst values from among the analyzed models in the table. These values can later be used for MIN-MAX normalization purposes for each application. Theoretically, our OpScore can take negative or relatively high values, but it will be between [0,5] for almost all practical cases. For example, a very poor performing model with  $\langle Acc:0,01; Size=1GB; Time=100sec \rangle$  has OpScore=5.0, whereas a relatively good model with  $\langle Acc:0,99; Size=10MB; Time=10sec \rangle$  has OpScore=0.0. OpScore values of all the real pretrained models for CIFAR10 are between [0,67-2,63] in Table 5.

Figure 7 shows the performance results of all models over CIFAR100 dataset using a bubble chart for visual comparison. Accuracy values have dropped for all the models due to classification complexity of CIFAR100 as expected. Yet, the general patterns and findings for CIFAR10 reported in Figure 6 are still valid. Inference times and model sizes are similar. Table 6 show the results ranked by OpScore, respectively. Due to the decrease in accuracy, OpScore values of all models with CIFAR100 are higher between [1,15-3,10]. There are only small changes in the ordering of the models.

Figure 8 presents the sensitivity analysis of clustering technique ( $K=512 \rightarrow 16,384$ ) on model accuracy. Note that the accuracy drops from ~0.9 to ~0.6 for CIFAR10 and ~0.7 to ~0.3 as the number of weight clusters decrease from  $K=8192$  to  $K=1024$ . This is due to the immense loss of information in NN weights. Table 7 presents the same analysis for effects on model size and model clustering time. Note that this clustering time is a "preparation" time that has to be spent before the model is used for training and testing. Therefore, it is an additional cost. Most of the related work do not report their model preparation times, which could become a significant limit in practice. The main goal of clustering is to reduce model size. Table 7 shows that this goal is achieved ( $103MB \rightarrow 39MB$ ), but at the expense of accuracy loss and added clustering times.

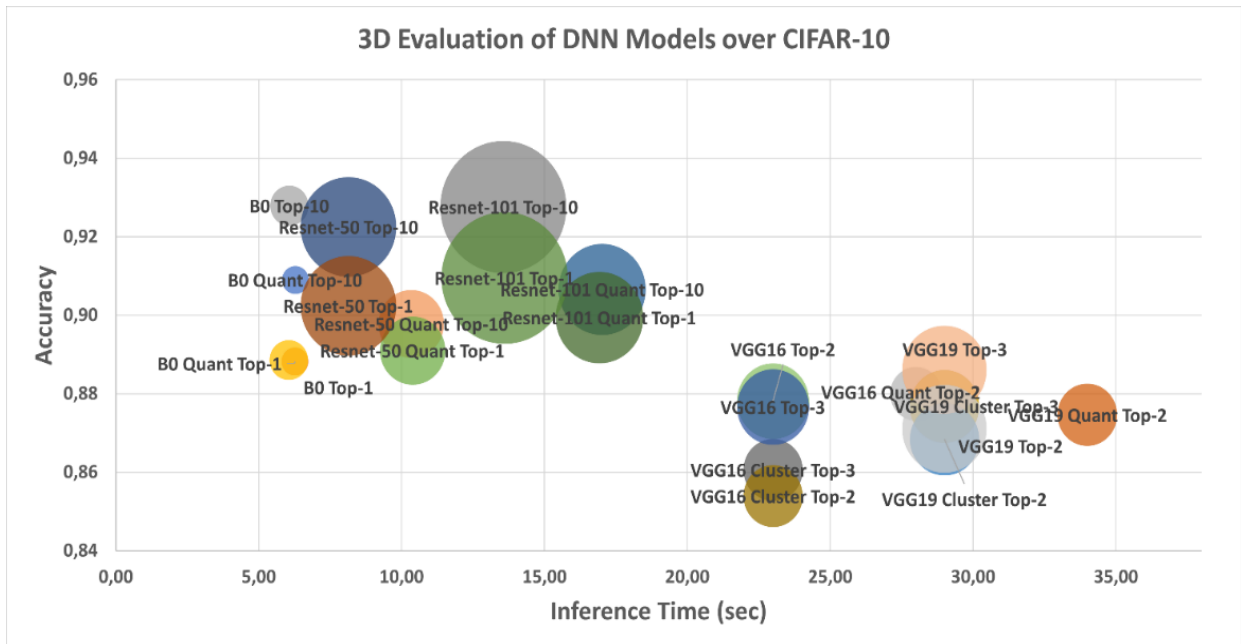


Figure 6. 3D Bubble chart comparison of selected DNN models with CIFAR10.

Table 5. Ranking of evaluated models based on 3D KPI and unified OpScore using CIFAR10 dataset. MIN-MAX values are obtained by selecting the best and worst value for each column and generating a score for this fictitious model as reference; this can be used for normalization. Note that best and worst OpScore values range between 0-5.

Model Name	Accuracy	Size (MB)	Training Time (sec)	Inference Time (sec)	Score
MIN	0,93	8,10	27,72	6,06	0,54
B0 Quant Top-10	0,91	8,10	31,77	6,28	0,67
B0 Quant Top-1	0,89	8,10	27,98	6,27	0,75
B0 Top-10	0,93	16,60	29,92	6,07	0,86
B0 Top-1	0,89	16,60	27,72	6,06	1,05
B0 Cluster Top-10	0,40	11,00	29,34	6,07	1,60
Resnet-50 Quant Top-10	0,90	48,00	56,41	10,34	1,71
B0 Cluster Top-1	0,21	11,00	27,75	6,07	1,72
Resnet-50 Quant Top-1	0,89	48,00	51,51	10,39	1,74
Resnet-50 Top-10	0,92	103,00	44,48	8,14	1,81
Resnet-50 Top-1	0,90	103,00	40,09	8,14	1,91
VGG16 Quant Top-2	0,88	29,50	136,00	28,00	2,00
Resnet-50 Cluster Top-10	0,75	61,00	44,48	8,14	2,09
VGG16 Cluster Top-3	0,86	39,90	118,00	23,00	2,11
VGG16 Cluster Top-2	0,85	39,90	114,00	23,00	2,13
Resnet-101 Quant Top-10	0,91	86,00	89,78	17,02	2,14
Resnet-101 Quant Top-1	0,90	86,00	84,29	16,92	2,17
Resnet-50 Cluster Top-1	0,70	61,00	40,16	8,15	2,17
VGG16 Top-2	0,88	58,90	114,13	23,00	2,22
VGG16 Top-3	0,88	58,90	117,78	23,00	2,22
VGG19 Quant Top-2	0,87	40,10	167,48	34,00	2,23
Resnet-101 Top-10	0,93	180,00	71,18	13,56	2,25
VGG19 Cluster Top-3	0,88	54,30	144,85	29,00	2,29
VGG19 Cluster Top-2	0,87	54,30	141,33	29,00	2,32
Resnet-101 Top-1	0,91	180,00	67,51	13,60	2,35
VGG16 Quant Top-3	0,73	29,50	140,00	28,00	2,35
VGG19 Top-3	0,89	80,20	144,80	29,00	2,42
VGG19 Top-2	0,87	80,20	141,23	29,00	2,48
VGG19 Quant Top-3	0,77	40,10	171,31	34,00	2,50
Resnet-101 Cluster Top-10	0,77	115,00	71,25	13,56	2,55
Resnet-101 Cluster Top-1	0,73	115,00	67,51	13,63	2,63
MAX	0,21	180,00	171,00	34,00	3,68

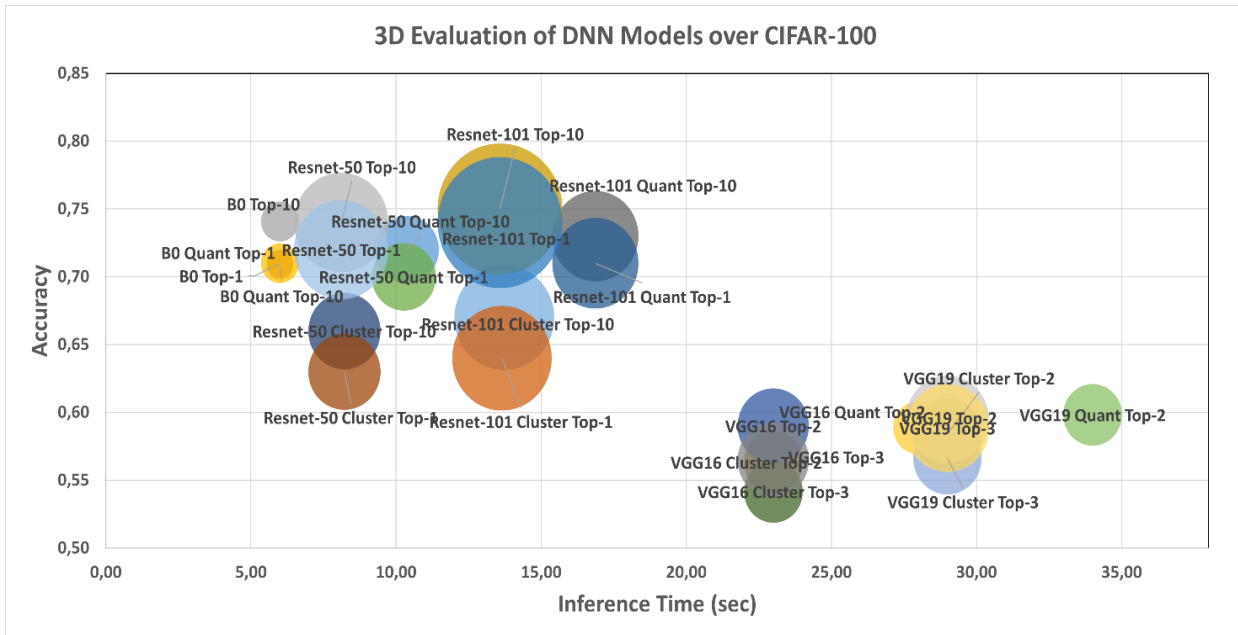
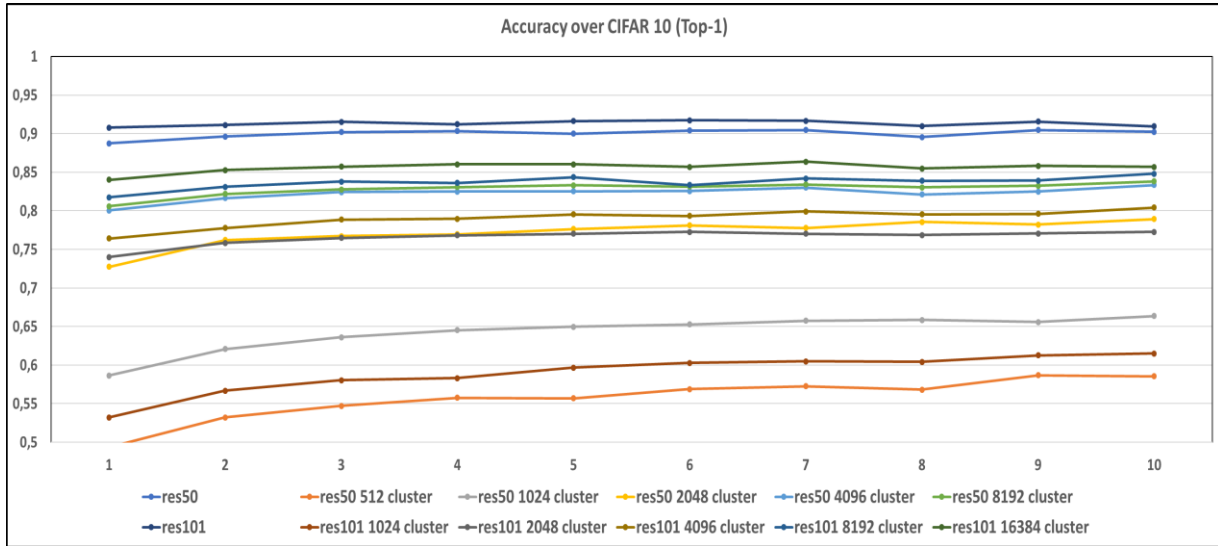


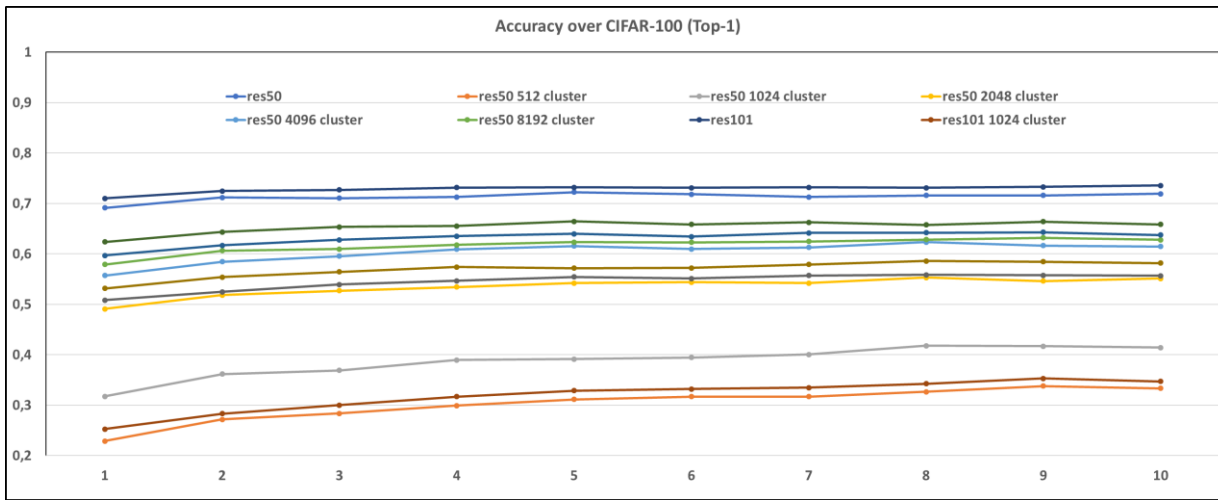
Figure 7. 3D Bubble chart comparison of selected DNN models with CIFAR100.

Table 6. Table 5. Ranking of evaluated models based on 3D KPI and unified OpScore using CIFAR100 dataset. MIN-MAX values are obtained by selecting the best and worst value for each column and generating a score for this fictitious model as reference; this can be used for normalization. Note that best and worst OpScore values range between 0-5.

Model Name	Accuracy	Size (MB)	Training Time (sec)	Inference Time (sec)	Score
MIN	0,75	8,10	27,28	6	1,08
B0 Quant Top-1	0,71	8,10	27,77	6,00	1,15
B0 Quant Top-10	0,71	8,10	30,94	6,00	1,15
B0 Top-10	0,74	16,60	29,63	6,00	1,41
B0 Top-1	0,71	16,60	27,36	6,00	1,46
B0 Cluster Top-10	0,12	11,00	29,59	6,00	1,76
B0 Cluster Top-1	0,05	11,00	27,28	6,00	1,80
Resnet-50 Quant Top-10	0,72	48,00	56,63	10,37	2,14
Resnet-50 Quant Top-1	0,70	48,00	51,05	10,27	2,17
Resnet-50 Cluster Top-10	0,66	61,00	44,61	8,22	2,23
Resnet-50 Cluster Top-1	0,63	61,00	40,50	8,22	2,27
Resnet-50 Top-10	0,74	103,00	44,47	8,11	2,34
VGG16 Quant Top-2	0,59	29,50	136,27	28,00	2,53
Resnet-101 Quant Top-10	0,73	86,00	88,43	16,87	2,59
VGG16 Cluster Top-2	0,56	39,90	113,84	23,00	2,60
Resnet-101 Quant Top-1	0,71	86,00	84,04	16,87	2,62
VGG16 Cluster Top-3	0,54	39,90	117,50	23,00	2,62
Resnet-101 Cluster Top-10	0,67	115,00	71,25	13,73	2,72
VGG19 Quant Top-2	0,60	40,1	167,37	34	2,74
VGG16 Top-2	0,59	58,90	113,92	23,00	2,74
Resnet-101 Cluster Top-1	0,64	115,00	67,50	13,65	2,75
VGG16 Top-3	0,56	58,90	117,63	23,00	2,77
Resnet-101 Top-10	0,75	180,00	70,18	13,58	2,79
Resnet-101 Top-1	0,74	180,00	67,3	13,58	2,80
VGG19 Cluster Top-2	0,58	54,3	141,1	29	2,82
VGG19 Cluster Top-3	0,57	54,3	144,5	29	2,83
VGG16 Quant Top-3	0,12	29,50	139,65	28,00	2,86
VGG19 Top-2	0,59	80,20	141,24	29,00	2,97
VGG19 Top-3	0,59	80,20	144,92	29,00	2,98
Resnet-50 Top-1	0,72	103,00	40,06	8,11	3,06
VGG19 Quant Top-3	0,07	40,1	171	34	3,10
MAX	0,07	180,00	167,37	29	3,69



(a) Sensitivity of ResNet accuracy to number of weight clusters over CIFAR10 dataset.



(b) Sensitivity of ResNet accuracy to number of weight clusters over CIFAR100 dataset.

Figure 8. Accuracy comparison of clustered Resnet50-101 pre-trained models over different cluster counts.

Table 7: Sensitivity of Resnet50-101 models per number of clusters over size (MB) and time (sec).

Model Name - Cluster#	Size (MB)	Clustering Time (sec)
Res50 – Original	103	-
Resnet50 -512	39	398
Resnet50 -1024	42	561
Resnet50 -2048	47	911
Resnet50 -4096	55	1556
Resnet50 -8192	61	2894
Res101 – Original	180	-
Resnet101 -512	76	1014
Resnet101 -1024	82	1560
Resnet101 -2048	93	2692
Resnet101 -4096	115	4808
Resnet101 -8192	130	7800

### 5. RELATED WORK

While increasing the depth in DNN, generally increases the accuracy of image classification, this accuracy gets saturated and then even drops as more layers are added. He *et al.* [17] proposed to reformulate plain or stacked layers as residual functions leading to deeper networks without increasing the costs of training. These networks were called Residual Networks or *ResNet* in short. The state-of-the-art DNN including ResNets [17], DenseNet [30], SqueezeNet [31] and EfficientNet [19,20] addressed models’ size, accuracy and inference time balance, but the outcomes were not operationalized in real-time field applications, because of resource limitations at the edge and the need for continuous architectural adaptation. Methods that only consider improvements in inference time, assume that model training or tuning would be done once at the beginning or infrequently. However, modern

distributed [32] or decentralized scenarios teach us that frequent model updates may be necessary. Our work quantifies the evaluation and ranking of DNN's for this purpose.

DeepCPU [33] developed by Microsoft also focuses on operational models, but only addresses model serving time reduction. They denote, users can tolerate long training times since this step is offline but improving serving time is crucial since this makes the biggest difference between *shippable vs. non-shippable* model. Models need to conform to Service-Level Agreements (SLA) in real-time and thus must make fast and accurate predictions. DeepCPU proposes cache-aware partitioning to optimize L2-L3 data movements and weight-centric streamlining. Specifically, they demonstrate 10x-60x speedups in serving time of different NLP models. There are other researchers that address model performance issues at the hardware level. HardNet [34] aimed to reduce the traffic to DRAM by reducing the number of accesses to feature maps in CNN. This operation reduced the number of floating-point operations (flops) and inference latency.

In this paper, we evaluated the performance benefits and potential hardware dependencies of quantization. While our work is orthogonal to scalar and vector-based quantization [35], it is still useful to summarize related work in this group. Fixed scalar quantization of weights stored in floating-point variables can reduce performance as the amount of information is reduced. Researchers analyzed resilience of DNN under quantization [36] and effects of retraining models to alleviate this problem. Zhang *et al.* [15] developed a strategy called Learned Quantization Networks (LQ-Nets) that can change the fixed point and adapt quantization level to balance the tradeoff between size and accuracy. Han, *et al.* [16] reduced model sizes further via deep compression while also trying to address power issues. They designed a three-stage pipeline of model pruning, trained quantization, and Huffman coding to reduce the storage requirement of NN without affecting their accuracy.

Nath, *et al.* [37] claim that most of these methods suffer from two strict requirements, which hinder their operational performance: (1) they require special hardware to be effective or (2) they have to modify their model structures and/or weights via pruning [38] and retraining which is quite costly to handle during operation. In their paper titled "better together" [37], they proposed "adjoint networks" where a large DNN continuously acts as a teacher to a smaller DNN. Their approach is orthogonal to quantization and neural architectural search [32] methods. A similar approach by Shen, *et al.* [39] called MealV2 uses an ensemble of teachers for "knowledge distillation" coupled with a good initialization of the student model. Our model-to-model layer transfer method was inspired by these approaches, but we wanted to further minimize the cost of retraining by quick layer transfers and stitching before retraining instead of weight initialization. We could resemble our new approach to creation of a "Frankenstein"

model, which borrows layers of the architecture model different related models.

Other orthogonal but related work include model generalization issues [40,41] and DNN performance surveys [42,43]. Recht, *et al.* [40] performed a longitudinal study for model generalization or robustness using ImageNet classifiers using CIFAR10 and ImageNet datasets. After creating a new labelled dataset called *CIFAR 10.1*, they found that (1) there was a significant performance drop in all models developed during 2013-2018 (a very active 5 years for ML research), (2) drops were due to models' inability to generalize to slightly harder images and not due to model adaptability to new data (*i.e.* overfitting to old datasets). Zhang, *et al.* proposed adversarial auto-augmentation [41] which is a data augmentation technique to help target NN learn hard features and improve model generalization

Li, *et al.* [43] gave a comprehensive survey of deep learning (DL) compilers. They listed properties of DL frameworks such as TensorFlow, PyTorch, Keras, *etc.* and DL accelerator hardware (by Google, Intel, Amazon). It is a useful survey for understanding general OpML issues.

## 6. CONCLUSION AND FUTURE WORK

We described a framework for multi-dimensional tradeoff assessment of DNN model performances based on accuracy, time, and size. Comparison of modern quantization and weight sharing (clustering) techniques revealed that quantization can provide size savings without loss of accuracy, yet it needs GPU support for mixed-precision float operations for achieving acceptable train & test times. If this can be provided, then its operational efficiency and viability is high. Yet, these high-end GPUs are currently not feasible for use in edge devices, especially at large-scale. NN weight sharing via clustering can save sizes by up to 60%, but its accuracy is sensitive to number of clusters used. While higher cluster counts give better accuracy, they eventually lose benefit in size and time dimensions. Finally, direct layer transfers among models requires careful selection. Otherwise, their accuracy can be lower than the original models.

In the future, we plan to extend our sensitivity analysis to different models and model parameters. Specifically, we are currently investigating the effect of decentralized, non-IID and unbalanced datasets in distributed federated learning settings [44,45,46]. Our work proposes a 3D evaluation scheme for these and other comparable techniques. We hope that DNN, AI, ML, OpML researchers can repeat these assessments over different DNN models and datasets.

## REFERENCES

- [1] F. Fabio, G. Lami, A. M. Costanza. "Deep learning in automotive software." *IEEE Software* 34(3), 56-63, 2017.

- [2] J. Villalba-Diez, D. Schmidt, R. Gevers, J. Ordieres-Meré, M. Buchwitz, W. Wellbrock, "Deep learning for industrial computer vision quality control in the printing Industry 4.0". *Sensors*, 19(18), 3987, 2019.
- [3] Z. Hu, Y. Zhao, M. Khushi. "A survey of forex and stock price prediction using deep learning." *Applied System Innovation* 4(1), 9, 2021.
- [4] J. Kim, Y. Shin, E. Choi. "An intrusion detection model based on a convolutional neural network." *Journal of Multimedia Information System* 6(4), 165-172, 2019.
- [5] Deng, L., Liu, Y. (Eds.). **Deep learning in natural language processing**. Springer, 2018.
- [6] Bashar, A. "Survey on evolving deep learning neural network architectures". *Journal of Artificial Intelligence*, 1(02), 73-82, 2019.
- [7] A. Collette, **Python and HDF5**, O'Reilly Media, Inc., November ISBN: 9781449367831, 2013.
- [8] Internet: Open Neural Network Exchange (ONNX), The open standard for machine learning interoperability, <https://onnx.ai>, 24.10.202.
- [9] Internet: Model Zoo, Open source deep learning code and pretrained models. <https://modelzoo.co>, 24.10.202.
- [10] Internet: Google Colaboratory, <https://colab.research.google.com/>, 24.10.202
- [11] M. Abadi, P. Barham, J.Chen, Z. Chen, A. Davis, J. Dean, M. Devin et al. "Tensorflow: A system for large-scale machine learning." **In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI)**, 265-283. 2016.
- [12] Yu, R., Li, P. "Toward resource-efficient federated learning in mobile edge computing". *IEEE Network*, 35(1), 148-155, 2021.
- [13] M. , Paulius, S. Narang, J. Alben, G. Diamos, E. Elsen, D. Garcia, B. Ginsburg, et al. "Mixed precision training." *arXiv preprint arXiv:1710.03740* (2017).
- [14] D. Lin, T. Sachin Talathi, A. Sreekanth "Fixed point quantization of deep convolutional networks." **In International Conference on Machine Learning, PMLR**, 2849-2858, 2016.
- [15] D. Zhang, J. Yang, D. Ye., G. Hua, "LQ-Nets: Learned quantization for highly accurate and compact deep neural networks". **In Proceedings of the European Conference on Computer Vision (ECCV)** 365-382, 2018.
- [16] S. Han, H. Mao, W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding". *arXiv preprint arXiv:1510.00149*, 2015.
- [17] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition". **In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)**, 770-778, 2016.
- [18] Simonyan, K., Zisserman, A. "Very deep convolutional networks for large-scale image recognition". *arXiv preprint arXiv:1409.1556*, 2014.
- [19] Internet: Keras Applications, <https://keras.io/api/applications>, 24.10.202
- [20] M. Tan, Q. Le. "EfficientNet: Rethinking model scaling for convolutional neural networks." **In International Conference on Machine Learning (ICML)**, 6105-6114. PMLR, 2019.
- [21] Tan, M., Le, Q. "EfficientNetv2: Smaller models and faster training". **In International Conference on Machine Learning**, 10096-10106, PMLR, 2021
- [22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, et al. "Scikit-learn: Machine learning in Python." *The Journal of Machine Learning Research* V.12, 2825-2830, 2011.
- [23] He, Y., Shen, Z., Cui, P. "Towards non-IID image classification: A dataset and baselines". *Pattern Recognition*, 110, 107383, 2021.
- [24] M.B. Çamlı, I. Ari, "Sensitivity Analysis of Federated Learning over Decentralized Data and Communication Rounds", **7. Ulusal Yüksek Basarimli Hesaplama Konferansı (BAŞARIM)**, no.14, 2022
- [25] Li, H., Kadav, A., Durdanovic, I., Samet, H., Graf, H. P. "Pruning filters for efficient ConvNets". *arXiv preprint arXiv:1608.08710*, 2016.
- [26] Glorot, X., Bengio, Y. "Understanding the difficulty of training deep feedforward neural networks". **In Proceedings of the 13th international conference on artificial intelligence and statistics**, 249-256, JMLR Workshop and Conference Proceedings, 2010.
- [27] Luo, J., Wu, X., Luo, Y., Huang, A., Huang, Y., Liu, Y., Yang, Q. "Real-world image datasets for federated learning". *arXiv preprint arXiv:1910.11089*, 2019
- [28] Zhu, H., Xu, J., Liu, S., Jin, Y. "Federated Learning on Non-IID Data: A Survey". *arXiv preprint arXiv:2106.06843*, 2021.
- [29] Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., Chandra, V. "Federated learning with non-IID data". *arXiv preprint arXiv:1806.00582*, 2018.
- [30] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, K. Keutzer, "DenseNet: Implementing efficient convnet descriptor pyramids" *arXiv preprint arXiv:1404.1869*, 2014.
- [31] F. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size." *arXiv preprint arXiv:1602.07360*, 2016.
- [32] H.R. Roth , D. Yang, W. Li, A. Myronenko, W. Zhu, Z. Xu, X. Wang, D. Xu, "Technique to perform neural network architecture search with federated learning", WO/2021-247338A1, WIPO Patent, 2021.
- [33] M. Zhang, S. Rajbhandari, W. Wang, Y. He, "DeepCPU: Serving RNN-based deep learning models 10x faster". **In 2018 USENIX Annual Technical Conference (ATC)**, 951-965, 2018.
- [34] P. Chao, C. Kao, Y. Ruan, C. Huang, Y. Lin. "Hardnet: A low memory traffic network" **In Proceedings of the IEEE/CVF International Conference on Computer Vision**, 3552-3561. 2019.
- [35] Sung, W., Shin, S., Hwang, K. "Resiliency of deep neural networks under quantization". *arXiv preprint arXiv:1511.06488*, 2015.
- [36] Gong, Y., Liu, L., Yang, M., Bourdev, L. "Compressing deep convolutional networks using vector quantization". *arXiv preprint arXiv:1412.6115*, 2014.



- [37] Nath, U., Kushagra, S. "Better Together: Resnet-50 accuracy with 13x fewer parameters and at 3x speed". *arXiv preprint arXiv:2006.05624*, 2020.
- [38] Ma, X., Yuan, G., Lin, S., Li, Z., Sun, H., Wang, Y. "ResNet can be pruned 60×: Introducing network purification and unused path removal (p-rm) after weight pruning". In **2019 IEEE/ACM International Symposium on Nanoscale Architectures**, 1-2, IEEE, 2019.
- [39] Shen, Z., Savvides, M. "Meal v2: Boosting vanilla Resnet-50 to 80%+ Top-1 accuracy on Imagenet without tricks". *arXiv preprint arXiv:2009.08453*, 2020.
- [40] B. Recht, R. Roelofs, L. Schmidt, L., V. Shankar, "Do Imagenet classifiers generalize to Imagenet?" In **International Conference on Machine Learning**, 5389-5400, PMLR, 2019.
- [41] Zhang, X., Wang, Q., Zhang, J., Zhong, Z. "Adversarial autoaugment". *arXiv preprint arXiv:1912.11188*, 2019.
- [42] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu, "A survey on deep transfer learning". In **International conference on artificial neural networks** 270-279, Springer, 2018.
- [43] M. Li, Y. Liu, X. Liu, Q. Sun, X. You, H. Yang, Z. Luan, L. Gan, G. Yang, D. Qian. "The deep learning compiler: A comprehensive survey." *IEEE Transactions on Parallel and Distributed Systems*, Vol. 32, No. 3, 708-727, 2020.
- [44] McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B. A. "Communication-efficient learning of deep networks from decentralized data". In *Artificial intelligence and statistics*, 1273-1282, PMLR, 2017.
- [45] T. Li, A. K. Sahu, A. Talwalkar, V. Smith, V. "Federated learning: Challenges, methods, and future directions". *IEEE Signal Processing Magazine*, 37(3), 50-60, 2020.
- [46] Maraş, A. Erol, Ç. "Emerging Trends in Classification with Imbalanced Datasets: A Bibliometric Analysis of Progression". *Bilişim Teknolojileri Dergisi*, 15(3), 275-288, 2022.

# COVID-19 Salgını Sürecinde Duygu Skorlarının ve Teknik İndikatörlerin Kullanılmasıyla BİST 100 Endeksi Trend Tahmini

*Araştırma Makalesi/Research Article*

 Meltem ALACA,  Aysun GÜRAN\*

Bilgisayar Mühendisliği Bölümü, Doğuş Üniversitesi, İstanbul, Türkiye

[20182105003@dogus.edu.tr](mailto:20182105003@dogus.edu.tr), [adogrusoz@dogus.edu.tr](mailto:adogrusoz@dogus.edu.tr)

(Geliş/Received:19.04.2022; Kabul/Accepted:19.08.2022)

DOI: 10.17671/gazibtd.1106017

**Özet**— Duygu analizi, metinsel verilerin yansıttığı olumlu, olumsuz ya da nötr duyguların tespit edilme aşamalarını içermektedir. Farklı konular hakkında yayınlanan metinsel verilerin yansıttığı olumlu ve/veya olumsuz duygular, insanların, küçük veya büyük ölçekli şirketlerin karar alma süreçlerini etkileyebilmektedir. Metinsel kaynakların yansıttığı duygular, duygu skorları ile sayısal hale getirilebilmekte ve bu skorlar zaman serilerinin ileriye yönelik kestirimlerinin yapılması aşamasında kullanılabilirler. Bilindiği gibi 1 Aralık 2019 tarihinde Çin'in Vuhan kentinde ortaya çıkan ve tüm dünyayı etkisi altına alan koronavirüs küresel bir pandemiye sebep olmuştur. Bu durum uluslararası ve yerel sermaye piyasaları üzerinde keskin düşümlere neden olmuştur. Bu çalışmada pandemi döneminde toplanan haber metinlerinden elde edilen duygu skorlarının Bist100 endeksi trendinin belirlenmesinde etkili olup olmadığının analizi yapılmıştır. Analizlerde borsa endeksleri kestirimlerinde büyük önem arz eden teknik indikatörlerden de yararlanılmıştır. Böylece pandemi döneminde borsa endeksinin artma veya azalma yönünün belirlenmesinde duygu skorlarının ve teknik indikatörlerin etkisi görülebilmektedir. Yapılan analizler sonucunda, haber metinlerinden elde edilen duygu skorlarının borsa endeksi trend tahminlemesi üzerinde bazı periyotlar için etkili olduğu gözlemlenmiştir.

**Anahtar Kelimeler**— duygu skoru, borsa teknik indikatörleri, covid-19, Bist100, duygu analizi

## Predicting BIST 100 Index Movement by using Sentiment Scores and Technical Indicators during the COVID-19 Pandemic

**Abstract**— Sentiment analysis includes the stages of identifying the positive, negative or neutral emotions contained in the text data. Positive and/or negative emotions reflected by the text data can affect the decision-making processes of people, small or large-scale companies. Emotions reflected by documents can be vectorized with sentiment scores and these scores could be useful to forecast time series models. As it is known, the coronavirus, which emerged in Wuhan, China on December 1, 2019 caused a global pandemic all around the world. The Covid-19 pandemic has brought sharp and sudden declines in global and domestic stock markets. Within the scope of our study, it was analyzed whether the sentiment scores obtained from the Covid-19 related news documents were effective in forecasting the trend of the Bist100 index. Technical indicators that have great importance in estimating stock market indices, were also used in the analysis. Hence, the effect of sentiment scores and technical indicators in determining the trend of the stock market index during the pandemic could be observed. As a result of the study, it was observed that the sentiment scores were effective to predict the price trend of stock market index for some periods.

**Keywords**— sentiment scores, technical indicators, covid-19, Bist100, sentiment analysis

## 1. GİRİŞ (INTRODUCTION)

Sermaye piyasalarında işlem gören endekslerin getiri yönlerinin tahmin edilebilmesi daha etkili yatırım stratejilerinin geliştirilebilmesine olanak sağlamaktadır. Fakat dinamik yapısından ötürü borsa endekslerinin değişim yönünün pek çok faktöre bağlı olması, endekslerin getiri değerlerinin ya da hareket yönlerinin tahminini zorlaştırmaktadır. Beklenmedik bir şekilde ortaya çıkan olaylar borsa endekslerinde dalgalanmalar yaratabilecek faktörler arasındadır [1]. Bilindiği gibi 1 Aralık 2019 tarihinde Çin'in Wuhan kentinde ortaya çıkan ve tüm dünyayı etkisi altına alan koronavirüs küresel bir pandemiye neden olmuştur. Pandemi nedeniyle ortaya çıkan belirsiz ortam ulusal ve uluslararası mali piyasalar üzerindeki dalgalanmaları arttırmıştır. ABD, İngiltere, İspanya ve İtalya gibi gelişmiş ekonomiler bu ani değişimden en çok etkilenen ülkeler arasına girmiştir [2]. ABD borsalarında işlem gören Dow Jones endeksi, 2020 yılı Mart ayında %12,94 düşüşle 1987'de yaşanan kara pazartesiden bu yana en büyük kaybı yaşamıştır. Ülkemizde de 11 Mart 2020 tarihinde görülen ilk vaka ile iktisadi faaliyetler sınırlanmaya başlamıştır. Mart ayı boyunca Borsa İstanbul'da hacmi en yüksek 100 adet hisse senedini kapsayan borsa endeksinde (BİST-100) %16,75 oranında değer kaybı yaşanmıştır. Pandemi koşullarının Borsa İstanbul'da yer alan sektörel endeksler üzerindeki etkisini inceleyen bir çalışma, salgın ile ilgili olaylar karşısında duyarlılığı en yüksek olan sektörlerin teknoloji, ulaştırma, turizm, KOBİ ve bilişim sektörleri olduğunu belirtmiştir [1]. Çalışma [3], salgın döneminde ulaştırma ve turizm sektörlerinin de olumsuz etkilendiğini ve bu sektörlerin hisse senedi getirilerinin düştüğünü ifade etmiştir.

Etkin piyasa hipotezi Farma tarafından 1970 yılında geliştirilen ve finans alanında en çok araştırılan hipotezlerden biri olmuştur [4]. Bu hipoteze göre yatırımcılar rasyoneldir. Piyasalardaki fiyatlar mevcut bilgiyi tam olarak yansıtır. Öte yandan davranışsal finans yaklaşımına göre menkul kıymetlerin fiyatlarının piyasa aktörlerince "algılanan" fiyatı yansıtmaktadır [5]. Yani yatırımcılar her zaman rasyonel olamayabilir. Bu iki yaklaşımın yakınlaştırılmasıyla ortaya çıkan Adaptif Piyasa Hipotezi'ne göre piyasaların etkinliği olaylar ve yapısal değişiklikler nedeniyle değişebilmektedir [6]. Bu yaklaşımlar doğrultusunda, son zamanlarda beklenmedik bir şekilde gelişen olayların, finansal haber dokümanlarının/makalelerin, sosyal medya platformlarında paylaşılan kısa iletilerin yansıtmış olduğu duygu durumları ile sermaye piyasaları arasındaki ilişki analizi önemli araştırma konuları arasına girmiştir.

Çetin çalışmasında [7], beklenmedik bir şekilde ortaya çıkan pandemi döneminde salgın oranının yüksek olduğu bir dönemde, Borsa İstanbul BİST-100 endeksinin kapanış, açılış, en yüksek ve en düşük fiyatları ile COVID-19 vaka sayılarının pozitif yönde ilişkili olduğunu belirtmiştir. Atan ve Çınar çalışmasında [8], finansal piyasalar ile ilgili haber metinlerinin yansıttığı olumlu, olumsuz duygu değerleri ile finansal değerler arasında anlamlı bir korelasyonun

olduğunu vurgulamışlardır. Ateş ve Güran çalışmaları ile [9], sosyal medyada paylaşılan kısa iletiler ile BİST30 endeksi arasında anlamlı korelasyon ilişkileri olduğunu belirtmiş ve ülkemiz adına önemli olayların olduğu bir dönemde duygu değerleri ve endeks arasında Granger nedensellik ilişkine rastlandığını ifade etmişlerdir. Pek çok çalışmanın deneysel olarak elde etmiş olduğu bulgular neticesinde günümüzde haber kaynaklarının beklenmedik olaylar ile ilgili yansıttığı olumlu ve/veya olumsuz algıların finansal piyasaları etkileyebileceği görülmektedir [10-15]. Bu çalışmalarda, olumlu ve olumsuz haber sayılarının kullanılması ile elde edilen duygu değerlerinin, gerek borsa endekslerinin gelecekteki değerlerinin ön görülmesinde; gerekse endekslerin pozitif/negatif yönlü hareketlerinin tahmin edilmesinde etkili olduğu belirtilmiştir.

Duygu değerleri haber metinlerinde yansıtılan olumlu, olumsuz genel kanının sayısallaştırılmasını sağlayan bir ölçümdür. Duygu değerlerinin endeksler üzerindeki etkinliğini araştıran çalışmaların çoğu yabancı sermaye piyasaları hakkındadır. Bu çalışmada 20.12.2019 – 16.02.2021 tarihleri arasında koronavirüs hakkında yayınlanan haber dokümanları VADER kütüphanesi yardımıyla olumlu, olumsuz ve nötr olarak etiketlendikten sonra üç farklı duygu değeri elde edilmiştir. Duygu değerleri elde edildikten sonra aynı tarih aralığındaki BİST100 endeksleri ele alınmış ve endeks değerlerinin günlük bazda artış veya azalış durumları, değer artışı olursa +1, tersi durumda 0 etiketleri ile belirtilerek analiz edilen durum bir sınıflandırma problemine dönüştürülmüştür. Sınıflandırma probleminde endeks trend tahmini için haber metinlerinden elde edilen duygu değerlerinin yanı sıra en sık kullanılan indikatörler arasında yer alan sekiz farklı adet teknik indikatör de kullanılmıştır. Teknik indikatörler yatırım araçlarının alımı/satımı konusunda karar verilmesine veya gelecekteki fiyatlarının/fiyat yönlerinin tahmin edilebilmesine katkı sunmaktadırlar. Belirtilen özelliklerin kullanımıyla farklı zaman periyotlarının da baz alınmasıyla sınıflandırma algoritması olarak Rassal Orman (RO) topluluk öğrenmesi algoritması kullanılmış ve elde edilen doğruluk değerleri paylaşılmıştır. Analizler sonucunda haber verilerinden elde edilen duygu değerlerinin ve borsa teknik indikatörlerinin trend tahminlemesine önemli ölçüde katkı sağladığı gözlemlenmiştir.

Çalışmamızın geri kalanı şu şekildedir. Bölüm 2 ile literatür araştırması ele alınmış; Bölüm 3 ile kullanılan veri kümesi tanıtılmış; Bölüm 4 ile sınıflandırma problemine ele alınan özellikleri ve RO algoritması sunulmuş; Bölüm 5 ile model parametreleri ve deneysel sonuçlar paylaşılmış ve son olarak sonuçlar başlığı elde edilen bulgular yorumlanmıştır.

## 2. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Bu bölümünde borsada işlem gören hisse senetleri getirilerinin tahmin edilmesine ya da getirilerin değişim yönünün tespit edilmesine yönelik yapılan çalışmalar incelenmiştir.

Pen ve ark. tarafından hazırlanan çalışmada hisse senedi fiyatlarının pozitif ya da negatif yönlü değişim durumu bir sınıflandırma problemine dönüştürülmüştür [10]. İlgili problemde derin ağlar kullanılmıştır. Algoritmaya girdi olarak geçmiş hisse senedi fiyatlarının yanı sıra haberlerden elde edilen BoK (Bag of Keywords), PS (Polarity Score), CT (Category Tag) değerleri de kullanılmıştır. Analizler sonucunda haberlerden elde edilen değerlerin tahminlemede başarımı yükselttiği görülmüş en düşük hata oranının %43.13'lük bir değer ile kapanış fiyatının yanı sıra BoK, PS ve CT değerlerinin beraber kullanılmasıyla elde edildiği görülmüştür.

Huynh ve ark. tarafından gerçekleştirilen analizde S&P 500 borsa endeksi üzerinde çalışılmıştır. Yapılan çalışmada finansal haberlerden elde edilen vektör değerleri ve geçmiş hisse senedi fiyatları bir araya getirilerek hisse senedi fiyatları yön tahminlemesi yapılmıştır [15]. Uzun kısa süreli bellek (LSTM), kapılı tekrarlayan hücre (Gated Recurrent Unit) ve makale tarafından önerilen BGRU (Bidirectional Gated Recurrent Unit) modelleri ilgili veriler üzerinde uygulanmıştır. Elde edilen sonuçlara göre BGRU modelinin diğer modellerden daha yüksek başarımlar ürettiği belirtilmiştir.

Bouktif ve ark. çalışmasında AMZN NASDAQ üzerinde hisse senedi artma azalma yönü tahmin edilmiştir [11]. Yapılan çalışmada DVM, lojistik regresyon (LR), ileri beslemeli sinir ağları (Feed Forward Neural Networks), RO ve ekstrem gradyan artırma (Extreme Gradient Boosting) modelleri kullanılmıştır. Algoritmaya girdi olarak ise hisse senedi fiyatlarının açılış, kapanış gibi özelliklerinin yanı sıra Twitter'dan ilgili hisse senedi ile toplanmış kısa iletilerden elde edilen duygu polarite özellikleri, duygu skorları, N-gramlar ve gecikmeli özellik değerleri kullanılmıştır. Elde edilen sonuçlarda en iyi sonucun yaklaşık %63 doğruluk oranıyla RO algoritması ile alındığı görülmüştür.

Mohan ve ark. çalışmasında S&P500 endeksine ait beş yıllık günlük endeks değerleri ve şirketler hakkındaki 265000 adet finansal makaleden elde edilen pozitif ve negatif polarite değerleri kullanılarak endeks tahminlemesi yapılmıştır [13]. Çalışmada, pozitif ve negatif polarite değerlerinin modellere katıldığı durumun, sadece hisse senedi fiyatlarının değerlendirildiği durumdan daha iyi sonuçlar verdiği belirtilmiştir. Model olarak Facebook Prophet, ARIMA ve tekrarlayan sinir ağları (Recurrent Neural Networks) kullanılmıştır. Sonuç olarak duygu değerleri ile kapanış fiyatlarının yönü arasında korelasyon olduğu görülmüştür.

Mehtap ve Şen çalışmasında, ulusal Hindistan borsasının NIFTY 50 endeksi üzerinde çalışılmıştır [19]. 2015 – 2018 tarihleri arasındaki günlük borsa verileri Yahoo finance web sitesi üzerinden elde edilmiştir. Kapanış fiyatı tahminlemesi ve trend tahminlemesi için sekiz adet regresyon, sekiz adet sınıflandırma modeli kullanılmıştır. Modellere girdi olarak açılış, kapanış gibi elde edilen borsa verileri kullanılmıştır. Sınıflandırma modelleri içerisinde en başarılı sonuçları %74.91 hassaslık oranı ile RO

algoritmasının, regresyon modelleri içerisinde ise uygulamalı evrişimsel sinir ağı (Applied Convolutional Neural Network) algoritmasının sergilediği görülmüştür.

Picasso ve ark. tarafından sunulan çalışmada NASDAQ100 endeksindeki en değerli 20 şirketin borsa verileri üzerinde çalışılmıştır. Çalışmanın amacı piyasa eğilimini tahminlemektir [12]. Piyasa eğilimini tahminlemede algoritma özellik girdisi olarak Basit Hareketli Ortalama (Simple Moving Average), Üstel Hareketli Ortalama (Exponential Moving Average), Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi, Göreceli Güç Endeksi, Bollinger Bantları, Stokastik Osilatör, Gerçek Aralık (True Range), Ortalama Gerçek Aralık, Williams %R Göstergesi ve CR teknik indikatörlerinin yanında haberlerden elde edilen duygu değerleri de işleme katılmıştır. Model olarak RO, DVM ve YSA kullanılmıştır. En başarılı sonuçların yaklaşık %70 doğruluk oranı ile haberlerden elde edilen duygu değerleri ve teknik indikatörlerin bir arada kullanılmasıyla elde edildiği görülmüştür.

Li ve ark. tarafından gerçekleştirilen çalışmada Hong Kong borsasında beş yıllık bir veri kümesi üzerinde, borsa trend tahminlemesi için çalışılmıştır [14]. Özellik değeri olarak duygu skorları ve Hareketli Ortalama (Moving Average), Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi (Moving Average Convergence Divergence), Göreceli Güç Endeksi (Relative Strength Index) ve Para Akış Göstergesi (Money Flow Index) borsa teknik indikatörleri kullanılmıştır. Çalışmada uzun kısa süreli bellek (Long Short Term Memory), çoklu çekirdek öğrenme (Multiple Kernel Learning) ve SVM modelleri karşılaştırılmıştır. Yapılan çalışmada uzun kısa süreli bellek algoritmasının yaklaşık %83 doğruluk oranı ile daha başarılı bir başarımlar ürettiği belirtilmiştir.

Patel ve ark. tarafından sunulan çalışmada Hindistan borsa hisse senetlerinin negatif/pozitif yönlü hareketleri bir sınıflandırma problemine dönüştürülerek tahmin edilmeye çalışılmıştır [16]. Çalışmada eğitim veri kümesi olarak on yıllık borsa verisi analiz edilmiş ve yapay sinir ağları (YSA), destek vektör makineleri (DVM), RO ve Naive-Bayes (NB) sınıflandırma algoritmaları kullanılmıştır. Sınıflandırma problemine dikkate alınan özellik değerleri popüler borsa teknik indikatörleridir. Kullanılan teknik indikatörler, “Basit hareketli ortalamalar (Simple Moving Average)”, “Hareketli ortalamalar (Weighted Moving Average)”, “Stokastik K% (Stochastic K%)”, “Stokastik D% (Stochastic D%)”, “Momentum”, “Göreceli güç endeksi (Relative Strength Index)”, “William R% (William's R%)”, Toplam Dağıtım Osilatörü (Accumulation/Distribution Oscillator)”, “Mal Kanal Endeksi (Commodity Channel Index)” göstergeleridir. Elde edilen sonuçlarda RO algoritmasının %91.48 doğruluk değeriyle en iyi başarımlar değerine sahip olduğu belirtilmiştir.

Khaidem ve ark., Apple, Microsoft ve Samsung hisse senetleri üzerinde çalışmıştır [17]. Model olarak RO algoritması kullanılmıştır. Modele girdi olarak Göreceli

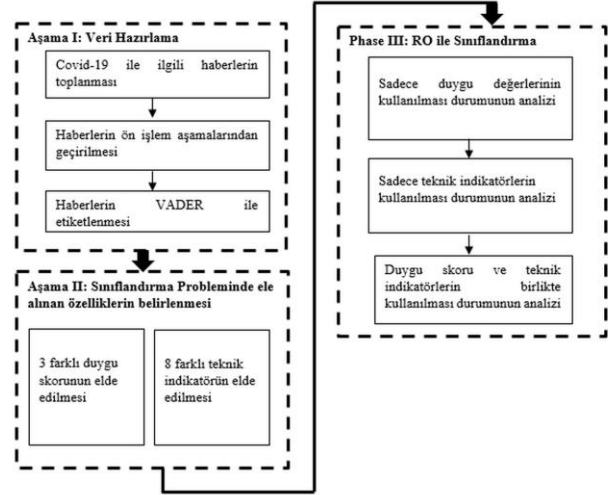
Güç Endeksi, Stokastik Osilatör, Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi, Williams %R Göstergesi, Fiyat Değişim Oranı (Price Rate of Change), Denge İşlem Miktarı (On Balance Volume) borsa teknik indikatörleri sunulmuştur. Hisse senetleri kapanış fiyatı artma azalma yönü tahminlemesi için alınan sonuçlarda RO algoritması ile %85 ve %95 aralığında başarımlar elde edilmiştir. RO algoritması ile daha önce kullanılan [18] referansı ile belirtilen DVM algoritmasından daha etkili başarımlara ulaşıldığı belirtilmiştir. Bununla beraber en yüksek başarımlar %92.08 ile 10 günlük kestirimler için Apple hisse senedi üzerinde elde edilmiştir.

Di., Apple, Amazon, Microsoft hisse senetleri üzerinde çalışmıştır. Sınıflandırma algoritması olarak destek vektör makinesi algoritması kullanılmıştır [18]. Sınıflandırma probleminde ilgili hisse senetlerinin artma azalma yönü tahminlemesi yapılmıştır. Sınıflandırma aşamasında borsa değerlerinin yanı sıra Momentum, Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi, Göreceli Güç Endeksi, Williams %R Göstergesi, Mal Kanalı Endeksi (CCI), Fiyat Değişim Oranı, Ortalama Yönel Gösterge (ADX), Üçlü Üstel Hareketli Ortalama, Denge İşlem Miktarı, Zaman Serisi Tahminlemesi, Ortalama Gerçek Aralık, Para Akış Göstergesi borsa teknik indikatörleri kullanılmıştır. Çalışmada teknik indikatörlerin trend tahminlemedeki başarılı sonuçları paylaşılmıştır. Belirtilen sonuçlara göre en yüksek doğruluk değerinin %77.125 ile Microsoft hisse senedi için 7 günlük kestirimlerde elde edildiği vurgulanmıştır.

Kullanımları 1933'lü yıllara uzanan teknik indikatörler hisse senetlerinin gelecekteki fiyat eğilimlerini belirlemek için geçmiş fiyatları, işlem hacimlerini ve diğer mevcut verileri kullanmaktadır [32]. Biz bu çalışmada özellikle duygu analizinin de katılarak gerçekleştirildiği hisse senedi eğilimlerini tahminleyen [12], [14], [16], [18] çalışmalarında kullanılan yüksek başarımlarına ulaşılmasını sağlayan "Ağırlıklı Hareketli Ortalama", "Göreceli Güç Endeksi", "Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi", "Williams %R Göstergesi", "Stokastik %K", "Stokastik %D", "Chaikin Toplama-Dağıtım Osilatörü", "Mal Kanalı Endeksi" göstergelerini kullandık. Bu teknik indikatörlerin yanına duygu skor değerlerini de ekleyerek Bist100 enkesinin trend tahminlemesi üzerine bir çalışma gerçekleştirdik. Böylece finansal metinlerin yansıttığı duygu durumlarının Türkiye piyasaları üzerinde etkili olabileceğini deneysel bir şekilde göstermiş olduk.

### 3. ÖNERİLEN METODOLOJİ (RECOMMENDED METHODOLOGY)

Bu çalışmada Şekil 1 ile belirtilen aşamalar takip edilmiştir. Bundan sonraki alt bölümlerde her bir aşama detaylıca anlatılmıştır.

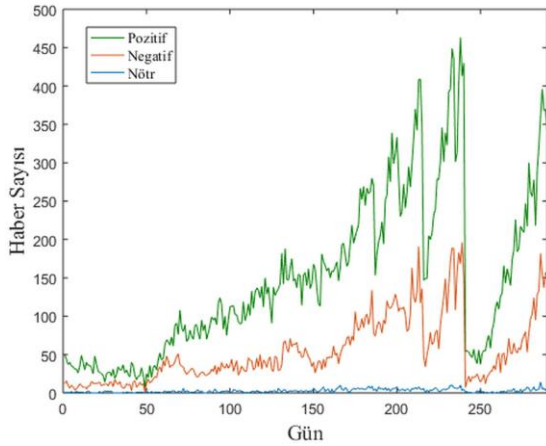


Şekil 1. Önerilen Metod Çerçevesi (Suggested Method Framework)

#### 3.1. Kullanılan Veri Kümesi (Dataset)

Bu çalışmada 20.12.2019-16.02.2021 tarihleri arasında farklı haber kaynakları tarafından Covid-19 ile ilgili olarak yayınlanmış olan ("korona", "koronavirüs", "coronavirus", "corona", "kovid19", "covid19", "kovid-19", "covid-19" vb. kelimeleri içeren) haber metinleri Java programlama dili ile yazılan bir yazılım ile toplanmıştır. Veri kümemiz bu tarihlardaki 290 işlem gününde paylaşılan 59313 haber dokümanını içermektedir. Veri kümesi üzerinde duygu analizi aşamalarını uygulamadan önce haber metinleri durak kelimelerinden arındırılmıştır. Ön işlem aşamasının ardından duygu analizinin gerçekleştirilmesi için sözlük tabanlı bir duygu analizi aracı olan VADER yazılımı kullanılmıştır. Bu araç ile -4 ile +4 arasında bir ölçekte kelimelerin duygu yoğunluk toplamlarına göre olumlu, olumsuz ve nötr olarak etiketlenebilmektedir. VADER yazılımından faydalanmak için Türkçe haber metinleri Google çeviri ara yüzü (Google translate api) kullanılarak İngilizce'ye çevrilmiştir. 20.12.2019-16.02.2021 tarih aralığında toplam 290 işlem günü mevcuttur. Bu günlerde duygu analizi gerçekleştirilen toplam haber sayısı 59313'tür. VADER yazılımı ile bu haber dokümanlarından 42892 tanesi olumlu, 15621 tanesi olumsuz ve 800 tanesi nötr olarak etiketlenmiştir. 290 iş günü boyunca yayınlanan olumlu, olumsuz ve nötr haber sayılarının yayılımı Şekil 2 ile belirtilmiştir.

Günlük bazda yayınlanan olumlu ve olumsuz haber sayılarının kullanılmasıyla Bölüm 4.1'de belirtilen  $S_1$ ,  $S_2$  ve  $S_3$  olarak adlandırılan 3 farklı duygu değeri oluşturulmuş ve bu değerlerin her biri sınıflandırma probleminde 8 adet teknik indikatörün yanında ayrı ayrı birer özellik olarak kullanılmıştır.



Şekil 2. 290 iş gününde paylaşılan olumlu, olumsuz ve nötr haber dağılımı  
(Distribution of positive, negative and neutral news shared in 290 working days)

Duygu değerleri dışında özellik değeri olarak kullanılan sekiz adet teknik indikatör “ağırlıklı hareketli ortalama”, “göreceli güç indeksi”, “hareketli ortalama yaklaşma uzaklaşma göstergesi”, “Williams %R göstergesi”, “stokastik %K”, “stokastik %D”, “Chaikin toplama dağıtım Osilatörü”, “mal kanal indeksi” göstergeleridir.

290 işlem günü için Bist100 endeksi kapanış fiyatlarının belirli bir güne göre artma ya da azalma durumunu bir sınıflandırma problemi olarak değerlendirebilmek adına BİST-100 endeksi kapanış değerleri günlük olarak baz alınmıştır. 290 işlem günü için günlere ait sınıf etiketleri, endekste belirli periyotta göre artış meydana geldiğinde +1, azalış meydana geldiğinde ise 0 etiketlerinin yaratılması ile belirlenmiştir. Böylece sınıflandırma probleminde ele alınacak özellikler ve değişecek periyotlara göre belirlenen sınıf etiketleri elde edilebilmiştir.

### 3.2. Modelde Kullanılan Özellikler (Features Used in the Model)

Bu bölümde sınıflandırma problemi olarak ele aldığımız endeks eğilimini öngörme problemimizde kullanılacak olan özellikler (duygu değerleri ve sıklıkla kullanılan 8 adet teknik indikatör) anlatılacaktır. Bilindiği gibi teknik indikatörler, yatırım araçlarının geçmiş ve güncel fiyat ve/veya hacim faaliyetlerine bağlı olarak yatırım araçlarının alımı/satımı konusunda karar verilmesine veya gelecekteki fiyatlarının ve/veya fiyat yönlerinin tahmin edilebilmesine olanak sağlarlar. Çalışmamızdaki amacımız pandemi döneminde yayımlanan olumlu ya da olumsuz haber sayılarından elde edilen duygu skor değerlerinin ve popüler teknik indikatörlerin BİST-100 endeksi kapanış değerlerinin pozitif ya da negatif yönlü değişim durumlarının tahmini üzerindeki başarımlarını incelemektir. Çalışmamızda kullanılan duygu skorlarına ve 8 adet teknik indikatöre ait olan açıklamalar sırasıyla aşağıdaki alt başlıklarda verilmektedir:

#### 3.2.1. Duygu Skorları (Sentiment Scores)

Duygu skorlarının oluşturulması için [9] çalışması baz alınmış ve VADER yazılımı ile olumlu ve olumsuz olarak etiketlenen haber sayıları kullanılmıştır. Bu kapsamda  $S_1$ ,  $S_2$ ,  $S_3$  olmak üzere üç adet duygu skoru elde edilmiştir.

$$S_1 = \frac{Nolumlu_t - Nolumsuz_t}{Nolumlu_t + Nolumsuz_t} \quad (1)$$

$$S_2 = \frac{Nolumlu_{t+1}}{Nolumsuz_{t+1}} \quad (2)$$

$$S_3 = \frac{Nolumlu_t}{Nolumlu_t + Nolumsuz_t} - \frac{Nolumlu_{t-1}}{Nolumlu_{t-1} + Nolumsuz_{t-1}} \quad (3)$$

Bu denklemlerde  $Nolumlu_t$  ve  $Nolumsuz_t$  incelenen günlerde paylaşılan olumlu ve olumsuz haber sayılarını belirtmektedir.  $Nolumlu_{t-1}$ ,  $Nolumsuz_{t-1}$  ise (t-1). günde paylaşılan olumlu, olumsuz haber sayısını ifade etmektedir.

#### 3.2.2. Ağırlıklı Hareketli Ortalama (Weighted Moving Average – WMA)

Ağırlıklı ortalama, verilen periyotta sondaki kapanış fiyatlarına daha çok ağırlık verilerek hesaplanan ortalamadır. Yakın zamandaki fiyattan, geçmiş verilere doğru önemlilik doğrusal olarak azalmaktadır. Pek çok çalışmada kısa dönem tahminlemelerinde 10 günlük periyot kullanıldığı için bu çalışmada da 10 günlük periyot baz alınmıştır [16], [20]. Ağırlıklı hareketli ortalama mevcut trendi teyit etmektedir.  $C_n$ , endeksin n. günündeki kapanış fiyatı olmak üzere ağırlıklı hareketli ortalama formülü denklem 4 ile belirtildiği gibidir.

$$WMA = \frac{(1) \times C_1 + (2) \times C_2 + (3) \times C_3 + \dots + (n) \times C_n}{1+2+3+\dots+n} \quad (4)$$

#### 3.2.3. Göreceli Güç Endeksi (Relative Strength Index – RSI)

Bu indikatör endeksin bünyesi içindeki gücünü kendisine göre karşılaştırır [21]. Endeksin geçmişte yükseldiği günlerle, azaldığı günlerin karşılaştırılmasına yardımcı olur. RSI hesaplanırken genellikle kısa dönem analizleri için on dört günlük periyot kullanılmaktadır [22-24]. RSI, istenen periyotlardaki artış ve azalış yaşanan günleri ve bu günlere ait ortalama değerleri alarak bu değerlerin birbirlerine oranını bulmaktadır.  $C_i$  ilgili gündeki kapanış fiyatı; n periyot olmak üzere, RSI hesabı denklem 5-8 ile belirtilen ara aşamalara bağlı olarak denklem 9 ile belirtilmiştir [16-10]:

$$PD_i = \begin{cases} C_i - C_{i-1}, & C_i > C_{i-1} \\ 0, & C_i \leq C_{i-1} \end{cases} \quad (5)$$

$$ND_i = \begin{cases} 0, & C_i > C_{i-1} \\ |C_i - C_{i-1}|, & C_i \leq C_{i-1} \end{cases} \quad (6)$$

$$OPD_i = \frac{\sum_{i=1}^n PD_i}{n} \quad (7)$$

$$OND_i = \frac{\sum_{i=1}^n ND_i}{n} \quad (8)$$

$$RSI_i = 100 - \left[ \frac{100}{1 + \left( \frac{OPD_i}{OND_i} \right)} \right] \quad (9)$$

### 3.2.4. Hareketli Ortalama Yaklaşma Uzaklaşma Göstergesi (Moving Average Convergence Divergence – MACD)

Hareketli ortalama yaklaşma uzaklaşma göstergesi, belirlenen periyotlardaki üssel hareketli ortalamaların farkı alınarak elde edilir. Bu gösterge endekste oluşabilecek trendin düşüş mü yükseliş mi olduğunu anlamamıza yardımcı olur. Genellikle, on iki ve yirmi altı günlük periyotlar için üstel hareketli ortalama sonuçlarının çıkarılmasıyla bulunur [22], [25]. On iki gün, hızlı hareketli ortalama için seçilen periyot iken, yirmi altı gün yavaş hareketli ortalama için seçilmiştir. MACD serisi belirtilen üstel hareketli ortalamaların birbirinden çıkarılmasıyla elde edilmektedir. Bu gösterge için ayrıca MACD serisinin 9 dokuz günlük üstel hareketli ortalamasının alınmasıyla MACD sinyal değeri hesaplanmaktadır. EMA, üstel hareketli ortalama; n, periyot ve  $C_n$  ilgili periyottaki kapanış fiyatı olmak üzere gösterge formülü denklem 10 – denklem 12 ile belirtilmiştir.

$$EMA_n = \left( \frac{2}{n+1} C_n \right) + \left( \frac{n-1}{n+1} EMA_{n-1} \right) \quad (10)$$

$$MACD = EMA_{12} - EMA_{26} \quad (11)$$

$$Sinyal = S = EMA_9(MACD) \quad (12)$$

### 3.2.5. Williams %R Göstergesi (Williams %R)

Williams %R göstergesi, kapanış fiyatının saptanan periyot içerisindeki en düşük ve yüksek fiyatlara göre yerini tespit edilebilmektedir. Bu göstergede hesaplamalar genellikle 14 gün üstünden yapılmaktadır [18], [22], [26], [27].  $L_{n(t)}$ , son n periyodun en düşük değeri (t. günde görüldüğü varsayılınsın),  $H_{n(t)}$  son n periyodun en yüksek değeri (t. günde görüldüğü varsayılınsın) olmak üzere gösterge formülü denklem 13 ile belirtilmiştir.

$$Williams \%R = \frac{H_{n(t)} - \text{güncel kapanış fiyatı}}{H_{n(t)} - L_{n(t)}} \times 100 \quad (13)$$

### 3.2.6. Stokastik Osilatör (Stochastic Oscillator – STOCH)

Stokastik Osilatör, kısa dönem için işlem yapanların en çok kullandığı alım satım göstergelerindendir [22], [28]. Bu gösterge hesabı %K ve %D değerlerine sahiptir. [26], [17] çalışmalarında belirtildiği gibi %K için on dört günlük, %D için ise üç günlük periyot ile çalışılmıştır. Çalışmamız kapsamında uygulanan modelde bu indikatörden gelen %K

ve % D değerleri iki ayrı teknik indikatör olarak kullanılmıştır.

$L_{n(t)}$ , son n periyodun en düşük değeri (t. günde görüldüğü varsayılınsın) ve  $H_{n(t)}$  son n periyodun en yüksek değeri (t. günde görüldüğü varsayılınsın) olmak üzere göstergenin kapsadığı %K ve %D formülleri denklem 14 ve denklem 15 ile belirtilmiştir.

$$\%K = \frac{\text{güncel kapanış fiyatı} - L_{n(t)}}{H_{n(t)} - L_{n(t)}} \times 100 \quad (14)$$

%D eğrisi, yavaşlatılmış %K'nın 3 günlük hareketli ortalamasının alınmasıyla elde edilmektedir.

$$\%D = \frac{\sum_{t=1}^n \%K_t}{n} \quad (15)$$

### 3.2.7. Chaikin Toplama-Dağıtım Osilatörü (Chaikin Accumulation/Distribution Oscillator – ADOSC)

Chaikin toplama-dağıtım osilatörü fiyat ve işlem hacminin analiz edilmesiyle yönsel bilgi vermeyi amaçlar ve toplama dağıtım endeksinin üç güne ait üstel hareketli ortalaması ile on güne ait üstel hareketli ortalamasının farkı alınarak elde edilir [22], [29].

CO, Chaikin Osilatörü; ADL, toplama dağıtım çizgisi (Accumulation-Distribution Line);  $ADL(EMA_3)$ , ADL'nin 3 günlük üstel hareketli ortalaması;  $ADL(EMA_{10})$ , ADL'nin 10 günlük üstel hareketli ortalaması; MFM, para akışı çarpanı (Money Flow Multiplier); MFV, para akışı hacmi (Money Flow Volume); n periyot;  $V_n$ , n. gündeki hacim;  $L_n$ , n. gündeki düşük fiyat;  $H_n$ , n. gündeki yüksek fiyatı olmak üzere indikatör formülü denklem 16 – denklem 21 ile belirtilmiştir:

$$CO = ADL(EMA_3) - ADL(EMA_{10}) \quad (16)$$

$$ADL(EMA_n) = [ADL_n \times k] + [ADL(EMA_{n-1}) \times (1 - k)] \quad (17)$$

$$k = \frac{2}{n+1} \quad (18)$$

$$ADL = ADL_{n-1} + MFV_n \quad (19)$$

$$MFV_n = MFM \times V_n \quad (20)$$

$$MFM = \frac{(C_n - L_n) - (H_n - C_n)}{(H_n - L_n)} \quad (21)$$

### 3.2.8. Mal Kanal Endeksi (Commodity Channel Index – CCI)

Mal kanal endeksi, fiyattaki döngüsel değişimleri belirtir [18]. İlgili indikatör için kullanılan periyot [30] çalışmasında olduğu gibi on dört gün alınmıştır. Gösterge formülü denklem 22- 25 ile verilmektedir [22].

$$CCI_i = \frac{A_i - B_i}{D \times 0.015} \quad (22)$$

$$A_i = \frac{H_i + L_i + C_i}{3} \quad (23)$$

$$B_i = \frac{\sum_{i=1}^n \frac{H_{(i)} + L_{(i)} + C_{(i)}}{3}}{n} \quad (24)$$

$$D_i = \frac{\sum_{i=1}^n |A_i - B_i|}{n} \quad (25)$$

#### 4. MODEL PARAMETRELERİ VE DENEYSSEL SONUÇLAR (MODEL PARAMETERS AND EXPERIMENTAL RESULTS)

Çalışmamızda BİST100 endeksi kapanış değerlerinin artış/azalış yönü (trend) tahminlemesi için Jupyter Notebook geliştirme ortamında Python dili kullanılmıştır. Yararlanılan kütüphaneler numpy, pandas, scikit-learn, talib kütüphaneleridir. Endeksin artış/azalış yönü tahminlemesi için ilgili günün sınıf etiketleri, endeks değeri bir önceki periyoda göre arttığında 1 ve diğer durumda 0 değerine sahip olacak şekilde oluşturulmuştur. Böylece trend tahminlemesi problemi ikili sınıflandırma problemine dönüştürülebilmştir. Sınıflandırma probleminde ele alınacak özneliker haber metinlerinden elde edilen duygu değerleri ve 8 adet teknik indikatördür. Tüm özellikler min-max normalizasyona göre 0-1 aralığında değer alacak şekilde normalize edilmiştir. Özneliklerin de oluşturulmasıyla beraber modelimizde RO sınıflandırma algoritması kullanılmıştır. Sınıflandırma algoritması olarak RO'nun seçilme sebebi, pek çok çalışma tarafından bu alanda iyi başarımlar elde ettiğinin belirtilmesidir [11], [16], [17], [19]. RO algoritması, eğitim aşamasında çok sayıda karar ağacının oluşturulmasıyla çalışan bir topluluk öğrenmesi yöntemidir. RO'da karar ağaçlarının inşa edilmesinde de yeniden örnekleme metodu ve bölme kriteri olarak ise Gini safsızlık (Gini Impurity) yöntemi temel alınmaktadır [17]. Çalışmamızda RO algoritmasında ağaç sayısı olarak [17] çalışmada kullanıldığı gibi 5 ağaç kullanılmıştır. Algoritmada endeksin artma azalma yönü tahminlemesi için mevcut gün ile belirli periyottaki önceki gün karşılaştırmasında [31] çalışmasında olduğu gibi 20 günlük periyoda kadar tahminlemeler yapılmıştır. Ele alınan her periyot için eğitim veri kümesi %75-%25 olarak bölünmüş ve bu kümenin baz alınmasıyla algoritma beş kez çalıştırılmıştır. Son olarak elde edilen doğruluk (accuracy) ve F-skor değerlerinin standart sapma ve ortalama değerleri hesaplanarak Tablo 1 ve Tablo 2 ile sonuçlar paylaşılmıştır.

Bir sınıflandırma algoritmasının tahmin değerlerini ve örneklemlerin gerçek sınıf etiketleri hakkında bilgi veren aşağıda belirtilen doğruluk matrisindeki TP (True Positive), TN (True neagtive), FP (False Positive) ve FN (False Negative) değerleri kullanılarak doğruluk değeri ve F-skor değeri denklem 26 - denklem 29 ile belirtildiği gibi hesaplanmaktadır.

	Tahmini Durum		
	+1	-1	
Gerçek	+1	TP	FN
Durum	-1	FP	TN

Şekil 3. Karmaşıklık Matrisi (Confusion Matrix)

$$\text{Doğruluk değeri} = (TP+TN)/(TP+FN+FP+TN) \quad (26)$$

$$P = TP/(TP+FP) \quad (27)$$

$$R = TP/(TP+FN) \quad (28)$$

$$F\text{-Skor} = 2PR / (P+R) \quad (29)$$

Tablo 1, farklı özellik kümelerinin kullanılmasıyla gerçekleştirilen deneylerin doğruluk değerlerini göstermektedir.

Tablo 1'e ait ilk üç kolon (Doğruluk  $S_1$ , Doğruluk  $S_2$  ve Doğruluk  $S_3$ ) özellik değeri olarak sadece  $S_1$ ,  $S_2$ ,  $S_3$  duygu değerlerinin ayrı ayrı kullanılması ile elde edilen doğruluk değerlerini göstermektedir. Endeksin artış azalış yönünün belirlenmesini ele alan bu sınıflandırma probleminde sadece  $S_1$ ,  $S_2$ ,  $S_3$  duygu değerlerinin bireysel olarak kullanılması 20 farklı periyot için sırasıyla ortalama (ilk üç kolon ortalamaları) 0,5771; 0,5659 ve 0,5620 doğruluk değerlerini üretmiştir. Bu sonuçtan yola çıkarak  $S_1$  duygu değerinin daha iyi bir başarımlar değeri ürettiği söylenebilir.

Tablo 1'in "Doğruluk BTİÖK" isimli dördüncü kolonu 8 adet borsa teknik indikatör özellik kümesinin (BTİÖK) dikkate alınmasıyla elde edilen sonuçları paylaşmaktadır. Buradaki sonuçlardan da görüldüğü gibi, tüm periyotlar için borsa teknik indikatörlerinin kullanımı, sadece  $S_1$ ,  $S_2$ ,  $S_3$  duygu değerlerinin kullanımına göre daha yüksek başarımlar değerlerinin elde edilmesini sağlamıştır. Ayrıca bu kolonun ortalaması alındığında elde edilen 0,8038 değeri  $S_1$ ,  $S_2$ ,  $S_3$  duygu değerlerinin ortalama değerlerinden daha yüksek çıkmıştır.

Tablo 1'in beşinci (Doğruluk BTİÖK +  $S_1$ ) kolonu, BTİÖK ve  $S_1$ 'in aynı anda kullanılması ile elde edilen doğruluk değerlerini göstermektedir. Belirli periyotlar için (3,5,6,7,11,12,14,15,16,17,18,20) yani toplam 12 deneyde BTİÖK ve  $S_1$ 'in birlikte kullanılması BTİÖK ve  $S_1$ 'in bireysel kullanımlarına göre daha yüksek doğruluk değerlerinin üretilmesini sağlamıştır.

Benzer şekilde Tablo 1'in altıncı kolonu (Doğruluk BTİÖK +  $S_2$ ) bazı periyotlar için (5,7,15,16,17,18,20), BTİÖK ve  $S_2$ 'nin birlikte kullanılması BTİÖK ve  $S_2$ 'nin bireysel kullanımlarına göre daha yüksek başarımlar değeri üretmiştir.



Tablo 1.  $S_1$ ,  $S_2$ ,  $S_3$ , BTİÖK, BTİÖK+ $S_1$ , BTİÖK+ $S_2$ , BTİÖK +  $S_3$  doğruluk değerleri  
( $S_1$ ,  $S_2$ ,  $S_3$ , BTİÖK, BTİÖK+ $S_1$ , BTİÖK+ $S_2$ , BTİÖK +  $S_3$  accuracy values)

Periyot	Doğruluk $S_1$	Doğruluk $S_2$	Doğruluk $S_3$	Doğruluk BTİÖK	Doğruluk BTİÖK + $S_1$	Doğruluk BTİÖK + $S_2$	Doğruluk BTİÖK + $S_3$
1	0,5126±0,0483	0,4952±0,0418	0,4638±0,0365	0,5579±0,0363	0,5557±0,0366	0,5499±0,0420	0,5422±0,0409
2	0,5357±0,0506	0,5138±0,0437	0,5153±0,0438	0,6142±0,0308	0,6067±0,0323	0,6100±0,0387	0,6085±0,0427
3	0,5801±0,0457	0,5538±0,0455	0,5291±0,0345	0,6709±0,0374	<b>0,6755±0,0405</b>	0,6602±0,0467	0,6656±0,0359
4	0,6068±0,0401	0,5894±0,0500	0,5707±0,0459	0,7248±0,0375	0,7188±0,0385	0,7147±0,0336	0,6959±0,0335
5	0,5978±0,0413	0,5906±0,0348	0,5456±0,0533	0,7412±0,0392	<b>0,7521±0,0338</b>	<b>0,7595±0,0369</b>	<b>0,7532±0,0427</b>
6	0,6018±0,0364	0,5761±0,0428	0,5256±0,0439	0,7935±0,0386	<b>0,7938±0,0396</b>	0,7795±0,0364	0,7862±0,0316
7	0,5992±0,0492	0,5694±0,0495	0,5612±0,0416	0,7981±0,0361	<b>0,8022±0,0289</b>	<b>0,8023±0,0305</b>	0,7886±0,0482
8	0,5571±0,0330	0,5559±0,0429	0,5315±0,0623	0,8216±0,0340	0,8042±0,0312	0,8047±0,0352	0,8201±0,0344
9	0,5696±0,0437	0,5657±0,0403	0,5681±0,0521	0,8232±0,0397	0,8161±0,0394	0,8177±0,0302	<b>0,8314±0,0269</b>
10	0,6087±0,0408	0,5844±0,0433	0,5537±0,0482	0,8368±0,0323	0,8306±0,0419	0,8292±0,0341	0,8251±0,0339
11	0,6017±0,0401	0,5778±0,0412	0,5283±0,0453	0,8495±0,0367	<b>0,8502±0,0336</b>	0,8375±0,0371	0,8315±0,0383
12	0,5901±0,0446	0,5911±0,0518	0,5635±0,0463	0,8699±0,0323	<b>0,8703±0,0339</b>	0,8556±0,0323	0,8587±0,0310
13	0,5710±0,0494	0,5597±0,0367	0,5814±0,0505	0,8766±0,0307	0,8530±0,0370	0,8645±0,0371	0,8546±0,0230
14	0,5836±0,0361	0,5699±0,0459	0,5542±0,0473	0,8627±0,0467	<b>0,8628±0,0337</b>	<b>0,8647±0,0320</b>	0,8546±0,0422
15	0,5735±0,0374	0,5614±0,0380	0,5802±0,0392	0,8702±0,0313	<b>0,8736±0,0396</b>	<b>0,8733±0,0367</b>	0,8620±0,0375
16	0,5774±0,0454	0,5744±0,0493	0,5900±0,0404	0,8704±0,0361	<b>0,8717±0,0293</b>	<b>0,8788±0,0297</b>	0,8698±0,0321
17	0,5458±0,0522	0,5882±0,0397	0,6384±0,0326	0,8817±0,0258	<b>0,8852±0,0278</b>	<b>0,8913±0,0338</b>	0,8740±0,0285
18	0,5559±0,0427	0,5595±0,0479	0,5811±0,0470	0,8613±0,0306	<b>0,8764±0,0325</b>	<b>0,8680±0,0337</b>	<b>0,8651±0,0411</b>
19	0,5868±0,0506	0,5612±0,0447	0,6351±0,0469	0,8756±0,0360	0,8718±0,0275	0,8729±0,0276	0,8689±0,0348
20	0,5865±0,0456	0,5810±0,0483	0,6233±0,0364	0,8767±0,0272	<b>0,8782±0,0338</b>	<b>0,8810±0,0409</b>	0,8749±0,0329

Son olarak Tablo 1'in yedinci kolonu (Doğruluk BTİÖK +  $S_3$ ) bazı periyotlar için (5,9,18), BTİÖK ve  $S_3$ 'ün birlikte kullanılması BTİÖK ve  $S_3$ 'ün bireysel kullanımlarına göre daha yüksek başarımlar üretmiştir. Genel olarak bu deneylerde de en iyi doğruluk değerlerine  $S_1$  duygu değerinin katkısıyla ulaşılmıştır.

Tablo 2 yukarıda da bahsi geçen deneylere ait F-Skor değerleri sonuçlarını yansıtmaktadır. Elde edilen F-ölçüm skorları sonuçlarına göre yine  $S_1$  başarımlarının kullanımı  $S_2$  ve  $S_3$  değerlerine göre daha verimli sonuçlar üretmiştir. Koyu renkle gösterilen hücre değerlerine denk gelen periyotlarda borsa teknik indikatörleri ile birlikte duygu değerlerinin de kullanılması başarımların daha yüksek sonuçların görülmesini sağlamıştır.

Tablo 1 ve Tablo 2'den edinilen sonuçlara göre BİST100 endeksinin artış ve azalış yönünün bir sınıflandırma problemi olarak ele alındığı bu çalışmada haber metinlerinden elde edilen duygu skorlarının problem üzerinde pozitif bir etkiye sahip olduğu görülmüştür.

## 5. DEĞERLENDİRME VE SONUÇ (EVALUATION AND CONCLUSION)

Günümüzde metinsel haber kaynaklarının farklı durumlar ile ilgili yarattığı olumlu ya da olumsuz algılar metin madenciliğinin araştırma alanlarından biri olan duygu analizi aşamalarıyla yapısal hale

getirilebilmektedir. “Olumlu” veya “olumsuz” etiketler ile etiketlenen haber sayılarına bağlı olan duygu skorları, haber kaynaklarının yarattığı olumlu/olumsuz algının pek çok durum üzerindeki etkisini analiz etmek için kullanılmaktadır. Özellikle koronavirüs salgını döneminde yayınlanan haber metinlerinin yansıttığı duygu durumlarının borsa üzerindeki bir etkiye sahip olup olmadığı merak konusu olmuştur. Bu çalışmada koronavirüs konulu haber metinlerinden elde edilen duygu skorlarının BİST100 endeksinin kapanış fiyatlarının trend tahmininde bir etkiye sahip olup olmadığı araştırılmıştır. Böylece haber metinlerinde yansıtılan olumlu ve olumsuz duygularının borsa gibi dinamik bir alanda geleceğe dair kestirimlerin yapılması için kullanılan popüler teknik indikatörler gibi rol alıp alamayacağı test edilebilmiştir. Bu amaçla çalışmamızda 20.12.2019 – 16.02.2021 tarihleri arasındaki bir dönemde derlenen koronavirüs konulu haber külliyatından elde edilen pozitif ve negatif haber sayılarının kullanılmasıyla üç farklı duygu skoru hesaplanmıştır. Ardından BİST 100 trendin belirlenmesi için kullanılacak olan popüler 8 farklı teknik indikatör hesabı gerçekleştirilmiştir. Böylelikle duygu değerlerinin ve teknik indikatörlerin RO algoritmasının kullanılmasıyla sınıflandırma problemine dönüştürülen BİST100 endeksi trend tahmini üzerindeki başarımların değerlendirilmesi kısıylandırılmıştır.

Tablo 2.  $S_1, S_2, S_3, BTiÖK, BTiÖK+S_1, BTiÖK+S_2, BTiÖK + S_3$  F-Skor değerleri $(S_1, S_2, S_3, BTiÖK, BTiÖK+S_1, BTiÖK+S_2, BTiÖK + S_3$  F-Score values)

Pe ri yo t	F-Skor $S_1$	F-Skor $S_2$	F-Skor $S_3$	F-Skor BTiÖK	F-Skor BTiÖK + $S_1$	F-Skor BTiÖK + $S_2$	F-Skor BTiÖK + $S_3$
1	0,5879±0,0491	0,5708±0,0466	0,5468±0,0353	0,6349±0,0360	<b>0,6356±0,0411</b>	0,6304±0,0436	0,6225±0,0443
2	0,6275±0,0471	0,6073±0,0446	0,6215±0,0423	0,6973±0,0326	0,6925±0,0359	<b>0,6989±0,0373</b>	0,6940±0,0432
3	0,6634±0,0500	0,6293±0,0370	0,6127±0,0399	0,7395±0,0366	0,7373±0,0396	0,7295±0,0440	0,7317±0,0332
4	0,7078±0,0359	0,6940±0,0473	0,6719±0,0423	0,8030±0,0316	0,7983±0,0328	0,7959±0,0288	0,7809±0,0285
5	0,7059±0,0346	0,6995±0,0311	0,6633±0,0491	0,8173±0,0311	<b>0,8252±0,0289</b>	<b>0,8296±0,0302</b>	<b>0,8245±0,0342</b>
6	0,7079±0,0354	0,6836±0,0457	0,6519±0,0413	0,8565±0,0299	0,8559±0,0300	0,8465±0,0264	0,8495±0,0254
7	0,7048±0,0515	0,6877±0,0472	0,6799±0,0391	0,8594±0,0280	<b>0,8607±0,0246</b>	<b>0,8627±0,0226</b>	0,8534±0,0367
8	0,6556±0,0350	0,6552±0,0410	0,6391±0,0591	0,8680±0,0282	0,8546±0,0259	0,8540±0,0266	0,8672±0,0306
9	0,6770±0,0393	0,6767±0,0388	0,6773±0,0485	0,8716±0,0309	0,8687±0,0317	0,8697±0,0218	<b>0,8769±0,0218</b>
10	0,7041±0,0387	0,6902±0,0402	0,6673±0,0455	0,8832±0,0246	0,8762±0,0340	0,8781±0,0278	0,8735±0,0249
11	0,6979±0,0372	0,6801±0,0473	0,6442±0,0425	0,8913±0,0289	<b>0,8943±0,0250</b>	0,8826±0,0288	0,8775±0,0324
12	0,6981±0,0462	0,7028±0,0491	0,6809±0,0401	0,9080±0,0255	<b>0,9109±0,0248</b>	0,8977±0,0249	0,8996±0,0228
13	0,6854±0,0451	0,6806±0,0342	0,6973±0,0487	0,9152±0,0219	0,8972±0,0297	0,9074±0,0283	0,8979±0,0192
14	0,6914±0,0333	0,6881±0,0386	0,6734±0,0503	0,9049±0,0331	0,9022±0,0267	0,9041±0,0246	0,8964±0,0309
15	0,6827±0,0401	0,6772±0,0320	0,6936±0,0353	0,9096±0,0242	<b>0,9099±0,0317</b>	<b>0,9109±0,0289</b>	0,9010±0,0294
16	0,6863±0,0470	0,6865±0,0450	0,7039±0,0338	0,9086±0,0275	<b>0,9104±0,0223</b>	<b>0,9157±0,0228</b>	0,9078±0,0237
17	0,6653±0,0503	0,7066±0,0317	0,7376±0,0265	0,9177±0,0184	<b>0,9215±0,0198</b>	<b>0,9241±0,0232</b>	0,9113±0,0217
18	0,6662±0,0366	0,6658±0,0475	0,6843±0,0446	0,9002±0,0240	<b>0,9087±0,0270</b>	<b>0,9037±0,0251</b>	<b>0,9043±0,0303</b>
19	0,6912±0,0497	0,6764±0,0405	0,7372±0,0405	0,9130±0,0263	0,9087±0,0203	0,9102±0,0230	0,9074±0,0263
20	0,6919±0,0385	0,6898±0,0445	0,7270±0,0306	0,9154±0,0200	<b>0,9157±0,0254</b>	<b>0,9190±0,0293</b>	0,9134±0,0252

Yapılan analizler sonucunda öncelikle popüler teknik indikatörlerinin borsa değeri artma azalma yönü tahminlemesi için oldukça etkili olduğu gözlemlenmiştir. Koronavirüs konulu haber verilerinden elde duygu skorlarının trend analizinde tek bir özellik olarak kullanılması iyi bir sınıflandırma doğruluk değeri ürettiği görülmüştür. Duygu değerleri ve teknik indikatörlerin birlikte kullanıldığı analizlerde bazı durumlarda, sadece teknik indikatörlerin kullanıldığı durumlara göre daha iyi başarımların elde edildiği görülmüştür. Böylece haber metinleri ile toplumda yaratılan olumlu, olumsuz algıların borsa üzerinde etkili olabileceği ve bu alanda gerçekleştirilecek olan çalışmalarda duygu skorlarının incelenebilecek bir özellik olabileceği görülmüştür. Bu çalışma ile elde edilen bulgular literatürdeki pek çok çalışma ile örtüşmektedir. Literatür taraması kısmında ele alınan yabancı sermaye piyasaları için gerçekleştirilen çalışmalarda ([16-19]), teknik indikatörlerin trend tahmininde kullanılabilecek etkili özellikler arasında olduğu bildirilmiştir. Paralel bir şekilde, bu çalışma sonucunda da kullanılan teknik indikatörlerin yüksek başarımların elde ettiği görülmüştür. Ayrıca literatürdeki pek çok çalışma, finansal makalelerden [13], [15], borsa ve finans ile ilgili haber metinlerinden [10], [12] ve popüler sosyal medya platformlarında paylaşılan kısa iletilerden elde edilen duygu skor değerlerinin [11], [14] borsa endekslerinin trend analizlerinde kullanılabileceğini belirtmiştir. Bu çalışmada da bazı periyotlar için haber metinlerinden elde edilen duygu değerlerinin BİST100 endeksi üzerinde etkili olduğu gözlemlenmiştir. Sonuç olarak finansal metin

kaynaklarının borsa analizlerinde kullanılabileceği yönünde bir bulguya erişilmiştir.

Gelecek çalışmalarda özellik seçim metodları kullanılarak sınıflandırma aşamasında kullanılan teknik indikatörler ve duygu skorlarının değişik kombinasyonlarla kullanılması planlanmaktadır. Ayrıca tahminleme aşamasında derin ağların (RNN, LSTM vb. gibi [33]) etkisi gözlemlenecek ve geleneksel makine öğrenmesi metodları ile kıyaslamalı analizler sunulacaktır.

#### KAYNAKLAR (REFERENCES)

- [1] G. G. Yağcılar, "Borsa İstanbul'da COVID-19 Etkisi: Kısa Dönemli Sektörel Piyasa Tepkilerinin Endeks Bazında Ölçülmesi", *Ekonomi Politika ve Finans Araştırmaları Dergisi*, 6(2), 439-463, 2021.
- [2] D. K. Pandey, V. Kumari, "Event study on the reaction of the developed and emerging stock markets to the 2019-nCoV outbreak", *International Review of Economics & Finance*, 71, 467-483, 2021.
- [3] Y. Peker, E. Demirhan, **Covid-19 küresel salgınının borsa İstanbul'daki sektörel etkileri**, Türkiye Ekonomi Politikaları Araştırma Vakfı (TEPAV), Türkiye, 2020.
- [4] E.F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work.", *The Journal of Finance*, 25(2), 383-417, 1970.
- [5] K. Eyüboğlu, S. Eyüboğlu, "Borsa İstanbul Endekslerinde Adaptif Piyasa Hipotezinin Geçerliliğinin Test Edilmesi", *Yaşar Üniversitesi E-Dergisi*, 15(59), 642-654, 2020.

- [6] A. W. Lo, "The Adaptive Markets Hypothesis." *The Journal of Portfolio Management*, 30(5), 15-29, 2004.
- [7] A. C. Çetin, "Koronavirüs (Covid-19) Salgınının Türkiye'de Genel Ekonomik Faaliyetlere ve Hisse Senedi Borsa Endeksine Etkisi", *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, 4(2), 341-362, 2020.
- [8] S. Atan, Y. Çınar, "Borsa İstanbul'da finansal haberler ile piyasa değeri ilişkisinin metin madenciliği ve duygu (sentiment) analizi ile incelenmesi", *Ankara Üniversitesi SBF Dergisi*, 74(1), 1-34, 2019.
- [9] A. Güran, E. Ateş, "Twitter iletileri duygu değerleri ve Bist 30 endeksi günlük değer değişimlerinin Pearson korelasyonu ve Granger nedensellik analizi", *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 36(3), 1687-1702, 2021.
- [10] Y. Peng, H. Jiang, "Leverage financial news to predict stock price movements using word embeddings and deep neural networks", *arXiv preprint arXiv:1506*, 2015.
- [11] S. Bouktif, A. Fiaz, M. Awad, "Stock market movement prediction using disparate text features with machine learning", **In 2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)**, IEEE, 1-6, October, 2019.
- [12] A. Picasso, S. Merello, Y. Ma, L. Oneto, E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction", *Expert Systems with Applications*, 135, 60-70, 2019.
- [13] S. Mohan, S. Mullapudi, S. Sammeta, P. Vijayvergia, D. C. Anastasiu, "Stock price prediction using news sentiment analysis", **In 2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)**, IEEE, 205-208, April, 2019.
- [14] X. Li, P. Wu, W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong", *Information Processing & Management*, 57(5), 102212, 2020.
- [15] H. D. Huynh, L. M. Dang, D. Duong, A new model for stock price movements prediction using deep neural network, **In Proceedings of the Eighth International Symposium on Information and Communication Technology**, 57-62, December, 2017.
- [16] J. Patel, S. Shah, P. Thakkar, K. Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques", *Expert systems with applications*, 42(1), 259-268, 2015.
- [17] L. Khaidem, S. Saha, S. R. Dey, "Predicting the direction of stock market prices using random forest", *arXiv preprint arXiv:1605.00003*, 2016.
- [18] X. Di, **Stock trend prediction with technical indicators using SVM**, Independent Work Report, Standford: Leland Stanford Junior University, USA, 2014.
- [19] S. Mehtab, J. Sen, "Stock price prediction using convolutional neural networks on a multivariate timeseries", *arXiv preprint arXiv:2001.09769*, 2020.
- [20] Y. Kara, M. A. Boyacioglu, Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange", *Expert systems with Applications*, 38(5), 5311-5319, 2011.
- [21] Ç. Özari, K. Turan, E. Demir, "Teknik indikatörlerin etkinliği: Bist30 ve bist100 endeksleri üzerine bir uygulama", *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, 6(1), 94-113, 2016.
- [22] C. Budak, **Teknik Analiz indikatörlerinin Performans karşılaştırması üzerine bir araştırma**, Doctoral dissertation, Marmara Üniversitesi, 2019.
- [23] S. Gumparthi, "Relative strength index for developing effective trading strategies in constructing optimal portfolio", *International Journal of Applied Engineering Research*, 12(19), 8926-8936, 2017.
- [24] E. A. S. Ahmed, S. B. Goyal, "**Impact of Technical Parameters for Short-and Long-term Analysis of Stock Behavior**", *Materials Today: Proceedings*, 2021.
- [25] P. Fernández-Blanco, D. J. Bodas-Sagi, F. J. Soltero, J. I. Hidalgo, Technical market indicators optimization using evolutionary algorithms, **In Proceedings of the 10th annual conference companion on Genetic and evolutionary computation**, 1851-1858, July, 2008.
- [26] M. R. Vargas, C. E. dos Anjos, G. L. Bichara, A. G. Evsukoff, Deep learning for stock market prediction using technical indicators and financial news articles, **In 2018 International Joint Conference on Neural Networks (IJCNN)**, IEEE, 1-8, July, 2018.
- [27] G. Yapıcı, **Teknik Analiz Yöntemi Ve Bist100'de İşlem Gören Ticari Bankacılık Payları Üzerine Bir Uygulama**, Yüksek Lisans Tezi, İstanbul Okan Üniversitesi, İşletme Anabilim dalı, İşletme Programı, 2020.
- [28] Y. Erdiñç, **Borsada Teknik Analiz El Kitabı, Siyasal Kitabevi**, Ankara, 1996.
- [29] M. N. Günak, **İleri Teknik Analiz Uygulamaları ve Bu Uygulamaların İMKB'de Test Edilmesi**, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, 2007.
- [30] V. Vaghela, R. Gor, N. Chavda, "Elliott Wave formation using combination of CCI and DEMA", *IOSR Journal of Mathematics (IOSR-JM)*, 2319-765X, 14-21, 2021.
- [31] P. Sadorsky, "A Random Forests Approach to Predicting Clean Energy Stock Prices", *Journal of Risk and Financial Management*, 14(2), 48, 2021.
- [32] A. Cowles, "Can stock market forecasters forecast?", *Econometrica Journal of the Econometric Society*, 309-324, 1933.
- [33] A. Ayaydın, M.A. Akçayol, "Deep Learning Based Forecasting of Delay on Flights.", *Bilişim Teknolojileri Dergisi*, 15(3), 239-249, 2022.

# Sürüş Davranış Tespiti ve Tasnifi için Takograflara Yönelik Yeni Bir Algoritma Tasarımı ve Uygulaması

*Araştırma Makalesi/Research Article*

 Cevat ALTUNKAYA<sup>1</sup>,  Ahmet ZENGİN<sup>2</sup>

<sup>1</sup> Elektrik-Elektronik Mühendisliği, Sakarya Üniversitesi, Sakarya, Türkiye

<sup>2</sup> Bilgisayar Mühendisliği, Sakarya Üniversitesi, Sakarya, Türkiye

[cevat.altunkaya1@ogr.sakarya.edu.tr](mailto:cevat.altunkaya1@ogr.sakarya.edu.tr), [azengin@sakarya.edu.tr](mailto:azengin@sakarya.edu.tr)

(Geliş/Received:26.02.2022; Kabul/Accepted:22.08.2022)

DOI: 10.17671/gazibtd.1079364

**Özet**— Günümüzde yollardaki araç sayısı arttıkça trafik güvenliğinin önemi de artmaktadır. Birden çok aracın karıştığı trafik kazaları sürücü ve yolcuların hayatını tehdit etmektedir. Takograf kullanımı zorunlu olan kamyon, otobüs gibi ağır vasıta araçlar, trafik kazalarında daha fazla ölüm, yaralanma ve maddi kayıplara yol açmaktadır. Geleneksel takograflar dönüş ve şerit değişiklikleri gibi yanıl manevraların şiddetiyle ilgili herhangi bir bilgi vermemektedir. Agresif sürücü davranışlarının trafik kazaları ve yakıt tüketimine etkisi önemli boyutlarda olabilmektedir. Bu çalışmada ivmeölçer ve jiroskop verilerini kullanarak sürücü davranışını tespit etmek ve derecelendirmek için yeni bir algoritma tasarlanmış ve yeni bir takograf cihazı geliştirilmiştir. Bu amaçla, yanıl manevraların algılanması için takograf cihazına düşük maliyetli bir IMU (Inertial Measurement Unit) sensör modülü eklenmiştir. Sağa-sola dönüşler ve şerit değişiklikleri manevralarını yüksek hassasiyetle tespit etmek için öncelikle jiroskop-Z verileri ile olayların kenarları yakalanır ve başlangıç ve bitiş noktaları belirlenir. Ardından aralıktaki ivmeölçer-X, jiroskop-Z ve hız verilerini bir formülle birleştirilerek yanıl manevraları puanlayan yeni bir algoritma önerilmiştir. Yapılan testlerde algoritmanın dönüşlerin algılanmasında %100, şerit değişikliklerinde %88 doğruluğa yaklaştığı gözlemlenmiştir.

**Anahtar Kelimeler:** sürüş manevra algılama, agresif sürüş, sürüş puanlama, takograf, sensörler

## Design and Implementation of a New Algorithm to Tachographs for Driving Behavior Detection and Classification

**Abstract**—Today, as the number of vehicles on the roads increases, the importance of traffic safety increases. Traffic accidents involving more than one vehicle threaten the lives of drivers and passengers. Heavy-duty vehicles such as trucks and buses, for which the use of tachographs is mandatory, cause more deaths, injuries and financial losses in traffic accidents. Conventional tachographs do not provide any information on the severity of lateral maneuvers such as turns and lane changes. The effects of aggressive driver behaviors on traffic accidents and fuel consumption can be significant. In this study, a new algorithm was designed and a new tachograph device was developed to detect and rate driver behavior using accelerometer and gyroscope data. For this purpose, a low-cost IMU (Inertial Measurement Unit) sensor module has been added to the tachograph device to detect lateral maneuvers. In order to detect right-left turn and lane change maneuvers with high precision, first the edges of the events are captured with the gyroscope-Z data and the start and end points are determined. Then, a new algorithm that scores lateral maneuvers is proposed by combining the accelerometer-X, gyroscope-Z and speed data in the range with a formula. In the tests performed, it was observed that the algorithm approached 100% accuracy in detecting turns and 88% accuracy in lane changes.

**Keywords:** driving maneuver detection, aggressive driving, driving scoring, tachograph, sensors

## 1. GİRİŞ (INTRODUCTION)

Günümüzde yollardaki araç sayısı arttıkça trafik güvenliğinin önemi de artmaktadır. Özellikle kamyon ve otobüs gibi ağır vasıta araçların sebep olduğu kazalardaki kayıplar hem ölüm ve yaralanma hem de maddi hasar bakımından daha yüksektir. İstatistiklere göre kazalar çok yüksek oranda sürücü kusurlarından kaynaklanmaktadır [1]. Karayollarında yük ve yolcu taşımacılığı yapan araçların büyük bir kısmında kullanılması zorunlu olan takograflar, trafik güvenliği için sürücülerin hız, araç kullanma ve dinlenme sürelerinin denetlenebilmesini sağlayan cihazlardır [2]. Mevcut takograflar, trafik kazalarının ve fazla yakıt tüketiminin en önemli nedenleri olan yorgunluk, dikkatsizlik, uykusuzluk, agresif sürüş ve benzeri sürücü davranışları hakkında herhangi bir veri sunmamaktadır. Halbuki birçok kazanın sebebi tehlikeli sürüş davranışlarıdır. Trafik kazaları ve yakıt tüketimini azaltmak amacıyla sürüş manevralarını tespit ederek düzeltmek için birçok çalışma yapılmıştır [3-6]. Doğrusal manevraların tespitinde GPS veya araçtan alınan hız bilgisi yeterli olurken [7] yanal manevraların tespitinde jiroskop, ivmeölçer ve manyetometre gibi sensörler kullanılmıştır [8-12]. Artık hayatımızın vazgeçilmez haline gelen akıllı telefonlar söz konusu sensörlerin hepsini ihtiva ettiği için son dönemde özellikle yanal manevra davranışlarıyla ilgili çalışmaların ana unsuru haline gelmiştir [5,6,13]. Bu çalışmanın amacı ağır vasıtalarda kullanımı zorunlu olan takograf cihazlarına yönelik ivmeölçer ve jiroskop verilerini kullanarak sürücünün yanal manevralarını tespit etmek ve puanlamak için yeni bir algoritma tasarlamaktır.

Çalışma için maliyet, boyut ve montaj kolaylığı nedeniyle üzerinde 3D ivme ölçer, 3D jiroskop, 3D manyetometre ve basınç sensörü içeren 10 DOF IMU sensor modülü takograf cihazına monte edilmiştir. Bu sensörler kullanılarak akıllı telefonlar üzerinde yanal manevraları algılamak için yapılan çalışmalar bulunmaktadır. Bu çalışmaların sürücüyü uyararak sürücü davranışlarını düzeltme [3,4], sürücü davranışlarını tespit ederek sürücüyü etiketleme veya puanlama [5,6,9,14-16,19], yolculuklardaki manevraları inceleyerek sürüş özelliklerini tespit etme [8,11,13,20,21], yolculuklardaki kaza riskini algılama [7,10,17,19], yakıt tüketimini azaltmak için sürüş verilerini istatistiksel olarak inceleme [12] ve sürüş sırasında sürücü dikkat dağınıklığı tespit etme [22] gibi çeşitli alanlara odaklandığı görülmektedir.

Revize edilerek sensör ve hız verilerini yolculuklarda hafızasına kaydetmesi sağlanan test takografından indirilen veriler analiz edilerek yanal manevraları tespit edip puanlayan yeni bir algoritma geliştirildi. Jiroskop-Z verileri taranarak öncelikle kenarlar bulunur. Önceki kenarların durumu da dikkate alınarak kenarın bir şerit değişikliği veya dönüş manevrasına ait olduğuna karar verilir. Ardından dönüş manevralarına jiroskop, ivmeölçer ve hız ortalamalarını içeren bir formülle, şerit değişikliklerine jiroskop eğimi ve hız ortalaması ile puan verilir.

## 2. İLGİLİ ÇALIŞMALAR (RELATED WORKS)

Son yıllarda sürüş davranışıyla ilgili çalışmalar otonom sürüş, sürücü destek ve agresif sürüş tespiti konularında yoğunlaşmıştır. Otonom sürüş çalışmalarında nesnelerin ve yol çizgilerinin algılanmasında genellikle kamera görüntüleri ve LIDAR dönüşleri kullanılmış, görüntü işleme tekniklerinden istifade edilmiştir [27,28]. Sürücü davranışlarına odaklanan çalışmaların bir kısmında biyomedikal sensörler ve görüntü işleme teknikleri de kullanılmıştır [29-31]. Görüntü işleme tekniklerinin yüksek işlem ve hafıza kapasiteli mikrodenetleyiciler veya bilgisayarlar gerektirmesi maliyeti artırmaktadır. Biyomedikal sensörlerin de çoğunun maliyeti yüksektir. Ayrıca birçoğu sürücü üzerine monte edilme gibi kullanım zorluklarına sahiptir.

Düşük maliyetli sensörlerle yapılan yakın alandaki agresif sürüş tespit çalışmaları daha çok sürüş stilini, sürücü davranışını ve kaza riskini algılama ve yakıt tüketimini azaltmaya yönelik yönlendirme ve bilgi vermeye odaklanmıştır. Literatürdeki çalışmalarda virajlar, eğimler, şerit değişiklikleri ve u dönüşleri algılanmaya ve tasnif edilmeye çalışılmıştır. Sadece yanal manevralara odaklanan çalışmalar olduğu gibi doğrusal ve yanal manevralara birlikte odaklanan çalışmalar da mevcuttur. Hız bilgisi genellikle GPS kaynaklarından alınmıştır. Dönüşler, şerit değişiklikleri ve u dönüşleri jiroskop, ivmeölçer ve manyetometrelerle algılanmıştır. Birçok çalışmada sadece jiroskop verileri kullanılmıştır. GPS'den alınan hız verileri yüksek binalar nedeniyle yansımaların çok olduğu şehir merkezlerinde, tünellerde ve sinyallerin zayıf olduğu durumlarda hatalara sebep olmaktadır. Diğer sensörler için araca koordinatları belirlenmiş şekilde monte edilmiş bir cihaz yerine akıllı telefonların kullanılması çalışmaların pratikte uygulanabilirliğini güçleştirmektedir.

Çalışmalarda manevraların başlangıç ve bitiş noktalarını algılamada daha çok belirli bir pencere için basit hareketli ve karesel ortalamaya dayanan uç nokta algılama sistemi [23] veya standart sapma ve benzeri değerlerin eşik değere bağlı kontrolü gibi yöntemler kullanılmıştır. Bazı çalışmalarda kontrol edilen verilerin zamana bağlı dalgaları eşik değer üzerinden incelenmiştir. Bu yöntemlerle özellikle şerit değişikliklerinin çoğunun ve art arda gelen dönüş manevraların bir kısmının algılanması zordur. Bu çalışma kapsamında yapılan testlerde de bu durum görülmüştür. Yanal manevraların başlangıç ve bitiş noktalarının tespitinden sonra manevranın türünün tespit edilmesinde çoklu dalga (art arda gelen dalgaların yorumlanması) [3], bir desen eşleştirme yöntemi olarak Dinamik Zaman Bükme (Dynamic Time Warping - DTW) [4-6,8,13,20] ve K-En Yakın Komşu (K-Nearest Neighbors - K-NN) [5,21] gibi yöntemlerin yanı sıra olasılık ve tahmine dayalı Çok Katmanlı İleri Beslemeli Sinir Ağı (Multilayer Feed Forward Neural Network - FFNN) [23], Gizli Markov Modeli (Hidden Markov Model - HMM) [14,19,20] ve Maksimum Benzerlik (Maximum Likelihood - ML) [6] gibi teknikler de kullanılmıştır.

2011 yılında Johnson ve Trivedi tarafından akıllı telefon ile olaylar jiroskopun dönme enerjisinin basit hareketli ortalamasını kullanan uç nokta algılama yöntemiyle belirli bir pencere için tespit edilmiştir [5]. Daha sonra her üç sensörden (jiroskop, ivmeölçer, manyetometre) her olay için 5 farklı şablon olmak üzere toplam 120 şablon oluşturulmuştur. Desen eşleştirme tekniği DTW ve sınıflandırma yöntemi olarak K-En yakın komşular yöntemlerini kullanarak manevraları ayırt etmişlerdir. Çalışmada dönüşler için maksimum %93 doğruluğa ulaşılmıştır. Birçok farklı şablonun kullanılması akıllı telefonun dışında bir cihaz söz konusu olduğunda gecikmelere ve hatalara sebep olabilecek bir durumdur. Manevra puanlaması normal ve agresif olarak iki kategoride yapılmıştır.

2015 yılında Pholprasit ve arkadaşları tarafından akıllı telefon ile doğrusal ve yanal manevralar algılanmaya çalışılmıştır [4]. İvmeölçer verilerinin standart sapması eşik değer kontrolüne tabi tutularak olaylar algılandıktan sonra DTW desen eşleştirme yöntemi ile manevralar ayrıştırılmış, sadece ivme ölçer kullanılarak maksimum %80 doğruluğa ulaşılmıştır. İvmeölçer verileri jiroskop verilerine göre daha fazla gürültü içerdiğinden standart sapma hesabıyla özellikle şerit değişikliği ve art arda gelen dönüş manevralarının algılanmasında zorluklar bulunmaktadır. Manevraların puanlanması normal ve ani olmak üzere sadece iki kategori için yapılmıştır.

2015 yılında Engelbrecht ve arkadaşları tarafından akıllı telefon vasıtasıyla yanal manevraları algılamak amacıyla olaylar belirli bir pencere için ortalama verilerini kullanarak uç nokta algılama sistemiyle tespit edilip desen eşleştirme (DTW) yöntemi ve bir olasılık yöntemi olan ML ile ayırt edilmiştir [6]. Çalışmada iki yöntemin sonuçları kıyaslanarak ML'nin daha iyi sonuçlar verdiği gösterilmiştir. ML yöntemiyle maksimum %89.1 doğruluğa ulaşılmıştır. Agresif manevraların sınıflandırma doğruluğu normal manevradan daha düşüktür. Agresif manevralar için ayrı şablonlar bulunduğu bazı manevraların kaçırılması veya yanlış kategoride değerlendirilmesi kaçınılmazdır.

2015 yılında Daptardar ve arkadaşları tarafından akıllı telefon vasıtasıyla jiroskop verileri Kalman filtresiyle düzeltildikten sonra saniyelik eğimler hesaplanarak HMM'nin giriş katmanı oluşturulmuş, manevra sınıflandırmada %95 doğruluğa ulaşılmıştır [14]. Manevraların çoğu bir saniyeden daha geniş olduğundan genişliği değişken eğim modeliyle manevra kenarlarının daha iyi yakalandığı bu çalışmada gözlenmiştir.

2016 yılında Ouyang ve arkadaşları tarafından akıllı telefon ile yanal manevraları algılamak için MultiWave ismi verilen yöntemle jiroskop verilerindeki dalgalar için eşik değer kontrolü yapıldıktan sonra art arda gelen dalga şekillerinden manevra türüne karar verilmiş, dönüşlerin algılanmasında %92 doğruluğa ulaşılmıştır [3]. Sıfıra göre eşik değere bakıldığından eşik değere bağımlılık yüksektir.

Art arda gelen dalga formlarından şerit değişikliklerinin algılanması yöntemi ani yön değiştiren veya sağa-sola dönüş sırasında gerçekleşen şerit değişikliklerini algılamada zorlanacaktır.

2016 yılında Kumtepe ve arkadaşları tarafından kameradan alınan görsel verilerle aracın CAN-bus (Controller area network-bus) hattından alınan verileri kullanarak belirli bir zaman aralığında sürücünün şerit takip ve öndeki araçlarla mesafesi araç hızı ve devir hızına göre değerlendirilerek davranışın agresif olup olmadığı tespit edilmiştir [29].

2018 yılında Najim Al-Din ve arkadaşları tarafından akıllı telefon vasıtasıyla olaylar kısa süreli enerji ortalamaları kullanılarak uç nokta algılama sistemiyle tespit edilip FFNN yöntemiyle ayırt edilmiştir [23].

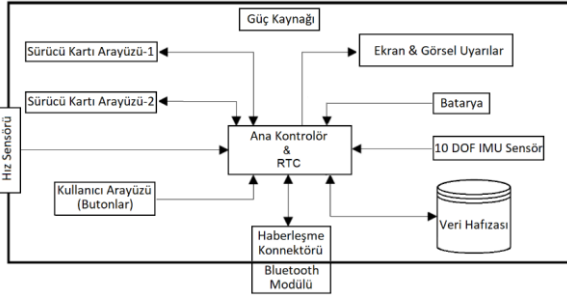
Tablo 1. Manevra algılama kullanılan yöntem ve sensörler (Methods and sensors used in maneuver detection)

Çalışma	Sensörler	Olay algılama tekniği	Manevra sınıflandırma tekniği	Manevra puanlama tekniği
Johnson – Trivedi [5]	GPS, Jiroskop, İvmeölçer, Manyetik	Uç nokta algılama	DTW, k-NN	DTW-şablonlarla (Normal-agresif)
Pholprasit [4]	GPS, Jiroskop, İvmeölçer, Manyetik	Eşik değer algılama	DTW, Desen eşleştirme	Eşik değer kontrolü (Normal-Ani)
Engelbrecht [6]	GPS, Jiroskop, İvmeölçer	Uç nokta algılama	DTW, ML	Hız+jiroskop+ ivme ile formül (Normal-agresif)
Daptardar [14]	GPS, Jiroskop, İvmeölçer,	Eğim kontrolü	HMM (Gizli Markov Modeli)	Doğrusal manevra için formül
Ouyang [3]	GPS, Jiroskop, İvmeölçer	Eşik değer algılama	Çoklu dalga kontrolü	Kullanılmamış
Kumtepe[29]	Kamera, OBD	Görüntü işleme	Görüntü işleme	Öznitelik kıyaslama (Sakin/Agresif)
Najim Al-Din [23]	Jiroskop, İvmeölçer, Manyetik, OBD	Uç nokta algılama	FFNN	Kullanılmamış

Tablo 1 çalışmalarda kullanılan sensör ve teknikleri özetlemektedir. Bu çalışmada yapılan test sonuçlarında özellikle şerit değişikliklerini algılamada manevra desenlerinin manevradan manevraya çok fazla değişiklik göstermesi nedeniyle desen eşleştirme yönteminin uygun olmadığı görülmüştür. Bunun dışında gerçekleşmiş olan veriler üzerinde inceleme yapıldığı için bu çalışmada tahmin ve olasılığa dayalı yöntemlerin kullanılmasına gerek görülmemiştir. Önceki çalışmalarda genellikle doğruluk sonuçlarında yakalanan manevra sayıları dikkate alınmıştır. Bu çalışmada yakalanan manevra sayısının yanı sıra manevra şiddet derecesi karşılaştırılması da yapılmıştır.

### 3. PROTOTİP TAKOGRAF TASARIMI (PROTOTYPE TACHOGRAPH DESIGN)

Saha testleri için piyasada kullanılan bir takograf üzerinde donanımsal ve yazılımsal değişiklikler yapıldı. Standart bir takograf cihazı, sürücü kimlik kartları vasıtasıyla aracı kullanan sürücülerin hız ve kullanım süre bilgilerini kaydetmektedir [17]. Takografin içine I2C haberleşme arayüzünü kullanarak takografin işlemcisiyle haberleşebilen 10 DOF IMU sensörü yere paralel olacak şekilde monte edildi. Test takografinin blok şeması Şekil 1’de gösterilmiştir.



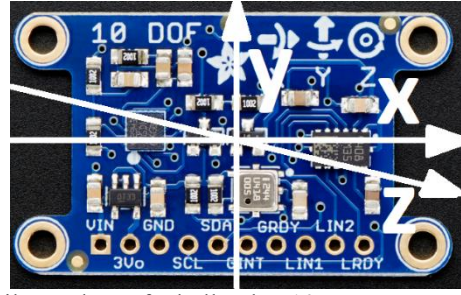
Şekil 1. Test takograf blok şeması  
(Test tachograph block diagram)

Test takograf araca titreşim ve darbelerden etkilenmemesi için Şekil 2’deki gibi araca sağlam bir şekilde monte edilmiştir. Ayrıca Android bir uygulama ile manuel olarak manevra başlangıç, bitişi ve manevra şiddetinin takograf kayıtlarına eklenebilmesi için bluetooth haberleşme donanımı seri port üzerinden takograf cihazına bağlanmıştır. Takografin kayıt kapasitesinin artırılması için ilave hafıza modülü (EEPROM) eklenmiş, böylece takograf ile ilgili donanımsal eklene ve bağlantılar tamamlanmıştır. Takografin yazılımı geliştirilerek her saniye ivmeölçer ve jiroskop verilerinin hafızaya kaydedilmesi sağlandı. Hafıza kapasitesinin 256 kbyte’a çıkarılmasıyla 8 saat civarında bir sürüş kayıt kapasitesine ulaşılmıştır.



Şekil 2. Araca monte edilen takograf cihazı  
(Mounted tachograph device to the vehicle)

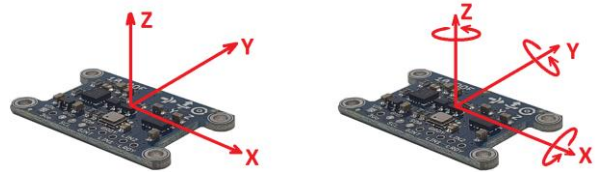
IMU modülü içinde 3D ivmeölçer ve 3D magnetometer bulunan LSM303DLHC [24] entegresi ile 3D jiroskop L3GD20 [25] entegresi ve BMP180 [26] basınç sensörü bulunmaktadır (Şekil 3). İvmeölçer saniyede 50 defa ölçüm yapacak şekilde, jiroskop saniyede 95 defa ölçüm yapacak şekilde ayarlanmıştır.



Şekil 3. Takografda kullanılan 10 DOF IMU sensör  
(10 DOF IMU sensors used in the tachograph)

Hız sinyali test yapılan HYUNDAI Accent Blue model aracın hız göstergesinden 2400 pulse/km şeklinde alınmakta, hız bu katsayıya göre hesaplanmaktadır. Hız dahil bütün sensör verileri saniyede bir okunarak hafızaya kaydedilmiştir. İvmeölçerin her bir ekseninin ölçüm hassasiyeti 12 mg/digit, jiroskopun 8.75 mdps/digit, hız ölçümünün hassasiyeti ise 1,5 km/h’dir. İvmeölçer sensöründen okunan ham veriler g’ye, jiroskop sensöründen okunan ham veriler dps (degree per second)’ye çevrilmektedir.

Şekil 4’ten görüldüğü gibi ivmeölçerin Y eksenini doğrusal ivmelenmeleri, X eksenini yanal ivmelenmeleri, Z eksenini ise yerçekimi veya tersi yöndeki ivmelenmeleri tespit etmektedir. Çalışmada yanal ivmelenmelerin tespiti için X ekseninin kullanılmıştır. X ekseninde sağa dönüş pozitif, sola dönüş negatif değer vermektedir.

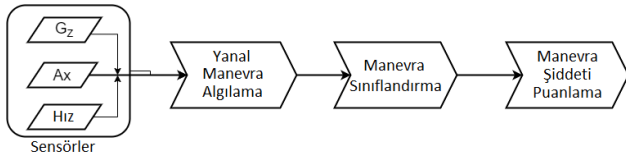


Şekil 4. İvmeölçer ve Jiroskop 3 eksen koordinatları  
(Accelerometer and Gyroscope 3 axis coordinates)

Sağa veya sola dönüş açıları jiroskopun Z eksenini tespit edildiğinden dönüş açıları için kullanılmıştır. Jiroskopun pozitif dönüş yönü saat yönünün tersi olduğu için sola dönüşler pozitif işaretli, sağa dönüşler negatif işaretlidir.

### 4. YANAL SÜRÜŞ DAVRANIŞ TESPİTİ VE PUANLAMASI İÇİN YENİ BİR ALGORİTMA (A NEW ALGORITHM FOR DETECTION AND SCORING OF LATERAL DRIVING BEHAVIOR)

Bu bölüm takografdaki hız, jiroskop ve ivmeölçer sensörlerinden alınan verileri kullanarak sürüşlerdeki yanal manevraların agresiflik şiddetinin belirlenmesindeki algoritmayı anlatır. Sensörlerden alınan veriler hız dahil senkronizedir. Hız verisinin GPS yerine araçtaki hız sensörlerinden darbe şeklinde alınması kesintisiz ve gecikmesiz olmasını sağlamaktadır. Şekil 5’te algoritmanın aşamaları gösterilmiştir.



Şekil 5. Algoritma aşamaları  
(Algorithm stages)

#### 4.1. Önışleme (Preprocessing)

Manevraların başlangıç ve bitiş noktalarını tespit etmede aracın sağa-sola dönüşlerinin ivme ölçere göre daha gürlütüsüz olarak kaydedildiği Jiroskopun Z eksenini kullanılmaktadır.  $G_z$  verisinin kullanılması için birçok filtre denenmiş, filtre kullanmanın şerit değişikliklerini yakalamayı olumsuz etkilediği ve sağa-sola dönüş manevralarının tespitinde filtrenin çok fazla önemi olmadığı gözlenmiştir. Bu nedenlerle çalışmada saniyede bir kaydedilen  $G_z$  verileri için filtre kullanılmamıştır.

Cihazın araca montaj pozisyonuna bağlı olarak verilerde kayma olabilmektedir. O nedenle ivmeölçer-X ve jiroskop-Z'in yolculuktaki genel ortalamaları hesaplanır (1)(3). Ardından ofset değerini sıfıra çekmek için okunan verilere aşağıdaki işlem uygulanır (2)(4). Jiroskop Z ve ivmeölçer-X'in sıfır ofset ayarı

$$A_{x-ort} = \frac{1}{N} \sum_{i=1}^N A_x(i) \quad (1)$$

$$A_x(i)_{ofset} = A_x(i) - A_{x-ort} \quad (2)$$

$$G_{z-av} = \frac{1}{N} \sum_{i=1}^N G_z(i) \quad (3)$$

$$G_z(i)_{ofset} = G_z(i) - G_{z-ort} \quad (4)$$

Burada,

$A_{x-ort}$  : Ortalama ivmeölçer-X, g

$G_{z-ort}$  : Ortalama jiroskop-Z, dps

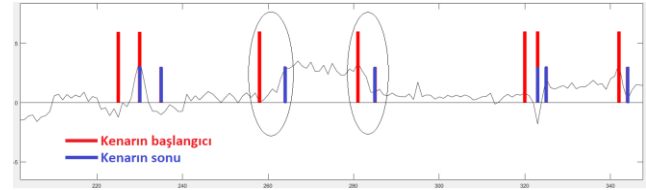
#### 4.2. Olay Algılama Algoritması (Event Detection Algorithm)

Geliştirilen yöntemde başlangıçta manevranın tamamı yerine kenarlara (edge) odaklanılmıştır. Bir dalganın (manevranın) öncelikle kenarlarını bulup daha sonra bu kenarın manevra başlangıcı, manevra bitışı veya müstakil bir şerit değişikliği olduğuna karar verme yoluna gidilmiştir.

##### 4.2.1. Kenar Algılama (Edge Detection)

Kenarların tespiti için sürekli olarak belirli bir aralık içinde maksimum ve minimum noktalar yakalanmaya çalışılmaktadır. Bahsi geçen belirli aralık olarak başlangıçta 7 sn baz alınmaktadır. 7 sn bir kenar için olabilecek geniş bir zaman olarak veriler incelenerek

bulunmuştur. Şekil 6.'da gerçek verilerden algılanan kenarların gösterimine ait bir bölüm bulunmaktadır.



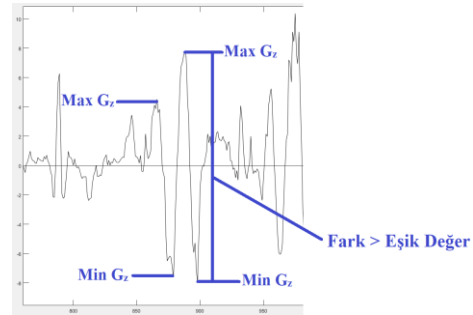
Şekil 6. Algılanan kenarlar  
(Detected edges)

7 sn içinde bulunan  $G_z$  değerlerinin maksimum ve minimum noktaları arasındaki fark için bir eşik değeri belirlenmiştir. Eşik değeri olarak da yapılan testler sonucu 2 dps değeri kullanılmaya karar verilmiştir. Bu eşik değeri o kadar hassas değildir. Biraz daha azaltılırsa sadece fazladan manevra tespiti yapılmış olur. Agresiflik derecelendirilmesinde fazladan tespit edilen manevralar elenecektir.  $G_z(t)_{ofset}$  ( $t = 1, 2, \dots, T$ ) zaman serisinin tamamı olarak varsayarsak, t eksenini boyunca kenar bulmak amacıyla  $T/k$  zaman serisi bölümlerine ayrılmıştır. Her bölümde bir k değeri vardır. Her bir kesit için aşağıdaki formülde gösterilen  $E(n)$  değeri bulunur (5).

n: Zaman serisindeki kesit numarası

$$E(n) = \max_{1 \leq t \leq k} G_z(t, n)_{ofset} - \min_{1 \leq t \leq k} G_z(t, n)_{ofset} \quad (5)$$

$E(n) >$  Üst Eşik Değer ( $T_{fark} = 2$  dps)



Şekil 7. Fark (Maksimum  $G_z$ -Minimum  $G_z$ )  
(Difference (Maximum  $G_z$ -Minimum  $G_z$ ))

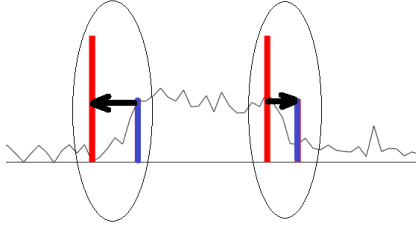
Maksimum ve minimum noktaları arasındaki fark  $E(n)$  eşik değerden büyükse ve maksimum-minimum noktalarının işaretleri farklıysa veya en az biri eşik değerden büyükse ileri 3 saniye boyunca maksimum-minimum kontrolüne devam edilir. Böylece maksimum 9 saniye genişliğinde bir kenarı yakalamak mümkündür. 7 sn'den daha kısa sürede fark eşik değerini iki katına ulaşırsa döngüden erken çıkılır. Böylece algoritmanın birinci aşaması olan kenar bulma tamamlanmış olur. Kenarların bulunmasında sıfıra olan mutlak uzaklık yerine maksimum-minimum farklarının dikkate alınması manevraların tespitinde önemli bir avantaj sağlamaktadır (Şekil 7).



#### 4.2.2. Kenar İçi Nokta (In-edge Point)

Maksimum-minimum noktaları bulduktan sonra bu noktalar arasında kalan sifıra en yakın kenar içi noktanın bulunmasında şöyle bir yöntem izlenmektedir. Kenarın maksimum ve minimum noktalarından mutlak değer olarak daha yüksek olandan daha düşük olana doğru ilerleyerek sifıra en yakın nokta bulunmaktadır (6).

$$G_z(t,n)_{\text{nokta}} = \begin{cases} 1, & \text{Eğer } \min_{1 \leq t \leq k} |G_z(t,n)_{\text{offset}}| < \text{Alt Eşik Değer (Tlow=1)} \\ 0, & \text{Diğer} \end{cases} \quad (6)$$

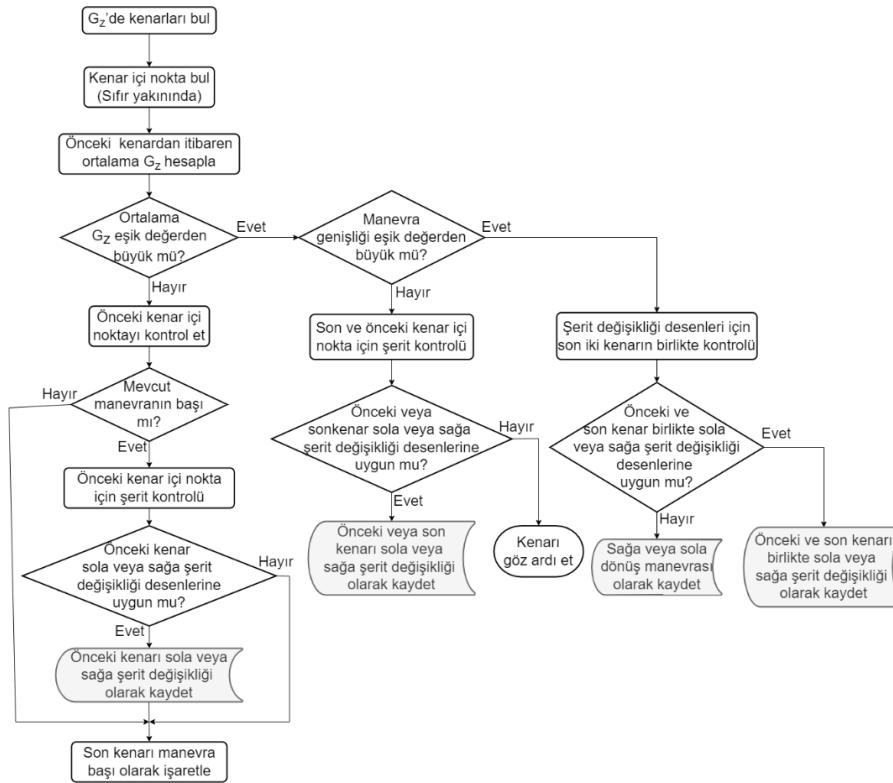


Şekil 8. Kenar içi nokta arama yönleri  
(In-edge point search directions)

Manevra başında ileriye doğru tarayarak sifıra yakın nokta bulunduğunda titreşimler nedeniyle kenar noktasının tespitinde hatalar olabilmekte, daha geride bir nokta bulunabilmektedir. Bu nedenle Şekil 8’de gösterildiği gibi kenar manevra başı ise geriye doğru, manevra sonu ise ileriye doğru tarama söz konusudur.

#### 4.3. Manevra Sınıflandırma (Maneuver Classification)

Şekil 9’da manevra sınıflandırma algoritmasının akış diyagramında gösterildiği gibi sifıra en yakın olarak bulunan kenar içi noktanın önceki değerlere bakılarak manevra başı veya manevra sonu olup olmadığı incelenir. Algoritma bulunan noktanın bir manevra kenarı olmayıp şerit değişikliği olduğunu tespit ederse şerit değişikliği parametresi olarak bulunan bu sifıra en yakın noktayı değil kenarın başlangıç ve bitiş noktalarını kaydeder. Dar (3 sn ve altı) manevralar ile öncesinde ve sonrasında belirli bir ortalama eşik değerin üstünde bir darbe olmayan kenarlar şerit değişikliği kategorisinde incelenmektedir.



Şekil 9. Manevra sınıflandırma algoritması akış diyagramı  
(Flowchart of maneuver classification algorithm)

Eş zamanlı olarak tespit edilen kenar içi sifıra yakın noktadan itibaren  $G_z$  değerleri toplanmaktadır. Böylece herhangi bir anda bir kenar içi noktaya ulaşıldığında bir

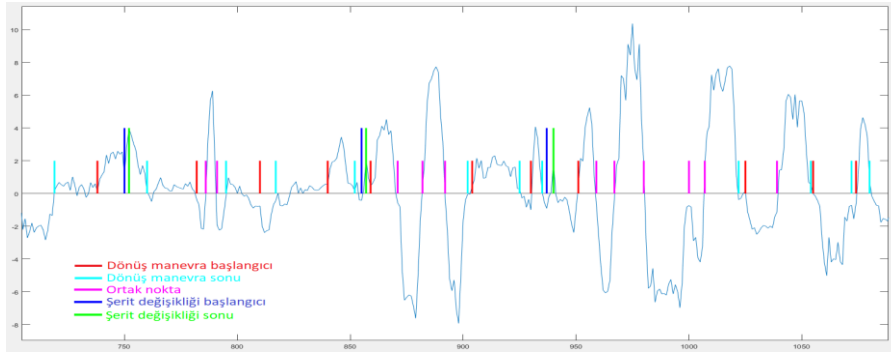
önceki kenar içi noktadan itibaren toplam  $G_z$  değeri ve nokta sayısı (N) bilinmektedir. Bu sayede aşağıdaki formüle görülen ortalama dönüş açısı ( $\theta_{ort}$ ) bulunur (7)

$$\theta_{ort} = \frac{\sum_i^N G_z(i,n)}{N} \quad (7)$$

Ortalama dönüş açısı  $\theta_{ort}$  belirli bir eşik değerinin altında ise kenar noktası manevra başı, üstünde ise manevra sonu olarak değerlendirilecektir (8).

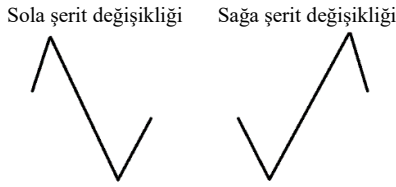
$$n = \begin{cases} \text{Manevra başı, } \theta_{ort} < \text{Eşik Değer} \\ \text{Manevra sonu, } \text{Diğer} \end{cases} \quad (8)$$

Çalışmada manevraların algılanması için önce kenar, ardından kenar içi nokta bulunduğundan sonra iki kenarın arasındaki ortalama değere bakılarak manevra tespit edilmektedir. Sağa ve sola şerit değişiklikleri zamana bağlı  $G_z$  değerlerinde en fazla 3-4 saniye içindeki bir değişikliğe yol açmaktadır. Şerit değişiklikleri iki kenarı olan manevralar yerine herhangi bir manevraya ait olmayan müstakil kenarların veya çok dar olan (3 sn ve altı) manevraların incelenmesiyle tespit edilmektedir. Şekil 10'da yolculuk verilerinde algoritma tarafından başlangıç ve bitiş noktaları işaretlenen manevraların bir bölümü gösterilmiştir.



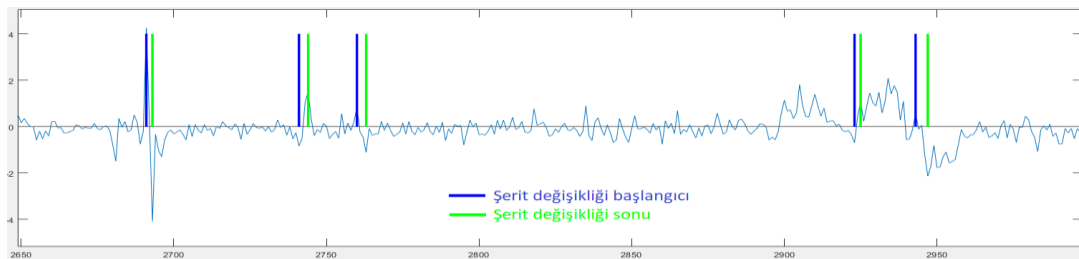
Şekil 10. Manevra başlangıç-bitiş noktalarının algılanması  
(Detection of maneuver start-end points)

*Şerit değişikliği deseni (Lane Change Pattern):* Manevraya ait olmadığı anlaşılan bir kenarın şerit değişikliği olup olmadığını tespit için başlangıç ve bitiş noktalarının ötesine bakılmaktadır. Yapılan incelemelerde sağa ve sola şerit değişikliklerinin Şekil 11'deki gibi bir desene sahip olduğu görülmüştür.



Şekil 11. Şerit değişikliği desenleri  
(Lane change patterns)

Kenarlar bulunduğundan hemen sonra manevra kontrolü yapılmadan önce kenarların 2-3 saniye devamı da kaydedilmektedir. Böylece kenarlara göre yapılan manevra kontrolünden sonra müstakil olduğu anlaşılan bir kenar için maksimum-minimum noktalarının devamına bakılarak şerit değişikliği kontrolü yapılmaktadır. Şekil 12'deki  $G_z$  yolculuk grafiğinde sadece şerit değişikliklerinin işaretlendiği bir bölüm görülmektedir.



Şekil 12. Şerit değişikliği kenar noktalarının algılanması  
(Detection of lane change edge points)

#### 4.4. Agresif Sürüş Puanlama (Aggressive Driving Scaling)

##### 4.4.1. Dönüş Manevralarının Puanlanması (Scoring Turn Maneuvers)

Yanal dönüş manevralarının agresiflik derecesinin saptanmasında sadece jiroskop-Z verisinin yeterli olmadığı gözlenmiştir. İlave olarak yanal ivmelenme ve hız verilerinin de dikkate alınması doğruluğu artırmaktadır. Hız bilgisinin dikkate alınması konusunda herhangi bir tereddüt yoktur. Çünkü yanal manevra şiddetinin yüksek hızlarda oluşturduğu tehlike ile düşük hızlarda oluşturduğu tehlike farklıdır. Yapılan testler ivmeölçer verilerinin hesaba katılmasının da manevra şiddetini belirlemede doğruluğu artırdığını göstermiştir. Bu amaçla her bir manevrada jiroskop-Z, ivmeölçer-X ve hızın ortalama değerleri aşağıdaki formülle birleştirilerek manevranın GAS (Gyro-Accelometer-Speed) değeri bulunmaktadır (9).

$$GAS(n) = (G_{z-ort}(n) - D \cdot A_{x-ort}(n)) \cdot F \cdot Hız_{ort}(n) \quad (9)$$

n: Manevra indexi

$G_{z-ort}(n)$ : İki kenar arasındaki manevranın ortalama  $G_z$  değeri

$A_{x-ort}(n)$ : İki kenar arasındaki manevranın ortalama  $A_x$  değeri

$Hız_{ort}$ : İki kenar arasındaki manevranın ortalama hız değeri

D: Manevranın ortalama  $A_x$  değerinin katsayısı. Büyüklüğü iki katına çıkarmak için  $G_z$  genel karekök ortalamasının  $A_x$  genel karekök ortalamasına (10) oranı olarak hesaplanmaktadır (11).

$$RMS(G_z) = \sqrt{\frac{\sum_i^n G_z(i)^2}{n}}, \quad RMS(A_x) = \sqrt{\frac{\sum_i^n A_x(i)^2}{n}} \quad (10)$$

$$D = RMS(G_z) / RMS(A_x) \quad (11)$$

F: Manevranın ortalama hız değerinin katsayısı. Yolculuktaki genel ortalama hız değeri dikkate alınarak GAS değerini 100'e ölçeklemek amacıyla ortalama için 50 değeri referans alınmak suretiyle F katsayısı aşağıdaki formülle bulunur (12).

$$F = (1/Hız_{gen-ort}) * 50 / (2 * RMS(G_z)) \quad (12)$$

GAS formülünde  $G_z$  ortalama değerleri ana referans olarak alınmış,  $G_z$ 'ye göre ters işaretli olan  $A_x$  ortalama değerleri eksi işaretli belirli bir katsayı ile çarpılarak  $G_z$  ortalamalarına eklenmektedir. Böylece yolculuktaki genel ortalamalar dikkate alındığında  $G_z$  büyüklükleri iki katına çıkmaktadır. Her bir manevra için bulunan GAS değerlerinin işareti yanal manevranın yönünü, büyüklüğü agresiflik derecesini belirlemektedir. GAS değerinin işareti pozitif ise sola dönüş, negatif ise sağa dönüş manevrasındır.

##### 4.4.2. Şerit Değişikliklerinin Puanlanması (Scoring Lane Changes)

Şerit değişikliği manevralarının büyüklüğü sağa-sola dönüş manevralarından farklı olarak bulunan kenarın eğimiyle tespit edilmektedir. Çalışmada GSS (Gyro-Slope-Speed) olarak adlandırılan büyüklük eğimin yüksekliğine ve aralığın ortalama hızına bağlı olarak değişmektedir (13). Eğim söz konusu olduğu için şerit değişikliklerine ait büyüklükler işaretsizdir.

$$GSS(n) = \frac{\max G_z(n) - \min G_z(n)}{\Delta t} \cdot F \cdot Hız_{ort}(n) \quad (13)$$

n: Şerit değişikliği indeksi

F: Manevranın ortalama hız değerinin katsayısı.

Şerit değişikliğindeki F katsayısı bulunurken büyüklük olarak maksimum minimum fark değerleri kullanıldığından uyumlu olması için yolculuktaki  $G_z$  verilerine ait RMSD ( $G_z$ ) standart sapma değeri referans alınmaktadır.

$$RMSD(G_z) = \sqrt{\frac{\sum_i^n (G_z(i) - G_{z,gen-ort})^2}{n}} \quad (14)$$

Yolculuktaki genel ortalama hız değeri dikkate alınarak GSS değerini 100'e ölçeklemek amacıyla ortalama için 50 değeri referans alınarak F katsayısı aşağıdaki formülle bulunur (15).

$$F = (1/Hız_{gen-ort}) \cdot 50 / (2 \cdot RMSD(G_z)) \quad (15)$$

Şerit değişikliklerinin yönü yukarıda anlatılan sağa ve sola şerit değişikliği desenlerinden bulunmaktadır. Algoritma bir kenarın veya dar bir manevranın şerit değişikliği olduğuna karar verdiğinde sağa veya sola hangi yöne olduğuna da karar vermektedir. Algoritma tarafından tespit edilen yanal manevralar ve manevraların puan aralığı üç kademe olarak Tablo 2'de verilmiştir.

Tablo 2'de sağa-sola dönüş manevraları ve sağa-sola şerit değişikliklerine ait büyüklüklerin aralığı üç kademe için verilmiştir. Elde edilen GAS ve GSS büyüklüklerini daha fazla kademelendirmek mümkündür. Algoritma ile yanal manevraların agresiflik derecesine odaklanıldığından U dönüşleri ayrı bir kategori olarak değil, sağa veya sola dönüş olarak kaydedilmiştir.

## 5. DENEYLER (EXPERIMENTS)

Bu bölümün amacı test takografında kaydedilen sensör verilerinin algoritmada denenmesidir.

Tablo 2. Algılanan yanal manevraların tasnifi  
(Classification of detected lateral maneuvers)

	TR	TL	LCR	LCL
Kullanılan sensörler	Jiroskop-Z, İvme ölçer-X, Hız		Jiroskop-Z, Hız	
Düşük	>-60	<60	<12	
Orta	<=-60 >-100	>=60 <100	>=12 <24	
Yüksek	<=-100	>=100	>=24	
Puanlama birimi	GAS		GSS	
TR : Sağa dönüş TL : Sola dönüş LCR: Sağa şerit değişikliği LCL: Sola şerit değişikliği GAS: Dönüş manevra puanı GSS: Şerit değişikliği puanı				

### 5.1. Deneysel kurulum (Experimental Setup)

İçine 10-DOF IMU sensör modülü eklenen test takografı HYUNDAI Accent Blue model araca Şekil 13'teki gibi monte edilmiş, hız sinyali araç göstergesinin de bağlı olduğu ABS müşirinden alınmıştır.



Şekil 13. Araca monte edilen takograf cihazı  
(Vehicle-mounted tachograph device)

Bluetooth haberleşme donanımı eklenen takograf, Android uygulama ile haberleşmeye uygun hale getirilmiştir. Şekil 14'te ara yüzü gösterilen uygulama ile araçta hareket halinde iken bir manevra gerçekleştirileceği zaman manevranın başlangıcında uygulamada ilgili tuşa basılarak takograf kayıtlarına manevra başlangıcının, manevra bittiği anda ise gerçekleşen manevranın şiddetine bağlı olarak normal, orta veya agresif olmak üzere üç farklı kademedeki bitiş noktasının etiketlenmesi sağlanmıştır. Böylece sürüş sırasında istenilen manevraların insan iradesiyle etiketlenebilmesi mümkün hale getirilmiştir.

Bu amaçla Yalova-Sakarya arasında 3833 saniyelik (yaklaşık 1 saat 5 dakika), 85 km'lik bir yolculuk güneşli havada, asfalt yolda yapılmıştır. Şekil 15'te gösterilen güzergahta virajların ve trafik ışıklarının olduğu bölümler de mevcuttur.



Şekil 14. Manevraların etiketlendiği uygulama ara yüzü  
(Application interface where maneuvers are labeled)



Şekil 15. Yalova-Sakarya yolculuk rotası  
(Yalova-Sakarya trip route)

Yolculukta hız ile birlikte jiroskop ve ivmeölçerin 3 eksen verileri her saniye hafızaya kaydedilmiştir. Yolculuk tamamlandıktan sonra bu veriler takografıdan bilgisayara aktarıldıktan ve ham veriler uygun birimlere çevrildikten sonra algoritmada denenmiştir.

Yapılan testlerde saniyede bir kaydedilen ham veriler filtrelenerek de denemeler yapılmış fakat o durumda özellikle şerit değişikliklerinin çoğunun kaçırıldığı gözlenmiştir. Veriler içinde yolculuk sırasında manuel olarak kaydedilen manevralar da bulunduğu için algoritmanın bulduğu sonuçlarla manuel olarak etiketlenen manevraların kıyaslanabilmesi mümkün olmuştur.

### 5.2. Sonuçlar (Results)

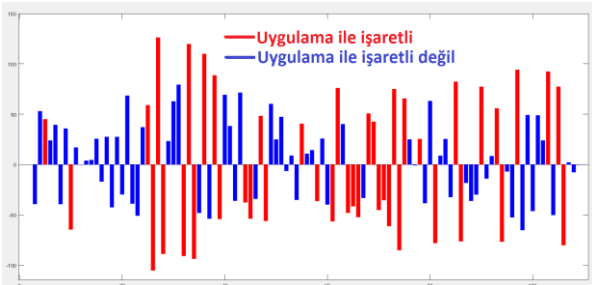
Yapılan yolculukta uygulama üzerinden bütün yanal manevralar etiketlenmemiş, belirli miktardaki bir manevra için etiketleme işlemi yapılmıştır. Yolculuk sırasında erken veya geç basmadan kaynaklanan etiketlemedeki kaymalar, bilgisayar ortamında veriler incelenerek kolayca düzeltilebilmiştir.

Tablo 3. Algoritmanın manevra sayısı tespit doğruluğu  
(Maneuver count detection accuracy of the algorithm)

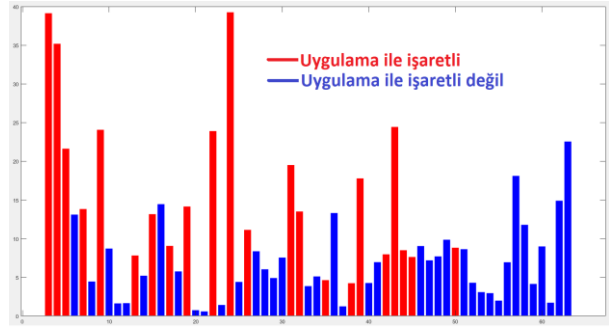
	Yanal Manevra Sayıları			
	TR	TL	LCR	LCL
Manuel kaydedilen	22	20	12	18
Algoritmanın yakaladıkları	22	20	8	13
Algoritma toplam	53	56	31	30
Yönü yanlış	0	0	0	2
Dönüş algılanan şeritler	3	4		
Şerit algılanan dönüşler			0	0
Algılanmayanlar	0	0	0	2
Doğruluk (%)	100	100	66,66	72,22

Algoritma test sonuçlarının detaylı verildiği Tablo 3'te manuel şerit değişikliği olarak etiketlendiği halde algoritmanın sağa-sola dönüş olarak kaydettiği 7 manevra bulunmaktadır. Bu manevraların dalga genişliği limit değerlerden fazladır. Dalganın geniş olması şerit değişikliği manevralarının agresif olmayacağı şeklinde yorumlanabilir. Kaçırılan şerit değişikliklerinin agresiflik açısından düşük puanlı oldukları hesaplanmıştır. Dolayısıyla söz konusu şerit değişikliklerinin sağa-sola dönüş manevraları olarak algılanması da doğru olarak kabul edildiğinde LCR için doğruluk oranı %100'e, LCL için ise %88.88'e çıkmaktadır. Fakat LCL sayılarından iki tanesinin yanlış bir şekilde LCR olarak kaydedildiğini de belirtmek gerekir. Tespit edilen manevra sayıları açısından sonuçlar değerlendirildiğinde algoritmanın sağa-sola dönüş manevralarında oldukça başarılı olduğu, şerit değişikliklerinde büyük oranda başarılı olduğu görülmektedir. Yapılan önceki çalışmalarda da yanal manevralar algılanırken şerit değişikliklerinde sorunlar yaşandığı anlaşılmıştır.

Yalova-Sakarya yolculuğunda algoritma tarafından yakalanan bütün manevraların puanları hesaplanmıştır. Hesaplanan puanlara ait grafikler Şekil 16 ve Şekil 17'de verilmiştir. Dönüş manevralarındaki GAS değerleri yönlü oldukları için grafikte eksenin üst tarafındaki (pozitif) çubuklar sola dönüş büyüklüklerini, alt tarafındaki (negatif) çubuklar sağa dönüş büyüklüklerini göstermektedir.



Şekil 16. Algoritmanın yakaladığı dönüş manevralarının GAS değerleri  
(GAS values of turning maneuvers detected by the algorithm)



Şekil 17. Algoritmanın yakaladığı şerit değişikliklerinin GSS değerleri

(GSS values of lane changes detected by the algorithm)

Şerit değişikliklerindeki GSS değerleri eğime bağlı olduklarından yönlü değildir. Fakat algoritma şerit değişikliklerinde söz konusu kenarların maksimum ve minimum noktalarının önce veya sonra olmasına göre yönü tespit ederek kaydetmektedir.

Grafiklerden uygulama ile işaretlenmeyen mavi çubuklu manevra büyüklüklerinin genel olarak kırmızılardan daha kısa olduğu görülmektedir. Dönüş manevralarının ve şerit değişikliklerinin puanlarının doğruluk değerlendirmesi yapılırken olay bitiminde üç kademedeki şiddet etiketlemesi sırasında araç hızının dikkate alınmadığını, sadece hissedilen dönüş sertliğine göre karar verildiğini belirtmek gerekir. Bu durum Tablo 4'de gösterilen sonuçlara yansımıştır.

Tablo 4. Algoritmanın manevra şiddet doğruluğu  
(Maneuvering intensity accuracy of the algorithm)

	Toplam Yanal Manevra Sayıları			
	TR	TL	LCR	LCL
Düşük	41	37	42	
Orta	11	16	14	
Yüksek	1	3	5	
Manuel kayıtlarda algoritmanın yakalama oranları				
Düşük	11/14	8/10	5/14	
Orta	5/7	4/7	2/7	
Yüksek	0/1	0/3	-	
Doğruluk(%)	72,72	60	33,33	
Puan birimi	GAS		GSS	

## 6. DEĞERLENDİRMELER (CONCLUSIONS)

Çalışma ile karayollarında insan ve yük taşımacılığı yapan ağır vasıta araçlarda kullanımı zorunlu olan takografların sürüşlerle ilgili hız ve süre denetiminin yanı sıra sürüşün nasıl yapıldığına dair agresiflik derecesi hakkında bilgi verilmesi sağlanarak sürücü kaynaklı kazaların ve yakıt tüketiminin azaltılması hedeflenmiştir. Çalışmada sürücünün inisiyatifinde olan akıllı telefon yerine ağır vasıta araçlarda kullanımı zorunlu olan takograf cihazlarının kullanılması çalışmaya pratik hayata uygulanabilirlik özelliği katmaktadır. Literatürdeki manevra sınıflandırma çalışmalarında genellikle olasılık, desen eşleştirme ve yapay zekâ teknikleri kullanılmıştır. Bu yöntemler çoğunlukla yüksek hafıza ve işlem

kapasitesine sahip cihazlar gerektirir. Bu çalışmada geliştirilen algoritma düşük işlem ve hafıza kapasiteli takograf cihazına bile kolaylıkla uygulanabilmektedir.

Önceki çalışmalardan farklı olarak manevra başlangıç ve bitiş noktalarının daha yüksek doğrulukla yakalanmasına yönelik yeni bir kenar algılama yöntemi sunulmuştur. Önceki çalışmalardaki en önemli problemlerden biri şerit değişikliklerinin algılanmasındaki yetersizlikti. Geliştirilen algoritma ile kenarların şerit değişikliği açısından öncelikle incelenmesi sayesinde şerit değişiklikleri daha yüksek oranda yakalanmaktadır. Şerit değişikliklerinin dönüş manevrası olarak kaydedilmesi mümkün olsa da kaçırılma oranı düşüktür. Dönüş manevraları çok yüksek doğrulukla yakalanmaktadır.

Yanal manevraların agresif olmayanları da algoritma tarafından yakalanabildiği için manevra agresiflik puan skalasının doğruluğu yüksektir. Manevra agresiflik derecesinin puanlamasında hız da formüle dahil edildiğinden kaza riski daha gerçekçi bir şekilde tespit edilmektedir. Manevraların agresiflik puanlamasının yapılması, sürüşlerin puanlanabilmesinin, sonrasında sürücünün puanlanabilmesinin yolunu açacaktır.

#### KAYNAKLAR (REFERENCES)

- [1] Internet: T.S. Institute, <https://data.tuik.gov.tr/Bulten/Index?p=Karayolu-Trafik-Kaza-Istatistikleri-2019-33628>, 19.5.2021.
- [2] Internet: C. I. R. (EU), Implementing Regulation (EU) No 165/2014 of the European Parliament and of the Council laying down the requirements for the construction, testing, installation, operation and repair of tachographs and their components, [https://www.legislation.gov.uk/eur/2016/799/pdfs/eur\\_20160799\\_2018-04-17\\_en.pdf](https://www.legislation.gov.uk/eur/2016/799/pdfs/eur_20160799_2018-04-17_en.pdf), 19.5.2021.
- [3] Z. Ouyang, J. Niu, Y. Liu ve J. Rodrigues, "Multiwave: A Novel Vehicle Steering Pattern Detection Method based on Smartphones", **IEEE International Conference on Communications (ICC)**, Kuala Lumpur, Malezya, 1-7, 22-27 Mayıs 2016.
- [4] T. Pholprasit, W. Choochaiwattana ve C. Saiprasert, "A Comparison of Driving Behaviour Prediction Algorithm Using Multi-Sensory Data on a Smartphone", **IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)**, Takamatsu, Japonya, 1-6, 1-3 Haziran 2015.
- [5] D. A. Johnson ve M. M. Trivedi, "Driving style recognition using a smartphone as a sensor platform", **14th International IEEE Conference on Intelligent Transportation Systems (ITSC)**, Washington, DC, ABD, 1609-1615, 5-7 Ekim 2011.
- [6] J. Engelbrecht, M. J. Booysen, G-J. van Rooyen ve F. Bruwer, "Performance Comparison of Dynamic Time Warping (DTW) and a Maximum Likelihood (ML) Classifier in Measuring Driver Behavior with Smartphones", **IEEE Symposium Series on Computational Intelligence**, Cape Town, Güney Afrika, 427-433, 7-10 Aralık 2015.
- [7] B. Fernandes, V. Gomes, J. Ferreira ve A. Oliveira, "Mobile Application for Automatic Accident Detection and Multimodal Alert", **IEEE 81st Vehicular Technology Conference (VTC Spring)**, Glasgow, Birleşik Krallık, 1-5, 11-14 Mayıs 2015.
- [8] H. Eren, S. Makinist, E. Akin ve A. Yılmaz, "Estimating Driving Behavior by a Smartphone", **Intelligent Vehicles Symposium**, Alcalá de Henares, İspanya, 234-239, 3-7 Haziran 2012.
- [9] Y. Wang, T. Zhao, F. Tahmasbi, J. Cheng, Y. Chen ve J. Yu, "Driver Identification Leveraging Single-turn Behaviors via Mobile Devices", **29th International Conference on Computer Communications and Networks (ICCCN)**, Honolulu, HI, ABD, 1-9, 3-6 Ağustos 2020.
- [10] F. Li, H. Zhang, H. Che ve X. Qiu, "Dangerous Driving Behavior Detection Using Smartphone Sensors", **IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)**, Rio de Janeiro, Brezilya, 1902-1907, 1-4 Kasım 2016.
- [11] J. Paefgen, F. Kehr, Y. Zhai ve F. Michahelles, "Driving Behavior Analysis with Smartphones: Insights from a Controlled Field Study" **Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia**, Ulm, Almanya, 1-8, 4 Aralık 2012.
- [12] Y. A. Alqudah ve B. H. Sababha, "A Statistical Approach to Estimating Driving Events by a Smartphone", **International Conference on Computational Science and Computational Intelligence**, Las Vegas, NV, ABD, 1021-1025, 15-17 Aralık 2016.
- [13] X. Liu, H. Mei, H. Lu, H. Kuang ve X. Ma, "A Vehicle Steering Recognition System Based on Low-Cost Smartphone Sensors", **MPDI sensors**, 17(3), 633, 2017.
- [14] S. Daptardar, V. Lakshminarayanan, S. Reddy, S. Nair, S. Sahoo ve P. Sinha, "Hidden Markov Model based Driving Event Detection and Driver Profiling from Mobile Inertial Sensor Data", **IEEE Sensors**, Busan, Kore (Güney), 1-4, 1-4 Kasım 2015.
- [15] A. Bhatt, V. Dave, Y. Panchamia ve P. P. Thakre, "Analyzing Behavioral Attributes of Drivers and Implementing Safe Driving Model", **IEEE International Conference on Vehicular Electronics and Safety (ICVES)**, Viyana, Avusturya, 228-232, 27-28 Haziran 2017.
- [16] H. R. Eftekhari ve M. Ghatee, "Hybrid of discrete wavelet transform and adaptive neuro fuzzy inference system for overall driving behavior recognition", **Transportation Research Part F: Traffic Psychology and Behaviour**, 58(1), 782-796, 2018.
- [17] G. Castignani, T. Derrmann, R. Frank ve T. Engel, "Driver Behavior Profiling Using Smartphones: A Low-Cost Platform for Driver Monitoring", **IEEE Intelligent Transportation Systems Magazine**, 7(1), 91-102, 2015.
- [18] R. Sun, Q. Cheng, F. Xie, W. Zhang, T. Lin ve W. Y. Ochieng, "Combining Machine Learning and Dynamic Time Wrapping for Vehicle Driving Event Detection Using Smartphones", **IEEE Transactions On Intelligent Transportation Systems**, 22(1), 194-207, 2019.
- [19] E. G. Mantouka, E. N. Barmounakis ve E. I. Vlahogianni, "Identification of driving safety profiles from smartphone data using machine learning techniques", **Safety Science**, 119, 84-90, 2019.

- [20] Y.-C. Tsai, W.-H. Lee ve C.-M. Chou, "A Safety Driving Assistance System by Integrating In-Vehicle Dynamics and Real-Time Traffic Information", **IEEE 8th International Conference on Awareness Science and Technology**, Taichung, Tayvan, 416-421, 8-10 Kasım 2017.
- [21] K. B. Ahmed, B. Goel, P. Bharti, S. Chellappan ve M. Bouhorma, "Leveraging Smartphone Sensors to Detect Distracted Driving Activities", *IEEE Transactions On Intelligent Transportation Systems*, 20(9), 3303-3312, 2019.
- [22] L. Lamel, L. Rabiner, A. Rosenberg ve J. Wilpon "An improved endpoint detector for isolated word recognition", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(4), 777-785, 1981.
- [23] M. S. N. Al-Din, "Calibration and Pre-Processing Techniques for A Smartphone-Based Driving Events Identification and Classification System", **IEEE Electron Device Kolkata Conference**, Kolkata, Hindistan, 396-402, 24-25 Kasım 2018.
- [24] Internet: STMicroelectronics, Ultra-compact high-performance eCompass module: 3D accelerometer and 3D magnetometer, <https://www.st.com/resource/en/datasheet/lsm303dlhc.pdf>, 19.5.2021.
- [25] Internet: STMicroelectronics, MEMS motion sensor: three-axis digital output gyroscope, <https://www.st.com/en/mems-and-sensors/l3gd20.html>, 19.5.2021.
- [26] Internet: B. S. GmbH, BMP180 Digital pressure sensor, <https://cdn-shop.adafruit.com/datasheets/BST-BMP180-DS000-09.pdf>, 19.5.2021.
- [27] S. Fadadu, S. Pandey, D. Hegde, Y. Shi, F. C. Chou, N. Djuric ve C. V. Gonzales, "Multi-View Fusion of Sensor Data for Improved Perception and Prediction in Autonomous Driving", **2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)**, Waikoloa, HI, ABD, 3292-3300, 3-8 Ocak 2022.
- [28] G. Yenikaya, E. Düven, A. Üzgeç ve E. Yürüklü, "Sürücü Davranış Karakteristiklerinin Tanımlanması İçin Görü Temelli Bir Sürüş Sisteminin Tasarlanması", **OTEKON 2010**, 3 Haziran 2010.
- [29] Ö. Kumtepe, E. Yüncü ve G. B. Akar, "Agresif Sürüş Tespiti İçin Çok Modlu Bir Yaklaşım", **2016 24th Signal Processing and Communication Application Conference (SIU)**, Zonguldak, Türkiye, 3292-3300, 729-732 Mayıs 2016.
- [30] M. H. Z. M. Fodli, F. H. K. Zaman, N. K. Mun ve L. Mazalan "Driving Behavior Recognition using Multiple Deep Learning Models", **2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA)**, Selangor, Malaysia, 138-143, 12 Mayıs 2022.
- [31] A.Ç. Seçkin, "Otonom Araçların Görsel Eğitimi için EEG, EMG ve IMU ile Etiketleme Sistemi", *Bilişim Teknolojileri Dergisi*, 12(4), 299-305, 2019.

# Afet Risk Yönetiminde Yapay Zekâ Kullanımının Rolü

## Literatür Makalesi/Review Article

 Nur Sinem PARTİGÖÇ\*

Şehir ve Bölge Planlama Bölümü, Pamukkale Üniversitesi, Denizli, Türkiye

[npartigoc@pauedu.tr](mailto:npartigoc@pauedu.tr)

(Geliş/Received:03.02.2022; Kabul/Accepted:23.08.2022)

DOI: 10.17671/gazibtd.1067831

**Özet**— Küreselleşme etkisi altında gözlenen büyüme süreçleri ve yoğun nüfus hareketliliği nedeniyle gittikçe karmaşıklaşan kentsel faaliyetlerin yarattığı sorunlara alışlagelen yöntemlerin yanıt vermediği açıkça ortadadır. Buna ek olarak, hızlı kentleşme süreçleri ve küresel iklim değişikliğine bağlı olarak yaşanan afet olaylarındaki artış, zaman içerisinde kentlerin temel hizmet alanlarında (çevre, sağlık, eğitim, altyapı, güvenlik, vb.) yaşanan problemleri önemli ölçüde tetiklemektedir. Dolayısıyla, çoklu bir ağa dönüşen yaşam alanlarında toplumun refah düzeyinin sürdürülebilir biçimde devam ettirilebilmesi ve etkin bir afet yönetim sürecinin ortaya konulabilmesi için bilgi teknolojilerinin etkin biçimde kullanılması artık bir zorunluluk haline gelmiştir. Bu noktadan hareketle, çalışmada afet öncesi döneme referans veren Risk Yönetimi alanında olası kayıpların azaltılması ve/veya bertaraf edilmesi konusunda yapay zekâ kullanımının öneminin vurgulanması amaçlanmıştır. Çalışmanın kapsamını yapay zekâ uygulamalarında risk yönetiminin yeri, yapay zekâ kullanımının afet risklerinin azaltılması sürecindeki avantajları ve dezavantajları, uygulama örneklerinin aktarılması gibi konular oluşturmaktadır. Yöntem olarak nitel araştırma yönteminin kullanıldığı çalışmada, yapılan araştırmalar sonucunda denilebilir ki, sürdürülebilir, uzun vadede etkin, çok paydaşlı ve disiplinler arası niteliğe sahip Modern Bütünleşik Afet Yönetim sürecinde Bilgi ve İletişim Teknolojileri (BİT) kullanımının karar alma süreçlerinin temel yapı taşlarından biri haline gelmiştir ve kentsel dirençliliğin artırılmasında yapay zekâ uygulamaları kritik bir rol oynamaktadır.

**Anahtar Kelimeler**— afet risk yönetimi, yapay zekâ, kentleşme, bilgi ve iletişim teknolojileri (BİT), kentsel dirençlilik

## The Role of Artificial Intelligence in Disaster Risk Management

**Abstract**— It is clear that the usual methods do not respond to the problems based on urban activities which are becoming increasingly complicated due to the growth processes observed under the influence of globalization and intensive population mobility. In addition, due to the increase in the number of disasters associated with rapid urbanization processes and global climate change, there has been a significant increase in the problems experienced in the main service areas of cities (environment, health, education, infrastructure, security, etc.). Accordingly, the usage of information technologies effectively become almost an obligation in order to sustain the level of well-being in settlements which have turned into a multi-network in a sustainable manner and to put forward an effective disaster management process. From this point, the aim of the study is to emphasize the importance of the usage of Artificial Intelligence (AI) for reducing and/or eliminating possible disaster losses associated with Disaster Risk Management (DRM) processes. The scope of the study includes the role of risk management in AI applications, the advantages and disadvantages of the usage of AI in the disaster risk reduction process and also application examples. The method of the study is the qualitative research method. As a result of research, it can be said that the use of Information and Communication Technologies (ICT) is necessary for DRM which is sustainable, effective in the long term, multi-stakeholder and inter-disciplinary. Moreover, AI plays a critical role in increasing urban resilience.

**Keywords**— disaster risk management (DRM), artificial intelligence (AI), urbanization, information and communication technologies (ICT), urban resilience



## 1. GİRİŞ (INTRODUCTION)

Kentsel alanlarda ilişkisel sistemler (kentsel işlevler, ekonomik etki alanı, üretim – tüketim ilişkileri, kentsel ağlar, vb.) karmaşık ve çok katmanlı bir yapıya sahiptir. Küreselleşme etkisi altında gözlenen büyüme süreçleri ve yoğun nüfus hareketliliği nedeniyle karmaşa, suç, trafik, sağlık, atık, enerji, su ve hava kirliliği, çarpık kentleşme, kaynaklara ulaşım gibi konuların yerel yönetimlerin gündeminde öncelikli konular arasında yer aldığı bilinmektedir ve bu konuların yol açtığı sorunların çözümüne alışlagelen yöntemlerin (kentsel dönüşüm uygulamaları, parçacıl planlama uygulamaları, sosyo-ekonomik düzeye göre hizmetlere erişim imkânı, vb.) yanıt vermediği gözlenmektedir. Hızlı kentleşme süreçleri ve küresel iklim değişikliğine bağlı olarak yaşanan afet olaylarındaki artış, zaman içerisinde kentlerin temel hizmet alanlarında (çevre, sağlık, eğitim, altyapı, güvenlik, vb.) yaşanan problemleri önemli ölçüde tetiklemektedir. Buna ek olarak, kırsal alandan kentsel alana göçlerin yaşanması, bölgeler arasında sosyo-ekonomik dengelerin bozulması, doğal kaynakların hızlı tüketimi, çevrenin tahribatı ve kentsel gelişme süreçlerinin planlama çerçevesinde yürütülmemesi gibi birçok olumsuz gelişme günümüz kentlerini kaotik bir ortama sürüklemektedir [1, 2].

Dolayısıyla, çoklu bir ağa dönüşen yaşam alanlarında toplumun refah düzeyini yükseltebilmek ve sürdürülebilirliğini sağlamak için, nüfus, kaynak, çevre ve bilgi teknolojilerinin karşılıklı olarak düzenlenmesi artık bir zorunluluk haline gelmiştir [3]. Vatandaşların yaşam kalitesini yükseltmek ve kamusal hizmetler hakkındaki memnuniyetlerini artırmak için Bilgi ve İletişim Teknolojileri (BİT)'nin yaşamın her alanına entegre edilmesi, yerel düzeyde ilişkisel sistemlere dair uygulanan yöntemler konusunda eksen değişikliği yaşandığının önemli bir kanıtıdır [4].

EM-DAT (Emergency Events Database) raporlarına göre, acil bir durumun afet olarak değerlendirilebilmesi için belli başlı kriterleri sağlaması gerekmektedir. Buna göre, meydana gelen afette 10 veya daha fazla kişinin ölmüş olması, 100 veya daha fazla kişinin durumdan etkilenmiş olması ve/veya ilgili devletlerin olağanüstü hal bildirisinde veya uluslararası yardım çağrısında bulunmuş olması kriterlerinden en az birinin gerçekleşmesi beklenmektedir. 1965-2020 yılları arasında farklı coğrafyalarda meydana gelen doğal ve beşeri afetlerde dünya çapında 4,5 milyon ölüme sebep olan 20.533 afet meydana gelmiştir. Bu afetlerin %62'si doğal afet niteliği taşıırken, %38'i ise beşeri afetlerdir [5, 6].

Geçmiş dönemlerden günümüze kadar dünya çapında önemli kayıplara sebep olan afetler yaşanmıştır ve halen devam etmektedir. Afetlerin birçoğu çok kısa bir zaman dilimi içerisinde ortaya çıktığından dolayı başarılı ve etkin bir afet yönetim sürecinde hızlı planlama, karar verme ve doğru uygulamaların hayata geçirilmesi kritik bir öneme sahiptir. Genel olarak, afet yönetimi süreci karmaşık ve dinamik bir yapıya sahiptir. Kontrolü zor

olan bu yapı için sistemlerden gelen verilerin depolanması, işlenmesi, haritalandırılması ve değerlendirilmesi için teknolojinin sağladığı olanaklara ihtiyaç vardır. Yenilikçi teknolojilerden yararlanılması ise, afet sonrasında olduğu kadar, afet öncesinde de riskin azaltılması adına oldukça önemli kazanımlar sağlayacaktır [5, 7-9].

Dünya genelinde yoğun nüfus ve yapılaşma pratiklerinin gözlemlendiği günümüzde, afetlerin sayısının ve sıklığının her geçen gün artmakta olduğu ve küresel iklim değişikliği sebebiyle farklı coğrafi bölgelerde çeşitlenen afet türlerinin görüldüğü açıkça ortadadır. Kentsel alanların afet riski söz konusu olduğunda çarpan etkisi yaratması sonucunda, etkin bir afet yönetimi ihtiyacı söz konusu olmuştur. Sürdürülebilir, uzun vadede etkin, çok paydaşlı ve disiplinler arası biçimde sürecin yönlendirildiği Modern Bütünleşik Afet Yönetimi sürecinin hayata geçirilebilmesi için dünya genelinde kabul görmüş bir genel kanı vardır: Afet yönetimi alanında BİT'in kullanımının gerekliliği [8, 10, 11].

Mekânsal ve mekânsal olmayan verilerden yararlanılarak, afetler karşısında olası risklerin öngörülmesi, gerekli analiz süreçlerinin yürütülmesi, hazırlıklı olma, müdahale ve iyileştirme aşamaları için gerekli eylem planları, politikalar ve uygulamaların hayata geçirilmesi gibi tüm aşamalarda BİT'in avantajlarından yararlanılmaktadır. Kentlerde gözlenen karmaşık yapıdaki ilişkisel sistemlere dair problemlerin çözülmesi ve afet riskinin minimum düzeye indirgenebilmesi için teknoloji tabanlı uygulamalar günümüzde sıkça kullanılan temel kaynaklar haline gelmiştir [5, 12-14].

Bu noktadan hareketle, bu çalışmada afet öncesi döneme referans veren Risk Yönetimi alanında olası kayıpların azaltılması ve/veya bertaraf edilmesi konusunda yapay zekâ kullanımının öneminin vurgulanması amaçlanmıştır. Afet risklerinin azaltılması konusunda güncel teknolojik uygulamaların genel ismi olan Yapay Zekâ uygulamalarında risk yönetiminin yeri, yapay zekâ kullanımının afet risklerinin azaltılması sürecindeki avantajları ve dezavantajları, uygulama örneklerinin aktarılması gibi konular çalışmanın genel kapsamını oluşturmaktadır. Yöntem olarak nitel araştırma yönteminin kullanıldığı çalışmada, yapılan araştırmalar sonucunda denilebilir ki, sürdürülebilir, uzun vadede etkin, çok paydaşlı ve disiplinler arası niteliğe sahip Modern Bütünleşik Afet Yönetim sürecinde BİT kullanımının karar alma süreçlerinin temel yapı taşlarından biri haline gelmiştir ve kentsel dirençliliğin artırılmasında yapay zekâ uygulamaları kritik bir rol oynamaktadır.

## 2. AFET RİSK YÖNETİMİNDE YAPAY ZEKÂ KULLANIMI (THE USAGE OF ARTIFICIAL INTELLIGENCE IN DISASTER RISK MANAGEMENT)

Etkin bir afet yönetim sürecinde, örgütsel yapılanmanın ve operasyonel kapasitenin hem merkezi düzeyde hem de yerel düzeyde gittikçe daha önemli bir hale geldiği açıkça

ortadadır. Önceki dönemlerde uydu sistemleri, televizyon, radyo, telsiz, SMS uygulamaları afet yönetiminde etkin araçlar iken; günümüzde Nesnelerin İnterneti (IoT), Yapay Zekâ, robotlar, akıllı sistemler gibi yeni teknolojiler devreye girmiştir. Afet yönetimi sürecinde taraflar arasında iletişimin sağlanması ve veri temini konusunda kapasitenin artırılması bakımından teknolojik olanaklardan yararlanılması oldukça önemli ve tercih edilen bir durum olmuştur. Covid – 19 salgın süreci, kriz yönetimi bakımından bu tespite verilebilecek başarılı örneklerden biridir [8, 15].

Afet öncesi aşamada afet risklerinin belirlenmesi ve azaltılması ile afet sonrası aşamada hızlı ve etkin bir iyileşme sürecinin yürütülmesinde yeni teknolojik araçların rolü ve sağladığı fırsatlar önemlidir. Modern Bütünleşik Afet Yönetimi kapsamında ele alınan bu konuların hayata geçirilmesi aşamasında Coğrafi Bilgi Sistemleri (CBS), Sosyal Ağ Servisleri (*Social Networking Services - SNS*) ve BİT gibi uygulama araçlarından yararlanılmaktadır [9, 11]. Afet öncesi ve afet sonrası aşamalarında başvuru alan teknolojik araçların kullanım alanları Tablo 1’de sunulmuştur.

Tablo 1. Afet yönetim süreçlerinde teknolojik araçların kullanım alanları (Yazar tarafından oluşturulmuştur) (The usage areas of technological tools in disaster management processes) (Prepared by the author)

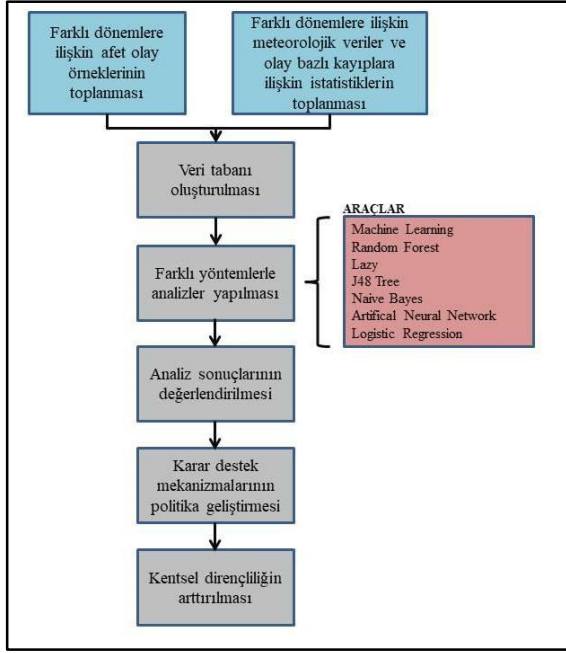
Afet Yönetim Süreci	Evre	Kullanım Alanları
Afet Öncesi Aşama	Zarar Azaltma	Risk ve tehlike analizleri
		Kırılganlığın ölçülmesi
		Kentsel sosyal parametreler
		Altyapı ve erişim olanakları
	Hazırlık	Tahliye ve kaçış planları
		Tatbikatlar ve eğitimler
		Erken uyarı sistemleri
		Acil durum iletişim sistemleri
		Taraflar arası koordinasyon
Afet Sonrası Aşama	Müdahale	Yönetim platformları oluşturulması
		Blok zincir tabanlı izleme
		Akıllı ses ve görüntü tarama
		5G teknolojisi
		Big Data uygulamaları
	İyileştirme	Hasar tespit projeleri
		Yeniden inşaat projeleri
		Veri analizleri

Yapay zekâ, insanın düşünme yeteneğini ve beynin çalışma özelliğini modellemeye çalışan, insan zekâsını bilgisayar aracılığıyla taklit etmeye çalışan, belirli bir durum veya problem karşısında gerekli bilgileri kullanarak çözüm üretebilen yöntemin genel adıdır. 1940’lı yıllarda üzerinde ilk çalışmalar yapıldığı yapay zekâ kavramı, 1990’lı yıllara gelindiğinde insan biçimli robotların yapılması gibi önemli bir noktaya taşınmıştır. Veriye ulaşmanın her zaman ve tamamen mümkün olmadığı durumlarda önerilen yapay zekâ çözümleri, veri tabanlarının kullanımı ve farklı uzmanlar tarafından geliştirilmesinde lokomotif görevi görmesi bakımından oldukça avantajlı bir tekniktir. Bu bakımdan denilebilir ki, yapay zekâ kullanımının artan afet riskleri karşısında etkili bir risk yönetim süreci sunması ve karmaşık ilişkisel sistemlerin çözümünde yardımcı olması bakımından önemli bir araçtır [7, 13].

İklim değişikliğine bağlı olarak gözlenen mevsimsel anomaliler (deniz seviyesinde artış, aşırı ve sağanak yağışlar, sıcaklık değişimleri, afet olaylarının artışı, vb.) sebebiyle pek çok açıdan (fiziksel, ekonomik, sosyal, çevresel, vb.) daha kırılgan hale gelen yapılı çevrenin afetler karşısında hazırlıklı olma ve uyum sağlama hali olarak tanımlanan kentsel dirençliliğin artırılmasında yapay zekâ kullanımlarının temel olarak 3 önemli işlevi vardır: (a) Afetlere ilişkin tahminlerin yapılması, (b) afetlere ilişkin etki değerlendirme yapılması ve (c) afetler karşısında toplumsal dirençliliğin artırılması. Bu işlevler doğrultusunda sıklıkla başvuru alan yapay zekâ araçları arasında Yapay Sinir Ağı (*Artificial Neural Network - ANN*), Tekrarlayan Sinir Ağı (*Recurrent Neural Network - RNN*), Konvolüsyonel Sinir Ağı (*Convolutional Neural Network - CNN*), İlişki Ağı (*Relation Network - RN*) ve Makine Öğrenmesi (*Machine Learning - ML*) yer almaktadır. Bu araçların uygulamada karşılık bulduğu alanlar incelendiğinde ise, Yapay Sinir Ağı’nın sosyal medya kullanımıyla ön plana çıktığı, Tekrarlayan Sinir Ağı’nın değişikliklerin izlenmesi yöntemiyle afet öncesi ve sonrasında karşılaştırılması alanında kullanıldığı, Coğrafi Bilgi Sistemleri ve Bilgi ve İletişim Teknolojileri’nin ise afetlerin etkilerinin tahmin edilmesi amacıyla kullanılması gibi örnekler verilebilir [12, 16, 17].

Afet risk yönetimi alanında yapay zekâ kullanımına yönelik uygulama örnekleri incelendiğinde, afet türlerine ve yapılmak istenen analize göre kullanılacak yöntemin farklılaştığı görülmektedir. Örneğin Makine Öğrenmesi yöntemi kullanılarak, geçmişten bugüne kadar meydana gelen afet olaylarına ilişkin verilerin toplanması, sınıflandırılması ve gelecek öngörülerinin ortaya konulması adına modelleme çalışmaları söz konusu olmaktadır. Bu yöntem, yapay zekâ kapsamında en çok başvuru alan yöntemlerden biridir. Bunun dışında, modelleme çalışmalarını yapabilmek adına Random Forest (Rassal Orman - RF), Lazy, J48 Decision Tree (J48 Karar Ağacı), Yapay Sinir Ağı (YSA), Naive Bayes (NB), Lojistik Regresyon (LR) gibi diğer yöntemlere de başvurulmaktadır [16,17]. Bağımlı ve bağımsız değişkenlerin belirlenerek afet türüne göre farklı tematik

analizlerin yapılması bu yöntemler kullanılarak mümkün olmaktadır. Afet risk yönetimi alanında yapay zekâ uygulamalarının kullanımı ve kullanılan yöntemlere yönelik akış şeması Şekil 1’de sunulmuştur.



Şekil 1. Afet risk yönetimi alanında yapay zekâ uygulamalarının kullanımı (Yazar tarafından oluşturulmuştur) (The use of artificial intelligence applications in disaster risk management) (Prepared by the author)

Farklı coğrafyalara özgü topografik, meteorolojik, doğal ve yapı çevre koşullarına bağlı olarak ortaya çıkan doğal ve/veya beşeri afetlere ilişkin başarılı yöntem süreçlerinin yürütülebilmesi adına yapılmakta olan risk azaltma çalışmalarında IoT, CBS, Uzaktan Algılama (UA), sensörler, Radyo Frekans Tanımlama Sistemleri (RFID) gibi pek çok teknolojik olanaktan yararlanılmaktadır. Bu olanakların kullanım alanları şu şekilde sıralanabilir [5, 55]:

- Nesnelerin İnterneti (Internet of Things) ve Uzaktan Algılama (Remote Sensing): Afet senaryolarının yerleşimler özelinde belirlenmesi, oluşturulan risk ve tehlike haritalarına göre mikrobölgeleme çalışmalarının yapılması, yapısal risk durumunun tespit edilmesi, olası afet durumlarına ilişkin maliyet tahminlerinin yapılması.
- Radyo Frekans Tanımlama Sistemleri (Radio Frequency Identification), Görüntü İşleme (Image Processing), Karar Destek Sistemleri (Decision Support Systems) ve Sosyal Medya (Social Media): Acil durum planlarının oluşturulması, geçici barınma alanlarının tespit edilmesi, ulaşım – sağlık – barınma – teknik ekipman sorunlarının çözülmesi, kurumlar arası koordinasyon ve eşgüdümün sağlanması.
- Coğrafi Bilgi Sistemleri (Geographic Information Systems): Hasar tespit çalışmaları, tahliye koridorlarının belirlenmesi, acil müdahale ekiplerinin

sahada doğru yönlendirilmesi, karar destek sistemleri oluşturulması.

Buna ek olarak, özellikle acil durum halinde Büyük Veri (Big Data) ve Veri Analitiği yöntemlerine başvurulduğu da bilinmektedir. Bu tür durumlarda veri temini üç farklı şekilde yapılmaktadır [9]:

- Direkt veriler** (Kullanıcı bazlı ve belirli bir bölgeden toplanan veriler),
- Otomatik veriler** (Sistem üzerinden pasif olarak elde edilen veriler)
- Gönüllü verileri** (Sosyal medyada aktif kişilerin paylaştığı veriler)

Her ne kadar son dönemlerde dünya genelinde sıklıkla gerçekleşmekte olan afetler karşısında bilgi teknolojilerinin kullanımının gerekliliğine yönelik farkındalık artmış olsa da, istatistiksel olarak bilgi teknolojilerine ve yapay zekâ uygulamalarına beklenen düzeyde başvurulmadığı görülmektedir. Dünya genelinde bu ortalama %7 seviyesinde iken, afet temelli uygulamaların web sitesi üzerinden kullanımı ortalama %20 ve akıllı telefonlar üzerinden kullanımı ortalama %7 olarak tespit edilmiştir. Ülke özelinde incelendiğinde ise, ABD ve Çin’de afet risklerinin azaltılması amacıyla teknolojinin olanaklarından yararlanma oranı %14 – 15 seviyelerinde iken, pek çok ülkede (Avustralya, Yunanistan, Tayvan, İngiltere, Almanya, İtalya, Japonya) bu ortalamanın %3 – 5 civarında olduğu gözlenmektedir [12, 19].

### 3. YAPAY ZEKÂ KULLANIMIYLA ÖNE ÇIKAN AVANTAJLAR VE DEZAVANTAJLAR (ADVANTAGES AND DISADVANTAGES STANDING OUT WITH THE USE OF ARTIFICIAL INTELLIGENCE)

Teknolojik yeniliklerin çeşitlenmesi ve yapay zekâ kullanımının yaygınlaşmasının doğal bir sonucu olarak, günümüzde dünya çapında ‘ikinci bilgi devrimi’ yaşanmaktadır. Bu devrimin zamanla kendini daha fazla gösterdiği alanlardan biri ise afet risk yönetim süreçleridir. Akademik yazında ve güncel uygulamalarda vurgulandığı üzere, web ve bulut tabanlı platformlar üzerinden gerçekleştirilen veri paylaşımıyla afet riskinin azaltılması ve sakınım planlaması sürecinin başarıyla yönetilmesi sağlanabilecektir. Bunun yanı sıra, sürdürülebilir kalkınma hedeflerine ulaşılmasının önemli araçları arasında Nesnelerin İnterneti, blok-zincir teknolojisi, çeşitli açık kaynaklar, dronlar ve diğer birçok teknolojik olanak sayılabilir. Bu noktadan hareketle, denilebilir ki, yapay zeka dahil olmak üzere, pek çok teknolojik yeniliğe akıllı kentlerin oluşturulması ve bu kentlerin sürdürülebilirliğinin sağlanması için ihtiyaç vardır [20].

Hem akademik yazında vurgulanan hem de uygulama alanlarında tecrübe edilen bilgilerden yararlanılarak, afet risk yönetim sürecinde yapay zekâ kullanımının avantajları ve dezavantajları üzerinde durulmuştur. Yapay zekâ uygulamalarına afet risk yönetim süreçlerinde

başvurulmasının sağladığı avantajlar şu şekilde listelenebilir [12, 55]:

- Afetin olası kayıplarının minimize edilmesi için kapsamlı ve sistematik bir yaklaşım geliştirilmesi
- Afetin olası kayıplarının minimize edilmesi için teknolojik olanaklardan yararlanılması
- Yerel ve merkezi yönetimlerin bilgi teknolojilerinden yararlanarak afet planlama ve risk yönetimi alanında politika ve stratejiler geliştirmesi
- Afet sürecinde insan davranışları hakkında daha fazla bilgi edinilmesi
- Afet sonrası arama ve kurtarma faaliyetlerinde ve enkaz çalışmaları için teknolojik olanaklardan yararlanılması (Drone, insansız hava taşıtları, vb.)
- Bilgi teknolojilerinin kentleşme süreçlerine yeni bakış açıları kazandırması
- Sürdürülebilir kentsel gelişmeye katkı sağlaması ve kamu hizmetlerinin kolaylaşması
- Kentin mevcut potansiyellerinin ve limitlerinin saptanması
- Büyük Veri (*Big Data*) yöntemiyle afet öncesi ve sonrasında ulaşım ağlarına ilişkin gerçek zamanlı ve anlık bilgiye ulaşılabilmesi
- Kaynak atama çalışmalarının afet öncesi ve sonrasında yapılabilmesi
- Öneri fayda – maliyet analizlerine olanak sağlaması
- Afet senaryolarının oluşturulması ve olası hasar tespitlerinin yapılabilmesi.

Yapay zeka uygulamalarına afet risk yönetim süreçlerinde başvurulmasının ortaya çıkardığı dezavantajlar ise şu şekilde sıralanabilir [11, 20, 55]:

- Kentsel dirençliliğin tesis edilmesinde teknolojik olanakların tek başına yetersiz kalması
- Teknolojilerin ilerleme hızı ile politika geliştirme hızının aynı olmaması
- Küresel ölçekte gözlenen çevresel sorunlara yerel ölçekte çözüm bulunamaması
- Mekansal veri temini ve çevresel politikalar konusunda yasal mevzuatın yetersiz kalması
- Yapay zeka uygulamalarında sınır ağları davranışlarının ve işleyişin tam olarak açıklanamaması
- Kullanıcı tecrübeleriyle sınırlandırılan bu uygulamaların afet risklerinin öngörülmesi konusunda yetersiz kalması
- Uygun parametre seçimi yapılmadığında sağlıklı çözüme ulaşamaması özelliği sebebiyle afet yönetim süreçlerinde eksiklikler ve aksaklıklar meydana gelmesi.

Bilindiği üzere, acil durum yönetimi, önceden kestirilebilen veya ani meydana gelen acil durumların gelişmesinden önce başlayan ve durumun eski haline döndürülebilmesi için yapılan çalışmalar ile bunların organizasyonunu kapsayan yönetim sürecidir. Büyük ölçekte acil durum yönetimi doğal afetler, terör saldırıları ve birçok insanı maddi manevi yönden etkileyen olayları içine almaktadır. Dinamik bir süreç olan acil durum yönetiminde birçok farklı açıdan olası durumlar doğru ve koordineli biçimde planlanmış olmalı. Başka bir deyişle, iyi işleyen bir sistemin tasarlanmış olması acil durum yönetimi için kritik öneme sahiptir.

Teknolojik olanakların kullanımı ve afet risklerinin önceden tahmin edilerek olası kayıpların azaltılması, her ne kadar yapay zekâ uzmanlık alanının konuları arasında yer alsada, listelenen avantaj ve dezavantajlar dikkatle incelendiğinde birkaç husus göze çarpmaktadır. Bunlardan ilki, afet yönetim süreçlerinin geçmişten günümüze kat etmiş olduğu mesafe ile yapay zeka uzmanlık alanının kat etmiş olduğu mesafe aynı değildir ve bu fark birtakım sistemsel aksamaları beraberinde getirmektedir. Bunun yanı sıra, 1990'lı yıllarda hız kazanan yapay zeka uygulamalarının 30 yıllık kısa bir zaman diliminde afet riskinin azaltılması konusunda yürütülen çalışmalara aktif ve etkin biçimde dahil edilebilmiş olması oldukça önemli bir kazanımdır. Kentleşme pratikleri ile teknolojik uygulamaların ortak bir özelliği olarak, sürecin işleyişi konusunda pek çok belirsizliğin olması ve sürdürülebilir bir yapı kurabilmek için çok değişkenli süreçlerin dikkatle izlenmesi ve yönetilmesi ön plana çıkmaktadır. Bu noktadan hareketle denilebilir ki, risk yönetim süreçlerinde yapay zekâ kullanımının avantajları ve dezavantajları dikkatli biçimde ve karşılaştırmalı olarak izlenmelidir.

#### 4. UYGULAMA ÖRNEKLERİ (THE APPLICATION EXAMPLES)

Literatürdeki çalışmalar incelendiğinde, günümüzde bilgisayar teknolojilerinin ve donanımlarının gelişimine paralel olarak, yapay zekâ uygulamalarının her alanda kullanımının yaygınlaştığı görülmektedir. Sistem analizi, tıp, savunma sanayi, haberleşme, üretim, otomasyon, afet yönetimi, ekonomi, karar destek sistemleri gibi pek çok alan buna örnek verilebilir [13, 16]. Afet risklerinin doğru ve etkin biçimde yönetilebilmesi konusuna odaklanan bu çalışmada, risk yönetimi sürecinde birçok sistemin (Nesnelerin İnterneti, Uzaktan Algılama, Coğrafi Bilgi Sistemleri, Radyo Frekansı Tanımlama Sistemleri, sensörler, vb.) kullanıldığı ifade edilmektedir [5]. Çalışma kapsamında, farklı içerikte üretilen çalışmalarda teknolojik olanaklardan nasıl ve ne ölçüde yararlandığı ortaya konulmuştur. Buna ek olarak, Tablo 2'de afet yönetimi aşamalarına ait faaliyetler ve bu amaçlara yönelik yapılan uygulamalara yer verilmiştir.

Tablo 2. Afet yönetimi aşamalarına ilişkin faaliyetler ve uygulamalar (Yazar tarafından oluşturulmuştur) (Activities and applications related to disaster management process) (Prepared by the author)

İlgili Aşama Faaliyetleri	Kullanılabilecek Sistemler ve Uygulamaları
<b>Zarar Azaltma Aşaması</b>	
Afet senaryolarının oluşturulması	Uzaktan algılama teknolojisi ile hasarlı veya yeterli dayanıklılığa sahip olmayan yapıların belirlenmesi  Nesnelerin interneti kullanımı ve sosyal medya verileri ile afet öncesinde erken uyarı sistemlerinin kurulması  Önceki afet kayıtları ve görüntülerinin işlenmesi ile hasar tahmin çalışmalarının yapılması
Afet durumunda etkileyecek kişilerin ve bölgelerin belirlenmesi	
İlgili bölgeler ve bu bölgelerdeki binaların, yapıların risk analizinin yapılması	
Afet senaryolarının üretilmesi ve çözüm yollarının geliştirilmesi	
Olası bir afet durumu için gerekebilecek bütçenin belirlenmesi	
Erken uyarı sistemlerinin geliştirilmesi	
Bina, yol, yapıların güçlendirilmesi	
<b>Hazırlık Aşaması</b>	
Acil durum planlarının oluşturulması	RFID etiketleme ile mevcut malzeme bilgilerine ulaşılması ve malzeme yönetimi yapılması  Görüntü işleme teknolojilerinin kullanımı ile yol ağlarının kontrol edilmesi  Özel geliştirilmiş karar destek sistemleri ile ilgili tüm paydaşları bir araya getiren, bilgi paylaşımını sağlayan platformlar oluşturulması  Sosyal medya veri işleme sistemi kurularak erken uyarı ve yer belirleme yapılabilmesi  Sosyal medya veri işleme sistemi kurularak nüfusun durumuna göre kaynak atama optimizasyonu yapılması  Afet sonrasına ilişkin insani yardımların tahminlerinin yapılması için karar ağacı yöntemi kullanılması ve karar verme sürelerinin kısaltılması
Acil tıp merkezleri için yer tespitlerini yapılması	
Geçici acil yardım merkezlerinin yer belirlemelerinin yapılması	
Afet senaryolarına göre gerekebilecek malzeme ihtiyaçlarının belirlenmesi	
Afet senaryolarına göre gerekebilecek insan kaynaklarının planlanması	
Yaralı taşıma, kurtarma operasyonlarının planlanması	
Afet durumları için gereken iletişim ağı alt yapılarının geliştirilmesi	
Gereken araç, ekipman, teknoloji seçimlerinin yapılması	
Personel eğitimleri, toplum bilinçlendirme çalışmalarının yapılması	
Gerekli bütçeleme çalışmalarının yapılması	
Sığınak yerlerinin belirlenmesi, ulaşım ağlarının oluşturulması	
Devlet, kurum, kuruluş koordinasyon sisteminin oluşturulması	

Peña-Mora vd. (2008), çalışmalarında afet yönetiminde hazırlık ve müdahale aşamalarını dikkate alarak bina değerlendirme çalışmaları üzerinde durmuşlardır. Çalışmalarında 'collaboration for preparedness, response and recovery (CP2R)' sistemi ile belirledikleri ihtiyaçları karşılayacak ve bina değerlendirme çalışmalarını destekleyecek bir bilgi sistemi önerisinde bulunmuşlardır [21]. Alamdar vd. (2016) çalışmalarında sensörlerin ve onların verilerinin afet yönetiminde farklı organizasyonlara nasıl karar desteği sağlayabileceğini incelemektedir. Çalışma Avusturya'da sel felaketini dikkate alan örnek bir çalışma üzerinden işlenmiştir. Afet yönetiminde karşılaşılan zorlukları üç maddeye ayıran çalışmada sensör web tipine dayalı yeni bir yaklaşım önermişlerdir [22].

Gill vd. (2016) ileri yaştaki insanları acil durumlarda bilgilendirebilecek IoT özellikli, yönlendirmeli bilgi sistemi yaklaşımı önerisinde bulunmaktadır. Çalışmanın amacı, acil durumların sıkça rastlandığı Avusturya'da araştırmacılar tarafından ileri yaştaki insanlar için bu gibi bilgilendiricilerin dikkate alınmasını sağlamaktır. Çalışmada farklı mimari yapılar önerilmiştir ve bir örnekle beraber prototip çalışması yapılmıştır [23]. Ray vd. (2017) çalışmalarında afet yönetiminde IoT'nin kullanılabilirliğini araştırmıştır. Afet türleri yönetim, izleme, analiz ve tahmin açılarından değerlendirilmiştir. Afet yönetiminde IoT destekli protokollerle ilgili geniş bilgi verilen çalışmada birçok afet türünü incelemiş ve var olan uygulamalardan örnekler vererek karşılaştırmalar yapmışlardır. Afet türüne göre sistem tasarımıyla ilgili bilgilerin verildiği çalışmada, afet yönetiminde IoT kullanımının zorlukları ve gelecek yönelimleri ele alınmıştır [24].

Ray ve Turuk (2017), afet durumunda zamanın önemini vurguladıkları çalışmalarında afet sonrası iletişim için bir öneride bulunmuşlardır. Ağın afet durumunda karşılaşılabileceği zorluklar, kısıtlamalar açıklanarak önerilen sistemin var olanlara göre karşılaştırması yapılmıştır [25]. Lv vd. (2017), çalışmalarında elektronik hükümet bulut hizmet modeli önerisinde bulunmuşlardır. Afet yönetimi için kapsamlı olarak tüm aşamalar üzerinde geliştirilmiş karar destek sistemi önermişlerdir. Önerilen sistem teknolojik yapılar ve uygulamalar olarak farklı katmanlar halinde ele alınıp her katmanın ayrıntılı açıklamasına yer verilmiştir [26].

Landwehr vd. (2016) ve Ai vd. (2016) karar destek sistemi önerdikleri çalışmalarında tsunami durumu için zarar azaltma aşamasını dikkate almıştır. Çalışma CBS ve sosyal medya üzerinde yoğunlaşmıştır [27, 28]. Chung ve Park (2016), afet olduğu durumda hızlı ve güvenli olarak veri elde edip depolayabilecekleri bir IoT tabanlı 'Peer to Peer (P2P)' bulut sistemi önermişlerdir [29]. Deak vd.(2013), Yang vd.(2013) ve Wang vd.(2015) müdahale aşamasını dikkate almışlardır. Bu çalışmalarda IoT, sensör ve CBS sistemlerinden yararlanılarak karar destek ve iletişim ağ sistemi önerisinde bulunulmuştur [30-32].

Yao vd. (2017), çalışmada gönüllü verilerinden yararlanılarak şehir nüfusunun binalara göre nasıl dağıldığını farklı algoritmalar uygulayarak belirlemiştir. Çalışmanın temel amacı nüfus dağılımlarının belirlenmesi ile kaynak atama faaliyetlerinin optimizasyonunu sağlamaktır. Afet öncesi, afet durumu ve sonrası için yapılan çalışmalarda erken uyarı, afet durumunda bölgelerdeki şiddetin büyüklüğü hakkında bilgi, afet sonrasında kişilerin güvenliği, ihtiyaçları bilgilerinin edinilmesinde sosyal medya verileri kullanılabilir. Elde edilen veriler veri analitiği çalışmaları ile analiz edilmektedir [27, 33, 34, 35, 36]. Landwehr vd. (2016), atılan metin mesajlarının (tweet) analizi ile tsunami için planlama ve erken uyarı üzerine çalışmışlardır. Bu mesajlardan anahtar kelimeleri yakalayıp sınıflandırmışlardır. Sınıflandırma sonrasında istatistik ve makine öğrenmesi yöntemlerini uygulamışlardır [28].

Ma ve Zhang (2017), Büyük Veri yöntemini kullanarak işbirliği, koordinasyon, haberleşme için karar destek sistem önerisinde bulunmuşlardır [37]. Yeum vd. (2018), çalışmalarında geçmiş afet görüntülerinden yararlanılarak afet sonrası araştırma çalışmalarından elde edilen görüntülerin işlenmesi ve sınıflandırılması için çalışma yapmışlardır. Çalışmalarında görüntü sınıflandırma ve hasar belirleme çalışmasında yapay sinir ağları yöntemini uygulamışlardır [38]. Reznik vd. (2017) üç farklı bilgi sisteminden alınan verilerin tarım alanlarında acil durumlarda nasıl kullanılabileceğine dair bir çalışma yapmışlardır. Hazırlık ve müdahale aşamalarında edinilen verilerin kullanımına dair değerlendirmede bulunmuşlardır [39]. Tak ve Soo (2017), farklı sensörlerden alınan bilgilerin değerlendirilmesi ve Coğrafi Bilgi Sistemleri (CBS)'nin buna entegrasyonu ile gaz hattı üzerinde değerlendirmeyi mümkün kılan gerçek zamanlı bir sistem önermişlerdir. Afet hazırlık aşamasını dikkate alan sistem için risk değerlendirme faaliyetleri üzerinde durulmuştur [40].

Griffith vd. (2017), çalışmalarında insani yardım lojistiği faaliyetleri için öneriler sunmuşlardır. İnternet üzerinde istenilen konuları dikkate alan harita oluşturma ve Karar Ağacı yöntemleri ile karar çevrim sürelerini düşüren sistem önerisinde bulunmuştur. Çalışmada verilerin elde edilebileceği kaynaklar özellikle belirtilmiştir [41]. Lele ve Lhua (2016) kimya endüstrisi alanında oluşabilecek acil durumlar için önleyici, tüm süreç içerisinde erken uyarı sistemini inceleyecek karar destek sistemi önermişlerdir [42].

Leiras vd. (2014), afet durumunda insani yardım lojistiği faaliyetlerini içeren literatür çalışmasında afet operasyonları yönetimindeki trendler ve gelecek çalışmalar için önerilerde bulunmuşlardır [43]. Özdamar ve Ertem (2015) müdahale ve iyileştirme aşamaları, Jain vd. (2017) zarar azaltma aşaması için literatür çalışması yapmışlardır. Bu çalışmalarda sistematik yaklaşımlar izlenmiştir ve geleceğe yönelik olarak çalışılabilir alanlarla ilgili önerilerde bulunulmuştur [44, 45]. Chen vd. (2017) afet yönetim sürecinin tamamının dikkate alındığı çalışmada acil durum yönetiminde akıllı sistemler

üzerine literatür çalışması yapmışlardır. Fosso vd. (2015) afet yönetiminde Büyük Veri uygulamalarını içeren sistematik bir literatür araştırması yapmış ve konuyla ilgili önerilerde bulunmuşlardır. Goswami vd.(2016) doğal afetlerde Veri Madenciliği ve Veri Analitiği yöntemlerini dikkate alan bir literatür çalışması hazırlamışlardır. Çalışmada istatistiksel yöntemler, yapay sinir ağları, kümeleme, metin madenciliği (text mining), zaman serileri gibi yöntemler ön plana çıkartılmıştır [46-48].

Ohio kentinde (ABD) ve Hindistan'ın belirli bölgelerinde yapay zekâ destekli modellerin kurulması ve meteorolojik verilerin kullanılarak Nesnelerin İnterneti yöntemiyle yağış rejimlerinin izlenmesi ve taşkın tahminlerinin yapılmasının amaçlandığı projeler hâlihazırda yürütülmektedir. Buna ek olarak, Google Acil Durum Uyarıları programı aracılığıyla, önceki dönemlerde meydana gelen taşkınların sisteme işlenmesi ve olası taşkın tahminlerinin yapılması için Google firmasının yapay zekâ kullanımına dair bir proje önerisi geliştirdiği bilinmektedir. Yapay zekâ uygulamalarının su ile bağlantılı afetlerin yönetiminde kullanılmasına dair geliştirilen bu projede belirlenen hedefler arasında şu maddeler dikkat çekmektedir [49]:

- Hava durumu ve yağış rejimlerinin izlenmesi
- Dinamik yapıli veri tabanlarıyla periyodik risk haritalarının oluşturulması
- Kentsel gelişme stratejilerine öngörü ve bulguların dâhil edilmesi
- Risk planlama süreçlerinde senaryoların geliştirilmesi.

Akademik yazında yer alan uluslararası uygulama örneklerinin yanı sıra, Türkiye'de hayata geçirilen uygulamalar incelendiğinde ise, 'Dijital Dönüşüm Ofisi' çatısı altında farklı uzmanlık alanlarında yapay zekâ uygulamalarına öncülük etmek ve kamu hizmetlerinde Büyük Veri kullanımının yaygınlaştırılması gibi önemli çalışmalar yürütülmektedir. Ayrıca, Çevre ve Şehircilik Bakanlığı tarafından hazırlanan '2020 – 2023 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı' kapsamında akıllı şehir uygulamalarının desteklenmesi planlanmıştır. İçişleri Bakanlığı Afet ve Acil Durum Yönetimi Başkanlığı (AFAD) tarafından Coğrafi Bilgi Sistemleri araçlarından yararlanılarak 'Afet Yönetimi ve Karar Destek Sistemi (AYDES)' kurulmuştur. Buna ek olarak, AYDES-UZAL uygulamasında ise Coğrafi Bilgi Sistemleri ve Uzaktan Algılama teknolojilerinden birlikte yararlanılması planlanmıştır [50].

Teknolojik olanaklardan yararlanılarak hayata geçirilen ve afet yönetim sisteminin önemli parçasını oluşturan uygulamalara ek olarak, Türkiye'de merkezi yönetimin önderliğinde yürütülen projeler şu şekilde sıralanabilir [50]:

- Kesintisiz ve Güvenli Haberleşme Sistemi (KGHS)
- Bütünleşik İkaz Alarm Sistemi (İKAS)
- Türkiye Deprem Tehlike Haritası

- Taşkın Tahmini ve Erken Uyarı Merkezi
- Taşkın Arıza Müdahale Bilgi Sistemi (TAMBİS)
- Meteorolojik Veri Bilgi Sunum ve Satış Sistemi (MEVBİS)
- Sayısal Hava Tahmini (SHT)
- Kent Bilgi Sistemleri (KBS)
- Türkiye Ulusal Coğrafi Bilgi Sistemleri Altyapısı (TUCBS).

Ayrıca, mimari tasarım alanında yapay zekâ kullanımına ilişkin uygulamalar son dönemde sıklıkla tercih edilmektedir. Yapı taşıyıcı sisteminin kurgusunun oluşturulması amacıyla, Python ve Image Artificial Intelligence (IAI) araçları kullanılarak görüntü tanıma, işleme ve derin öğrenme tekniklerinin geliştirilmesi ve 2 boyutlu çalışmalar üretilmesi sağlanmaktadır. Bina Bilgi Sistemi (BBS) uygulamasında ise temel olarak hedeflenen, 2 boyutlu vektörel çizimlerden ziyade fotoğraf üzerinden hızlıca değerlendirmeler yapma olanağının sağlanması ve güncel veri tabanlarının oluşturularak yapı envanterinin çıkartılmasıdır. Diğer bir yenilikçi uygulama ise Düzensizlik Kontrol Asistanı (DKA) uygulamasıdır. Deprem Yönetmeliğinde yer alan plan düzlemindeki düzensizliklerin ve proje taşıyıcı sistemine dair risk oluşturacak düzensizliklerin uzmanlar tarafından incelenmesi için bu uygulama geliştirilmiştir [51].

Şehir planlama uzmanlık alanı gibi, mimarlık ve inşaat mühendisliği alanlarında da yapay zekâ kullanımının gittikçe yaygınlaştığı gözlemlenmektedir. Bu alanlarda yapay zekâ kullanımına örnekler şu şekilde sıralanabilir [51]:

- Binaların tasarım süreçlerinin erken aşamalarında yapay zekâ yöntemleri kullanılarak maliyet tahminlerinin yapılması
- İnşaat sürecinde süre tahmininin yapılması
- Geleneksel mühendislik yöntemlerinin sınırlı kaldığı yerde riski azaltmak adına karar optimizasyonu ve risk analizleri yapılması
- İmar Yönetmeliği'ne göre arsa için en uygun bina tasarımının belirlenmesi ve vaziyet planı üretilmesi
- Mimari Akıllı Tasarım (MAT) uygulamasının geliştirilmesi ve kullanılması
- Yapısal risk değerlendirmelerinin teknolojik olanaklardan yararlanılarak yapılması.

## 5. SONUÇ VE DEĞERLENDİRME (CONCLUSION AND DISCUSSION)

Dünya genelinde yoğun nüfus hareketleri ve yapılaşma pratiklerinin gözlemlendiği günümüzde, afetlerin sayısının ve sıklığının her geçen gün artmakta olduğu ve küresel iklim değişikliği sebebiyle farklı coğrafi bölgelerde çeşitlenen afet türlerinin görüldüğü açıkça ortadadır. Kentsel alanların afet riski söz konusu olduğunda çarpan etkisi yaratması sonucunda, etkin bir

afet yönetimi ihtiyacı söz konusu olmuştur. Sürdürülebilir, uzun vadede etkin, çok paydaşlı ve disiplinlerarası biçimde sürecin yönlendirildiği Modern Bütünleşik Afet Yönetim sürecinin hayata geçirilebilmesi için dünya genelinde kabul görmüş bir genel kanı vardır: Afet yönetimi alanında BİT'in kullanımının gerekliliği [8, 10, 11].

Geçmiş dönemlerden günümüze kadar dünya çapında önemli kayıplara sebep olan afetler yaşanmıştır ve halen devam etmektedir. Afetlerin birçoğu çok kısa bir zaman dilimi içerisinde ortaya çıktığından dolayı başarılı ve etkin bir afet yönetim sürecinde hızlı planlama, karar verme ve doğru uygulamaların hayata geçirilmesi kritik bir öneme sahiptir. Genel olarak, afet yönetimi süreci karmaşık ve dinamik bir yapıya sahiptir. Kontrolü zor olan bu yapı için sistemlerden gelen verilerin depolanması, işlenmesi, haritalandırılması ve değerlendirilmesi için teknolojinin sağladığı olanaklara ihtiyaç vardır. Yenilikçi teknolojilerden yararlanılması ise, afet sonrasında olduğu kadar, afet öncesinde de riskin azaltılması adına oldukça önemli kazanımlar sağlayacaktır [5, 7-9].

Günümüzün hızlı gelişme ortamında verinin toplanması, derlenmesi, saklanması, güncelleştirilmesi, organizasyonu, analizi, yönetimi ve sunumu amacıyla uzay teknolojileri temelli sistemlerin özellikle karar verme süreçlerinde tam entegrasyonlarının sağlanarak kullanılması gerekmektedir [1, 6, 52].

Afet risk değerlendirmesi, Risk Yönetim sürecinin temel yapı taşlarından biri olup, aynı zamanda karar alma süreçlerinin temel dayanaklarından biri sayılmaktadır. Afetlere ilişkin risklerin yönetilmesi konusu oldukça zorlayıcı olabilmektedir. Bunun gerekçesi, afetler gerçekleşene kadar risklerin görünmez kalması ve tam anlamıyla olası etkilerinin öngörülememesidir. Doğal bir afetin gerçekleşmesi durumunda, başarılı bir risk yönetim süreci ortaya koymak kavramsal, bilimsel ve teknik açıdan hem karar mekanizmalarını hem de afetten etkilenen tüm kesimleri zorlayıcı bir unsur haline gelmektedir. Bu sebeple, özellikle afet öncesi dönemde afet risklerine dair ortaya konulan tüm göstergeler kentin tüm paydaşlarının erişimine açık, şeffaf, kolayca anlaşılır ve güvenilir olmak durumundadır [53, 54].

Afet yönetimi içerisinde riskin belirlenmesi ve olası hasarların tahmininde genelde fiziksel faktörlerin değerlendirmeye alınması ve sosyal faktörlerin devre dışı bırakılması önemli bir sorundur. Afet temelli çalışmaların verimliliği bu sebeple olumsuz yönde etkilenmektedir. Nüfus, istihdam, sosyal ve ekonomik yapı, göç olgusu gibi parametreler Yapay Sinir Ağları (YSA) yöntemiyle risk analiz çalışmalarına dâhil edilmektedir [5,15]. Yalnızca risk ve hasar tespitinde değil, aynı zamanda mekansal ve mekansal olmayan bilgilerin afet öncesinde zarar azaltma çalışmalarında kullanılması bakımından yapay zeka uygulamalarının kritik bir rolü olduğu söylenebilir.

Sonuç olarak, teknolojik yeniliklerden yararlanılarak Risk Yönetim süreçlerinde modern yöntemler geliştirilmesi ve her türlü kaynağın (zaman, bütçe, nitelikli insan, teknik destek vb.) temin edilmesi gerekmektedir. Buna ek olarak, olası bir afet sonrasında geleneksel yöntemler kullanılarak Kriz Yönetim süreçlerine odaklanmak yerine, olası bir afet öncesinde modern yöntemlerle Risk Yönetim süreçlerine odaklanmak yerel ölçekte karar mekanizmaları için önemli kazanımlar sağlayacaktır. Başka bir deyişle, yoğun nüfus ve yapı stoğunun bulunduğu kentsel alanlarda afet risklerinin doğurabileceği olumsuz sonuçları minimize edebilmek ve/veya bertaraf edebilmek adına önemli bir eksen değişikliğine ihtiyaç vardır.

Kentsel planlama ve sürdürülebilir kalkınma hedeflerinin gerçekleştirilmesi süreçlerinde vatandaşların sağlıklı, güvenli ve nitelikli fiziksel çevrelerde ikamet edebilmeleri adına, günümüzün teknolojik olanaklarından (Yapay Zekâ, Büyük Veri, sensörler, vb.) yararlanmak yerel yönetimler tarafından önceliklendirilmesi gereken konulardan biridir. Yalnızca afet risk yönetim süreçleri değil, aynı zamanda karmaşık ve çok boyutlu kentsel planlama süreçlerinin başarılı biçimde yönetilebilmesi için yerel ölçekte kentsel altyapıya yönelik bu uygulamaların hayata geçirilmesi önemli, gerekli ve acildir. BİT temelli uygulamalara ağırlık verilmesi, mevcut afet risklerinin azaltılması ve/veya bertaraf edilmesi sürecinde gerekli politika ve eylemlerin geliştirilmesine önemli katkılar sağlayacağı açıktır.

#### KAYNAKLAR (REFERENCES)

- [1] S.S. Durduran, A. Geymen, "Türkiyede Afet Bilgi Sistemi Çalışmalarının Genel Bir Değerlendirmesi", **2. Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu (UZAL-CBS 2008)**, 344 – 352, Kayseri, 2008.
- [2] L. Lin, A. Nilsson, J. Sjolín, M. Abrahamsson, H. Tehler, "On the perceived usefulness of risk descriptions for decision-making in disaster risk management", *Reliability Engineering and System Safety*, 142, 48–55, 2015.
- [3] E. Örselli, C. Akbay, "Teknoloji ve Kent Yaşamında Dönüşüm: Akıllı Kentler", *Uluslararası Yönetim Akademisi Dergisi*, 2 (1), 228-241, 2019.
- [4] C. Harrison, I.A. Donnelly, "A Theory of Smart Cities", **Proceedings of the 55th Annual Meeting of the ISSS - 2011**, Hull, UK, 1-15, 2011.
- [5] N. Çağlayan, Ş.I. Satoğlu, E.N. Kapukaya, "Afet Yönetiminde Büyük Veri Ve Veri Analitiği Uygulamaları: Literatür Araştırması", **7. Ulusal Lojistik ve Tedarik Zinciri Kongresi (ULTZK 2018)**, Bursa, 2018.
- [6] Internet: Emergency Events Database (EM-DAT), [http://emdat.be/sites/default/files/adsr\\_2016.pdf](http://emdat.be/sites/default/files/adsr_2016.pdf), 05.01.2022.
- [7] H. Kemper, G. Kemper, "Sensor Fusion, GIS and AI Technologies for Disaster Management", **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, Volume XLIII-B3-2020, XXIV ISPRS Congress, 2020.
- [8] Y.M. More, "Disaster Management Using Artificial Intelligence", *Journal of Xi'an University of Architecture and Technology*, Volume XI, Issue XII, Issn No : 1006-7930, 2019.
- [9] L. Tan, J. Guo, S. Mohanarajah, K. Zhou, "Can we detect trends in natural disaster management with artificial intelligence? A review of modeling practices", *Natural Hazards*, 107, 2389–2417, 2021.
- [10] G. Abhijeet, D. Samir, "Information Based Approach for Disaster Risk Management", **20th International Symposium on Logistics (ISL 2015)**, Bologna, Italy, 5-8 Temmuz, 2015.
- [11] R. Corrado, "ICTs and AI-Driven Solutions for Disaster Management", *Cambodia Development Center*, 3 (10), 2021.
- [12] W. Sun, P. Bocchini, B.D. Davison, "Applications of artificial intelligence for disaster management", *Natural Hazards*, 103, 2631–2689, 2020.
- [13] D. Sürmeli, **Yapay Sinir Ağları İle Afet Yönetiminde Sosyal Zarar Görebilirlik Riskinin Belirlenmesi**, Sakarya Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi, Sakarya, 2011.
- [14] T. Yiğitcanlar, K.C. Desouza, L. Butler, F. Roozkosh, "Contributions and Risks of Artificial Intelligence (AI) in Building Smarter Cities: Insights from a Systematic Review of the Literature", *Energies*, 13, 1473, 2020.
- [15] L. Memiş, C. Babaoğlu, **Afet Yönetimi ve Teknoloji: Farklı Boyutlarıyla Afet Yönetimi** (Edt. M. Yaman ve E. Çakır), Nobel Yayınevi, Ankara, Türkiye, 2020.
- [16] World Bank, **Machine Learning for Disaster Risk Management**, International Bank for Reconstruction and Development/International Development Association, GFDRR, Washington, 2018.
- [17] A. Ayaydın, M. A. Akçayol, "Deep Learning Based Forecasting of Delay on Flights", *Bilişim Teknolojileri Dergisi*, 15 (3), 3-5, 2022.
- [18] C. Şen, İ. S. Mert, M. Abubakar, "Büyük Veri Yönetimi, Bilgi Aramada Sosyal Medya Kullanımı ve T-Yetenek Üzerindeki Etkileri", *Bilişim Teknolojileri Dergisi*, 14 (4), 3-5, 2021.
- [19] V. Nunavath, M. Goodwin, "The Role of Artificial Intelligence in Social Media Big Data Analytics for Disaster Management - Initial Results of a Systematic Literature Review", **5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)**, 2018.
- [20] S. Pirasteh, M. Varshosaz, "Geospatial Information Technologies in Support of Disaster Risk Reduction, Mitigation and Resilience: Challenges and Recommendations", **Sustainable Development Goals Connectivity Dilemma**, 1st Edition, ImprintCRC Press, 2019.
- [21] F. Peña-Mora, Z.U.H. Aziz, A. Chen, A. Plans, S. Foltz, "Building assessment during disaster response and recovery", *Proceedings of the Institution of Civil Engineers*, 161(4), 183–195, 2008.
- [22] F. Alamdar, M. Kalantari, A. Rajabifard, "Towards multi-agency sensor information integration for disaster management", *Comput. Environ. Urban Syst.*, 56, 68–85, 2016.






- [23] A.Q. Gill, N. Phennel, D. Lane, V.L. Phung, "IoT-enabled emergency information supply chain architecture for elderly people: The Australian context. Information Systems", *Information Systems*, 58, 75–86, 2016.
- [24] P.P. Ray, M. Mukherjee, L. Shu, "Internet of Things for Disaster Management: State-of-the-Art and Prospects", *IEEE Access*, 5, 18818–18835, 2017.
- [25] N.K. Ray, A.K. Turuk, "A framework for post-disaster communication using wireless ad hoc networks", *Integration, the VLSI Journal*, 58(Supplement C), 274–285, 2017.
- [26] Z. Lv, X. Li, K. Choo, "E-government multimedia big data platform for disaster management", *Multimedia Tools and Applications*, 1–13, 2017.
- [27] F. Ai, L. K. Comfort, Y. Dong, T. Znati, "A dynamic decision support system based on geographical information and mobile social networks: A model for tsunami risk mitigation in Padang, Indonesia", *Safety Science*, 90, 62–74, 2016.
- [28] P.M. Landwehr, W. Wei, M. Kowalchuck, K.M. Carley, "Using tweets to support disaster planning, warning and response", *Safety Science*, 90, 33–47, 2016.
- [29] K. Chung, R.C. Park, "P2P cloud network services for IoT based disaster situations information. Peer-to-Peer Networking and Applications", *Peer-to-Peer Networking and Applications*, 9 (3), 566–577, 2016.
- [30] G. Deak, K. Curran, J. Condell, E. Asimakopoulou, N. Bessis, "IoTs (Internet of Things) and DfPL (Device-free Passive Localisation) in a disaster management scenario", *Simulation Modeling Practice and Theory*, 35, 86–96, 2013.
- [31] W. Wang, C. Hu, N. Chen, C. Xiao, C. Wang, Z. Chen, "Spatio-temporal enabled urban decision-making process modeling and visualization under the cyber-physical environment", *Science China Information Sciences*, 58(10), 1–17, 2015.
- [32] L. Yang, S.H. Yang, L. Plotnick, "How the internet of things technology enhances emergency response operations", *Technological Forecasting and Social Change*, 80(9), 1854–1867, 2013.
- [33] S. Linardi, "Peer coordination and communication following disaster warnings: An experimental framework", *Safety Science*, 90(Supplement C), 24–32, 2016.
- [34] V.K. Neppalli, C. Caragea, A. Squicciarini, A. Tapia, S. Stehle, "Sentiment analysis during Hurricane Sandy in emergency response", *International Journal of Disaster Risk Reduction*, 21(Supplement C), 213–222, 2017.
- [35] T. Papadopoulos, A. Gunasekaran, R. Dubey, N. Altay, S.J. Childe, S. Fosso-Wamba, "The role of Big Data in explaining disaster resilience in supply chains for sustainability", *Journal of Cleaner Production*, 142(2), 1108–1118, 2017.
- [36] Y. Yao, X. Liu, X. Li, J. Zhang, Z. Liang, K. Mai, Y. Zhang, "Mapping fine-scale population distributions at the building level by integrating multi source geospatial big data". *International Journal of Geographical Information Science*, 31(6), 2017.
- [37] Y. Ma, H. Zhang, "Enhancing Knowledge Management and Decision-Making Capability of China's Emergency Operations Center Using Big Data", *Intelligent Automation and Soft Computing*, 24(1), 1–8, 2017.
- [38] C.M. Yeum, S.J. Dyke, J. Ramirez, "Visual data classification in post-event building reconnaissance", *Engineering Structures*, 155, 16–24, 2018.
- [39] T. Řezník, V. Lukas, K. Charvát, K. Charvát, Z. Krivánek, M. Kepka, L. Herman, H. Řezníková, "Disaster Risk Reduction in Agriculture through Geospatial (Big) Data Processing", *ISPRS International Journal of Geo-Information*, 6(8), 238, 2017.
- [40] P.J. Tak, K.C. Soo, "A Study on the Construction of City-Gas Smart Disaster Prevention System Based on GIS". *International Journal of Control and Automation*, 10, 2017.
- [41] D.A. Griffith, B. Boehmke, R.V. Bradley, B.T. Hazen, A.W. Johnson, "Embedded analytics: improving decision support for humanitarian logistics operations", *Annals of Operations Research*, 1–19, 2017.
- [42] Q. Lele, K. Lihua, "Technical Framework Design of Safety Production Information Management Platform for Chemical Industrial Parks Based on Cloud Computing and the Internet of Things", *International Journal of Grid and Distributed Computing*, 9(6), 299–314, 2016.
- [43] A. Leiras, I. De Brito, E. Queiroz, T. Bertazzo, H. Yoshida, "Literature review of humanitarian logistics research: trends and challenges", *J. Humanitarian Logistic. Supply Chain Manage.*, 4, 95–130, 2014.
- [44] G. Jain, A. Kulshrestha, N.L. Vyas, "Radio Frequency Identification Technology application for disaster and rescue: a review", *International Archive of Applied Sciences and Technology*, 8, 64–73, 2017.
- [45] L. Özdamar, M.A. Ertem, "Models, solutions and enabling technologies in humanitarian logistics", *European Journal of Operational Research*, 244(1), 55–65, 2015.
- [46] N. Chen, W. Liu, R. Bai, A. Chen, A., "Application of computational intelligence technologies in emergency management: a literature review", *Artificial Intelligence Review*, 1–38, 2017.
- [47] S. Fosso Wamba, S. Akter, A. Edwards, G. Chopin, D. Gnanzou, "How "big data" can make big impact: Findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, 165, 234–246, 2015.
- [48] S. Goswami, S. Chakraborty, S. Ghosh, A. Chakrabarti, B. Chakraborty, "A review on application of data mining techniques to combat natural disasters", *Ain Shams Engineering Journal*, In press, 2016.
- [49] Su Politikaları Derneği, **Yapay Zeka ve Su Yönetimi**, Rapor No: 30, Ankara, 2020.
- [50] L. Memiş, C. Babaoğlu, "Acil Durum ve Afet Yönetiminde Süreç Yaklaşımı ve Teknoloji", *Academic Review of Economics and Administrative Sciences*, 13(4) 776–791, 2020.
- [51] K. Bingöl, E.A. Akan, H.T. Örmecioglu, A. Er, "Artificial intelligence applications in earthquake resistant architectural design: Determination of irregular structural systems with deep learning and Image AI method", *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35(4), 2197–2209, 2020.

- [52] A. Maskrey, “Revisiting community-based disaster risk management”, *Environmental Hazards*, 10, 42–52, 2011.
- [53] M.L. Carreno, O.D. Cardona, A. H. Barbat, A.H., “Urban seismic risk evaluation: a holistic approach”, *Nat. Hazards*, 40 (1), 137–172, 2007.
- [54] N. Lantada, M.L. Carreno, N. Jaramillo, “Disaster risk reduction: A decision-making support tool based on morphological analysis”, *International Journal of Disaster Risk Reduction*, 42, 2020.
- [55] Inter-American Development Bank, “**Indicators of Disaster Risk and Risk Management**”, Technical Notes, No. IDB-TN-169, 2010.

# Information System Proposal to Improve Warehouse Operations in a Production System

*Araştırma Makalesi/Research Article*

 Volkan ÇETİN<sup>1</sup>,  Başak GÖK<sup>2</sup>,  Hadi GÖKÇEN<sup>3</sup>

<sup>1</sup> STM Defense Technologies and Engineering Inc., Ankara, Turkey; Dept. of MIS, Institute of Informatics, Gazi University, Ankara, Turkey

<sup>2</sup> Department of Management Information Systems, Faculty of Applied Sciences, Gazi University, Ankara, Turkey

<sup>3</sup> Department of Industrial Engineering, Faculty of Engineering, Gazi University, Ankara, Turkey

[volkan.cetin3@gazi.edu.tr](mailto:volkan.cetin3@gazi.edu.tr), [basakgok@gazi.edu.tr](mailto:basakgok@gazi.edu.tr), [hgokcen@gazi.edu.tr](mailto:hgokcen@gazi.edu.tr)

(Geliş/Received:25.02.2022; Kabul/Accepted:28.08.2022)

DOI: 10.17671/gazibtd.1079208

**Abstract**—The purpose of lean thinking is to get rid of as much waste as possible. Eliminating wasteful activities such as transport, overproduction, waiting, defects, overprocessing, inventory and movement is one of the most important prerequisites in a successful company. This concept is an integral part of lean thinking, and it helps increase profitability. Material picking costs constitute approximately 55%-60% of the total costs related to warehouse operations [1]. The purpose of this article is to develop an information system that finds the shortest path for the responsible personnel to collect the materials required for production from the stock addresses in the warehouse, determines the number of picking boxes required for these materials, and provides the material sequences within the box (in-box sequences). With this developed information system, it will be possible to reduce material picking costs by eliminating wasteful activities such as transport, waiting and movement. The implementation of the developed information system is carried out in a company that produces electromechanical products in Ankara. With the Material Picking Information Systems (MPIS) presented in this article, up to 50% improvement has been achieved in terms of the distance and time spent by the personnel and the model has been validated.

**Keywords**— information system, warehouse management, determining of material picking route, determining of material picking box quantity

## Bir Üretim Sisteminde Depo Operasyonlarını İyileştirmeye Yönelik Bilgi Sistemi Önerisi

**Özet**— Yalın düşüncenin amacı, mümkün olduğunca fazla israftan kurtulmaktır. Taşıma, aşırı üretim, bekleme, uygunsuz malzemeler, aşırı işleme, envanter ve hareket gibi katma değeri olmayan faaliyetleri ortadan kaldırmak, başarılı bir şirketin en önemli ön koşullarından biridir. Bu kavram, yalın düşüncenin ayrılmaz bir parçasıdır ve karlılığı artırmaya yardımcı olur. Malzeme toplama maliyetleri, depo operasyonları ile ilgili toplam maliyetlerin yaklaşık %55-60'ını oluşturmaktadır [1]. Bu makalenin amacı, sorumlu personelin üretim için gerekli malzemeleri depodaki stok adreslerinden toplaması için en kısa yolu bulan, bu malzemeler için gerekli olan toplama kutusu sayısını belirleyen ve kutu içindeki malzeme dizilimlerini sağlayan bir bilgi sistemi geliştirmektir. Geliştirilen bu bilgi sistemi ile taşıma, bekleme ve hareket gibi israfli faaliyetler ortadan kaldırılarak malzeme toplama maliyetlerinin düşürülmesi mümkün olacaktır. Geliştirilen bilgi sisteminin uygulaması Ankara'da elektromekanik ürünler üreten bir firmada gerçekleştirilmiştir. Bu yazıda sunulan Malzeme Toplama Bilgi Sistemleri (MPIS) ile personelin harcadığı mesafe ve zaman açısından %50'ye varan iyileştirme sağlanmış ve model valide edilmiştir.

**Anahtar Kelimeler**— bilgi sistemi, depo yönetimi, malzeme toplama rotasının belirlenmesi, malzeme toplama kutusu miktarının belirlenmesi

## 1. INTRODUCTION

Reducing the cost provides an improvement in the product price for the businesses and this provides an advantage in the marketplace. Reducing the costs in storage and material transfer movements reduces the product cost and this provides a competitive edge.

Any organization or companies have one or many warehouses. These warehouses considered most important and complex part which responsible for:

- Minimizing costs.
- Increasing the customer service level [2].

The warehouse management system purpose is the materials arrangement, management and controlled.

The preparation and distribution of raw materials, parts, semi-materials, and products for manufacture or sale is one of the core business of warehouse management. Material picking activities account for 55 percent to 65 percent of total costs associated with warehouse operations [1]. Rapid and accurate material preparation operations for the production line or customers ensure that not only does the production line or customer not have to wait for supplies, but that warehouse labor is also used effectively. The fundamental operation of a warehouse is the safe storage (stocking) of materials and their transit (handling) from one location to another. The warehouse activities for the preparation of raw materials, parts, semi-materials, and products for production or sale from the warehouse, which is one of the handling operations, begin with the picking activity of goods in the warehouse stock areas. This procedure accounts for the majority of warehouse activities. Improvements in material selection will have a direct impact on the overall operation as a big improvement.

In this article, a Material Picking Information System (MPIS) has been developed for responsible personnel who will collect the materials required for production from the warehouse, which determines the number of picking boxes (how many boxes are required for the materials), finds the order in which the materials will be collected by the shortest route, and lists how these materials are placed in the boxes.

Determining the material picking order is expressed as a Traveling Salesman Problem (TSP) problem and a solution method based on evolutionary algorithms presented by MS Excel Solver is used to find the shortest path (shortest material picking order). CLP Spreadsheet Solver, which is developed by Erdoğan [3] and which solves the container loading problem based on the large neighborhood search algorithm, is used to find the number of the picking box, calculate its size, and determine the material arrangement within the box.

The developed information system is implemented in a company that produces electromechanical products in Ankara. In the current system, the number of boxes, the shortest material picking order and the arrangement of the material in the box (in-box sequences) are realized with the experience and intuition of the responsible personnel. With the proposed MPIS, an improvement of approximately 50% has been achieved in the distance traveled and the time spent by the personnel for material collection.

With the Material Picking Information Systems (MPIS) presented in this article, up to 50% improvement has been achieved in terms of the distance and time spent by the personnel and the model has been validated.

Here, there are analytical models used by the information system, but the study has not been called a decision support system (DDS). A DSS usually deals with semi-structured or unstructured decisions and often provides decision support to tactical and operational level managers. However, the reports produced in this study are daily routine reports and reports at non-strategic levels. For this reason, the developed system can be considered as a management information system even though it uses an analytical model.

## 2. WAREHOUSE MANAGEMENT

In the literature, storage issues are classified as warehouse design and warehouse operations. Warehouses are an essential component of any supply chain. Their major roles include: buffering the material flow along the supply chain to accommodate variability caused by factors such as product seasonality and/or batching in production and transportation; consolidation of products from various suppliers for combined delivery to customers; and value-added-processing such as kitting, pricing, labeling, and product customization [4].

Warehouse design encompasses the overall structure, scaling and dimensioning, department layout, equipment selection, and operating strategy. Receiving and shipping, as well as storage and order picking, are all part of the warehouse operation.

Warehouse management encompasses the principles and processes involved in running the day-to-day operations of a warehouse. At a high level, this includes accepting and organizing warehouse space, scheduling labor, managing inventory and fulfilling orders.

### 2.1. Literature

Traditional discrete optimization issues such as the traveling salesman problem (TSP) and the bin (box) packing problem (BPP) are examples of warehouse challenges. Due to their NP-hard nature, these two issues are simple to formulate yet difficult to solve. BPP is frequently used to determine where products should be stored in a warehouse or on a pallet. Order picking is the

most labor-intensive function in the warehouse. Many studies on order selection policies have been published [5].

Armstrong, Cook, and Saipe (1979) devised an order picking algorithm that uses a mixed integer linear programming model to reduce overall picking time [6]. For precise programming, they employed artificial orders. Field picking procedures for the distribution center have been studied. There are numerous published studies that highlight the utility of the heuristic algorithm in warehouse picking [7-9]. The goal of their study was to reduce the overall material pickup route.

Ratliff and Rosenthal (1983) provided a picking route algorithm for a warehouse system with two cross aisles, front and back [10], that minimizes picking time. Using a basic step structure, they created a reliable process for selecting a picking route.

Daniels et al. (1998) formulated a model for simultaneously determining the assignment and sequencing decisions, and compared it to previous models for order picking. They discussed the complexity of the order picking problem and established an upper bound on the number of feasible assignments. They presented a tabu search algorithm and experimentally tested [11].

De Koster et al. (2007) presented a literature overview on typical decision problems in design and control of manual order-picking processes. They focused on optimal layout design, storage assignment methods, routing methods, order batching and zoning [12].

Tsai, Liou, and Huang (2008) proposed a batch picking model that considers not only travel cost but also an earliness and tardiness penalty to fulfil the current complex and quick response oriented environment and solved this model using a multiple-GA method for generating optimal batch picking plans [13].

Lin, Kang, Hou, and Cheng (2016) suggested a model that incorporates order grouping and order choosing into one model. They used the Particle Swarm Optimization (PSO) technique to create a model for order grouping and choosing [14].

Erdoğan (2017) used the big neighborhood search heuristic technique to create an MS Excel VBA macro. This research aids in determining the quantity of each box type and the method for placing them in the appropriate boxes [3].

As seen by the literature, efforts to identify the optimal solution continue, as corporations seek the most appropriate solutions in order to gain a competitive advantage. There are many studies on warehouse operations (material collection route, bin packing, etc.). In this regard, especially review articles can be examined [4,5,12,15]

### 3. METHOD

#### 3.1. Definition of the Problem and System Expectations

Preparing raw materials, parts, semi-finished products, and finished items for manufacturing or sale, referred to as stock materials, is one of the daily routine operations in the warehouse system with parallel aisles and starting and ending points in the same location. In these routine activities carried out, the determination of the picking route is based on the intuition and experience of the warehouse personnel. For this reason, losses of time and wastes will occur in material collection activities based on intuition and experience. Due to the fact that the responsible personnel cannot fully determine the amount and dimensions of the picking box they need before starting the picking activity, the warehouse personnel return to the picking starting point many times during the collection activity and take a new picking box or replace the existing box with a different size box. These repetitive transactions increase the time and labor losses of the enterprise.

#### 3.2. Description of Current System and Limitations

In the company where the implementation is realized, the existing storage (warehouse) system has a width of 26 meters (m) and a length of 16m. There are five parallel lanes, four perpendicular aisles, and 72 designated stock locations. Each warehouse consists of 6 floors (level) and the warehouse has a total height of 5 m. Warehouse corridors are 2 m wide. The width of the stock areas is 2x2 m. In the coding of stock address, as shown in Figure 1, the first subscript indicates the column number, while the second subscript represents the shelf row number. The sketch of the Warehouse System is given in Figure 2.

The shortest material picking path, number of picking boxes  $x$ , and box positioning simulation are all scheduled to function with 20 different materials to be picked from stock regions in the warehouse system. The purpose for this is to avoid the solution time being extended due to the MS Excel Solver infrastructure that will be used in the proposal. Different software infrastructures are anticipated to be able to boost this variety and solution time. There will be 72 different sorts of materials at each inventory location in the system's infrastructure.

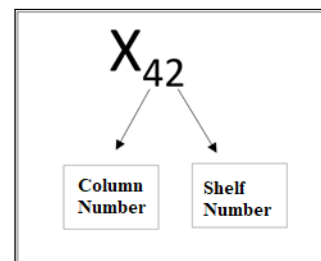


Figure 1. Stock address coding

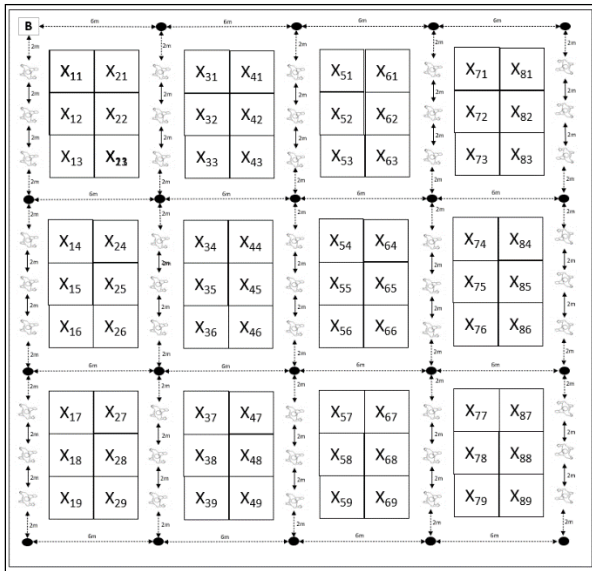


Figure 2. Warehouse's map

Personnel who will begin picking materials from point (B) given in top left corner in figure 2 will return to point (B) once the picking is completed.

The required number of amount of boxes, the stuff of the box, and how it will be positioned will be displayed to warehouse personnel using simulation in proposed information system.

The materials and quantities of the picking list will be combined into a single lot, and the picking will be completed. In other words, all quantities of a substance to be selected should be in a single box. There are 3 types of picking box used in the warehouse: (0.40m x 0.60m x 0.28m), (0.60m x 0.80m x 0.32m) and (0.61m x 0.81m x 0.52m). Values in parentheses represent width, length and height, respectively. Pallets are used to convey materials that do not fit into any of the picking boxes. There are 5 picking boxes for each type and a total of 15 picking boxes. Figure 3 shows models of picking boxes.

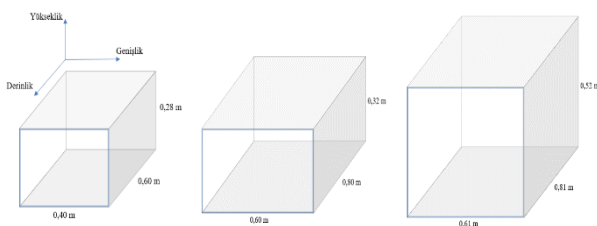


Figure 3. Picking boxes types

In the creation of the information system, Data Flow Diagrams (DFD) are used. DFDs are a widely used tool in the development of structural systems that gives a graphical depiction of the system. Any complicated system can be described using DFDs by utilizing four symbols: external entity, data flow, data store, and process. DFDs are

drawn as context diagram, 0-level diagram, and sub-level diagrams [16].

### 3.3. Creating Context, 0-level, and Sub-level Diagrams

The Context Diagram of the system given in Figure 4 was designed with the system's constraints and expectations in mind. Picking box types information, warehouse location and distance between locations, material stock location information, material size information, sequence information according to pieces, required shipping box and in-box layout calculation request, 3D in-box layout request, and report are requests for the Material Picking Information System. The sequence of picking materials from stock locations, the number of boxes, the 3D in-box layout list, and other reports that indicate the shortest path are among the outputs of the MPIS. The external entities are warehouse supervisors and warehouse operators. The warehouse supervisor is the warehouse's administrative officer and is in charge of the warehouse's operations. The warehouse supervisor's key responsibilities include planning and arranging the day's activities. Warehouse operators are the employees who carry out the actions outlined in the warehouse supervisor's plan.

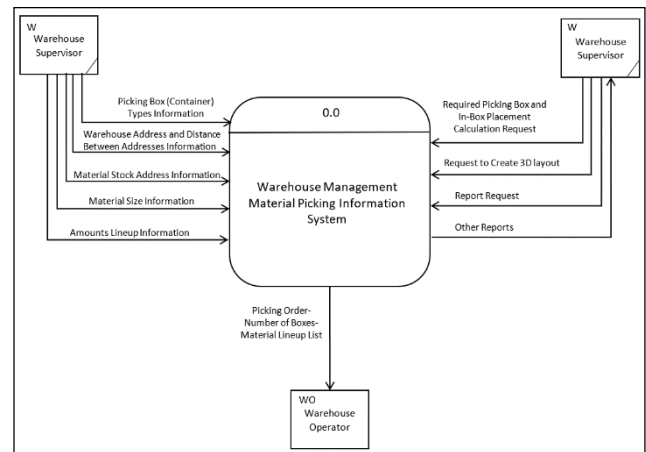


Figure 4. Context diagram

### 3.4. Creating 0-Level (Overview) Data Flow Diagram

0-level or overview diagram is given in detail in annex (a). The system has 9 processes and nine datastores. Sub-level diagram of the process 2.0 is shown in annex (b).

### 3.5. Creating System Data Stores and Data Model

In order to provide data entries such as transport/picking box type data, material stock location data, material size data, warehouse location-distance data, box sequence data, and warehouse operator data, the user interface given in Figure 5 is designed.

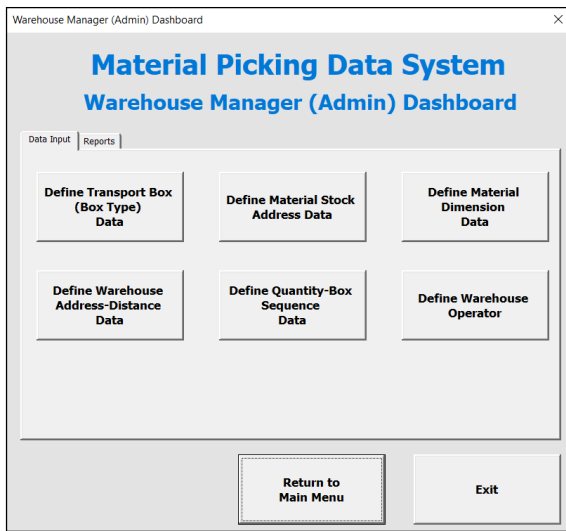


Figure 5. Data entry interface

The Entity Relationship (ER) Diagram and data model created for MPIS are given in figure 6 and figure 7 respectively.

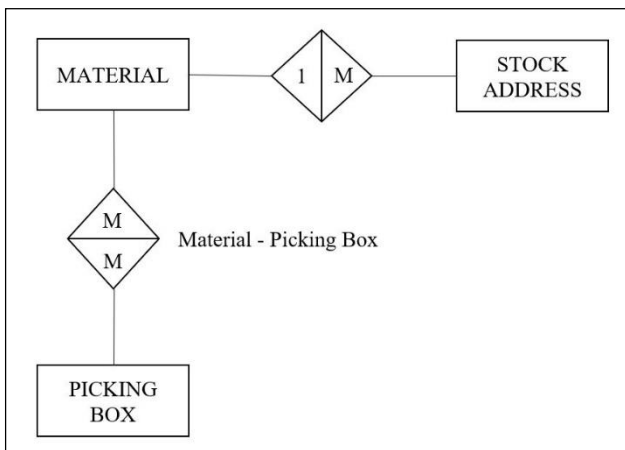


Figure 6. ER Diagram of the system

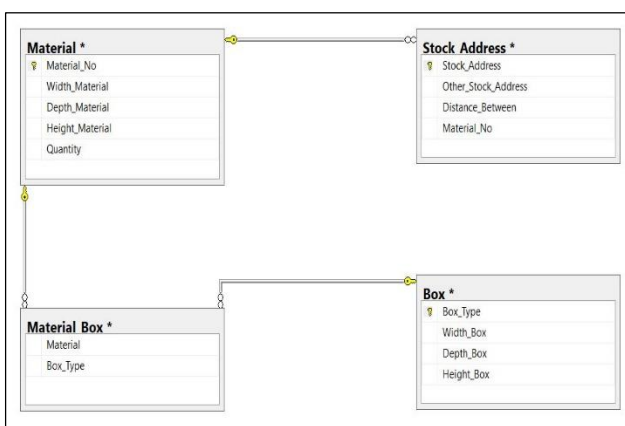


Figure 7. Data Model of the system

### 3.6. Solution Method

#### 3.6.1. Calculating the Shortest Path

It is aimed to reach a solution by applying the the Traveling Salesman Problem (TSP) model to find the shortest picking path in the process 5.0 in the 0-level DFD. Practical applications of the TSP are in the field of distribution and transport logistics. In addition to optimal solutions, heuristic and genetic algorithms can be applied to solve TSP. Although TSP is easy to formulate, it is among the NP-hard natured problems.

The TSP can be formulated as an integer linear program. Several formulations are known but two notable formulations are the Miller–Tucker–Zemlin (MTZ) formulation [17] and the Dantzig–Fulkerson–Johnson (DFJ) formulation [18]. The DFJ formulation is stronger, though the MTZ formulation is still useful in certain settings.

Here, nodes (e.g.cities) are labeled with the numbers 1, ..., n and define:  $X_{ij} \in \{0,1\}$ :  $X_{ij}$  expresses whether to go from point i to point j. It takes the value 1 in case of going, and 0 in the otherwise. Take  $C_{ij} > 0$  to be the distance from city i to city j. Then TSP can be written as DFJ formulation the following integer linear programming problem [18]:

$$\min \sum_{i=1}^n \sum_{j \neq i, j=1}^n c_{ij} x_{ij} \tag{1}$$

s. t.

$$\sum_{i=1, i \neq j}^n x_{ij} = 1 \quad j = 1, \dots, n; \tag{2}$$

$$\sum_{j=1, j \neq i}^n x_{ij} = 1 \quad i = 1, \dots, n; \tag{3}$$

$$\sum_{i \in Q} \sum_{j \neq i, j \in Q} x_{ij} \leq |Q| - 1 \quad \forall Q \subset \{1, \dots, n\}, |Q| \geq 2 \tag{4}$$

TSP; It is the problem of determining the least costly tour that goes through n cities, passes through each one only once and returns to the starting point, and the objective function can be expressed as given in function (1)

Equation (2) and equation (3) are constraints to ensure that each point is visited once.

The last constraint (Equation (4)) of the DFJ formulation ensures no proper subset Q can form a sub-tour, so the solution returned is a single tour and not the union of smaller tours. Because this leads to an exponential number of possible constraints, in practice, it is solved with delayed column generations. There are many exact algorithms (dynamic programming, branch-and-bound algorithm, Integer programming, cutting plane, etc.) and heuristics and approximation algorithms (greedy algorithms, Saving

heuristics, meta heuristics etc.) for the TSP problem [15,19,20,21,22,32].

In this article, MS Excel Solver was used to find the material collection route in the TSP problem. The MS Excel Solver that will be used to solve the problem includes the "Evolutionary Solver" option. This solution method is based on genetic or evolutionary algorithms. MS Excel Solver offers a good or near-optimal solution for TSP in possible conditions. MS Excel Solver was chosen because it is a widely used freeware. However, Container/Cargo Loading Problem (CLP) Spreadsheet Solver, which will be used in planning the required number of boxes and in-box placement, is based on MS Excel and works in harmony with each other in both programs.

A distance matrix (Stock Locations Distance Matrix) of 73x73 size is created showing the distance between the stock locations needed to find the shortest path. Using the "Index" formula, which is one of the Excel formulas, the distance between two stock locations is drawn from the "Stock Locations Distance Matrix" with the help of the formula. The "Index" formula is applied in line with the constraint given for the picking of 20 kinds of materials specified within the system limitations. In the TSP to be implemented in MS Excel Solver, an additional constraint has been entered for the warehouse personnel to return to the starting point. MS Excel Solver TSP application screen is given in Figure 8.

The user interface designed for the user to select materials and run MS Excel Solver is given in Figure 9. This part covers the 4.0 and 5.0 processes on the 0-level DFD.

Figure 9. Shortest path calculation interface

As an example, a material picking list is created as in Table 1. The solution was reached in 70 seconds with a computer with Microsoft Windows 10 operating system and Intel(R) Core (TM) i5-8265U CPU @ 1.60GHz, 1800 Mhz, 4 Cores, 8 Logical Processors. A shortest picking route/path has been found as a solution. The shortest picking route is as follows.

Queue	Address	Distance
1	B	2
2	X11	8
3	X14	2
4	X15	6
5	X17	12
7	X38	12
10	X49	2
13	X58	8
12	X55	0
9	X45	12
14	X63	12
18	X82	6
17	X84	10
18	X88	2
19	X89	10
20	X69	4
21	X77	6
22	X75	12
23	X53	4
24	X41	12
25	X22	10
26	1	

**Start Point Kontrol**

1    1

Total Distance: **152**

**Çözücü Parametreleri**

Hedef Ayarla: SK\$7

Hedef:  En Büyük  En Küçük  Değeri: 0

Değişken Hücreleri Değiştirerek: \$A\$5:\$A\$25

Kısıtlamalara Bağlıdır:

\$A\$5:\$A\$25 = TümFark  
SK\$5 = 1

Kısıtlanmamış Değişkenleri Pozitif Yap

Çözme Yöntemi: Açılım

Çözüm Yöntemi: Düzgün doğrusal olmayan Çözücü Problemleri için GRG Doğrusal Olmayan altyapısını seçin. Doğrusal Çözücü Problemleri için Basit LP altyapısını seçin ve düzgün olmayan Çözücü problemleri için Açılım altyapısını seçin.

Yardım    **Çöz**    Kapat

Figure 8. MS Excel Solver TSP application screen



**B** → X15 → X17 → X38 → X29 → X36 → X45 → X46 → X66 → X76 → X74 → **B**

they will perform material collection according to this list. Total distance of this route is calculated as 90 m.

The representation of the shortest path found on the warehouse sketch is as in Figure 10.

Table 1. Material Picking List

Material Part Number	Stock Address
P/N - 5	X15
P/N - 7	X17
P/N - 18	X29
P/N - 24	X36
P/N - 26	X38
P/N - 32	X45
P/N - 33	X46
P/N - 51	X66
P/N - 58	X74
P/N - 60	X76

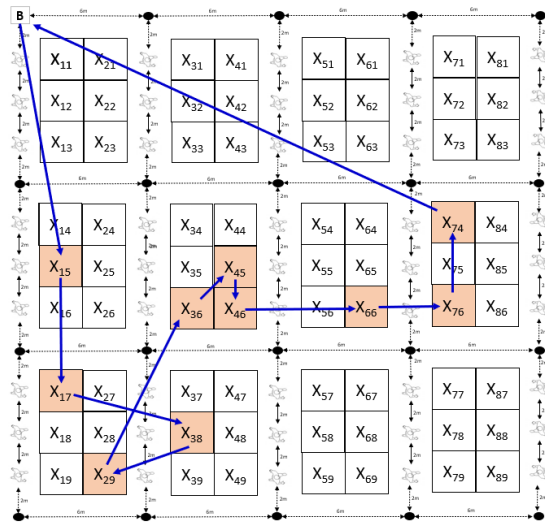


Figure 10. The shortest path on the warehouse sketch

The result report screen created with the MS Excel Visual Basic for the information system is given in Figure 11. This report will be listed to the warehouse operators and

Figure 11. The result report on the information system

Calculating Shortest Path
X

## Material Picking Information System

### Calculating Shortest Path

**Material Picking Lineup**

1. Stock Place to Destination: <input type="text" value="B"/>	10. Stock Place to Destination: <input type="text" value="X76"/>	19. Stock Place to Destination: <input type="text"/>
2. Stock Place to Destination: <input type="text" value="X15"/>	11. Stock Place to Destination: <input type="text" value="X74"/>	20. Stock Place to Destination: <input type="text"/>
3. Stock Place to Destination: <input type="text" value="X17"/>	12. Stock Place to Destination: <input type="text"/>	21. Stock Place to Destination: <input type="text"/>
4. Stock Place to Destination: <input type="text" value="X38"/>	13. Stock Place to Destination: <input type="text"/>	
5. Stock Place to Destination: <input type="text" value="X29"/>	14. Stock Place to Destination: <input type="text"/>	
6. Stock Place to Destination: <input type="text" value="X36"/>	15. Stock Place to Destination: <input type="text"/>	
7. Stock Place to Destination: <input type="text" value="X45"/>	16. Stock Place to Destination: <input type="text"/>	
8. Stock Place to Destination: <input type="text" value="X46"/>	17. Stock Place to Destination: <input type="text"/>	
9. Stock Place to Destination: <input type="text" value="X66"/>	18. Stock Place to Destination: <input type="text"/>	

Return to Main Menu

Exit

3.6.2. Calculating Picking Box Requirement

The problem of placing items of various volumes in boxes of various volumes in order to minimize the number of boxes used is known as the container loading problem (CLP)/3D bin packing problem. The another problem to be solved in this study is to place the materials in the box by minimizing the number of boxes.

CLPs belong to the NP-hard nature due to the dimensions of the boxes and the varying capacities of the containers. For this reason, it is not possible to solve the problem optimally as the problem size grows, so it is a necessity to use heuristic algorithms. The closed mathematical form of the container loading problem is given below:

$$\max \sum_{i=1}^n v_i p_i$$

s. t.

$$0 \leq x_i \leq W - w_i \quad (1)$$

$$0 \leq y_i \leq L - l_i \quad (2)$$

$$0 \leq z_i \leq H - h_i \quad (3)$$

$$W, L, H, w_i, l_i, h_i, n \in Q^+$$

$$x_i, y_i, z_i \geq 0$$

$$p_i \in \{0,1\}$$

$$i = 1, 2, \dots, n$$

Here  $Q^+$ : Positive integer,  $W, L, H$ : Container dimensions,  $w_i, l_i, h_i$ : Box dimensions,  $n$ : Total number of boxes,  $p_i$ : Binary code showing whether box  $i$  satisfies the constraints,  $v_i = w_i \cdot l_i \cdot h_i$ :  $i$  Indicates the volume of the box. In addition, the boxes will not overlap each other and will be supported by the container floor or other boxes.

There are many studies on the CLP/3D bin packaging problem in the literature. For some of these references, see [23-31]. The MS Excel workbook "CLP Spreadsheet Solver" is an open-source tool for representing, solving, and visualizing the results of Container Loading Problems developed by Erdogan [3] is used to calculate the amount and size of the required picking box and to determine the layout within the box. He implemented the Large Neighborhood Search heuristic within the CLP Spreadsheet Solver. An outline of the algorithm is given below [14]:

- **Step 1 (Initialization)**: Sort the items with respect to their priority, size, and profit. Sort containers with respect to their size and cost.
- **Step 2 (Constructive step)**: Use the First-Fit-Decreasing heuristic to pack the items into the containers.
- **Step 3 (Perturbation)**: Randomly remove items from containers, and randomly empty a number of containers. Sort the containers in decreasing order of the volume packed into them.
- **Step 4 (Reoptimization)**: Use a constructive heuristic to repack the removed items into the containers.
- **Step 5 (Solution update)**: If the new solution is better than the best-known solution, update the best-known solution. Otherwise, revert back to the best-known solution. If the time limit is not exceeded, go to Step 3.

The large neighborhood search algorithm used in Excel CLP Spreadsheet Solver is positioned as meta-heuristic in the literature. Meta-heuristic is a high-level problem-free algorithmic framework that provides a set of guidelines or strategies for developing heuristic optimization algorithms [32].

In the system, it is necessary to combine the materials to be collected on the basis of quantity so that they do not separate from each other. For example; Suppose 30 pieces of "P/N - 5" material are to be collected. We see that this amount is factored into 2-5-3. For this reason, these 30 items have been made into a single lot with the 2-5-3 arrangement. Dimension properties of 72 kinds of materials are recorded in meters as width-depth-height. An interface is designed to select the materials to be collected and enter their quantities. Within the limitations of the system, a data entry has been developed for 20 types of material selection. The interface design in accordance with the 6.0 process shown on the 0-level DFD is as in Figure 12.

Material Part List	Quantity	Material Part List	Quantity	Material Part List	Quantity
P/N - 1	10	P/N - 32	7	P/N - 59	5
P/N - 4	12	P/N - 36	14	P/N - 61	10
P/N - 5	16	P/N - 39	21	P/N - 65	20
P/N - 7	20	P/N - 41	16	P/N - 67	15
P/N - 11	4	P/N - 44	12	P/N - 71	12
P/N - 26	6	P/N - 48	8	P/N - 72	16
P/N - 28	10	P/N - 54	9		

Figure 12. Number of boxes calculation interface

The solution was reached in 45 seconds with a computer with Microsoft Windows 10 operating system and Intel(R) Core (TM) i5-8265U CPU @ 1.60GHz, 1800 MHz, 4 Cores, 8 Logical Processors. After the solution, the result report that will guide the user in the Material Picking Information System has been created with an interface as in Figure 13.

Material List	Amount	Lineup	Box Type
		Width Depth Height	
P/N - 71	12	2 x 3 x 2	Box 1 / Type 3
P/N - 59	5	5 x 1 x 1	Box 1 / Type 3
P/N - 5	16	4 x 2 x 2	Box 1 / Type 3
P/N - 48	8	2 x 2 x 2	Box 1 / Type 3
P/N - 4	12	2 x 3 x 2	Box 1 / Type 3
P/N - 36	14	7 x 2 x 1	Box 1 / Type 3
P/N - 32	7	7 x 1 x 1	Box 1 / Type 3
P/N - 28	10	5 x 2 x 1	Box 1 / Type 3
P/N - 26	6	3 x 2 x 1	Box 1 / Type 3
P/N - 1	10	5 x 2 x 1	Box 1 / Type 3
P/N - 72	16	4 x 2 x 2	Box 2 / Type 3
P/N - 61	10	5 x 2 x 1	Box 2 / Type 3
P/N - 54	9	3 x 3 x 1	Box 2 / Type 3
P/N - 44	12	2 x 3 x 2	Box 2 / Type 3
P/N - 11	4	2 x 2 x 1	Box 2 / Type 3
P/N - 7	20	5 x 2 x 2	Transport with Pallet
P/N - 67	15	5 x 3 x 1	Transport with Pallet
P/N - 65	20	5 x 2 x 2	Transport with Pallet
P/N - 41	16	4 x 2 x 2	Transport with Pallet
P/N - 39	21	7 x 3 x 1	Transport with Pallet

Figure 13. Picking box requirement and in-box placement result interface

In MPIS, a 20-second simulation is performed that shows the material sequences in the picking box (in-box placement) one by one. The result is presented to the warehouse operator as part of the 7.0 process in 0-level DFD. According to the Sample Report, 2 of the type 3

material picking boxes (0.61m. x 0.81m. x 0.52m) are found to be sufficient and the result is reached. The layout image of the materials inside the box created by the Information System is given in Figure 14.

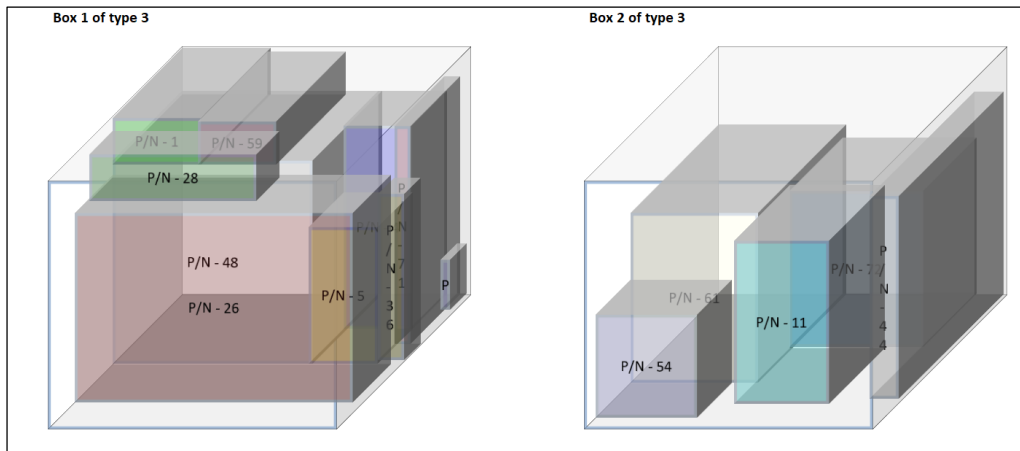


Figure 14. The layout image of the materials inside the boxes

3.6.3. Sequential Calculation of Shortest Path and Picking Box Requirement

The shortest path and the number of picking boxes calculated separately in the previous headings are run sequentially here, and it is aimed to meet the user expectations by adding the shortest path function first and then the picking box need calculation function. Even if the

use of this method increase the solution time, it can fully meet the user's needs. It took 70 seconds to find the shortest path, 45 seconds to calculate the required for the picking box, and 20 seconds to simulate the box. Thus, the solution could be reached in 135 seconds with approximately 25000 sub-problem solutions. The solution is presented to the warehouse operator as shown in Figure 15. Thus, the user can get information and waste can be prevented.

Report Form of Both of Method

**Picking Box Quantities and Types Results**

Material List	Qty.	Lineup			Box No		
		width	depth	length			
P/N - 71	12	2	x	3	x	2	Box 1 / Type 3
P/N - 59	5	5	x	1	x	1	Box 1 / Type 3
P/N - 5	16	4	x	2	x	2	Box 1 / Type 3
P/N - 48	8	2	x	2	x	2	Box 1 / Type 3
P/N - 4	12	2	x	3	x	2	Box 1 / Type 3
P/N - 36	14	7	x	2	x	1	Box 1 / Type 3
P/N - 32	7	7	x	1	x	1	Box 1 / Type 3
P/N - 28	10	5	x	2	x	1	Box 1 / Type 3
P/N - 26	6	3	x	2	x	1	Box 1 / Type 3
P/N - 1	10	5	x	2	x	1	Box 1 / Type 3
P/N - 72	16	4	x	2	x	2	Box 2 / Type 3
P/N - 61	10	5	x	2	x	1	Box 2 / Type 3
P/N - 54	9	3	x	3	x	1	Box 2 / Type 3
P/N - 44	12	2	x	3	x	2	Box 2 / Type 3
P/N - 11	4	2	x	2	x	1	Box 2 / Type 3
P/N - 7	20	5	x	2	x	2	Transport with
P/N - 67	15	5	x	3	x	1	Transport with
P/N - 65	20	5	x	2	x	2	Transport with
P/N - 41	16	4	x	2	x	2	Transport with
P/N - 39	21	7	x	3	x	1	Transport with

**The Shortest Path Result**

Material Picking Route

1. Stock Place to Destination	B	12. Stock Place to Destination	X82
2. Stock Place to Destination	X11	13. Stock Place to Destination	X84
3. Stock Place to Destination	X14	14. Stock Place to Destination	X88
4. Stock Place to Destination	X15	15. Stock Place to Destination	X89
5. Stock Place to Destination	X17	16. Stock Place to Destination	X69
6. Stock Place to Destination	X38	17. Stock Place to Destination	X77
7. Stock Place to Destination	X49	18. Stock Place to Destination	X75
8. Stock Place to Destination	X58	19. Stock Place to Destination	X53
9. Stock Place to Destination	X55	20. Stock Place to Destination	X41
10. Stock Place to Destination	X45	21. Stock Place to Destination	X22
11. Stock Place to Destination	X63		

Box Type 1

Box Type 2

Box Type 3

Run Simulator

Return to Main Menu

Exit

Figure 15. Result form interface for calculating shortest path and number of picking boxes

3.6.4. Reporting to the Warehouse Officer (Manager)

The “Warehouse Manager Interface” in the main menu of the information system (MPIS) has been developed so that

the Warehouse Officer can manage the work and information system during the day. With this interface, many reports such as Box Type Report, Warehouse Address Distance Report, Material Inventory Addresses

Report, Material Dimension Report, Piece-Box Arrangement Report, Warehouse Operators Report, Materials to be Purchased and Quantity Report, Number of Required Boxes and Placement Report and Shortest Path Report, can be obtained. The visual of the Material Dimension Report, which is one of these reports, is given in Figure 16.

### 3.7. Validation of proposed MPIS

Implementation of the developed information system is carried out in a company that produces electromechanical

products in Ankara for the verification of the proposed MPIS. In order to verify the MPIS and see its performance, it is compared with the existing system where warehouse operations is traditionally done. The comparison is made over 10 different material collection work orders. The material collection processes performed by the personnel without using MPIS, with their experience and intuition, and the material collection processes using MPIS for the same work order are compared in terms of number of return to the starting point, total distance traveled (m) and Total time spent (min). Comparison results are given in Table 2.

Part No	Stock Address	1 Piece of Material size (Width) (metre)	1 Piece of Material size (Depth) (metre)	1 Piece of Material size (Height) (metre)
P/N - 1	X11	0,08	0,09	0,04
P/N - 2	X12	0,07	0,08	0,03
P/N - 3	X13	0,10	0,10	0,01
P/N - 4	X14	0,21	0,07	0,21
P/N - 5	X15	0,11	0,07	0,14
P/N - 6	X16	0,27	0,04	0,30
P/N - 7	X17	0,23	0,29	0,25
P/N - 8	X18	0,27	0,16	0,34
P/N - 9	X19	0,32	0,12	0,12
P/N - 10	X21	0,21	0,07	0,21
P/N - 11	X22	0,20	0,17	0,28
P/N - 12	X23	0,05	0,11	0,09
P/N - 13	X24	0,60	0,04	0,20
P/N - 14	X25	0,21	0,21	0,09
P/N - 15	X26	0,03	0,01	0,23
P/N - 16	X27	0,10	0,21	0,07
P/N - 17	X28	0,21	0,08	0,03
P/N - 18	X29	0,27	0,10	0,08
P/N - 19	X31	0,07	0,21	0,35
P/N - 20	X32	0,34	0,09	0,07
P/N - 21	X33	0,05	0,21	0,18
P/N - 22	X34	0,04	0,30	0,28
P/N - 23	X35	0,27	0,14	0,27
P/N - 24	X36	0,21	0,18	0,34
P/N - 25	X37	0,60	0,10	0,32
P/N - 26	X38	0,23	0,20	0,34
P/N - 27	X39	0,03	0,23	0,16
P/N - 28	X41	0,09	0,07	0,20
P/N - 29	X42	0,27	0,07	0,21

Figure 16. Material dimension report

Table 2. Comparison of Current and Proposed System

Work order ID	# of material items to be collected	# of stock address point to be visited	# of return to the starting point		Total distance traveled (m)		Total time spent (in minute)	
			Without Using MPIS	Using MPIS	Without Using MPIS	Using MPIS	Without Using MPIS	Using MPIS
10455	11	11	0	0	124	74	83,6	45,1
10456	10	10	1	0	152	126	90,2	41
10457	20	20	1	0	196	126	166,2	82
10458	20	20	0	0	160	126	152	82
10459	3	3	0	0	48	148	22,8	12,3
10460	18	18	1	0	148	148	151	73,8
10461	16	16	3	0	220	48	164,2	65,6
10462	15	15	2	0	192	74	142,4	61,5
10463	17	17	1	0	160	74	143,4	69,7
10464	6	6	0	0	52	48	45,6	24,6
<b>Total</b>	<b>136</b>	<b>136</b>	<b>9</b>	<b>0</b>	<b>1452</b>	<b>992</b>	<b>1161,4</b>	<b>557,6</b>
Average value per unit					10,68	7,29	8,54	4,10

As can be seen in Table 2, 136 items of materials were collected from 136 stock addresses. When MPIS was not used, a total of 9 return to the starting point occurred during the activity. When MPIS is used, this value is 0, meaning that there was no return to the starting point. With the use of MPIS, the total distance traveled has decreased from 1452m to 992m. When these values are proportioned to a total of 136 items of material, 10,68m and 7,29m per item are obtained, respectively. That is, the total distance traveled decreased by 3,39m per unit. Similarly, the total time spent decreased from 1161,4 min. to 557,6 min. When these values are proportioned to a total of 136 items of material, 8,54min and 4,10min per item are obtained, respectively. That is, the total time spent decreased by 4,44min per unit. In summary, with the use of MPIS, it can be seen that approximately 50% improvement has been achieved in the material collection and preparation activities of the company

#### 4. CONCLUSION

Many businesses focus on reducing their costs because they cannot raise their product prices to sustain the economy of their businesses. Taking into account all the costs carried out in the warehouse, the cost portion of 55%-

65% consists of the activities of collecting and preparing the materials from the warehouse from the stock areas. Reducing these costs means increasing profitability. Reducing these costs means increasing profitability. This is the main objective of the study carried out in this article.

In this article, Material Picking Information System (MPIS) that finds the shortest path for the responsible personnel to collect the materials required for production

from the stock addresses in the warehouse, determines the number of picking boxes required for these materials, and provides the material sequences within the box (in-box sequences) is developed. With MPIS, it will be possible to reduce material picking costs by eliminating wasteful activities such as transport, waiting and movement.

Implementation of the developed information system (MPIS) is carried out in a company that produces electromechanical products in Ankara for the verification of the proposed MPIS. In order to verify the MPIS and see its performance, it is compared with the existing system where warehouse operations is traditionally done (In the current system, the number of boxes, the shortest material picking order and the arrangement of the material in the box (in-box sequences) were realized with the experience and intuition of the responsible personnel). The comparison is made over 10 different material collection work orders. The material collection processes performed by the personnel without using MPIS, with their experience and intuition, and the material collection processes using MPIS for the same work order are compared in terms of number of return to the starting point, total distance traveled (m) and Total time spent (min). With the proposed

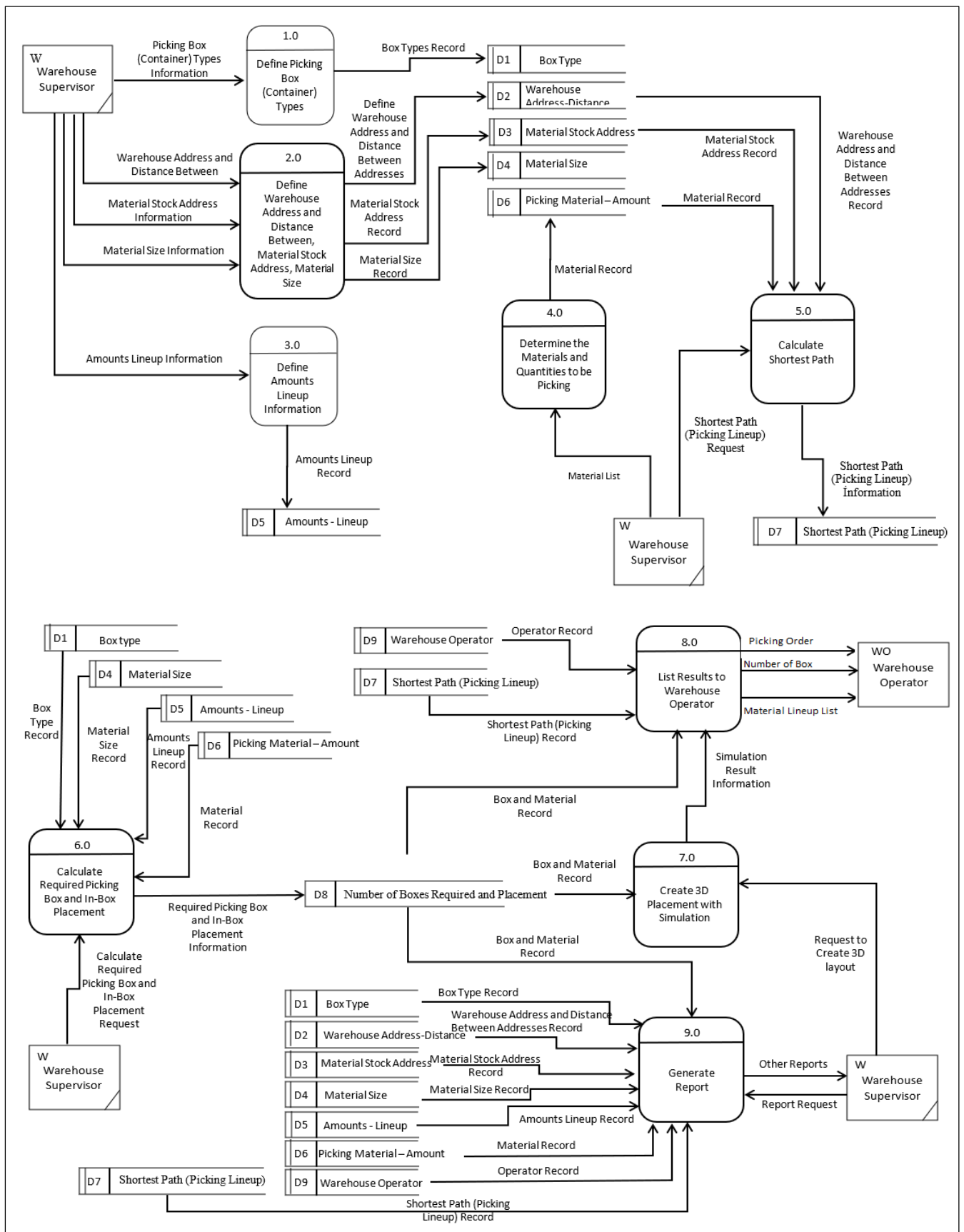
MPIS, an improvement of approximately 50% has been achieved in the distance traveled and the time spent by the personnel for material collection.

#### REFERENCES

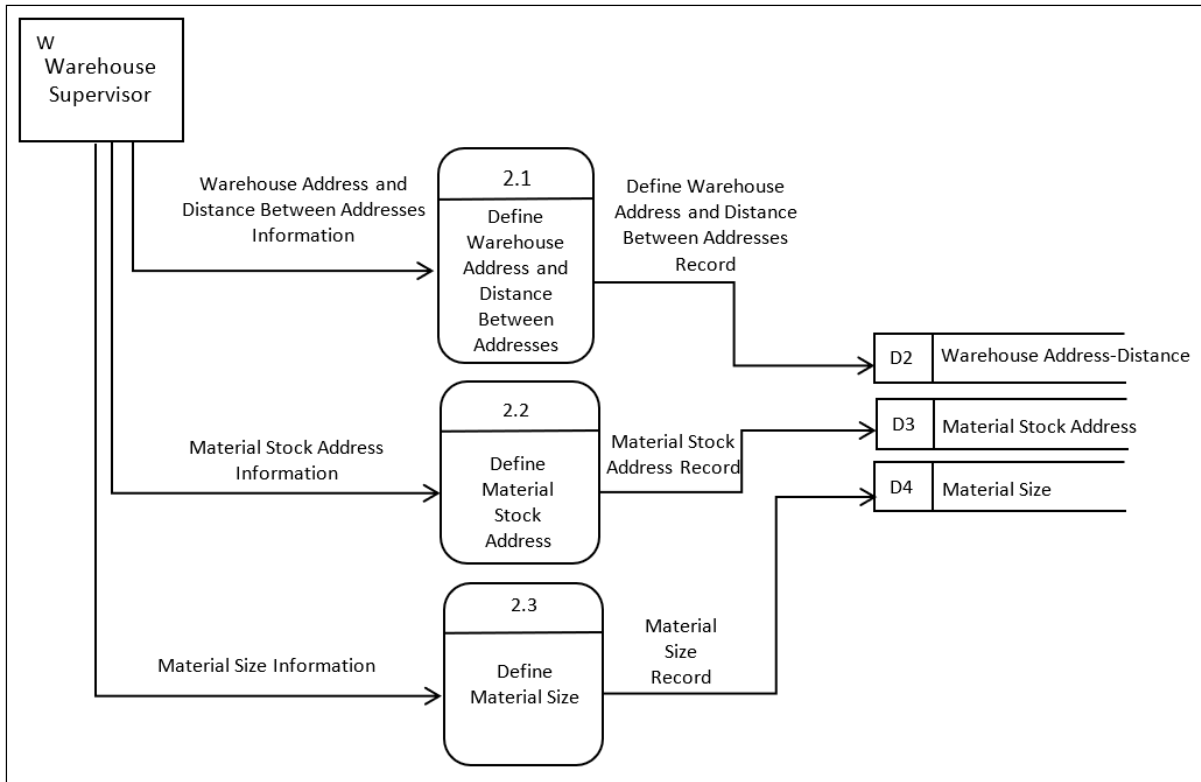
- [1] C. Theys, O. Braysy, W. Dullaert and B. Raa, "Using a TSP Heuristic for Routing Order Pickers in Warehouses", *European Journal of Operational Research*, 200(3), 755–763, 2010.
- [2] N. D. K. Al-Shakarchy, "Warehouse Management System", *International Journal of Science and Research (IJSR)*, 4(10), 1253–1260, 2015.
- [3] Internet: G. Erdoğan, CLP Spreadsheet Solver, Released on 07.09.2017, <https://researchportal.bath.ac.uk/en/publications/clp-spreadsheet-solver>, 31.05.2021.
- [4] J. Gu, M. Goetschalckx, and L.F. McGinnis, "Research on Warehouse Operation: A Comprehensive Review", *European Journal of Operational Research*, 177(1), 1–21, 2007.
- [5] F. Yener and H. R. Yazgan, "Optimal Warehouse Design: Literature Review and Case Study Application", *Computers & Industrial Engineering*, 129(March), 1-13, 2019.
- [6] R. D. Armstrong, W. D. Cook and A.L. Saipé, "Optimal Batching in a Semi-Automated Order Picking System", *Journal of the Operational Research Society*, 30(8), 711–720, 1979.
- [7] E. A. Elsayed, "Algorithms for Optimal Material Handling in Automatic Warehousing Systems", *International Journal of Production Research*, 19(5), 525–535, 1981.
- [8] G. R. Stern and E. A. Elsayed, "Computerized Algorithms for Order Processing in Automated Warehousing Systems", *International Journal of Production Research*, 21(4), 1983.
- [9] E. A. Elsayed and O. Unal, "Order Batching Algorithms and Travel-Time Estimation for Automated Storage/Retrieval Systems", *International Journal of Production Research*, 27(7), 1097-1114, 1989.
- [10] H. D. Ratliff and A. S. Rosenthal, "Order-picking in a Rectangular Warehouse: A Solvable Case of the Traveling Salesman Problem", *International Journal of Operations Research*, 31(3), 507–521, 1983.
- [11] R. L. Daniels, J. L. Rummel and R. Schantz, "A Model for Warehouse Order Picking," *European Journal of Operational Research*, 105(1), 1-17, 1998.
- [12] R. De Koster, T. Le-Duc and K. J. Roodbergen, "Design and Control of Warehouse Order Picking: A Literature Review", *European Journal of Operational Research*, 182(2), 481-50, 2007.
- [13] C. Y. Tsai, J. J. H. Liou and T. M. Huang, "Using a Multiple-GA Method to Solve the Batch Picking Problem: Considering Travel Distance and Order Due Time", *International Journal of Production Research*, 46(22), 6533–6555, 2008.
- [14] C. C. Lin, J. R. Kang, C. Y. Cheng, "Joint Order Batching and Picker Manhattan Routing Problem", *Computers and Industrial Engineering*, 95, 164–174, 2016.
- [15] M. Masae, C. H. Glock, ve E. H. Grosse, "Order Picker Routing in Warehouses: A Systematic Literature Review", *International Journal of Production Economics*, 224, 2-22, 2020.

- [16] H. Gökçen, **Yönetim Bilgi Sistemleri: Analiz ve Tasarım**, Afşar Matbaacılık, Ankara, 2011.
- [17] C. E. Miller, A. W. Tucker, R. A. Zemlin, "Integer Programming Formulation of Traveling Salesman Problems", *Journal of the ACM*, 7(4), 326-329, 1960.
- [18] G. Dantzig, R. Fulkerson, S. Johnson, "Solution of a Large-Scale Traveling-Salesman Problem", *Journal of the Operations Research Society of America*, 2(4), 1954.
- [19] R. Matai, S. Singh, M. L. Mittal, **Traveling Salesman Problem: An Overview of Applications, Formulations, and Solution Approaches, Traveling Salesman Problem, Theory and Applications**, Editor: D. Davendra, 2010.
- [20] F. Valdez, F. Moreno, P. Melin, **Hybrid Intelligent Systems in Control, Pattern Recognition and Medicine, A Comparison of ACO, GA and SA for Solving the TSP Problem**, Editors: O. Castillo, P. Melin, 181-189, Springer, 2020.
- [21] C. Theys, O. Bräysy, W. Dullaert, W., B. Raa, "Using a TSP Heuristic for Routing Order Pickers in Warehouses", *European Journal of Operational Research*, 200(3), 755-763, 2010.
- [22] V. Raman, N. S. Gill, "Review of Different Heuristic Algorithms for Solving Travelling Salesman Problem", *International Journal of Advanced Research in Computer Science*, 8(5), 423-425, 2017.
- [23] C. Chen, S. M. Lee, Q. Shen, "An Analytical Model for the Container Loading Problem", *European Journal of Operational Research*, 80(1), 68-76, 1995.
- [24] H. Gehring, A. Bortfeldt, "A Genetic Algorithm for Solving the Container Loading Problem", *International Transactions in Operational Research*, 4(5-6), 401-418, 1997.
- [25] A. Bortfeldt, H. Gehring, "A Hybrid Genetic Algorithm for the Container Loading Problem", *European Journal of Operational Research*, 131(1), 143-161, 2001.
- [26] D. Pisinger, "Heuristics for the Container Loading Problem", *European Journal of Operational Research*, 141(2), 382-392, 2002.
- [27] M. Hifi, "Exact Algorithms for Unconstrained Three-Dimensional Cutting Problems: A Comparative Study", *Computers & Operational Resource*, 31(5), 657-674, 2004.
- [28] J. Liu, Y. Yue, Z. Dong, C. Maple, M. Keech, "A Novel Hybrid Tabu Search Approach to Container Loading", *Computer Operations Research*, 38(4), 797-807, 2011.
- [29] M. M. Baldi, G. Perboli, R. Tadei, "The Three-Dimensional Knapsack Problem with Balancing Constraints", *Applied Mathematics and Computation*, 218(19), 9802-9818, 2012.
- [30] A. Bortfeldt, G. Wäscher, "Constraints in Container Loading—A State-of-the-Art Review", *European Journal of Operational Research*, 229(1), 1-20, 2013.
- [31] H. M. Ghomi, B. G. S. Amour, W. Abdul-Kader, "Three-dimensional container loading: a simulated annealing approach", *International Journal of Applied Engineering Research*, 12(7), 1290-1304, 2017.
- [32] B. Özoğlu, E. Çakmak, T. Koç, "Clarke & Wright's Savings Algorithm and Genetic Algorithms Based Hybrid Approach for Flying Sidekick Traveling Salesman Problem", *European Journal of Science and Technology, Special Issue*, 185- 192, 2019.

ANNEX



(a) 0-Level Data Flow Diagram





(b) The Sublevel Diagram of Process 2.0



# COVID-19 Tespitinde Nesnelerin İnterneti Kullanımı: Bibliyometrik Analiz ile Bir İnceleme

*Araştırma Makalesi / Research Article*

 Burcu KOCARIK GACAR<sup>1</sup>,  Selay GİRAY YAKUT<sup>2</sup>

<sup>1</sup>Ekonometri Bölümü, Dokuz Eylül Üniversitesi, İzmir, Türkiye

<sup>2</sup>Ekonometri Bölümü, Marmara Üniversitesi, İstanbul, Türkiye

[burcu.kocarikgacar@deu.edu.tr](mailto:burcu.kocarikgacar@deu.edu.tr), [selay.giray@marmara.edu.tr](mailto:selay.giray@marmara.edu.tr)

(Geliş/Received:29.04.2022; Kabul/Accepted:03.09.2022)

DOI: 10.17671/gazibtd.1111392

**Özet**— Pandemi süreçleri aylar hatta seneler sürebilir ancak oluşumlara hızlı yanıtlar verilerek pandeminin üstesinden gelinir. İnternete bağlanabilen cihazlar sayesinde gerçek zamanlı ve anlık olarak veri toplanması ve yayınlanması sağlık hizmetlerinin daha iyi yürütülmesine olanak tanımaktadır. Örneğin ateş ölçümü, virüsün sebep olduğu çeşitli semptomların gözlenmesi, sosyal mesafe ölçümü, sosyal mesafenin korunması ve hastalarda izlem yapan araçlar ile mobil uygulamalar, nesnelerin interneti altyapısı ile geliştirilmiştir. İnternet sayesinde birbiri ile gerçek zamanlı veri ve bilgi transferi yapan cihazlar, tıp alanında da kullanılmasıyla birlikte sağlık sektöründe oldukça faydalı hale gelmiştir. Bu çalışmanın amacı COVID-19 sürecinde nesnelerin interneti kullanımı üzerine yayınlanan akademik araştırmaların bir bibliyografisini elde etmektir. Bu amaçla Web of Science (WoS) veri tabanında yayınlanan 922 çalışma bibliyometrik olarak analiz edilmiş ve bilimsel haritalama yapılarak incelenmiştir. Bu araştırma COVID-19'un başlangıç yılı olan 2019 yılı sonundan önlemlerin esnetilmeye başlandığı 2022 yılının ilk çeyreğine kadar olan süreci kapsamaktadır. Bulgular en çok atıf alan anahtar kelimeler, araştırma alanları, yazarlar, dokümanlar, kurum/kuruluşlar ve ülkeler ışığında değerlendirilmiştir.

**Anahtar Kelimeler**—nesnelerin interneti, COVID-19, pandemi, bibliyometri, bilimsel haritalama

## Use of Internet of Things in Detection of COVID-19: A Review with Bibliometric Analysis

**Abstract**— Pandemic processes can take months or even years, but pandemic can be overcome by giving quick responses to occurrences. Thanks to devices that can be connected to the Internet, real-time and instantaneous data collection and publication enable better health services. For example, fever measurement, observation of various symptoms caused by the virus, social distance measurement, social distance protection and monitoring tools for patients and mobile applications have been developed with the internet of things infrastructure. Upon used in the field of medicine, devices transferring real-time data and information with each other thanks to the Internet have become very useful in the health sector. The aim of this study is to obtain a bibliography of academic research published on the use of the Internet of Things in the COVID-19 process. For this purpose, 922 studies published in the Web of Science (WoS) database were analyzed bibliometrically and examined by scientific mapping. This research covers the period from the end of 2019, the beginning of COVID-19, to the first quarter of 2022, when the measures began to be stretched. The findings were evaluated in the light of the most cited keywords, research areas, authors, documents, institutions/organizations and countries.

**Keywords**—internet of things, COVID-19, pandemic, bibliometrics, scientific mapping

## 1. GİRİŞ (INTRODUCTION)

İlk olarak Çin Halk Cumhuriyeti'nin Wuhan kentinde 2019 yılında ortaya çıkan Koronavirüs (COVID-19) salgını, küresel düzeyde tüm ülkeleri etkileyen önemli bir gündem haline gelmiştir. Dünya Sağlık Örgütü (WHO) tarafından 2019 yılının sonlarında pandemi olarak ilan edilen COVID-19 salgını tüm dünyayı hızla etkisi altına almıştır. Pandemi, dünyanın hemen hemen her yerinde görülen ve etkisi çok sayıda insan tarafından hissedilen salgın hastalıklar olarak tanımlanabilir. İnsan sağlığını oldukça olumsuz etkileyen bu salgın nedeniyle dünya genelinde çok sayıda insan, yaşamını yitirmiştir. Hükümetler, pandemiyi yayılımını azaltmak ve insan sağlığı üzerindeki olası olumsuz etkilerini en aza indirebilmek için çaba göstermektedir. Bu nedenle ülkeler, COVID-19 salgını önleyebilmek ve bulaş riskini azaltabilmek için çok sayıda önlem almış ve bu önlemler pandemi sürecince geçerliliğini korumuştur. Halen etkisi hafiflese de önemi devam etmektedir.

Bir pandeminin toplumlar üzerindeki kötü etkileri hükümetler tarafından iyi hazırlanarak azaltılabilir. Bu süreçte hazırlıklı olma ve olabilecek değişimleri öngörebilmek için planlama, hasta sayısındaki artış kapasitesinin tahmin edilmesi, iletişim, eğitim öğretim faaliyetleri, tıbbi yönetim, gözetim ve organizasyonel hazırlıklı olma gibi çeşitli başlıklar altında değerlendirme yapılabilmektedir [1].

COVID-19 salgını etkisinin en çok sağlık alanında olduğu açıkça bellidir. Sağlık hizmetleri yönetimi, bir ülke için oldukça önemlidir. Sağlık yönetimi alanındaki birimler, ülkenin pandemi sürecinin takibi ve yönetimi, sonuçların değerlendirilerek ilgililerle şeffaf bir biçimde paylaşılması, önleyici çalışmaların ve uyarıların yapılması ile tanı, tedavi ve süreç kontrolüyle ilgilenmektedir [2]. Pandemi süreci, uzun vadeli bir planlamayı gerektirmektedir. Bu anlamda pandemi sürecinde sağlık sektöründe yer alabilecek adımlar [3];

- Pandemi süreci ile ilgili proaktif olabilmek,
  - Meydana gelebilecek olaylar için hızlı karar alabilmek,
  - Alınan kararları uygulamayabilmek,
  - Ortaya çıkan sonuçları değerlendirebilmek,
  - Değerlendirmeler sonucunda gerekli önlemler alabilmek,
  - Normalleşme sürecini planlayabilmek
- şeklinde belirtilebilir.

Geniş kitlelere yayılabilen bir virüs söz konusu olduğunda gerek süreç takibi ve virüs bulaşan hasta tespiti gibi konularda gerekse istatistiksel olarak süreç analizi gibi konularda çeşitli cihaz ve araçlardan yardım almak kaçınılmazdır. Hükümetler açısından salgın yönetiminde büyük veri akışının sağlanması, ağ yapılarının belirlenmesi ve sıklıkla güncellenmesi bununla birlikte salgın yayılımının senkron bir halde izlenmesi ve salgın yayılımının tahminlenmesi amacıyla çeşitli uygulamalar kullanmak fayda sağlamıştır. Benzer şekilde hastalar ve toplum için de yönlendirici ve bilgilendirici sistemlerin kullanımı ve gerekli önlemlerin alınması gibi çeşitli

uygulamaların kullanımına olanak tanıyan cihazlara başvurulmuştur. Bu verilerin gerçek zamanlı ve anlık olarak sürekli toplanmasıyla veri tabanlarının oluşması ve yayınlanması internet sayesinde meydana gelebilmektedir. Bu aşamada internete bağlanabilen cihazlar sayesinde sağlık hizmetlerinin yürütülmesi çalışmalarında ilerleme kaydedilebilmektedir [4-6].

İlk olarak 1990'lı yılların sonunda ortaya çıkan nesnelerin interneti kavramı (Internet of Things – IoT), internet ortamında birbiriyle etkileşim halinde olan cihazlar veya daha genel anlamda nesneler olarak tanımlanabilir [7, 8]. Yüksek hacimli veri sağlayan bir kaynağa dönüşen nesnelerin interneti, büyük verinin hem kaydedilmesi hem de işlenmesinde esas oluşturmaktadır. Büyük veri teknolojileri kapsamında hayatımıza giren bilgisayar dışındaki akıllı nesneler de internete bağlanabilme özelliğiyle tanınmaktadır. Teknoloji sayesinde birbiriyle iletişim kuran, veri ve bilgi transferi yapan ve adeta internet aracılığıyla birbiriyle konuşan nesnelere bahsetmek mümkündür. Artık günlük hayatta dahi cep telefonlarından sonra akıllı araba, saat, ev, buzdolabı, çamaşır makinesi, yüzük, temizlik robotu hatta çay ve kahve makineleri ve daha birçok cihaz internete bağlanabilme yeteneğine sahiptir. Hepsini bir çeşit arayüz sunarak kullanıcıların veri aktarımına izin vermektedir. Her geçen gün bu tip cihazların çeşitliliği ve sayısı büyük bir hızla artıyor ve bununla beraber tıp alanında da kullanılmasıyla sağlık sektöründe de oldukça faydalı hale geliyor. Günümüzde, teknoloji araştırmaları yapan Gartner şirketinin çalışmalarına göre dünya çapında milyarlarca akıllı nesne internete ve dolayısıyla birbirlerine bağlanmaktadır. Nesnelerin interneti sayesinde teknolojinin kullanım düzeyi artarak insan yaşamını daha kolay hale gelebilmektedir [9-11].

COVID-19 sürecinde nesnelerin interneti başta sağlık alanında olmak üzere, üretim, tüketim, alışveriş, güvenlik, tarım, akıllı şehircilik gibi birçok alanda hızla yaygınlaşmaya başlamıştır. Nesnelerin interneti sayesinde internete bağlı cihazlar, COVID-19 salgının önlenmesine ve salgınla baş edilmesine yönelik olarak başvuru alan en önemli araçlardan biri olmuştur. Bu bağlamda, insanlarda ateş ölçümü, hastalığın çeşitli semptomlarının gözlenmesi, sosyal mesafe ölçümü ve sosyal mesafenin korunması ve hastaların izlenimini yapan araçlar ile mobil uygulamalar, nesnelerin interneti altyapısı ile geliştirilmiştir [12]. Nesnelerin interneti aracılığıyla geliştirilen araçların varlığı, sağlık yönetimindeki olası hataların en aza indirilmesinde önemli bir rol oynamaktadır.

Literatür incelemelerinin sonucunda, bu araştırmanın motivasyonu literatürde eksik olduğu görülen COVID-19 pandemisi ve nesnelerin interneti ilişkisinin daha detaylı ve geniş kapsamlı bir analizini yapmaktır. Böylece pandemi sürecinde COVID-19'un tespitinde nesnelerin internetinin kullanımı konusundaki bilimsel ilerlemeleri ortaya koyabilmek için bu konudaki yayın akışı yıllar bazında karşılaştırmalı olarak incelenmektedir.

Bu çalışmada bireylerde COVID-19 tespitinde, nesnelerin / sağlık nesnelere internetinin kullanımı üzerine yapılan araştırmalar ve yayınlar incelenmiştir. Bu çalışmada WoS veri tabanında COVID-19'un başlangıç yılı olan 2019 yılından günümüze bir literatür taraması yapılmıştır. Uluslararası dizinde "COVID-19 / Coronavirus" ve "Internet of Things / IoT" anahtar sözcükleri ile tarama yapılmıştır. Bu sözcükler topik içerisinde (çalışmanın başlığında, özetinde, anahtar kelimelerinde) geçiyorsa ilgili çalışmalar filtrelenmiştir. Bu çalışmalar gözlem birimi olarak kabul edilmiş ve künyelerindeki bilgiler veri olarak derlenmiştir. Yeni tip koronavirüsün etkisinin azalmasıyla 2022 yılının ilk çeyreğinden itibaren pandemi önlemlerinin esnetilmeye başlanması sebebiyle analiz bu zaman aralığını kapsamaktadır. Böylece toplam 922 çalışma incelemeye alınmıştır ancak bulgular karşılaştırma amacıyla yıllar bazında gruplanmıştır. Bu çalışmanın amacı toplum sağlığı açısından önemli olan hükümet politikalarına dahi yön veren COVID-19 sürecinin değerlendirilmesinin nesnelerin interneti ile olan ilgisi üzerine akademik araştırmaların bir istatistiksel bibliyografisini elde etmektir. Bu amaçla COVID-19 tespitinde nesnelerin interneti kullanımı üzerine yapılan araştırma eğilimlerini değerlendirebilmek için bibliyometrik bir inceleme gerçekleştirilmiştir. Veri setinin oluşturulmasında Bibexcel yazılımı kullanılarak bibliyometrik analiz yapılmıştır. Bu noktada son birkaç yılda COVID-19'un seyrinde akıllı nesnelere üzerine çalışmalar yayınlayan gerek araştırmacılar gerek kuruluş veya üniversiteler gerekse ülkeler bakımından tanımlayıcı istatistikler verilmiştir. Ayrıca böyle bir çalışmada bağlantılar arasında önemli fikirler verebilecek olan anahtar sözcük ilgilere de odaklanılmıştır. Ayrıca bu olgular arasındaki atf ilişkileri bilimsel haritalama ile sergilenmek amacıyla bir sosyal ağ uygulaması olan VOSViewer yazılımı ile haritalandırılmıştır.

## 2. LİTERATÜR TARAMASI (LITERATURE RESEARCH)

Karanfiloğlu ve Kara, 2020 yılında yaptıkları çalışmada Çin örneği incelemesi yaparak enformasyon ve iletişimin dijitalleşmesinin COVID-19 pandemi sürecindeki önemini yapılan literatür ve çeşitli kaynak taraması ile ortaya koymaktadır [13]. Aydın ve Karaarslan (2020), COVID-19 belirtilerini tespit etmek için dijital bir sağlık sistemi önermektedir. Her bireyin dijital bir kopyasının oluşturulduğu dijital ikiz tabanlı bir sistem tasarlanmıştır. Bu kopya, bulut üzerinden bireyin tıbbi bilgilerine sahiptir. Verileri sağlık nesnelere interneti (IoHT) gibi bir ölçümden elde etmektedir [14]. Yiğitöl ve Sarı (2020), küresel salgınlarla mücadelede endüstri 4.0 teknolojilerinin rolünü araştırmaktadır. Nesnelerin interneti, otonom robotlar gibi yeni teknolojilerin, dünya üzerinde salgınla mücadelede ne şekilde yardımcı olduğu, örneklerle açıklanmıştır [15]. Cicioğlu ve Çalhan (2021), bulut destekli medikal nesnelerin interneti tabanlı uzaktan sağlık izleme sistem tasarlamıştır. Kalp ritmi, kan basıncı, oksijen miktarı, vücut sıcaklığı ve solunum oranı verileri anlık olarak birey ve sağlık personeli arasında paylaşılmış olup özellikle pandemi süreçlerinde kullanılabilir bir uzaktan sağlık izleme sistemi kurulmuştur [16]. Aydemir ve

Arslan, 2021 yılında yaptıkları çalışmada COVID-19 pandemi sürecinde çocukların el yıkama alışkanlığının nesnelerin interneti tabanlı sistem ile izlemesi üzerine bir araştırma yapmıştır. Bu kapsamda lavabonun bulunduğu yere yerleştirilen bir IoT düğümü ile çocuğun lavaboda el hijyeni el ve ses tanıma fonksiyonları ile tespit edilmiştir. El yıkanmıyorsa ebeveyne uyarı maili gönderilmektedir [17]. Deringöz vd. (2021), Covid-19 takibinde giyilebilir sağlık teknolojileri yöntemleri ile değerlendirmektedir. Covid-19 hastalığına yakalanmış veya yakalanma riski bulunan kişiler için en önemli faktörlerden biri olan sosyal izolasyon için giyilebilir sağlık teknolojileri ele alınmıştır [18]. Kılıç ve Tosun (2021), akıllı sağlık ekosistemi ve güncel uygulama örneklerini incelemiştir. Çalışma sonucunda mSağlık, teletıp, robotik sağlık, giyilebilir sağlık teknolojileri, büyük veri, nesnelerin interneti gibi akıllı sağlık araçlarının birbirleriyle etkileşimi ve senkronize olmasıyla "Akıllı Sağlık Ekosistemi" oluşmaya başladığı tespit edilmiştir [19]. Köse ve Kurtalan (2021), sağlık hizmetlerinde nesnelerin interneti uygulamalarının bibliyometrik analizini gerçekleştirmiştir. WoS veritabanında 2001-2019 yılları arasında yer alan makalelerin R programlama dili kullanılarak bibliyometrik analizi yapılmıştır. Yayınlar kategorize edilerek türleri, yıllara göre eğilimleri incelenmiş; ülkelere, kurumlarına ve yazarlarına göre değerlendirilmiştir [20]. Tütüncü ve Esen (2021), COVID-19 salgın yönetiminde nesnelerin internetinden hangi alanlarda faydalandığını bilim haritalarıyla tespit etmeyi amaçlanmıştır. Bu amaçla Scencedirect ve Pubmed veri tabanlarında Ocak 2020–Ocak 2021 arasında bir sene sürecinde yayınlanan 150 makale incelenmiştir. Çalışma sonucunda, nesnelerin internetinin salgın yönetiminde önemli bir role sahip olduğu görülmüştür [21]. Mısırlıoğlu, 2022 yılında yaptığı çalışmada sağlık sektöründe hizmet kalitesine ilişkin yıllar içerisinde yeni uygulamalar, teknolojik gelişmeler, standartlardaki değişimler, sosyal ve kültürel faktörler, küresel çapta krizler vb. unsurların etkisiyle büyük bir revizyon yaşandığı görüşüyle WoS veri tabanında 1991-2022 yılları arasında yer alan ve sağlık sektöründeki hizmet kalitesine yönelik olarak yapılan 3825 makalenin bibliyometrik analizini gerçekleştirmiştir [22].

## 3. METODOLOJİ (METHODOLOGY)

### 3.1. Bibliyometrik Analiz (Bibliometric Analysis)

Bilimde ilerlemek isteyen her toplumun yapılan bilimsel araştırmalar ve yayınlar gibi bazı belli başlı ölçütlerle niceliğini ve niteliğini artırabilecek girişimlerde bulunması gerekmektedir. Bu konuda en iyi yol gösterici üretilen yayınların istatistiksel yöntemlerle incelendiği bibliyometrik araştırmalar olmaktadır. Bibliyometrik çalışmalar, hem ilgili konudaki en çok atf alan araştırma ve araştırmacıları belirlerken hem de bunların arasındaki etkileşimleri gösterdiği için bilim dünyasına ve ülkelere bilimsel çalışmaların geleceği hakkında yön veren araştırmalardır [23]. Başlarda istatistiksel bibliyografi olarak değerlendirilen bir kavram; Pritchard ile birlikte, matematiksel ve istatistiksel tekniklerin yayınlanmış dergi, kitaplar ya da diğer bilgi paylaşım ve iletişim yöntemlerine uygulanabilmesi yöntemi anlamında bibliyometri olarak

tanımlamıştır [24]. Güncel literatürde görüldüğü üzere günümüzde disiplinlerarası çalışmalarla bibliyometri araştırmalarına olan ilgi oldukça artmıştır.

Kitap (biblio) ve ölçüm (metric) sözcüklerinin birleşiminden Bibliyometri kelimesi oluşmuştur. Böylece bibliyometrinin bilimsel dergilerin, makalelerin ve araştırma kuruluşlarının etkilerini belirlemeye yarayan bir ölçüm birimi olduğunu belirtilebilir. Ayrıca yayınların konu, yazar bilgisi, atıf yapılan yazar/kaynaklar, kurum, ülke gibi verilerin istatistiksel olarak incelenmesi ile ilgilenmektedir [25]. Bibliyometrik araştırmalar, yayınların kaynaklarında bulunan referansların ayrıntılı olarak incelendiği atıf analizi araştırmalarıdır. Bibliyometrik yöntemlerle ilgilenilen konuda, hangi ülkelerde ve hangi konu başlıklarında çalışıldığı, bu konuyu çalışan yazarlar, yazarlar arası iş birliği ve ortaklık ve sayıca çokluk belirlenmektedir [26, 27]. Bu çalışmada yapılan bibliyometrik analiz sonucunda yanıt aranan araştırma soruları genel olarak şu şekilde özetlenebilir:

- COVID-19 sürecinde nesnelerin interneti kullanımı konusunda yayın yapan araştırmacıların dağılımı nasıl olmaktadır?
- COVID-19 sürecinde nesnelerin interneti kullanımının yıllara göre yayın sayıları nasıl değişmektedir?
- COVID-19 sürecinde nesnelerin interneti kullanımına sağladıkları katkı bakımından ülkelerin durumu nedir?
- COVID-19 sürecinde nesnelerin interneti kullanımında hangi yazarlar ne açıdan daha fazla katkı sağlamıştır?
- COVID-19 sürecinde nesnelerin interneti kullanımı konusunda araştırma makaleleri, hangi dergilerde ne ölçüde yer almaktadır?
- COVID-19 sürecinde nesnelerin interneti kullanımı konusunda kişi başına atıf sayısı dağılımı nasıl değişmektedir?

Bibliyometrinin ilgilendiği konuların başında atıf konusu gelir. Atıf verme, düşüncenin dayanağını açıklama başka bir ifadeyle atıf yapan çalışma ile atıf yapılan çalışma arasında etkileşim ilgisini kurmaya dayanmaktadır. Bu durum atıf verilen çalışmanın içeriğinden akademik düşünce olarak etkilenildiğini gösterir. Orijinal kaynağın tanıtılması, çalışma konusuna saygı gösterilmesi, daha önce çalışma yapmış kişilere saygı, konuyla ilgili olarak daha önce yapılmış yayınların okunmasının sağlanması, araştırmada kullanılan verinin doğruluk ve gerçekliğinin ispatlanması vb. amaçlarla atıf yapılmaktadır [28].

### 3.2. Bilimsel Haritalama ve Sosyal Ağ (Scientific Mapping and Social Networking)

Sosyal ağ, ağdaki temel birim olan aktörler ve onların aralarındaki benzerlik, iş ortaklığı, arkadaşlık gibi sosyal bağlantılardan oluşan ağları niteleyen, ilişki şekillerini ortaya çıkaran bir ağ analizi alt dalıdır. Yani birbiri ile herhangi bir şekilde ilişki içinde olan aktörler sosyal ağ yapıları ortaya çıkarmaktadır. Veri yapısındaki karmaşıklığın anlaşılması ve ölçülerek sayısal olarak ifade edilmesi gerekmektedir. Karmaşıklığın görselleştirilmesi ile ağ yapıları elde edilmektedir. Görselleştirme teknikleri,

ağları daha sade ve açık, anlaşılır hale getirmek için kullanılan bir ağ analizi yöntemidir [29]. Çeşitli yazılımlar yardımıyla elde edilen ağ çizimlerinden sonra istatistiksel özellikler belirlenerek diğer ağlarla karşılaştırmalar yapılabilmektedir. Matematik, istatistik ve bilişim teknolojilerinden yardım alan sosyal ağ analizi de bu amacı gütmektedir [30]. Ağda aktörler, az sayıda bağlantıya sahip çok sayıda aktör ve çok sayıda bağlantıya sahip az sayıda aktör bulunması anlamına gelen ‘‘kuvvet yasası’’ dağılımına sahiptir. Bilimsel haritalama yöntemi bibliyometrik analiz yöntemlerinden biridir. Veri tabanlarından elde edilen yayın bilgilerine ait veri kümelerinin oluşturulması, bibliyometrik verinin analiz edilmesi, mevcut ağ yapılarının haritalandırılması ve sonuçların uzman kişiler tarafından değerlendirilmesi aşamalarını içerir [31,32]. Analiz sonuçlarının görselleştirilmesi avantajlı bir sunum sağlamaktadır ve bibliyometrik bilimsel haritalandırma yöntemi çeşitli yazılımlar kullanılarak elde edilebilmektedir. Bu çalışmada VOSViewer kullanılmıştır. Bir Sosyal Ağ Analizi aracı olan VOSViewer da analize dahil edilen veri, çalışmalara ait kaynakça ve atıf verisi olduğunda elde edilen görsel haritalar bilimsel haritalama olarak elde edilmektedir. Başka bir ifadeyle atıf verisi ile sosyal ağ analizi yapıldığında elde edilen sonuçlar bibliyometrik (bilimsel) haritalama ile sonuçlanmaktadır [33].

## 4. ANALİZ SONUÇLARI (RESULTS)

### 4.1. Veri Seti ve Tanımlayıcı İstatistikler (Data Set and Descriptive Statistics)

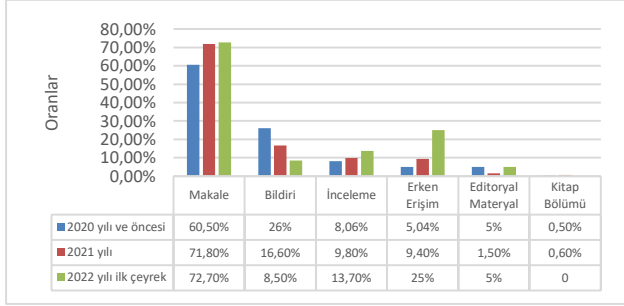
Veri seti olarak COVID-19 ile nesnelerin interneti üzerine yayınlanan 2020 yılı ve öncesi (2019 senesinde sadece 4 yayın) toplam 220 yayın, 2021 yılı 585 yayın ve 2022 yılı ilk çeyrekte (ocak, şubat, mart) 117 olmak üzere 922 çalışma incelemeye alınmıştır.

Tablo 1. Doküman çeşidi (Document type)

Dokümanlar	2020 yılı ve öncesi	2021 yılı	2022 yılı ilk çeyrek
Makaleler (Articles)	121	365	56
Konferans Bildirileri (Proceedings Papers)	57	97	10
İnceleme Yazısı (Review Articles)	19	57	16
Erken Erişim (Early Access)	12	55	29
Editoryal Materyaller (Editorial Materials)	10	9	6
Kitap Bölümü (Book Chapters)	1	1	-
Toplam	220	585	117

Çalışmaların sırasıyla en çok makale (article), bildiri (proceeding paper), inceleme yazısı (review), erken erişim (early 430ccess: tamamlanmış ama henüz bir cilt ve sayı bilgisi almamış makaleler) ve en az editoryal materyal ile kitap bölümü (book chapter) olduğu Tablo 1. Ve Şekil 1.'de görülmektedir. Ayrıca doküman tiplerinin yıllar bazında değişimi ve COVID-19 ile nesnelerin interneti üzerine en çok makale yayınlandığı görülmektedir. Tablo 2.'de COVID-19 ile nesnelerin interneti üzerine yapılan

çalışmaların üç yıl içinde ilk üçte bilgisayar bilimi (computer science), mühendislik (engineering), telekomünikasyon (telecommunications) alanlarında olduğu görülmektedir. Yayınların üç dönemde yaklaşık yarısının bilgisayar bilimleri alanında olduğu söylenebilir. Ayrıca bu konuda sağlık bilimleri (health care, microbiology, medical) alanında da yayın yapıldığı görülmektedir. Ek olarak 2022 yılından itibaren fizik (physics) ve otomasyon kontrol sistemleri alanında da bu konuda yayınların üretildiği görülmektedir.



Şekil 1. Doküman çeşidi (Document type)

Tablo 2. Araştırma alanı (Research area)

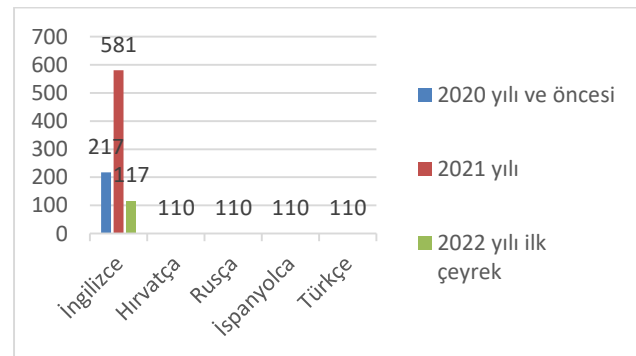
Araştırma Alanı	2020 yılı ve öncesi	2021 yılı	2022 yılı ilk çeyrek
Bilgisayar Bilimi (Computer Science)	%47.7	%47.5	%40.1
Mühendislik (Engineering)	%34.5	%40.6	%38.5
Telekomünikasyon (Telecommunications)	%24.5	%27	%12.8
Kimya (Chemistry)	%6.4	%5.5	%10.3
Malzeme Bilimi (Materials Science)	%5.5	%3	%8.5
İş Ekonomisi (Business Economics)	%4.5	%4.1	%7.7
Sağlık Bilimi Servisleri (Health Care Sciences Services)	%5	%1.7	%6
Enstrümantasyon (Instruments Instrumentation)	%3.6	%5.6	%9.4
Biyoteknoloji Uygulamalı Mikrobiyoloji (Biotechnology Applied Microbiology)	%3.6	%2	Fizik (Physics) %6
Tıp Bilişimi (Medical Informatics)	%3.1	%1.5	Otomasyon Kontrol Sistemleri (Automation Control Systems) %3.4

Tablo 3.'te COVID-19 ile nesnelerin interneti üzerine yapılan çalışmaların hangi ülkelere yayınlandığı bilgileri özetlenmiştir. 2020 yılı ve öncesinde Hindistan (India), Amerika (USA), Çin (Peoples R China) ve Suudi Arabistan (Saudi Arabia); 2021 yılında Hindistan, Çin, Amerika ve Suudi Arabistan; 2022 yılı ilk çeyrekte Hindistan, Amerika, Çin ve Suudi Arabistan yer almaktadır. 2022

yılında önceki yıllara nispeten Hindistan ve Suudi Arabistan da çalışmalar artmışken Amerika ve Çin'de daha az olmuştur. Ayrıca Türkiye'nin de her geçen yıl artan oranda yayına sahip olduğu ve yeni ülkelerin de katıldığı görülmektedir. Şekil 2.'ye göre yayınların genelde İngilizce olarak yayınlandığı görülmektedir.

Tablo 3. Ülkeler (Countries)

Ülke	2020 yılı ve önce	2021 yılı	2022 yılı ilk çeyrek
Hindistan (India)	%19	%20	%25.6
Amerika Birleşik Devletleri (USA)	%15.5	Peoples R China %14.5	USA %10.3
Çin Halk Cumhuriyeti (Peoples R China)	%10.5	USA %13.5	Peoples R China %8.6
Suudi Arabistan (Saudi Arabia)	%8.6	%11.6	%18
İngiltere (England)	%7.7	%8.5	%5.1
İspanya (Spain)	%5.9	%3.2	%2.6
Avustralya (Australia)	%5	%5.8	%7.7
İtalya (Italy)	%4.5	%6.2	%9.4
Romanya (Romania)	%4.5	-	Norveç (Norway) %6
Malezya (Malaysia), Kanada (Canada)	%3.6, %3.6	%4.6, %6	%3.4, %3.4
Brezilya (Brazil), Güney Kore (South Korea)	%3.6, %3.6	%2.5, %6.5	-, -
Japonya (Japan)	%3.2	-	%6
İran (Iran)	%2.7	Tayvan (Taiwan) %3	İran (Iran) %3.4
Türkiye (Turkey)	%2.7	%3	%6
Pakistan (Pakistan)	%2.3	%6.5	%6.8



Şekil 2. Yayın dili (Publication language)

Tablo 4. En çok atf alan çalışmalar  
(Most cited works)

Başlık	Yazar	Yıl	Yayıncı	Atf
-The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? ( <i>SARS, MERS ve yeni koronavirüs (COVID-19) salgınları, en yeni ve en büyük küresel sağlık tehditleri: Hangi dersleri aldık?</i> )	Peeri vd. [34]	2020	INTERNATIONAL JOURNAL OF EPIDEMIOLOGY	602
-A comprehensive review of the COVID-19 pandemic and the role of IoT, Drones, AI, Blockchain, and 5G in managing its impact ( <i>COVID-19 Pandemisinin Kapsamlı Bir İncelemesi ve Etkisini Yönetmede IoT, Drones, AI, Blockchain ve 5G'nin Rolü</i> )	Chamola vd. [35]	2020	IEEE ACCESS	250
-A drone-based networked system and methods for combating coronavirus disease (COVID-19) pandemic ( <i>Drone tabanlı ağ bağlantılı bir sistem ve koronavirüs hastalığı (COVID-19) pandemisiyle mücadele yöntemleri</i> )	Kumar vd. [36]	2021	FUTURE GENERATION COMPUTER SYSTEMS	57
-An intelligent framework using disruptive technologies for COVID-19 analysis ( <i>COVID-19 analizi için yokedici teknolojileri kullanan akıllı bir çerçeve</i> )	Abdel-Basset vd. [37]	2021	TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE	50
-Internet of medical things (IoMT)-integrated biosensors for point-of-care testing of infectious diseases ( <i>Bulaşıcı hastalıkların bakım noktası testi için tıbbi nesnelerin interneti (IoMT)-entegre biyosensörler</i> )	Jain vd. [38]	2021	BIOSENSORS & BIOELECTRONICS	40
-Realizing an effective COVID-19 diagnosis system based on machine learning and IoT in smart hospital environment ( <i>Akıllı hastane ortamında makine öğrenimi ve IoT'ye dayalı etkili bir COVID-19 teşhis sistemi gerçekleştirme</i> )	Abdulkareem vd. [39]	2021	IEEE INTERNET OF THINGS JOURNAL	36
-A lightweight and robust secure key establishment protocol for internet of medical things in COVID-19 patients care ( <i>COVID-19 hasta bakımında tıbbi nesnelerin interneti için hafif ve sağlam güvenli anahtar oluşturma protokolü</i> )	Masud vd. [40]	2021	IEEE INTERNET OF THINGS JOURNAL	35

Tablo 5. Yayıncı kuruluş  
(Publisher organization)

Yayıncı Kuruluş	2020 yılı ve öncesi	2021 yılı	2022 yılı ilk çeyrek	
IEEE	%33.2	%27	%14.5	
MDPI	%12.3	%15.9	%19.7	
Emerald Group Publishing	%7.3	%2.9	%4.3	
Springer Nature	%7.3	%10.1	%24.8	
Elsevier	%5.9	%15.6	%4.2	
Wiley	%2.7	%2.6	%4.3	
Mary Ann Liebert, Inc	%2.3	%1.7	Tech Science Press	%8.5
Ios Press	%1.8	-	%1.7	
Taylor & Francis	%1.8	%1.2	%3.4	
Sage	%1.4	%1	Frontiers Media	%2.6

WoS kategorilerine göre COVID-19 ile nesnelerin interneti konusu üzerine yapılan çalışmalar arasında en çok atf alanlar Tablo 4.'te detaylandırılmıştır. Buna göre 2020 yılında Peeri vd. tarafından yapılan yeni ve büyük küresel sağlık tehdidi salgınların incelendiği çalışmanın bilimsel katkısı en yüksek olmuş ve Uluslararası Epidemiyoloji Dergisi'nde yayınlanmıştır. Ardından COVID-19'un etkisini yönetmede nesnelerin interneti, Drones, yapay zeka, Blockchain ve 5G'nin rolünün incelendiği Chamola vd. tarafından yapılan IEEE ACCESS'de yayınlanmış ve çalışmaya ilgi yüksek olmuştur. 2021 yılında Kumar vd. tarafından yapılan Drone tabanlı ağ bağlantılı bir sistemle COVID-19 ile mücadele çalışması Gelecek nesil bilgisayar sistemleri dergisinde yayınlanmış ve 2021 yılından en fazla atf alan yayın olmuştur. Burada son olarak değinilecek olan IEEE nesnelerin internetiyle COVID-19 teşhis sistemi

ve Masud vd.'nin COVID-19 hasta bakımında tıbbi nesnelerin interneti için güvenli anahtar oluşturma çalışmaları ilgi çeken diğer bazı çalışmalardır. Bu çalışmaların yüksek miktarda atf alan çalışmalar olmaları sebebiyle diğer çalışmalar tarafından en çok etkilenen ve en sık başvuru alan başyapıt niteliğinde içerikler olduğu söylenebilir. Tablo 5.'e göre 2020 yılı ve öncesi çalışmaların %33.2'si IEEE; 2021 yılı %27'si IEEE, %15.9'u Multidisipliner Digital Publishing Institute (MDPI) ve %15.6'si Elsevier; 2022 yılı ilk çeyrek de %24.8'i Springer Nature, %19.7'si MDPI tarafından yayınlanmıştır. Bu konuya IEEE dışında da yayıncıların ilgisinin artmaya başladığı görülmektedir.

#### 4.2. Bulgular (Findings)

Yıllara göre değişim baz alınarak anahtar kelimelere, dokümanlara, yazarlara, yayıncı kurum/kuruluşlara ve ülkelere göre yapılan bibliyometrik analizin sonuçları sırasıyla aşağıda paylaşılmaktadır.

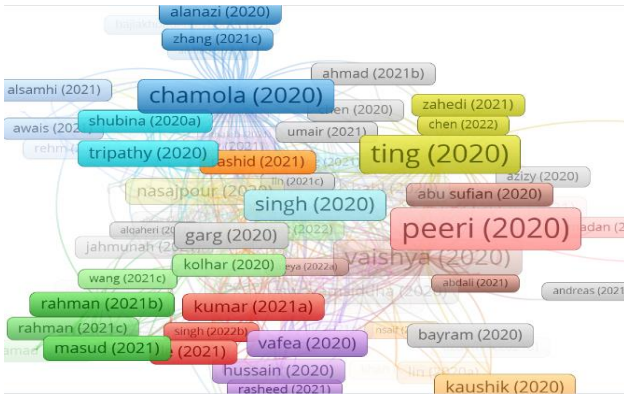
Tablo 6.'ya göre COVID-19 ve nesnelerin interneti gibi kilit anahtar kelimelerin yanında başta makine öğrenmesi ve yapay zeka olmak üzere blok zincir (blockchain), gizlilik ve güvenlik, derin öğrenme, sağlık bakımı (healthcare), bulut bilişim (cloud computing), biyomedikal izleme (biomedical monitoring), sensörler, teletıp/telesahlık, sağlık nesnelerinin interneti, büyük veri, robotik, aşı (vaccines), sosyal mesafe (social distancing) gibi anahtar kelimeler en sık kullanılanlar listesine girmiştir.







Tablo 6.'daki toplam bağlantı gücü ile Şekil 3.'de verilen sosyal ağ analizi sonuçlarında aktörlerin (düğüm noktaları veya birimler) birbirleriyle bağlantı sayısından elde edilmektedir ve ağdaki merkezi ve arasında konumlandırılmada bu değerlere başvurulmaktadır. Ağ yapılında COVID-19 anahtar kelimesinin merkezi konumu diğer kelimelerle kombinasyonlarından kaynaklanmaktadır. Yukarıdaki kelimelere ek olarak 2021 yılında COVID-19 ve nesnelere interneti konu ile ilgili olarak elde edilen verilerden yeni bilgiler elde etme amacıyla kullanılan derin öğrenme (deep learning), akıllı cihazların üretilen veriyi yerel noktalarda analiz etmelerini temel alan sis bilişim (fog computing), medikal robotlar ve güvenlik (security) daha çok çalışılan kelimeler olmuş ve ağda da görünür olmuştur. Benzer şekilde 2022 yılı itibari ile medikal nesnelere interneti (internet of medical things) göze çarpmaktadır (Şekil 3 b. ve c.).



Şekil 4. Dokümanlar arası atıf ağı  
(Co-document citation network)

Şekil 4.'deki analize göre COVID-19 ve nesnelere internetiyle ilgili olup en çok atıf alan doküman daha önce de bahsedildiği üzere ilk yazarı ile kısaltılmış olan ve boyutu en yüksek olan Peeri vd. (2020), [28]'dir. Bu eser çalışma konusu ile diğer yazarları etkilemekte ve atıf almaktadır. Bu analiz belgelerin kaynakçalarında referans verdikleri diğer kaynaklarla olan bağlantılarından ve diğer kaynaklardan aldıkları atıflardan hesaplanarak elde edilen toplam bağlantı sayılarının üzerinden ağda konumlandırma yapılmaktadır. Benzer renkler birbirine atıf yapan kümelenme yapılarına bir atıf yapar.

Şekil 5'teki analize göre (Tablo 7.'de görülebilir) Haque, Ubydul bu konuda en çok atıf alan fikir önderi yazardır. Peeri (2020) çalışmasında da ortak yazarlardan biri olup ayrıca bu konuda başka bir yayına da sahiptir. Diğer en çok atıfa sahip olan Aghamohammadi Nasrin, Baghbanzadeh Mahdi, Bibi Saana, Peeri Noah yakın bir atıf ağı kümesi oluşturuyorken; sonraki yüksek atıfa sahip olan Aavid Mohd ağda merkezi bir konumda yer alarak çevre kümelerle atıf ilişkisi içinde bulunuyor. Ardından azalan sırada Wong Tien; Guizzani Mohsen; Hossain, M. Shamim; Gupta Deepak'ın diğer atıf ağı kümelerini oluşturdukları görülüyor. Bu yazarlar bu konuda fikrine en çok başvuru öncü yazarlar olarak anılabilir. En çok yayına

sahip olan yazar ise Hossain, Shamim (7)'dir. Hossain, Shamim bu konuda yayın sayısı en çok olan yazar olması sebebiyle görece daha yüksek bağlantı sayısına sahiptir. Toplam atıf sayısı bakımından ise henüz ilk ona girememektedir.

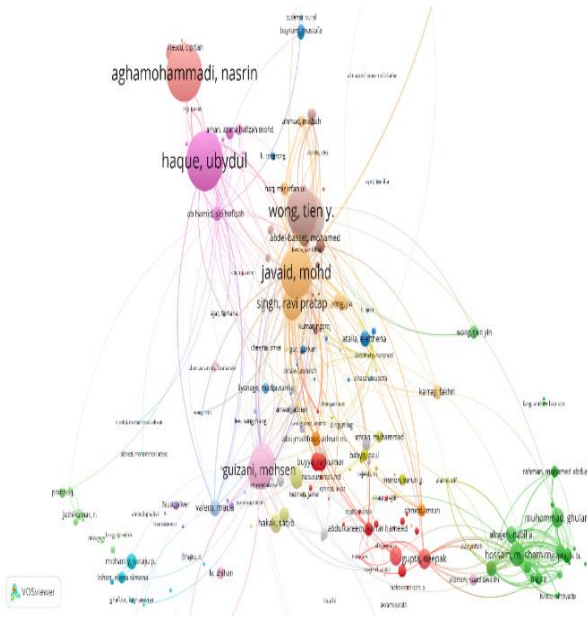
Tablo 7. Yazarların yayın dağılımı  
(Distribution of authors' publications)

Yazar	Doküman sayısı	Atıf sayısı	Toplam bağlantı kuvveti
Haque, Ubydul	2	641	168
Aghamohammadi, Nasrin	1	602	77
Baghbanzadeh, Mahdi	1	602	77
Bibi, Saana	1	602	77
Javaid, Mohd	6	489	518
Haleem, Abid	5	454	504
Dzau, Victor	1	403	217
Khan, Ibrahim Haleem	2	303	208
Guizzani, Mohsen	5	262	311
Chamola, Vinay	3	256	304
Hossain, M. Shamim	7	103	258
Chakraborty, Chinmay	5	83	121
Gupta, Deepak	5	79	92
...	...	...	...
Kumar, Adarsh	1	57	78

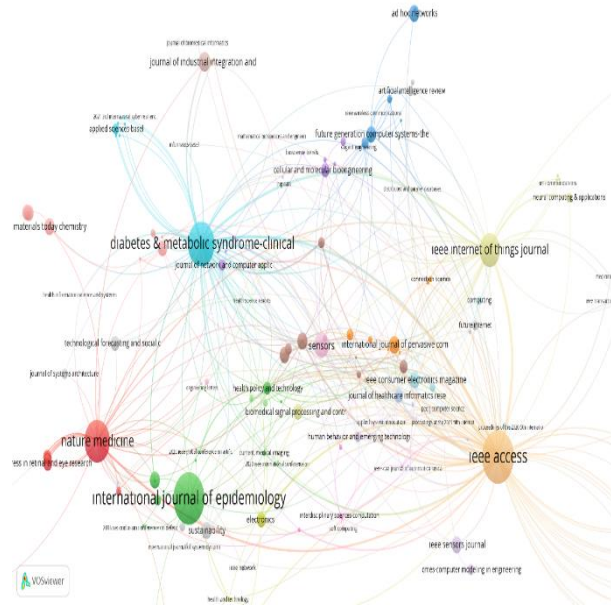
Şekil 6.'daki analize göre yayıncı kurum/kuruluşlar arasındaki atıf ağı yapısı ve ağ üzerinde konumlanma görülmektedir. İlgili konuda en çok sayıda yayına sahip olan IEEE ACCESS'in yüksek atıf sayısından dolayı ağda merkezi konumda olması beklenen bir durumdur. Benzer şekilde çevrelerindeki diğer yayınlardan oldukça atıf alan beş ana kümeden (renkler: yeşil, turuncu, mavi, kırmızı, sarı) bahsedilebilir.

O halde atıf sayısının çok oluşu bakımından değerlendirildiğinde kümelerin merkezinde (büyük aktör noktalarını temsil eden) International Journal of Epidemiology, IEEE ACCESS, Diabetes & Metabolic Syndrome-Clinical, Nature Medicine ve IEEE Internet of Things Journal dergileri bulunmakta olup bu konuda başta küme elemanları olmak üzere diğer yayıncılar tarafından en çok başvuru yapılan kaynaklar olmuşlardır. Bu konuda IEEE'nin potansiyel olarak çalışma yayınlanabilecek yerlerin başında geldiği söylenebilir.

İlk 20 kurum/kuruluşla sınırlandırılan Tablo 8.'e göre COVID-19 ve nesnelere interneti konusunda yayın yapan ve doküman sayısı en yüksek olan Kral Saud Üniversitesi (28), Taif Üniversitesi (16), Kral Abdulaziz Üniversitesi (14) (Suudi Arabistan) olup bu durum Şekil 7.a.'daki doküman yoğunluğu haritasında da büyük boyutta olan ve sarı renkle merkezi yoğunluğa sahip olan analiz sonucunda görülmektedir.



Şekil 5. Yazarlar arası atıf ağı  
(Co-author citation network)



Şekil 6. Yayıncı kurum/kuruluşlar arası atıf ağı  
(Institutions/organizations citation network)

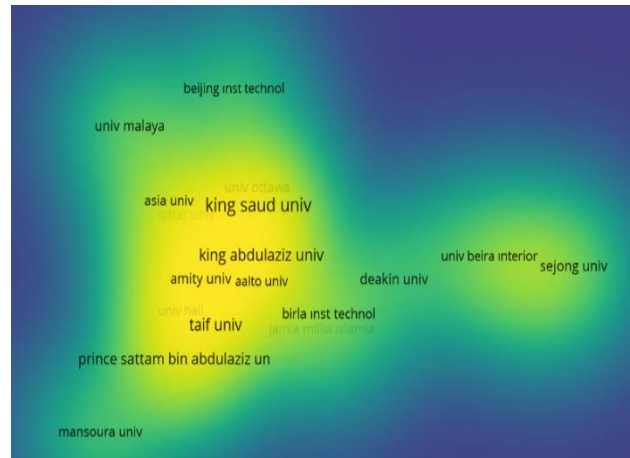
Atıf sayısı en yüksek olanlar ise Kuzey Texas (USA) ve Malaya (Malezya) Üniversiteleri ardından Jamia Millia Islamia (Hindistan) Katar, Kral Saud, Deakin (Avustralya) ve Kral Abdulaziz Üniversiteleri olup bu durum Şekil 7.b.'deki atıf yoğunluğu haritasında da büyük boyutta olan ve sarı renkle merkezi yoğunluğa sahip olan analiz sonucunda görülmektedir. Şekillerde Kuzey Texas ve Malaya Üniversiteleri üst üste çakıştığından sadece Kuzey Texas görülmektedir. Şekil 7.c.'de kurumlar arası atıf sosyal ağı görülmektedir. Benzer renkler birbirine ve diğerlerine atıflar bakımından birlikte kümelenme yapılarını temsil etmektedir. İlk 20 ülkeyle sınırlandırılan Tablo 9.'a göre COVID-19 ve nesnelerin interneti

konusunda yayın yapan ve doküman sayısı en yüksek olan ülkeler Hindistan (191), Amerika (125), Çin (118), Suudi Arabistan (108), İngiltere (73), Güney Kore (South Korea) (56), Avustralya (Australia) (54) ve Pakistan (51)'dir.

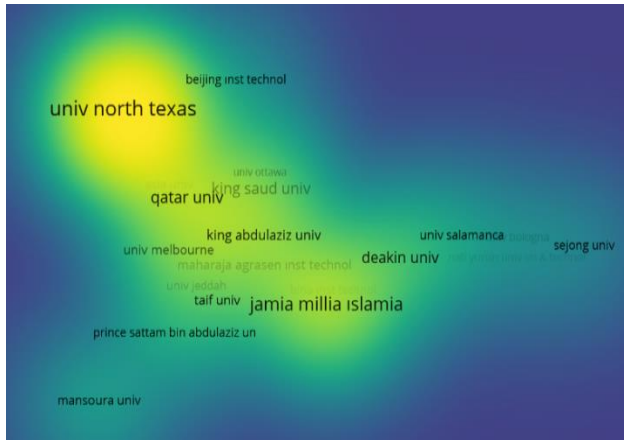
Tablo 8. Kurum/Kuruluşların yayın dağılımı  
(Publication distribution of institutions/organizations)

Kurum/Kuruluş	Doküman sayısı	Atıf sayısı	Toplam bağlantı kuvveti
Kuzey Texas Üni.	5	<b>687</b>	2
Malaya Üni.	8	673	4
Jamia Millia Islamia Ü.	6	457	2
Katar Üni.	7	263	3
Kral Saud Üni.	<b>28</b>	235	16
Deakin Üni.	9	193	3
Kral Abdulaziz Üni.	14	106	19
Melbourne Üni.	8	85	1
Maharaja A.Tekn.Ens.	5	79	8
Vellore Tekn. Ens.	10	75	1
Taif Üni.	16	72	20
Salamanca Üni.	5	65	3
Asia Üni.	7	59	12
Mansoura Üni.	6	59	5
Petr Üni. Energy Araş.	6	59	8
Zagazig Üni.	6	59	3
Lovely Profess Üni.	6	58	10
Kennesaw Devlet Üni.	5	52	1
Brno Tekn. Üni.	5	51	4
National Tekn. Ens.	10	47	2

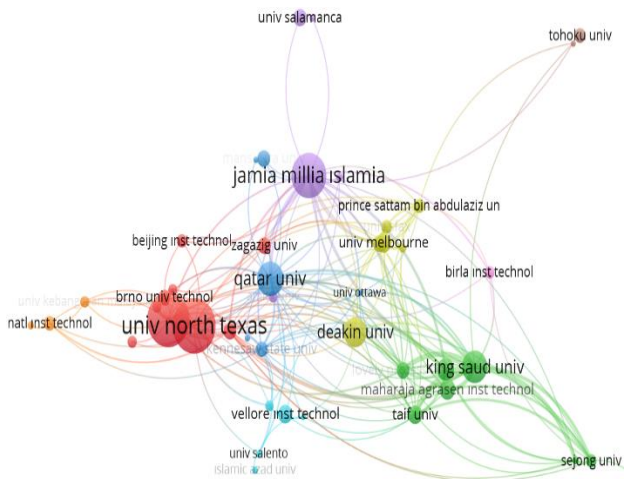
Türkiye (31)'de bu listeye girmiştir. Atıf sayısı en yüksek olanlar arasında ise Amerika, Hindistan, Çin, Pakistan, Malezya (Malaysia), Bangladeş (Bangladesh), İran ve Singapur (Singapore) gelmektedir.



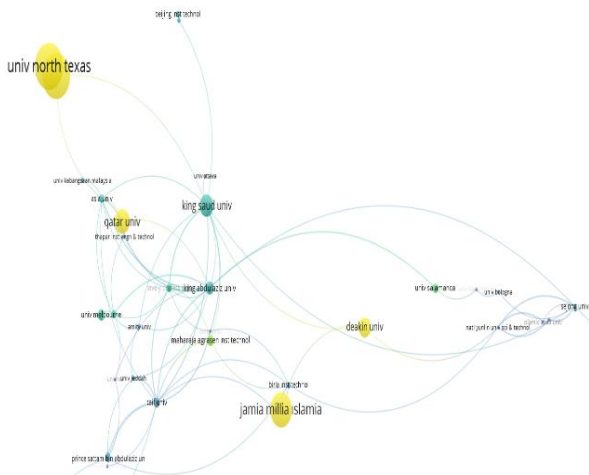
a. Doküman sayısı yoğunluk haritası  
(Number of Documents Density Map)



b. Atıf sayısı yoğunluk haritası  
(Citation Density Map)



c. Atıf ağı  
(Citation network)



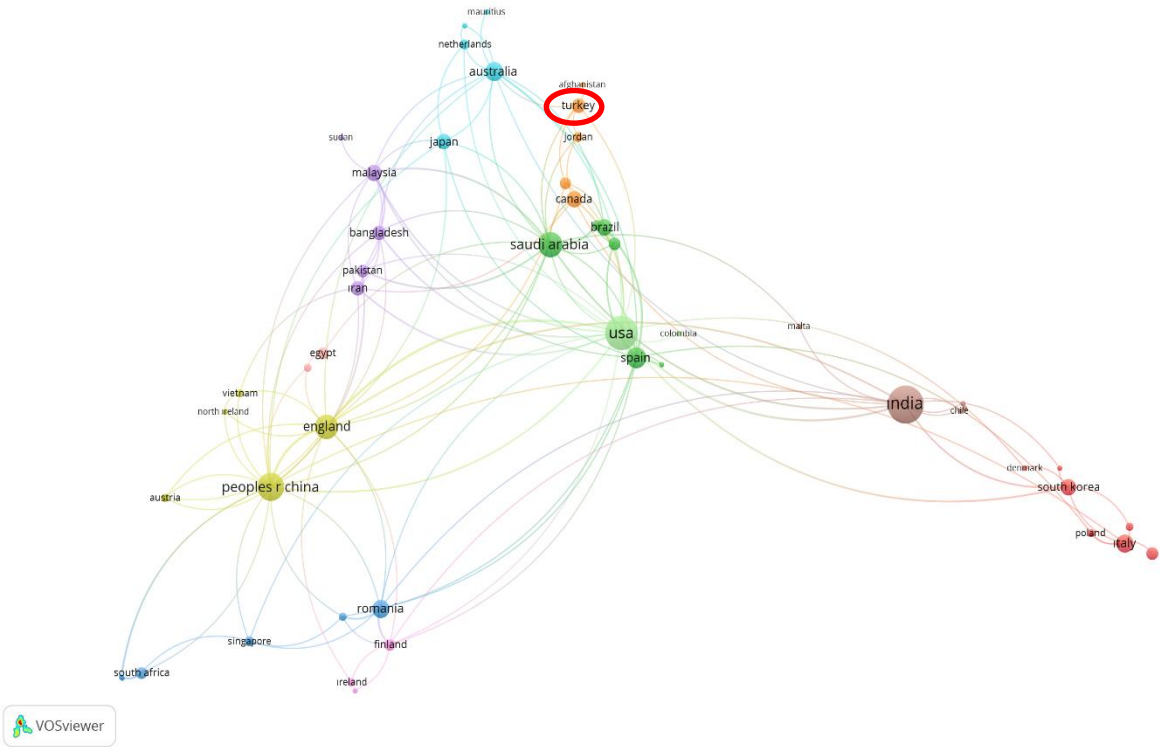
d. Ortak yazarlık atıf ağı  
(Co-authoring citation network)

Şekil 7. Kurum/Kuruluşların atıf sayılarına göre sosyal ağ  
(Social network according to the number of citations of institutions / organizations)

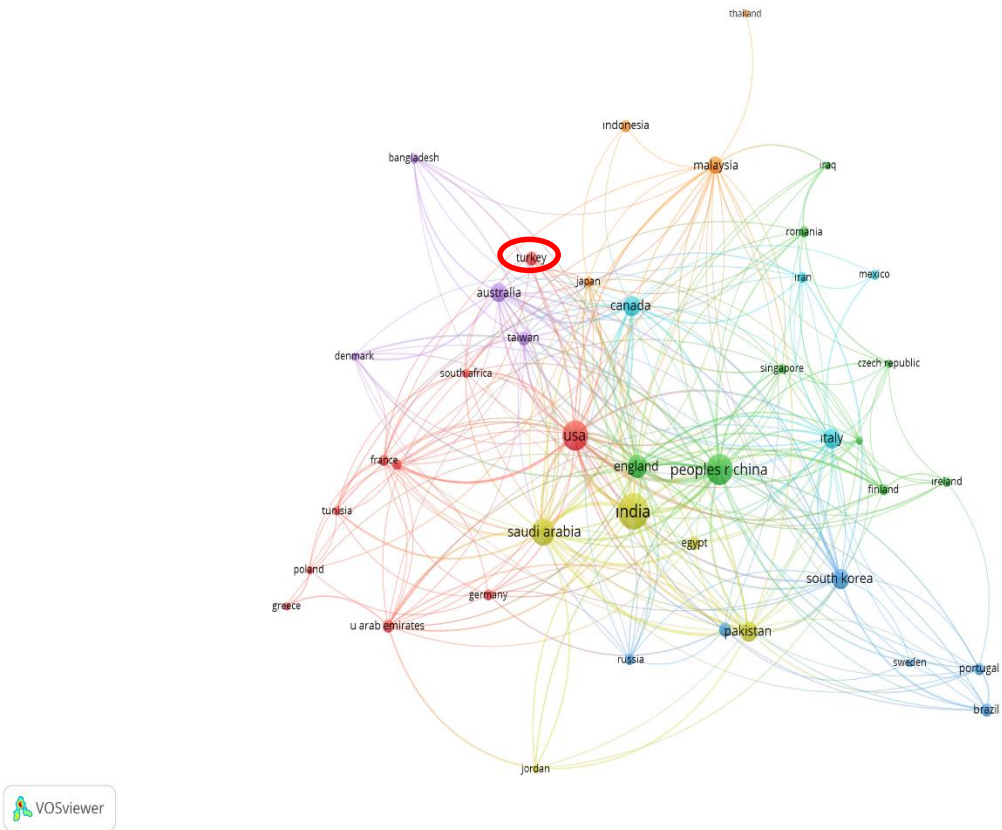
Tablo 9. Ülkelerin dağılımı  
(Distribution of countries)

Ülkeler	Doküman sayısı	Atıf sayısı	Toplam bağlantı kuvveti
Amerika Birleşik Devletleri (USA)	125	<b>1897</b>	174
Hindistan (India)	<b>191</b>	1597	174
Çin Halk Cumhuriyeti (Peoples R China)	118	1373	146
Pakistan (Pakistan)	51	822	108
Malezya (Malaysia)	39	779	56
Bangladeş (Bangladesh)	18	683	28
İran (Iran)	22	682	33
Singapur (Singapore)	14	613	28
Suudi Arabistan (Saudi Arabia)	108	584	172
İngiltere (England)	73	543	121
Avustralya (Australia)	54	528	100
Güney Kore (South Korea)	56	292	86
Katar (Qatar)	13	291	22
Kanada (Canada)	47	290	76
Japonya (Japan)	24	203	39
İtalya (Italy)	56	201	64
Mısır (Egypt)	26	146	34
Tayvan (Taiwan)	27	130	40
İspanya (Spain)	35	126	46
Türkiye (Turkey)	31	114	38

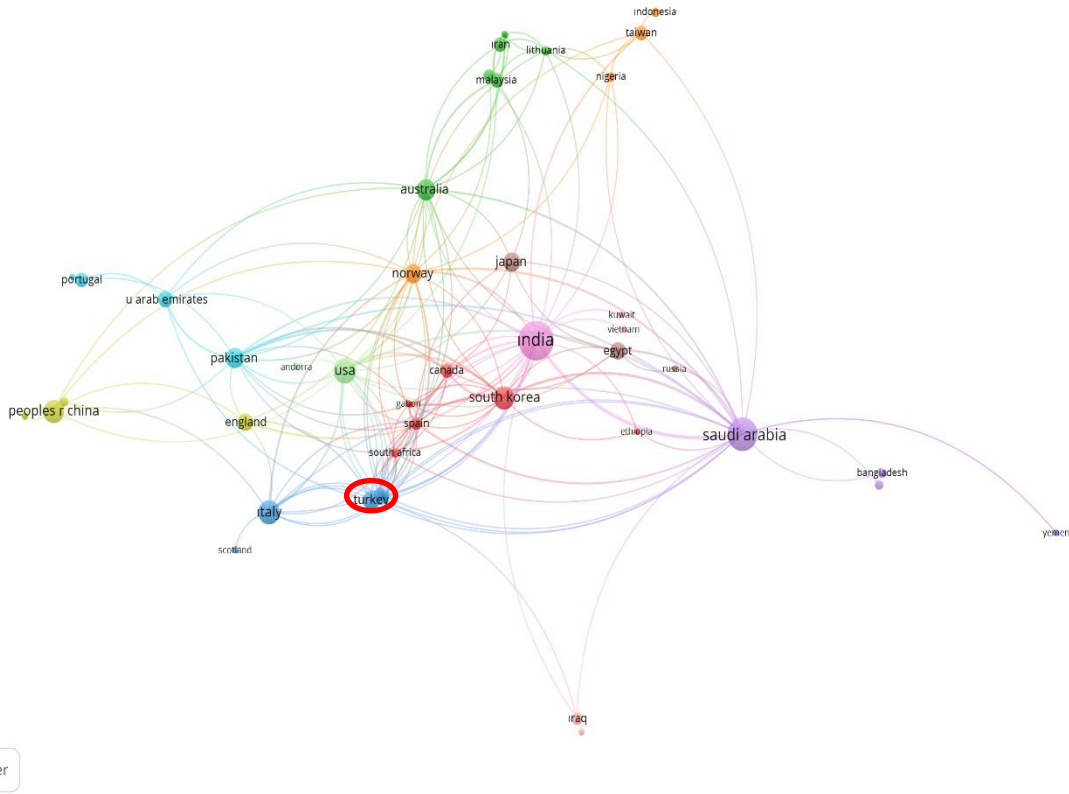
Şekil 8.'de yıllara göre ülkeler arası sosyal ağ yapıları atıflara göre oluşturulmuştur. Şekil 8. a. b. ve c. karşılaştırıldığında doküman sayısı 2021 yılında daha fazla olduğu için yoğun bir ağ ile karşılaşılmaktadır. 2020 yılına göre hem 2021 hem de 2022'de ülkeler arası atıf ağı yoğunluğu, karşılıklı iş birliği ve referans gösterme sıklığı artmıştır. Benzer renklerden oluşan kümeler belirginleşmiş ve ilgili konuda ülkeler arası bilimsel etkileşim çoğalmıştır. Şekil 8. d.'de verilen tüm yıllar için atıf yoğunluğu haritası üzerinde sarı renkli kısım en çok atıf alan ülkelerin olduğu USA ve Hindistan'a (India) dikkat çekmektedir. Ayrıca ağ yapıları üzerinde Türkiye'nin konumu işaretlenmiştir. Türkiye'nin ağ üzerinde yakın konumlandığı ülkelerle daha fazla atıflaşma yaptığı söylenebilir.



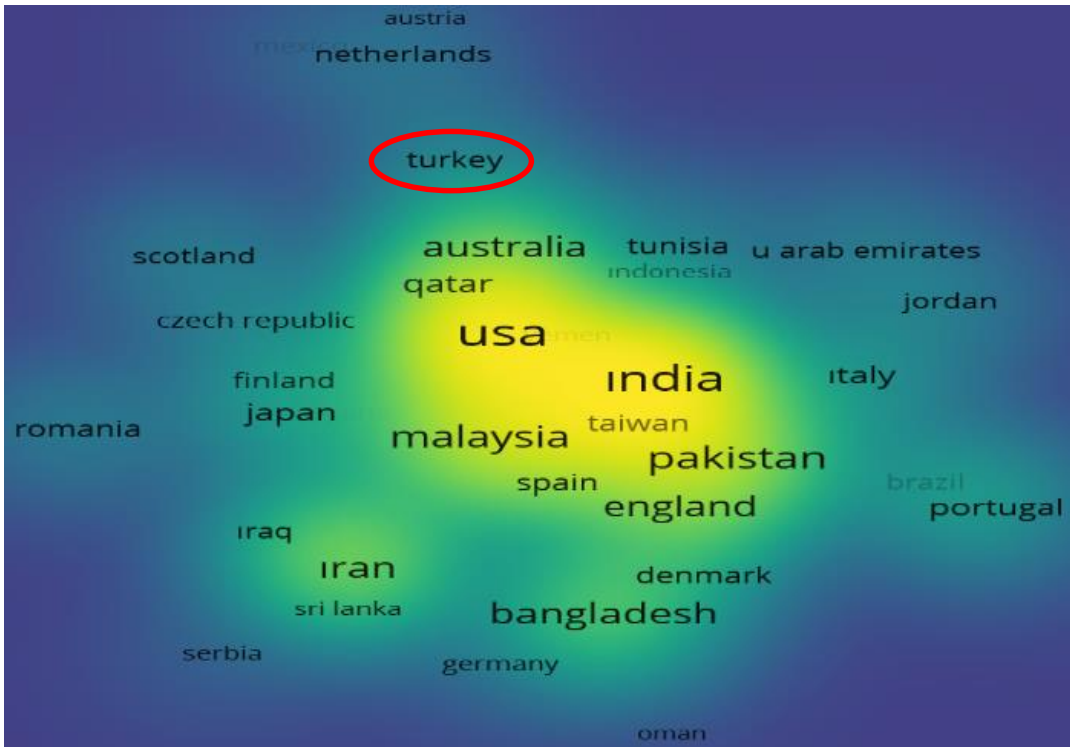
a. 2020 ve öncesi (2020 and before)



b. 2021 yılı (2021 year)



c. 2022 ilk çeyrek (2022 first quarter)

d. Tüm yıllar birlikte atıf yoğunluğu haritası  
(Citation density map for all years together)Şekil 8. Yıllara göre ülkeler arası sosyal ağlar  
(Co-country social networks by year)

## 5. SONUÇLAR VE TARTIŞMA (CONCLUSIONS AND DISCUSSION)

Anahtar kelimelere, dokümanlara, yazarlara, yayıncı kurum/kuruluşlara ve ülkelere göre yapılan bibliyometrik analizin sonuçlarını aşağıda özetlenmiştir.

- Pandemi sürecinde yayınlanan yayınların anahtar kelimelerinde; COVID-19 ve nesnelere interneti ana kelimelerinin yanında başta makine öğrenmesi ve yapay zeka olmak üzere blok zincir, gizlilik ve güvenlik, derin öğrenme, sağlık bakımı, bulut bilişim, biyomedikal izleme, sensörler, teletıp/telesaglık, sağlık nesnelere interneti, büyük veri, robotik, aşı, sosyal mesafe gibi anahtar kelimeler en sık kullanılanlar olduğu görülmüştür. Ek olarak 2020'den 2021'e geçince sosyal ağ içinde göze çarpan derin öğrenme, sis bilişim, medikal robotlar ve güvenlik kelimeleri olmuştur. 2022 yılı itibari ile medikal nesnelere interneti ile ilgili çalışmalar görülmektedir.
- COVID-19 ve nesnelere internetiyle ilgili olup en çok atf alan doküman ilk yazarı ile kısaltılmış olan Peeri (2020)'dir. Bu çalışmanın bu konuda önemli bir eser olduğu diğer yayınlara öncü olduğu söylenebilir.
- Haque, Ubydul bu konudaki çalışmaları ile en çok atf alan yazardır. Aynı zamanda Peeri (2020) çalışmasında da ortak yazarlardan biri olup ayrıca bu konuda başka yayına da sahiptir. Bu yazarın bu konuda başyazar olduğu diğer yazarları etkilediği söylenebilir. Bu konuda en çok çalışma yapan, en çok yayına sahip olan yazar ise Hossain, M. Shamim (7 doküman) olmuştur.
- Yayıncı kurum/kuruluşlar arasında bu konuda en çok dokümanı yayınlayan IEEE ACCESS iken en çok atf alan ise International Journal of Epidemiology dergisidir. Bu dergilerin bu konuda başvurulan temel dergiler olduğu söylenebilir.
- Doküman sayısı en yüksek olan kurumlar Kral Saud, Taif ve Kral Abdulaziz Üniversiteleridir (Suudi Arabistan). Bu kurumların bu konuda en çok çalışma sağlayan kurumlar olduğu söylenebilir.
- Atf sayısı en yüksek olan kurumlar ise Kuzey Texas (USA) ve Malaya (Malezya) Üniversiteleri ardından Jamia Millia Islamia (Hindistan) Katar, Kral Saud, Deakin (Avustralya) ve Kral Abdulaziz Üniversiteleridir.
- Doküman sayısı en yüksek olan ülkeler Amerika, Hindistan, Çin, Suudi Arabistan, İngiltere, Güney Kore, Avustralya ve Pakistandır. Türkiye de ilk yirmilik listeye gitmiştir.
- Atf sayısı en yüksek ülkeler ise Amerika, Hindistan, Çin, Pakistan, Malezya, Bangladeş, İran ve Singapur gelmektedir.
- Doküman sayısı pandeminin en yoğun olduğu 2021 yılında daha fazla olduğu için yayınlanan akademik çalışmalar arasındaki ağ yapısı da yoğunlaşmıştır. 2020 yılına göre hem 2021 hem de 2022'de ülkeler arası atf ağı yoğunluğu, karşılıklı iş birliği ve referans gösterme sıklaşmıştır. Benzer renklerden oluşan kümeler belirginleşmiş ve ilgili konuda ülkeler arası bilimsel etkileşim çoğalmıştır. Ayrıca ağ üzerinde Türkiye'nin konumu ve bu alanda çalışmalar yayınladığı da gözle görülmektedir.

Pandemi sürecinde nesnelere internetinin kullanımı konusu oldukça yeni ve güncel olmasına rağmen disiplinlerarası bir alana yayılmış olduğu görülmektedir. Bu durum bu konuda bilimsel araştırma ve çalışma yapma sayısının da gün geçtikçe artmasına olanak tanımaktadır. Öyle ki 2020'den 2021'e yayın sayısı iki katından daha fazla yükselmiştir. Pandemi sürecinde edinilen tecrübeler bu tür süreçlere hazırlıklı olmayı gerekli kılmaktadır. Dolayısıyla bu durum da bu tür araştırmaların ve bilimsel gelişmelerin devam edebileceğini göstermektedir. Konu hakkında hızlı okumalar yaparak bilgi sahibi olmak isteyen araştırmacılar, yapılan incelemeler sonucunda alan literatürü genel itibariyle özetlendiğinden ve konu hakkında araştırma yapılırken takip edilebilecek anahtar kelimeler, yazarlar, ülkeler, yayın ve yayıncılara yer verildiğinden dolayı bu çalışmaya kaynak olarak başvurabilirler. Yeni çalışmalarda COVID-19 sürecinin diğer büyük veri teknolojileri (makine öğrenmesi, derin öğrenme, bulut bilişim vb.) ile ilişkisi incelenebilir.

## KAYNAKLAR (REFERENCES)

- [1] M. Aminizadeh, M. Farrokhi, A. Ebadi, GR. Masoumi, P. Kolivand, HR. Khankeh, "Hospital management preparedness tools in biological events: A scoping review", *J Educ Health Promot*, 8:234, 2019.
- [2] B. Yücesan ve Ö. Özkan, "COVID 19 Pandemi Sürecinin Sağlık Yönetimi Açısından Değerlendirilmesi", *Avrasya Sağlık Bilimleri Dergisi, COVID-19, Special Issue*, 134-139, 2020.
- [3] Internet: T. Liang, Handbook of COVID-19 Prevention and Treatment, <https://files.alicdn.com/tps/service/ee33bb33f3583ea521f1e17e0800357e.pdf?spm=a3c0i.14138300.8102420620.download.6df3647f7wa4H, 15.01.2022>.
- [4] D. Tütüncü, M. F. Esen, "Salgın Hastalıkların Yönetiminde Nesnelere İnterneti (NİT) kullanımı: COVID-19 Örneği", *Sağlık Akademisyenleri Dergisi*, 8(2), 169-177, 2021.
- [5] A. Şenol, Y. Canbay, M. Kaya, "Makine Öğrenmesi Yaklaşımlarını Kullanarak Salgınları Erken Evrede Tespit Etme Alanındaki Eğilimler", *Bilişim Teknolojileri Dergisi*, 14(4), 2021.
- [6] U. Bayram, "Revealing the Reflections of the Pandemic by Investigating COVID-19 Related News Articles Using Machine Learning and Network Analysis", *Bilişim Teknolojileri Dergisi*, 15(2), 2022.
- [7] R. H. Weber, "Internet of Things- New Security and Privacy Challenges", *Computer Law&Security Review*, 26(1), 23-30, 2010.
- [8] D. Çulha, "Nesnelere İnterneti Aygıtlarının Görüşmelerinin Blok Zinciri", *Bilişim Teknolojileri Dergisi*, 14(2), 2021.
- [9] B. Kocank, "Büyük Veri Üzerine Uluslararası Literatürün Bibliyometrik Analizi", Yayınlanmamış Yüksek Lisans Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, 2018.
- [10] İnternet: Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016, <https://www.gartner.com/newsroom/id/3598917, 15.02.2022>.

- [11] A. Gacar, “Veriye Dayalı Kurumsal Risk Yönetimi”, **Veriye Dayalı İşletme Yönetimi**, Editör: Nermin Kişi, Gazi Kitabevi, Birinci Baskı, Ankara, 2021.
- [12] M. Nasajpour, S. Pouriye, R.M. Parizi, “Internet of Things for Current COVID-19 and Future Pandemics: an Exploratory Study”, *Journal of Healthcare Informatics Research*, (4), 325–364, 2020.
- [13] M. Karanfiloğlu, N. Kara, “İletişimin Dijitalleşmesi: Pandemi (COVID-19) ve Enformasyon Teknolojileri”, *AJIT-e: Online Academic Journal of Information Technology*, 11(42), 87-99, 2020.
- [14] İnternet: Ö. Aydın, E. Karaarslan, “Covid-19 Belirtilerinin Tespiti İçin Dijital İkiz Tabanlı Bir Sağlık Bilgi Sistemi”, Dokuz Eylül Üniversitesi Kurumsal Akademik Açık Arşivi, <http://hdl.handle.net/20.500.12397/13900>, 12.02.2022.
- [15] B. Yiğitöl, T. Sarı “Küresel Salgınlar ile Mücadelede Endüstri 4.0 Teknolojilerinin Rolü”, *Pamukkale University Journal of Social Sciences Institute*, (41), 53-73, 2021.
- [16] M. Cicioğlu, A. Çalhan, “Bulut Destekli Medikal Nesnelere İnterneti Tabanlı Uzaktan Sağlık İzleme Sistemi”, *Uludağ University Journal of The Faculty of Engineering*, 26(3), 1083-1096, 2021.
- [17] F. Aydemir, S. Arslan, “Covid-19 Pandemi Sürecinde Çocukların El Yıkama Alışkanlığının Nesnelere İnterneti Tabanlı Sistem ile İzlenmesi”, *Mühendislik Bilimleri ve Araştırmaları Dergisi*, 3(2), 161-168, 2021.
- [18] A. Deringöz, T. Danışan, T. Eren, “Covid-19 Takibinde Giyilebilir Sağlık Teknolojilerinin ÇKKV Yöntemleri ile Değerlendirilmesi”, *Politeknik Dergisi*, C-S, 1-1, 2021.
- [19] T. Kılıç, N. Tosun, “Akıllı Sağlık Ekosistemi ve Güncel Uygulama Örnekleri”, *The Journal of Business Science*, 9(3), 543-564, 2021.
- [20] G. Köse, M.N. Kurtalan, “Sağlık Hizmetlerinde Nesnelere İnterneti Uygulamalarının Bibliyometrik Analizi”, *Avrupa Bilim ve Teknoloji Dergisi*, (27), 412-432, 2021.
- [21] D. Tütüncü, M. F. Esen, “Salgın Hastalıkların Yönetiminde Nesnelere İnterneti (NİT) kullanımı: COVID-19 Örneği”, *Sağlık Akademisyenleri Dergisi*, 8(2), 169-177, 2021.
- [22] A. Mısırhoğlu, “Sağlık Sektöründe Hizmet Kalitesine Yönelik Bibliyometrik Bir Analiz”, **İktisadi ve İdari Bilimlerde Güncel Araştırmalar**, Editörler Doç. Dr. Mustafa Mete, Dr. Öğr. Üyesi Aytaç Toptaş, Birinci Basım, Gece Kitaplığı, Ankara, 2022.
- [23] U. Al, “Türkiye’nin Bilimsel Yayın Politikası: Atf Dizinine Dayalı Bibliyometrik Bir Yaklaşım”, Doktora Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara, 2008.
- [24] A. Pritchard, “Statistical Bibliography or Bibliometrics?”, *Journal of Documentation*, 25 (4), 348-349, 1969.
- [25] G. F. Önal, “Bildirilerin Bibliyometrik Profili”, **IX. Ulusal Müzik Eğitimi Sempozyumu**, KEFAD, 18(3), 1079-1097, 2017.
- [26] D. F. Wallace, “Bibliometrics and Citation Analysis”, **Principles and Applications of Information Science: For Library Professionals**, Editör J. N. Olsgaard, Chicago and London: American Library Association, 10-26, 1989.
- [27] İnternet : E. Erkut, “Bibliyometrik Analiz”, <http://eresearch.ozyegin.edu.tr/bitstream/handle/10679/285/Bibliyometrik%20Analiz.pdf?sequence=1>, 05.01.2022.
- [28] E. Garfield, “Can Citation Indexing Be Automated? In Statistical Association Methods for Mechanized Documentation”, **Symposium Proceedings**, 189-192, 1965.
- [29] N. Gürsakal, **Sosyal Ağ Analizi**. 1.Baskı, Dora Yayıncılık, Bursa, 2009.
- [30] S. Wasserman, K. Faust, **Social Network Analysis: Methods and Applications**, Cambridge University, Press 4. Cambridge, (35), 1994.
- [31] L. Kıdak, H. Demir, E. Özdemir, “Bilimsel Haritalama Yöntemiyle Hastane Yönetimi Alanındaki Tematik Değişimin Analizi”, *Genel Tıp Dergisi*, 27(2), 43-50, 2017.
- [32] C. López-Herrera, E. Herrera-Viedma ve F. Herrera “Science Mapping Software Tools: Review, Analysis, and Cooperative Study Among Tools”, *Journal of The American Society for Information Science and Technology* 62(7), 1382-402, 2011.
- [33] H. Demir, G. Erigüç, “Bibliyometrik Bir Analiz ile Yönetim Düşünce Sisteminin İncelenmesi.” *İş ve İnsan Dergisi*, 5(2), 91-114, 2018.
- [34] N. C. Peeri, N. Shrestha, M.S. Rahman, R. Zaki, Z.Q. Tan, S. Bibi, M. Baghbanzadeh, N. Aghamohammadi, W.Y. Zhang, U. Haque, “The SARS, MERS and Novel Coronavirus (COVID-19) Epidemics, The Newest and Biggest Global Health Threats: What Lessons Have We Learned?”, *International Journal of Epidemiology*, 49(3), 717-726, 2020.
- [35] V. Chamola, V. Hassija, V. Gupta, M. Guizani, “A Comprehensive Review of the COVID-19 Pandemic and the Role of IoT, Drones, AI, Blockchain, and 5G in Managing its Impact”, *IEEE ACCESS*, (8), 225-265, 2020.
- [36] A. Kumar, K. Sharma, H. Singh, S.G. Naugriya, S. S. Gill, R. “A Drone-Based Networked System and Methods for Combating Coronavirus Disease (COVID-19) Pandemic”, *Future Generation Computer Systems-The International Journal of Escience*, (115), 1-19, 2021.
- [37] M. Abdel-Basset, V. Chang, N. A. Nabeeh, “An Intelligent Framework Using Disruptive Technologies For COVID-19 Analysis”, *Technological Forecasting And Social Change*, (163), Article Number 120431, 2020.
- [38] S. Jain, M. Nehra, R. Kumar, N. Dilbaghi, T.Y. Hu, S. Kumar, A. Kaushik, C.Z. Li, “Internet of Medical Things (IoMT)-integrated Biosensors for Point -of -Care Testing of Infectious Diseases”, *Biosensors & Bioelectronics*, (179), Article Number 113074, 2021.
- [39] K. H. Abdulkareem, A.S. Bader, O. Geman, M. A. Mohammed, “Realizing an effective COVID-19 diagnosis system based on machine learning and IoT in smart hospital environment”, *IEEE Internet of Things Journal*, (99), 2021.
- [40] M. Masud, G.S. Gaba, S. Alqahtani, G. Muhammad, B. B. Gupta, BB P. Kumar, A. Ghoneim, “A Lightweight and Robust Secure Key Establishment Protocol for Internet of Medical Things in COVID-19 Patients Care”, *IEEE Internet of Things Journal*, 8(21), 15694-15703, 2021.

# Kelime Temsil Yöntemlerinin Hedef Tabanlı Duygu Analizine Etkisi

*Araştırma Makalesi/Research Article*

 Mesut POLATGİL<sup>1</sup>,  Murat Fatih TUNA<sup>2</sup>,  Oğuz KAYNAR<sup>2</sup>

<sup>1</sup>Sarıışla Uygulamalı Bilimler Yüksekokulu, Sivas Cumhuriyet Üniversitesi, Sivas, Türkiye

<sup>2</sup>Yönetim Bilişim Sistemleri, Sivas Cumhuriyet Üniversitesi, Sivas, Türkiye

[mesutpolatgil@cumhuriyet.edu.tr](mailto:mesutpolatgil@cumhuriyet.edu.tr), [mftuna@cumhuriyet.edu.tr](mailto:mftuna@cumhuriyet.edu.tr), [okaynar@cumhuriyet.edu.tr](mailto:okaynar@cumhuriyet.edu.tr)

(Geliş/Received:10.05.2022; Kabul/Accepted:21.09.2022)

DOI: 10.17671/gazibtd.1114901

**Özet**—Klasik duygu analizi yöntemlerinden farklı olarak hedef tabanlı duygu analizi (HTDA), birden fazla kategorinin olduğu karmaşık yapıdaki çevrimiçi tüketici geribildirimlerini değerlendirmede daha başarılı bir performans ortaya koyabilmektedir. Nitekim bir platformda yer alan tüketici geri bildirimleri bir ürüne ilişkin birden farklı hedefe atfedilebilmektedir ve standart duygu analizleri bu geribildirimleri analiz etmede yetersiz kalmaktadır. Literatürdeki gelişmeler gözden geçirildiğinde, HDTA çalışmalarının, duygu analizine odaklanan diğer çalışmalar içinde oldukça popüler olduğu anlaşılmaktadır. SemEval ABSA-2016 yarışmasında, HTDA için 8 farklı dilde veri setleri yayınlanmış ve ekipler duygu analizi için yarışmışlardır. Yarışmada hedef terim, kategori ve duygu sınıfı tespit etmek gibi farklı alt görevler bulunmaktadır. Bu alt görevlerin içindekilerden biri, hedef terimin tespit edilmesidir. Türkçe dili için HTDA çalışmaları oldukça sınırlıdır. Farklı diller ve farklı kelime temsil yöntemleri kullanan çalışmalar vardır. SemEval Absa 2016 yarışması Türkçe veri seti için kelime temsil yöntemlerinin etkisini inceleyen çalışma bulunmamaktadır. Bu çalışma, müşteri yorumlarındaki hedef terimlerin tespitinde farklı kelime temsil yöntemlerinin başarısının incelenmesi amacıyla gerçekleştirilmiştir. Word2Vec, Glove ve Fasttext kelime temsil yöntemleri analiz kapsamında incelenmiş ve hedef terimi en başarılı tespit edebilen yöntemin Fasttext kelime temsil yöntemi olduğu görülmüştür. Çalışmada ayrıca F-1 sınıflandırma ölçütü açısından %77 başarı oranı ile Türkçe veri seti için literatürdeki en yüksek sınıflandırma başarısı elde edilmiştir.

**Anahtar Kelimeler**— hedef tabanlı duygu analizi, hedef terim, kelime temsil yöntemleri, müşteri geribildirimleri, Türkçe metinler, Word2Vec, Glove, Fasttext.

## The Effect of Word Representation Methods on Aspect-Based Sentiment Analysis

**Abstract**—Unlike classical sentiment analysis methods, Aspect-Based Sentiment Analysis (ABSA) can demonstrate a more successful performance in evaluating complex online consumer feedbacks including more than one category. As a matter of fact, consumer feedbacks on a platform can be referred to more than one aspect regarding a product, and standard sentiment analysis method is insufficient to analyse these comments. When the developments in the literature are reviewed, it is understood that HDTA studies are very popular among other studies focusing on sentiment analysis. In the SemEval ABSA-2016 competition, datasets were published in 8 different languages for HTDA and the teams competed for sentiment analysis. There are different subtasks in the competition, determining sub-categories such as aspect term, category and sentiment class. One of these subtasks is to determine the aspect term. HTDA studies for Turkish language are quite limited. There are studies using different languages and different word representation methods. There is no study examining the effect of word representation methods for the Turkish data set of SemEval Absa 2016 competition. This study was carried out to examine the success of different word representation methods in identifying aspect terms in customer comments. This study was carried out with the aim of examining the success of different word representation methods in identifying target terms in customer comments. Word2Vec, Glove and Fasttext word representation methods were examined within the scope of the analysis and it was seen that the method that could detect the aspect term most successfully was the Fasttext word representation method. The highest classification success for Turkish dataset in the literature with a success rate of 77% in terms of the F-1 score was also achieved in the study.

**Keywords**— aspect based sentiment analysis, aspect term, word representation methods, consumer feedbacks, Turkish texts, Word2Vec, Glove, Fasttext



## 1. GİRİŞ (INTRODUCTION)

Web 2.0 ve Endüstri 4.0 kapsamında yaşanan değişimler, kullanıcı temelinde değişen içeriğin üretildiği bir yeni bir dünya oluşturmuştur. Eski dünyadan farklı olarak, sadece seçkin bir zümre değil artık her zümreden insanın kendini ifade edebildiği ve milyonlarca insana ulaşabildiği bir ortama geçilmiştir [1]. Bu ortam bireylerin duygularını ve ruh hallerini özgürce ve kendilerince ifade edebildiği bir mecaz halini almıştır. Hatta normal hayatlarında gerçek duygu ve düşüncelerini paylaşamayan kişilerin, sosyal medya üzerinden kendilerini rahatça ifade edebilmeleri, sosyal medya ve diğer ortamları, bireylerin düşünce, tutum ve davranışlarının ölçülmesinde daha gerçekçi sonuçlar verebilecek bir veri kaynağına dönüştürmüştür [2]. Dolayısıyla burada üretilen içerikler ve paylaşılan görüşler, işletmelerin tutmuş oldukları büyük verilere eklenerek işletmelerin karar verme süreçlerine katılmaya başlanmıştır [3].

Zaltman ve Zaltman [4], tüketicilerden gelen geri bildirimlerin altını kurcalamanın, onları daha net anlamada yardımcı olacağını vurgulamıştır. Günümüzde sosyal medya, bunu gerçekleştirebilmenin en az maliyetli yolu olarak düşünülebilir. Ayrıca Westerman vd. [5], sosyal medyanın sağladığı verilerin, tüketici ve işletme üzerine yapılan çalışmalar açısından güvenilir bir veri kaynağı oluşturduğunu saptamışlardır. İşletmeler (kâr amacı gütsün ya da gütmesin) ürün ve süreçlerle ilgili yorum ve görüşleri internet ve sosyal medya ortamından toplamakta, bunları inceleyerek elde ettikleri bilgileri pazarlama ve diğer işletme birimlerine yönelik strateji oluşturmada bir argüman olarak kullanmaktadır [6]. İletişim imkânlarının karşılıklı hale gelmesi, tüketicilerin anlık geribildirimlerini firmaya iletebilecekleri koşulların sağlanmasına neden olmuştur. Gerek fiziki ürünlerin gerekse hizmetlerin tüketiciyle buluşturulmasından hemen sonra tüketiciler sosyal medya ve internet platformlarını kullanarak ürün ya da hizmete ilişkin geribildirimlerini paylaşmaktadır. Dahası yapılan paylaşımlar, sosyal medyadaki yoğun etkileşimden dolayı hızlı bir şekilde yayılabilmektedir. İşletmeler bu viral etkinin ve satışlara olan katkısının farkında olduğundan hem tüketicilerin geribildirimlerini hızlı işleme hem de işlenen geribildirimlerden elde edilen göstergeleri hızlı bir şekilde potansiyel müşterilerle paylaşılabilirliği amaçlamaktadır [7]. Sedanter eğilime sahip [8] ve zaman yönetiminde zorlanan [9] günümüz tüketicisi, yoğun şekilde kullandığı internet ortamını daha yaygın etkileşimli bir pazar haline getirmektedir. Ayrıca artan çevrimiçi firma-müşteri etkileşiminden doğan sözel müşteri verilerinin anlık analiz edilmesinin gerekliliği, herhangi bir ürün ya da hizmeti ve bunların alt bileşenlerine ilişkin müşteri görüşlerini özet bir şekilde ortaya koyan uzman sistemlere olan ihtiyacı artırmaktadır [10].

Müşteri odaklı yaklaşımın bir sonucu olarak işletmelerin sürdürülebilirlik düzeyleri ve rekabet güçleri, müşterileri ile olan iletişimlerine ve onlardan gelen geri bildirimleri değerlendirme kapasitelerine bağlıdır [11]. Bu kapasite

aynı zamanda bir ürün ya da hizmetin kalitesine müşteri dokunuşunu sağlamanın da bir yoludur [12]. Dolayısıyla günümüz tüketicileri, sosyal medya ve internet platformlarıyla inşa sürecine katıldığı ürün, hizmet ve bilgiye daha yoğun ilgi göstermektedir [13]. Bu noktada Tuna vd. [14] bu geri bildirimleri bilgisayar ve teknoloji destekli süreçlerle hızlı bir şekilde analiz edebilme kapasitesini müşteri sadakatini sağlamada etkin bir yol olarak değerlendirmişlerdir. Bu durum işletmeler tarafından da bilindiğinden işletmelerin sosyal medya gönderilerine gösterdikleri önemi artırmıştır. Nitekim ürün ya hizmetleriyle alakalı geri bildirim alan ve bunları doğru şekilde değerlendiren kurumlar güçlü sayılmaktadır [15]. Bugün Uber\*, Amazon\*\*, TripAdvisor\*\*\* firmaları, ilgili alanda faaliyet gösteren işletme ve satıcılara ilişkin yorumlarla sırasıyla 100 milyon\*, 300 milyon\*\* ve 450 milyon\*\*\* civarı kullanıcının yoğun iletişim ve ilişki kurmasına imkân tanımaktadır [16].

Bu konjonktürde işletmelerin metinsel ifadelerden oluşan bu verileri olabildiğince hızlı ve detaylı bir şekilde analiz edebilmesi gerekmektedir ki insan gücü ile günümüz işletmelerinde bu ilave maliyet ve zaman kaybı anlamına gelmektedir [1]. Ayrıca yorumların nicel olmayışı ve sayısının fazlalığı, klasik analiz metodlarının etkinliğini azaltmaktadır. Bahsedilen durumların gerektirdiği yöntemlerden biri duygu analizidir. Diğer bir adı fikir madenciliği olan duygu analizi, metinsel verinin içindeki ifadenin tamamının ya da bir kısmının içerdiği duygunun metin analizi, doğal dil işleme, hesaplamalı dilbilimle işlenmesidir [17]. Duygu analizi yapılırken belirli bir alana özgü ve içerisinde farklı unsurları konu edinen metin derlemi (corpus) kullanılmakta, derlem içerisinde bir test ve eğitim veri seti oluşturulmaktadır. Analiz sonucunda metinsel ifadeler temelde pozitif ve negatif olarak ayrıştırılabilmektedir. Oluşturulan bu veri seti genellikle makine öğrenmesi teknikleriyle bir bütün doküman olarak (document-based) analiz edilebildiği gibi konu tabanlı (topic-based), cümle tabanlı (sentence-based) ve hedef tabanlı (aspect-based) olarak da incelenebilmektedir. Bu çalışmaya konu olan hedef tabanlı duygu analizinin ilk alt görevi olan hedef terimin tespitidir. Al-Smadi vd. [18]'ne göre bu analiz tipinde bir metinsel verinin temasında yer alan farklı hedef terimlere yönelik duygu yönelimi ve şiddetine odaklanılmaktadır.

Hedef tabanlı duygu analizinde hedef terimin neyle ilgili olduğunun anlaşılması, başka bir ifadenin içerdiği duygunun hangi kapsamda olduğunun tespit edilmesiyle mümkün olabilmektedir [19]. İnce taneli yöntemlerden biri olan bu analize ilişkin Wang vd. [20] tarafından verilen örnek, duygulardaki hedeflerin belirlenmesinin anlaşılır bir örneğini oluşturmaktadır. Buna göre, bir restoran yorumunda geçen “şehirdeki en hızlı teslim zamanlarından birine sahip olduklarını söylemek zorundayım” şeklinde bir ifade hedef terimin “teslim zamanları”, fikir teriminin ise “en hızlı” olduğu anlaşılabilir. Nitekim teslim zamanları ifadesi en hızlı sıfatıyla nitelendirilmektedir [21].

## 2. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Literatür incelendiğinde, ilk duygu analizi çalışmasının Pang vd. [22] tarafından yapıldığı görülmektedir. Klasik duygu analizi kullanarak yapılan çalışma ve uygulamaların; metin analizinde yeterli derinliğe inemediği [20], ince taneli duygu analizi gerçekleştirebilecek kapasiteye sahip olmadığı [23] ve bütün bir metnin içerisindeki baskın duygu sınıfını belirlemek üzerine odaklandığı [1] görülmektedir. Ayrıca bir ürün ya da hizmetin farklılaştığı esas noktaları belirlemede klasik duygu analizi türlerinin yeterli olmadığı savı üzerinde görüş birliği sağlandığı görülmektedir [15], [24].

Hedef tabanlı duygu analizine temel teşkil eden çalışmalar, hedef derecelemesi [25] ve hedef tarama [26] konularıyla başlamıştır. Hedef tabanlı duygu analizine yapılan devam çalışmalarının; elektronik ürünlere [27], filmlere [28], otellere [14], [29] ve restoranlara [30] yönelik yorumları kapsayacak şekilde geliştirildiği görülmektedir. Literatür incelendiğinde, hedef tabanlı duygu analizine ilişkin çalışmalara 2014 yılı itibarıyla rastlanmaktadır. Bu alandaki çalışmalardan birinde Blinov ve Kotelnikov [31], dizüstü bilgisayar ile restoranlara ilişkin İngilizce yorumlardan oluşturulan iki farklı veri seti üzerinde çalışmışlardır. Semeval-2014 yarışmasının verilerini kullandıkları çalışmalarında, hedef terim çıkarımı ve hedef terim duygu polaritesine ilişkin yöntem önerisinde bulunmuşlardır. Bu çalışmanın sonucunda kategori tespitinde %50,51; kategori polarite tespitinde dizüstü bilgisayar ve restoran veri setleri için sırasıyla %37,38 ve %43,34 şeklindeki F-1 değerlerine ulaşmışlardır.

Literatürde çeşitli dillerde hazırlanan veri setleri üzerinde yapılmış çalışmalar bulunmaktadır [18, 32]. Ayrıca LSTM tabanlı farklı kelime temsil yöntemlerini kullanan çalışmada bulunmaktadır [16]. Bununla birlikte bu çalışmada Türkçe veri setleri üzerinde çalışılacağından ötürü, Türkçe veri setleri üzerinde hedef tabanlı duygu analizi yürüten çalışmalara odaklanılacaktır. Farklı dillerle birlikte Türkçe veri setinin de birlikte konu edinildiği bir çalışmada Pontiki vd. [33], SemEval-2016 yarışmasının 5 numaralı görevini yerine getirmek için çok katımlı bir hedef tabanlı duygu analizi çalışması yürütmüşlerdir. Çalışmalarında İngilizce, Arapça, Çince, Hollandaca, Fransızca, Rusça, İspanyolca ve Türkçe dillerindeki restoran ve otel veri setleri üzerinde çalışmışlardır. Hedef tabanlı duygu terimlerini çıkardıkları çalışmanın sonucunda, kullandıkları dillerden Türkçe dili için ulaştıkları en yüksek F1 performans ölçüm değerinin %41,86 olduğu ortaya konulmuştur.

Kama vd. [34], online arama motorunu kullanarak frekans tabanlı nitelik çıkarma performansını artıracak denetimsiz bir model kullanarak hedef tabanlı duygu analizi gerçekleştirmişlerdir. Donanımhaber isimli teknoloji forumunu kullandıkları çalışmalarında önerdikleri web arama tabanlı nitelik çıkarma (WSBFE) yönteminin hedef terim çıkarmada performansın F1 değerinin %69,79 olduğunu iddia etmişlerdir.

ABSA 2016 yarışmasında sunulan Türkçe restoran veri setinin kullanıldığı bir çalışmada Çetin ve Eryiğit [1]; hedef kategori belirleme, hedef terim belirleme, hedef kategoriyle terimi aynı anda belirleme ve duygu sınıfı belirleme şeklinde yarışmaya açılan dört görevi gerçekleştirmeyi amaçlamışlardır. Bu dört görev için sırasıyla %66,7; %53,2; %46,7 ve %76,1 F1 değerleri elde edilmiştir. Çalışmada kullanılan duygu analizi bir lineer sınıflandırma ile gerçekleştirilmiştir.

İstatistiksel, linguistik ve kural tabanlı yaklaşımları (gizli dirichlet ataması "LDA", c-değeri, WSBFE) bir arada kullandıkları çalışmalarında Bayraktar vd. [35], ABSA 2016'da yayınlanan restoran verilerini kullanmışlardır. Çalışmanın sonucunda performans elde etmenin amaçlandığı hedef terim çıkarımı ve hedef-duygu eşleşmesine ilişkin F1 değerleri sırasıyla %56,28 ve %52,05 olarak hesaplanmıştır.

Özyurt ve Akçayol [36], yapmış oldukları çalışmalarında ürünlere ilişkin hedef terimlerin çıkarılmasında ve ürün terimlerinin çıkarılmasında hedef tabanlı duygu analizi kullanmışlardır. Semeval-2016 yarışmasındaki görevlerin tamamlanması için cümle segmentasyonu ve LDA yöntemlerinin birlikte kullanıldığı çalışmada, hedef terimlerin çıkarılması ve ürün terimlerinin çıkarılması görevlerine ilişkin F1 ölçüm değerleri sırasıyla %62,25 ve %82,39 olarak hesaplanmıştır.

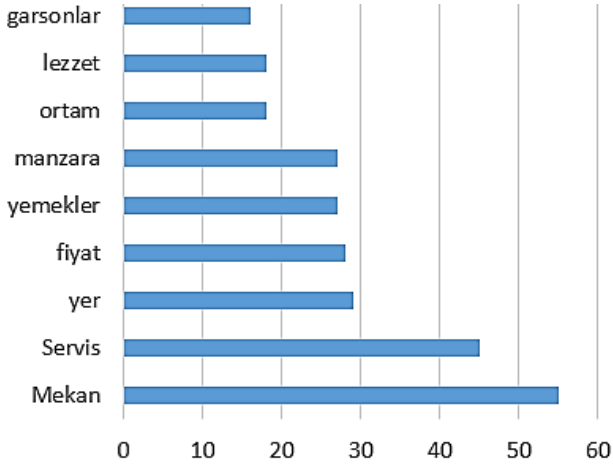
Salur ve Aydın [37], Güneydoğu'daki turistik merkezlere ilişkin ziyaretçi yorumlarına hedef tabanlı duygu analizi uygulamışlardır. Turistik merkezlerin hedef terimlerini çıkartıldığı çalışmada Türkçe etiketlenmiş bir veri setini de kullanıcıların erişimine sunmuşlardır. TF-IDF ve LDA yöntemlerini kullandıkları çalışmalarında F-1 ölçüm değerini 0,5931 olarak hesaplamışlardır.

Salur vd. [15], ABSA 2016 yarışmasında sunulan Türkçe restoran yorumlarını kullandıkları çalışmalarında yorumlarından hedef terimlerini çıkarmak için istatistiksel (TF-IDF), konu modelleme (LDA ve NMF) ve kural-tabanlı denetimsiz yöntemleri bir arada kullanan bir sınıflandırma yaklaşımı önerilmiştir. Önerdikleri 'ensemble' yaklaşımının %61,08'lik bir F1 değerine ulaştığını ileri sürmüşlerdir.

## 3. YÖNTEM (METHOD)

### 3.1. Veri seti (Data Set)

Çalışmada SemEval ABSA 2016 kapsamında oluşturulmuş ve birçok çalışmada da kullanılmış olan Türkçe Restoran yorumlarını içeren bir veri seti bulunmaktadır. Bu veri setinde 1415 tane cümle bulunmaktadır. Bu cümlelerden 148 cümle bir duygu durumunu içermemektedir. Veri setinde yer alan cümlelerde hedef terim barındırmayan ('Null') cümleler değerlendirmeye alınmamıştır. Veri setinde en çok yer alan hedef terimler Şekil 1'de verilmiştir.



Şekil 1. Veri setinde en çok geçen hedef terimler [15]  
(Most common target terms in the dataset)

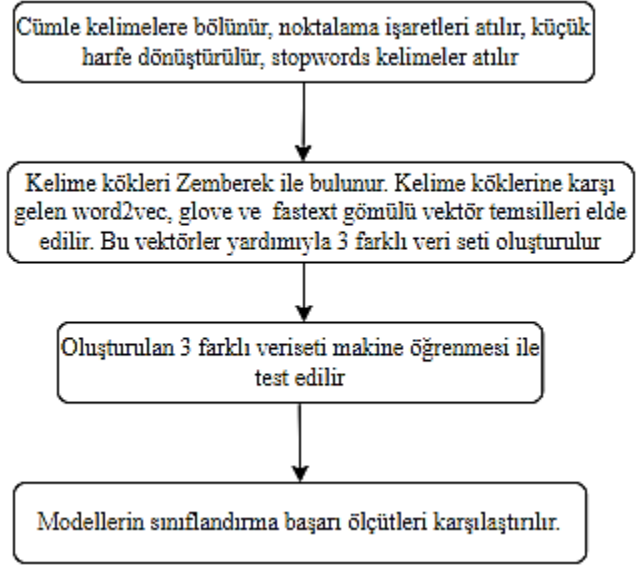
Şekil 1’de en çok geçen hedef terimlerin mekân, servis, yer ve fiyatlar olduğu görülmektedir. Restoran ile ilgili olabilecek ifadelerin hedef terim olarak öne çıktığı görülmektedir.

### 3.2. Önerilen Yöntem (Proposed Method)

Bu çalışmada ABSA 2016 yarışması kapsamında görevlerden birisi olan hedef terimin tespit edilmesi için bir model önerilmiştir. SemEval ABSA yarışmasında 3 temel görev olduğu belirtilmiştir. Bunlar; Sentence-level ABSA, Text-level ABSA ve Out-of-domain ABSA olarak belirlenmiştir. Sentence-level ABSA görevi 3 alt bölüme ayrılmıştır. Bunlar; hedef terimin belirlenmesi, hedef kategorinin belirlenmesi ve duygu sınıfının belirlenmesi. Burada tespit edilen tüm tahminler cümle düzeyinde olmaktadır. Text-level ABSA kategorisinde ise analiz bir yorum üzerinden yapılmaktadır ki bu yorum genellikle birden fazla cümle içerebilir. Out-of-domain ABSA kategorisinde ise ekiplere daha önce hiç görmedikleri bir alanda eğitim verisi sağlanmadan sistemlerini test etme imkânı verilmektedir. Yarışma sonuçları değerlendirilmesinde 2 aşama söz konusudur. İlk aşamada hedef terim, hedef kategori çiftinin bulunması, diğer aşamada ise duygu sınıfının bulunması yer almaktadır. Sonuçlar F-1 sınıflandırma ölçütüne göre değerlendirilmektedir.

Önerilen modelde kelimenin kendisi ve kök bilgisi cümleden elde edilerek önceden eğitilmiş modellerden elde edilen gömülü kelime temsil vektörleri ile desteklenmesi sonucunda hedef terimin yüksek başarı ile tespit edilebileceği hipotezi test edilmek istenmiştir. Özetle bu araştırma, ABSA 2016 yarışmasında cümle-düzeyi (sentence-level) ABSA kategorisinin hedef terim ifadesinin tespit edilmesi alt görevine odaklanmıştır.

Bu hipotez önerilirken daha önce aynı veri seti üzerinde yapılan hedef terim tespiti çalışmalarında önerilen modeller incelenmiştir [1, 37]. Bu çalışmalarda kelime kök bilgisi, morfolojik analiz, kelimenin kendisi, komşu kelimeler, bağıllık ayrıştırıcı, LDA analizi ve kural tabanlı olmak üzere farklı yaklaşımlar önerilmiştir. Fakat önerilen yöntemler hem veri setinden çekilecek özellikler bakımından karmaşık hem de bu verilerin elde edilmesi için kullanılacak sistem ve programlar açısından zahmetlidir. Bu çalışmada ise sadece kelime kök bilgisinden elde edilecek gömülü vektör temsilleri yardımıyla böyle bir zorlu görevin başarılabileceğini göstermektedir.

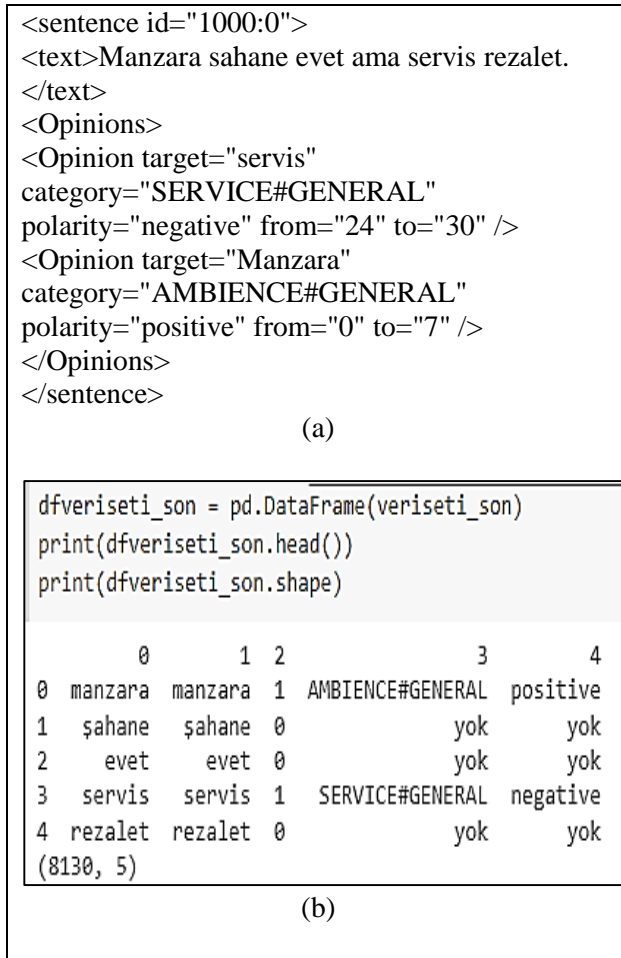


Şekil 2. Çalışmada önerilen işlem adımları (Suggested steps in the study)

Şekil 2’de verildiği gibi hedef terimi elde etmek için önerilen model öncelikle ilgili cümlelerin kelimelere bölünmesi, sonrasında küçük harfe dönüştürülmesi, ardından noktalama ve durak(stopwords) kelimelerin atılması gibi bir takım metin ön işleme adımlarının uygulanması ile başlamaktadır.

Daha sonra her bir kelimenin kökü Zemberek Kütüphanesi yardımı ile elde edilerek bulunur. Zemberek kütüphanesi ile veri setindeki bir örnek cümle için kök bulma işlemi şekil 3’de verilmiştir.

Kelime kökleri elde edildikten sonra her kelimenin Word2Vec, Glove ve Fasttext yöntemleri ile daha önce hazır olarak eğitilmiş modellerden sırasıyla 400, 300 ve 300 boyutlu gömülü vektör temsil modelleri elde edilmiştir. Bu sayısal vektörler veri seti için giriş bilgisini oluştururken, ilgili terimin hedef terim olup olmadığı bilgisi ise çıkış bilgisini oluşturmaktadır. Oluşturulan veri setinden ilk cümle için örnek Şekil 3’de verilmiştir.



Şekil 3. ABSA veri setinde bulunan örnek cümle. (a). Xml formatındaki veri (b). Python ve Zemberek ile işlenmesi (Example sentence found in the ABSA dataset. (a). Data in xml format (b). Processing with Python and Zemberek)

Tablo 1. Veri setinin yapısı örnek cümle gösterimi (Structure of the dataset example sentence representation)

Kelime	Kelime Vektörü	Hedef terim
Manzara	(0.43, 0.47, 0.01 .....0.22)	1
Şahane	(0.12, 0.41, 0.59 .....0.45)	0
Evet	(0.15, 0.52, 0.07 .....0.74)	0
Servis	(0.59, 0.48, 0.02 .....0.36)	1
Rezalet	(0.92, 0.14, 0.85 .....0.62)	0

Word2Vec kelime temsillerini elde etmek için daha önceden eğitilmiş bir modelden yararlanılmıştır. Türkçe

Wikipedia derlemi kullanılarak eğitilen modelde Cbow yöntemi tercih edilmiş, kelime vektör uzunluğu 400, pencere boyutu ise 5 olarak belirlenmiştir [38]. Glove yöntemi ile oluşturulan diğer model ise 21 GB boyutundaki Commoncrawl verileri ile eğitilerek, 300 uzunluğundaki kelime vektörleri ile oluşturulmuştur [39]. Fasttext vektör temsillerini elde etmek için daha önce Grave vd tarafından yapılmış bir çalışmada kullanılan ve Türkçe için eğitilmiş 300 uzunluklu kelime vektörüne sahip bir modelden yararlanılmıştır [40].

#### 4. BULGULAR (RESULTS)

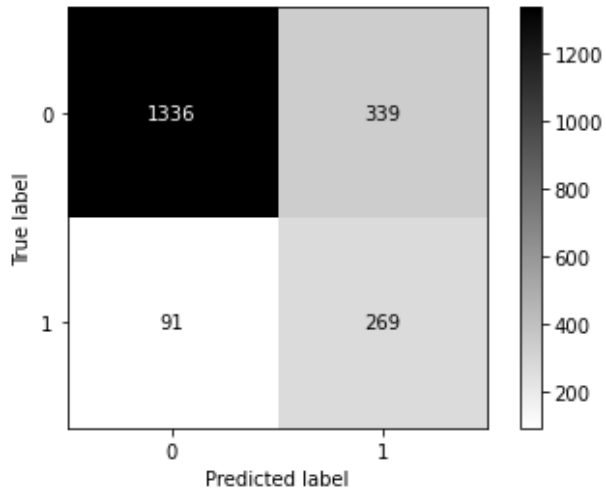
Makine öğrenmesi algoritmaları uygulanmadan veri setine 5 fold çapraz geçerlilik uygulanmıştır. Veri seti 5 parçaya bölünerek 4 parçasıyla eğitim 1 parçasıyla test işlemi gerçekleştirilmiştir. Bu işlem her seferinde test için kullanılan birim değiştirilerek 5 kere tekrar edilmiş ve elde edilen test skorlarının ortalaması modelin genel performansı olarak hesaplanmıştır. Böylelikle veri setini eğitim ve test kümelerine ayırma işlemi sırasında doğacak avantajlı durumların önüne geçilerek model performansını gösteren metriklerin daha güvenilir bir şekilde elde edilmesi sağlanmıştır. Tablo 2'de belirtilen makine öğrenme yöntemleri kullanılarak analizler gerçekleştirilmiş, F-1 metrikleri yardımıyla algoritmaların başarımları kıyaslanmıştır. Tablo 2'de görüleceği üzere model performansları kullanılan algoritmadan bağımsız olarak birbirine yakın F-1 skoru değerleri elde edilmiştir. Bu nedenle F-1 skoru en yüksek olan lojistik regresyon yönteminin kullanılmasına karar verilmiş analiz detayları bu yönetime göre verilmiştir.

Lojistik Regresyon yöntemi kullanırken verilerin dengesiz bir dağılım göstermesi sebebi ile 'class\_weight' parametresi kullanılarak bu durumun öğrenme sırasında göz önünde bulundurulması sağlanmıştır.

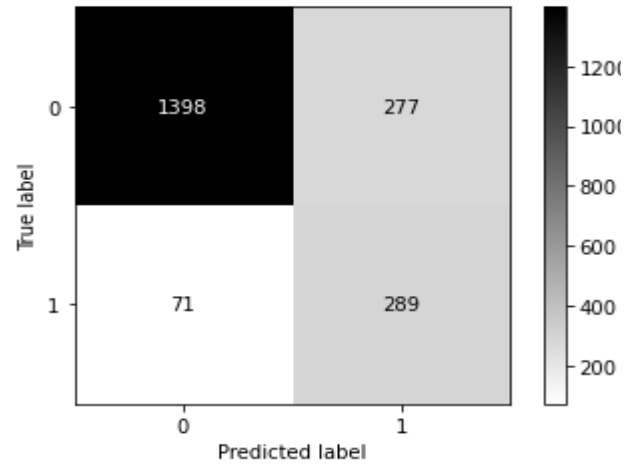
Tablo 2. Algoritmaların F-1 başarı ölçütleri (F-1 scores criteria of algorithms)

	Word2Vec	Glove	Fasttext
KNN	72	74	75
Lojistik Reg.	73	75	<b>77</b>
Destek Vektör	72	73	75
Rassal Orman	74	74	76

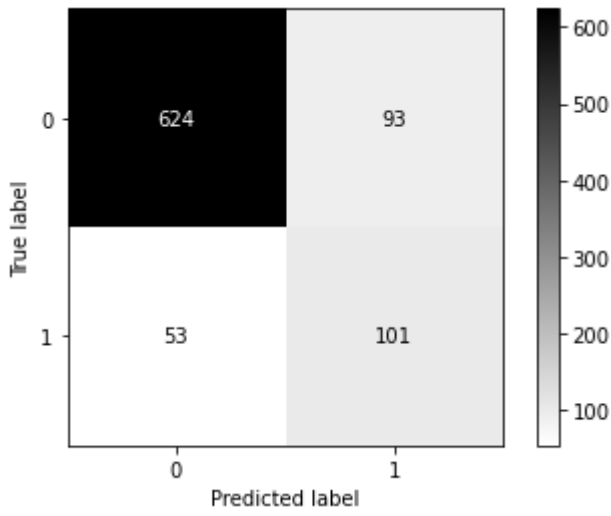
Lojistik Regresyon yönteminin eğitim ve test kümesi üzerindeki karmaşıklık matrisleri Şekil 7-9'da verilmiştir.



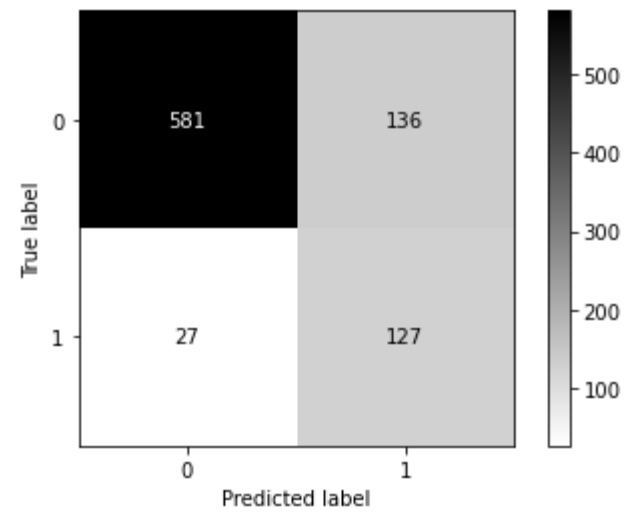
(a)



(a)



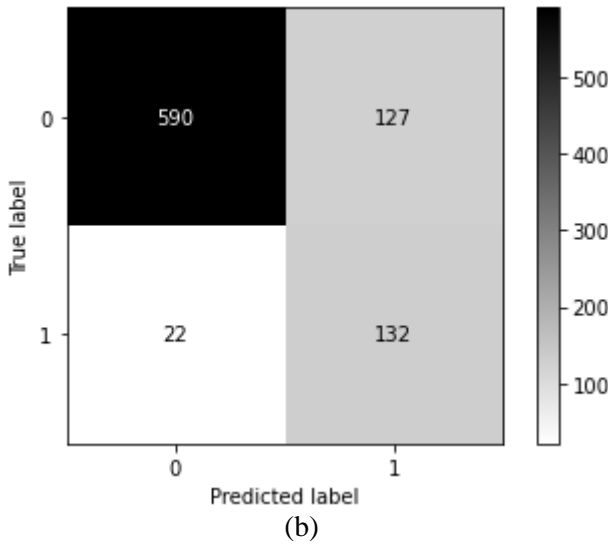
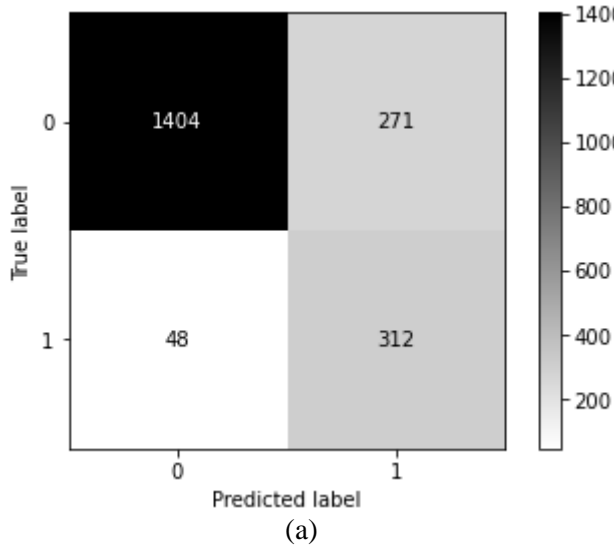
(b)



(b)

Şekil 4. Word2Vec veri seti hedef terim tespiti karmaşıklık matrisi (a) eğitim aşaması , (b) test aşaması (Word2Vec dataset aspect term detection confusion matrix (a) training phase , (b) testing phase)

Şekil 5. Glove veri seti hedef terim tespiti karmaşıklık matrisi (a) eğitim aşaması , (b) test aşaması (Glove dataset aspect term detection confusion matrix (a) training phase , (b) testing phase)



Şekil 6. Fasttextveri seti için hedef terim tespiti karmaşıklık matrisi (a) eğitim aşaması , (b) test aşaması (Aspect term detection confusion matrix for fasttext data set (a) training phase, (b) testing phase)

Tablo 3 Word2Vec test veri kümesi sınıflandırma ölçütleri (Word2Vec test dataset classification scores)

	Sınıflar		Ortalama
	0	1	
Kesinlik	0.92	0.52	0.72
Duyarlılık	0.87	0.66	0.765
F-1	0.90	0.58	0.74

Tablo 4. Glove test veri kümesi sınıflandırma ölçütleri (Glove test dataset classification scores)

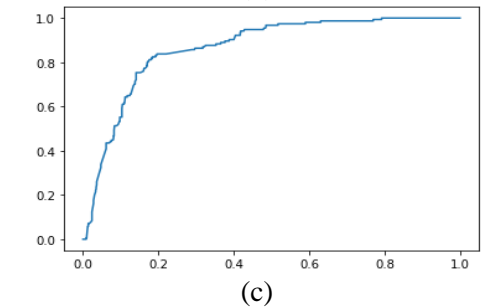
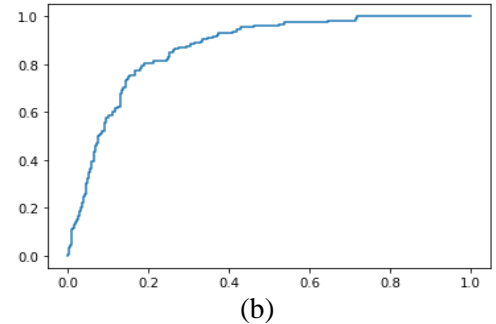
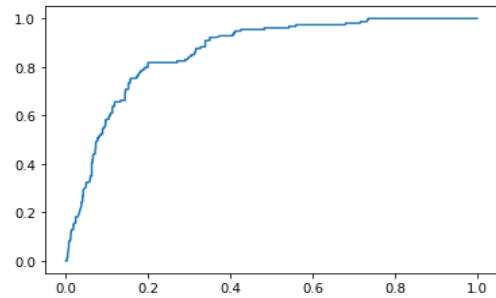
	Sınıflar		Ortalama
	0	1	
Kesinlik	0.96	0.48	0.72
Duyarlılık	0.81	0.82	0.815
F-1	0.88	0.61	0.745

Tablo 5. Fasttext test veri kümesi sınıflandırma ölçütleri (Fasttext test dataset classification criteria)

	Sınıflar		Ortalama
	0	1	
Kesinlik	0.96	0.51	0.735
Duyarlılık	0.82	0.86	0.84
F-1	0.90	0.64	0.77

ABSA yarışmasında F-1 ölçütü kullanıldığı için çalışma sonuçları da bu ölçüt üzerinden değerlendirilmekle birlikte diğer ölçütler de Tablo 3-5’de verilmiştir.

Tablo 3-5’de verilen değerler kullanılarak modelin test kümesi üzerinde F-1 ölçütü bakımından sırasıyla Word2Vec için ortalama %73, Glove için ortalama %75 ve Fasttext için ortalama %77 başarı elde edilmiştir. Sonuçlardan da anlaşılacağı üzere en iyi başarı 0.77 F-1 skoru ile Fasttext vektör temsili kullanılması durumunda elde edilmiştir. Çalışmada kullanılan veri setinin dengesiz dağılımı göz önünde bulundurularak çalışma sonuçlarının güvenilirliğinin tam olarak yansıtılabilmesi için ROC (Receiver operating characteristic) eğrisi çizdirilerek eğri altında kalan alan AUC (Area under curve) değerleri de hesaplanmıştır.



Şekil 7. ROC eğrisi grafiği (a). Word2Vec (b). Glove (c). Fasttext (ROC curve graph (a). Word2Vec (b). Glove (c). fasttext)

Şekil 7’de verilen ROC eğrileri önerilen her üç kelime temsil yönteminin hedef terim tespitinde başarılı bir şekilde kullanılabileceğini göstermektedir. ROC eğrisi altında kalan alanı temsil eden AUC değerleri sırasıyla Word2Vec için 0.83, Glove için 0.86 ve Fasttext için ise 0.88 olarak bulunmuştur.

Çalışmanın son aşamada ise 3 farklı kelime vektörü yönteminden elde edilen vektörler birleştirilerek hedef terim tespiti yapılmıştır. Her üç yöntemden elde edilen kelime temsillerinin birleştirilmesiyle 1000 boyutlu bir kelime vektörü kullanılmıştır. Fakat farklı kelime temsil vektörlerinin birleştirilmesinin sınıflandırma başarısı üzerinde bir iyileştirme sağlamadığı tespit edilmiştir.

Çalışma kapsamında her ne kadar hedef terim tespiti problemi ele alınmış olsa da kelime vektörlerinin hedef kategori ve duygu sınıfını belirlenmesi problemleri de test edilmiştir. Fakat elde edilen sonuçlar literatür de kullanılan yöntemlere oranla daha yüksek bir başarı sağlayamamıştır. Bu durum kategori sınıflarındaki örnek sayısının dengesiz olmasından kaynaklanmaktadır. Bazı kategoriler çok sayıda örnek cümle içerirken bazı kategorilerdeki örnek cümle sayısı oldukça sınırlıdır.

## 5. TARTIŞMA (DISCUSSION)

Bu çalışmada ABSA yarışmasında oluşturulmuş olan Türkçe restoran veri kümesi üzerinde hedef terim tespiti için yeni bir model önerilmiştir. Önerilen modelde kelimenin sadece kökü kullanılarak hedef terimi başarılı bir şekilde tespit edebileceği hipotezi araştırılmıştır. Aynı veri kümesi üzerinde daha önce farklı yöntemlerle yapılan çalışma sonuçları Tablo 6’da verilerek karşılaştırılmıştır.

Tablo 6. Literatürdeki hedef terim tespiti yapılan çalışmalar ile kıyaslama (Comparison with studies on aspect term detection in the literature)

Kaynak	Yöntem	F1 (%)
[33]	Sözlük tabanlı	41,86
[34]	WSBFE	69,79
[1]	CRF	53,12
[35]	Kural-Tabanlı Yaklaşım	56,28
[36]	SS-LDA	62,25
[37]	TF-IDF+LDA	59,31
[15]	Ensemble Yaklaşımı	61,08
<b>Önerilen yöntem</b>	<b>Lojistik Regresyon</b>	<b>77</b>

Tablo 6 incelendiği zaman hedef terim tespitine yönelik Türkçe Restoran veri seti üzerinde farklı çalışmaların yapıldığı görülmektedir. SemEval ABSA tarafından hedef terim tespiti için temel bir yöntem zaten önerilmiş ve yarışma ekiplerinden daha başarılı sonuçlar beklenmiştir. ABSA tarafından önerilen temel modelde kategoriler için bir sözlük oluşturulmuştur. Kategorilerde geçen terimler bu sözlük içine konulmuş ve bir terim ilgili kategori

sözlüğünde yer alıyorsa hedef terim olarak sınıflandırılmıştır. Bu yöntemin başarısı ise %41.8 olarak belirtilmiştir. Yarışmada Türkçe verisi ile ilgili olarak başka başvuru olmamıştır. Bu çalışma kapsamında önerilen modelin hedef terim tespit etmede ABSA temel yöntemine göre oldukça başarılı olduğu görülmektedir.

Hedef terim tespit edilmesi için literatürde farklı yöntemler [1,33-35] ve farklı yöntemlerin birlikte kullanıldığı ensemble yöntemler [15,36,37] önerilmiştir. Önerilen yöntemlerin sınıflandırma başarısında özellikle ABSA temel yöntemine göre iyileştirmeler yapıldığı gösterilmiştir. Bu çalışmalar temel yöntemine göre başarı göstermiş olsalar da bu çalışmada önerilen kelimelerin farklı gömülü temsillerinin hedef terim tespitinde kullanımı diğer yöntemlerden daha başarılı sonuçlar elde etmiştir. Hedef terim tespitinde önceden eğitilmiş gömülü vektör temsillerinin kullanılması sınıflandırma algoritmasından bağımsız olarak literatürde önerilen diğer yöntemlerden daha üstün performans göstermiştir. Sonuç olarak gömülü temsil yöntemlerinin hedef terim belirlemede sınıflama algoritmalarından bağımsız olarak iyi bir öznelik belirleyici olarak kullanılabileceğini göstermektedir.

## 6. SONUÇ (CONCLUSION)

Müşterilerden gelen geribildirimlerin analizi, onların zihinlerinde bir ürün ya da hizmete yönelik oluşacak algıların temel belirleyicileri konumuna yükselmiştir. Hedef tabanlı duygu analizi, sadece bir ürün ya da hizmete değil aynı zamanda ürün ya da hizmetin alt boyutlarına yönelik kullanıcı yorumlarını kategori bazında resmetmeyi sağlayan bir araçtır. Bu analizi etkin kullanmayı başaran firmalar, daha önceki tüketicilerin deneyimlerini çarpıcı biçimde potansiyel müşterilerin bilgisine sunmanın alternatif bir yoluna da sahip olmaktadır. Nitekim günümüz pazarlama çabaları büyük ölçüde müşterilerden gelen geri bildirimlere dayandığı gibi, bu geri bildirimler müşterilerin karar verme süreçlerinde vazgeçilmez bir referans noktası haline almıştır. Deneysel pazarlamanın bir getirisi olarak kabul edilebilecek bu durum, tüketicilerden gelen sözel geri bildirimlerin içerdiği duyguları tespit etmenin işletmeler açısından neden elzem olduğuna işaret etmektedir. Bu durum, tüketicinin ziyade kullanıcının deneyimine ve deneyim sonrasındaki geri bildirimlere üretici ya da satıcının anlatılarından daha fazla güvendiğine de işaret etmektedir. Dolayısıyla çalışmanın, işletmelerin alt hizmet kalemlerine ilişkin müşteri duygularının sınıflandırılmasında ve ilgili kalemlere yönelik özet duygu gösterimlerinde gerek işletme sahiplerine gerekse bu alanda akademik çalışma yapan araştırmacılara efektif ve yüksek performanslı yöntem teşkil edeceği düşünülmektedir. İlerleyen çalışmalarda hedef terimin tespit edilmesine ek olarak, kategori ve duygu sınıfının ayrı ayrı ve birlikte tespit edilebileceği farklı gömülü kelime temsil yöntemlerini tek tek ya da hibrit şekilde kullanan yeni modeller gerçekleştirilebilir ve bu modellerin başarımları farklı veri setleri üzerinde test edilebilir.

**KAYNAKLAR (REFERENCES)**

- [1] F. S. Çetin, G. Eryiğit, “Türkçe Hedef Tabanlı Duygu Analizi İçin Alt Görevlerin İncelenmesi – Hedef Terim, Hedef Kategori ve Duygu Sınıfı Belirleme”, *Bilişim Teknolojileri Dergisi*, 11(1), 43–56, 2018.
- [2] O. Kaynar, Y. Görmez, M. Yıldız, A. Albayrak, “Makine öğrenmesi yöntemleri ile Duygu Analizi”, **International Artificial Intelligence and Data Processing Symposium (IDAP’16)**, Malatya, 234–241, September 17-18, 2016.
- [3] Z. Khan, T. Vorley, “Big data Text Analytics: An Enabler of Knowledge Management”, *Journal of Knowledge Management*, 21(1), 18–34, 2017.
- [4] G. Zaltman, L. H. Zaltman, **Marketing Metaphoria: What Deep Metaphors Reveal About the Minds of Consumers (1st edition)**, Harvard Business Review Press, Boston, 2008.
- [5] D. Westerman, P. R. Spence, B. Van Der Heide, “Social Media as Information Source: Recency of Updates and Credibility of Information”, *Journal of Computer-Mediated Communication*, 19(2), 171–183, 2014.
- [6] V. Ahuja, Y. Medury, “Corporate Blogs as e-CRM Tools – Building Consumer Engagement through Content Management”, *Journal of Database Marketing & Customer Strategy Management*, 17(2), 91–105, 2010.
- [7] P.-Y. Chen, S. Wu, J. Yoon, “The Impact of Online Recommendations and Consumer Feedback on Sales”, in **ICIS 2004 Proceedings**, 58, 2004.
- [8] A. A. Thorp, G. N. Healy, E. Winkler, B. K. Clark, P. A. Gardliner, N. Owen, D. W. Dunstan, “Prolonged Sedentary Time and Physical Activity in Workplace and Non-Work Contexts: A Cross-Sectional Study of Office, Customer Service and Call Centre Employees”, *International Journal of Behavioral Nutrition and Physical Activity*, 9(128), 1–9, 2012.
- [9] J. Cotte, S. Ratneshwar, D. G. Mick, “The Times of Their Lives: Phenomenological and Metaphorical Characteristics of Consumer Timestyles”, *Journal of Consumer Research*, 31(2), 333–345, 2004.
- [10] T. Y. Lee, E. T. Bradlow, “Automated Marketing Research Using Online Customer Reviews”, *Journal of Marketing Research*, 48(5), 881–894, 2011.
- [11] S. A. Bhat, M. A. Darzi, “Service, People and Customer Orientation: A Capability View to CRM and Sustainable Competitive Advantage”, *Vision*, 22(2), 163–173, 2018.
- [12] A. Stelzer, F. Englert, S. Hörold, C. Mayas, “Improving Service Quality in Public Transportation Systems Using Automated Customer Feedback”, *Transportation Research Part E: Logistics and Transportation Review*, 89, 259–271, 2016.
- [13] V. Barger, J. W. Peltier, D. E. Schultz, “Social Media and Consumer Engagement: A Review and Research Agenda”, *Journal of Research in Interactive Marketing*, 10(4), 268–287, 2016.
- [14] M. F. Tuna, O. Kaynar, M. Ş. Akdoğan, “Otellere İlişkin Çevrimiçi Geribildirimlerin Makine Öğrenmesi Yöntemleriyle Duygu Analizi”, *İşletme Araştırmaları Dergisi*, 13(3), 2232–2241, 2021.
- [15] M. U. Salur, İ. Aydın, M. Jamous, “An ensemble approach for aspect term extraction in Turkish texts”, *Pamukkale University Journal of Engineering Sciences, Ahead of Print*, 2022.
- [16] Ł. Augustyniak, T. Kajdanowicz, P. Kazienko, “Comprehensive Analysis of Aspect Term Extraction Methods using Various Text Embeddings”, arXiv:1909.04917 [cs], 2020.
- [17] S. G. Barbounaki, K. Gourounti, A. Sarantaki, “Advances of Sentiment Analysis Applications in Obstetrics/Gynecology and Midwifery”, *Materia Socio Medica*, 33(3), 225–230, 2021.
- [18] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, B. Gupta, “Deep Recurrent Neural Network vs. Support Vector Machine for Aspect-Based Sentiment Analysis of Arabic Hotels’ Reviews”, *Journal of Computational Science*, 27, 386–393, 2018.
- [19] B. Liu, “Sentiment Analysis and Opinion Mining”, *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167, 2012.
- [20] Y. Wang, M. Huang, X. Zhu, L. Zhao, “Attention-Based LSTM for Aspect-Level Sentiment Classification”, **Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing**, Austin (Texas):Association for Computational Linguistics, 606–615.
- [21] W. Wang, S. J. Pan, D. Dahlmeier, X. Xiao, “Recursive Neural Conditional Random Fields for Aspect-based Sentiment Analysis”, **Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing**, Austin (Texas):Association for Computational Linguistics, 616–626.
- [22] B. Pang, L. Lee, S. Vaithyanathan, “Thumbs up? Sentiment Classification using Machine Learning Techniques”, **Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)**, Pennsylvania(Philadelphia): Association for Computational Linguistics, 79–86, 2002.
- [23] T. T. Thet, J.-C. Na, C. S. G. Khoo, “Aspect-Based Sentiment Analysis of Movie Reviews on Discussion Boards”, *Journal of Information Science*, 36(6), 823–848, 2010.
- [24] S. Brody, N. Elhadad, “An Unsupervised Aspect-Sentiment Model for Online Reviews”, Program No: HLT-NAACL, 2010. URL: [https://openreview.net/forum?id=HJ-8d7-\\_bH](https://openreview.net/forum?id=HJ-8d7-_bH), 04.04.2022.
- [25] I. Titov, R. McDonald, “A Joint Model of Text and Aspect Ratings for Sentiment Summarization”, **Proceedings of ACL-08: HLT**, Columbus(Ohio): Association for Computational Linguistics, 308–316, 2008.
- [26] J. Zhu, H. Wang, B. K. Tsou, M. Zhu, “Multi-Aspect Opinion Polling from Textual Reviews”, **Proceedings of the 18th ACM Conference on Information and Knowledge Management**, New York: ACM Digital Library, 1799–1802, 2009.
- [27] J. Wang, B. Xu, Y. Zu, “Deep Learning for Aspect-Based Sentiment Analysis”, **International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)**, Chongqing: IEEE, 267–271, 2021.
- [28] D. Anand, D. Naorem, “Semi-Supervised Aspect Based Sentiment Analysis for Movies Using Review Filtering”, *Procedia Computer Science*, 84, 86–93, 2016.
- [29] T. Tran, H. Ba, V.-N. Huynh, “Measuring Hotel Review Sentiment: An Aspect-Based Sentiment Analysis Approach”, In **Integrated Uncertainty in Knowledge Modelling and Decision Making**, Cham, 393–405, 2019.



- [30] D. Ekawati, M. L. Khodra, "Aspect-Based Sentiment Analysis for Indonesian Restaurant Reviews", **International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA-17)**, Kuta: IEEE, 1–6, 2017.
- [31] P. Blinov, E. Kotelnikov, "Blinov: Distributed Representations of Words for Aspect-Based Sentiment Analysis at SemEval 2014", **Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)**, Dublin: Association for Computational Linguistics, 140–144, 2014.
- [32] W. Wang, G. Tan, H. Wang, "Cross-Domain Comparison of Algorithm Performance in Extracting Aspect-Based Opinions from Chinese Online Reviews", *International Journal of Machine Learning & Cybernetics*, 8(3), 1053–1070, 2017.
- [33] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. Al-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. D. Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S. M. Jiménez-Zafra, G. Eryiğit, "SemEval-2016 Task 5: Aspect Based Sentiment Analysis", **Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)**, San Diego (California): Association for Computational Linguistics, 19–30, 2016.
- [34] B. Kama, M. Ozturk, P. Karagoz, I. H. Toroslu, O. Ozay, "A Web Search Enhanced Feature Extraction Method for Aspect-Based Sentiment Analysis for Turkish Informal Texts", **In Big Data Analytics and Knowledge Discovery**, Cham, 225–238, 2016.
- [35] K. Bayraktar, U. Yavanoglu, A. Ozbilen, "A Rule-Based Holistic Approach for Turkish Aspect-Based Sentiment Analysis", **IEEE International Conference on Big Data (Big Data)**, Los Angeles: IEEE, 2154–2158, 2019.
- [36] B. Ozyurt, M. A. Akcayol, "A New Topic Modeling Based Approach for Aspect Extraction in Aspect Based Sentiment Analysis: SS-LDA", *Expert Systems with Applications*, 168, 114231, 2021.
- [37] M. U. Salur, İ. Aydın, "An Annotated Turkish Aspect Based Sentiment Analysis Corpus for Smart Tourism", **2021 Innovations in Intelligent Systems and Applications Conference (ASYU)**, Elazığ: IEEE 1–6, 2021.
- [38] Internet: A. Köksal, Github, <https://github.com/akoksal/Turkish-Word2Vec>, 15.02.2022.
- [39] O. Çiftçi, GitHub, <https://github.com/inzva/Turkish-GloVe>, 15.02.2022.
- [40] E. Grave, P. Bojanowski, P. Gupta, A. Joulin, T. Mikolov, "Learning Word Vectors for 157 Languages", **Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)**, Miyazaki, European Language Resources Association (ELRA), 3483–3487, 2018.

# Performance Analysis of Object-Relational Mapping (ORM) Tools in .Net 6 Environment

*Araştırma Makalesi/Research Article*

 Abdullah Eren GÜVERCİN<sup>1</sup>,  Bilgin AVENOĞLU<sup>2</sup>

<sup>1</sup>Yazılım Mühendisliği, Ahmet Yesevi Üniversitesi, Ankara, Türkiye

<sup>2</sup>Yazılım Mühendisliği, Ankara Üniversitesi, Ankara, Türkiye

[egvrcn@gmail.com](mailto:egvrcn@gmail.com), [bavenoglu@gmail.com](mailto:bavenoglu@gmail.com)

(Geliş/Received:18.01.2022; Kabul/Accepted:01.10.2022)

DOI: 10.17671/gazibtd.1059516

**Abstract**— ORM tools are frequently used in projects developed by object-oriented programming paradigm. Software developers generally look at the performances of these tools when they select an ORM tool. Most of the performance studies on ORM tools are limited to processing time and RAM usage information, and CPU usage information is not included. Moreover, no ORM performance study has been found in the literature, conducted in .NET 6, which is an open-source and platform-independent new generation .NET platform. In this study, to close the mentioned gap in the literature and guide the software developers, we conduct research for analyzing performances of certain ORM tools in .NET 6. Our study includes CPU usage information as well as processing time and RAM usage information. We develop a software for measuring processing time, RAM and CPU usage while performing read, insert, update, delete, search and sort operations with Dapper, NHibernate and Entity Framework Core (EF Core) ORM tools. As a result, while Dapper is best in terms of processing time for read, delete, search and sort operations, EF Core has the best results for insert and update operations. We conclude that Dapper has the best performance in terms of resource usage, while the rankings of EF Core and NHibernate vary among themselves according to the number of records and operation type.

**Keywords**— orm, .net 6, dapper, nhibernate, entity framework core

## Nesne-İlişkisel Eşleme (ORM) Araçlarının .NET 6 Ortamında Performans Analizi

**Özet**— Nesneye yönelik programlamada, Nesne-İlişkisel Eşleme (Object-Relational Mapping – ORM) araçları sıklıkla kullanılmaktadır. Yazılım geliştiricilerin ORM seçimi yaparken en önemli seçim kriterlerinden birisi bu araçların sağladığı performanstır. ORM araçları üzerine yapılan performans araştırmalarının çoğu işlem süresi ve Rastgele Erişimli Bellek (Random-Access Memory – RAM) kullanım bilgileriyle sınırlı kalmış, Merkezi İşlem Birimi (Central Processing Unit – CPU) kullanım bilgilerine yer verilmemiştir. Ayrıca literatürde, platform bağımsız ve açık kaynak olarak üretilen yeni nesil .NET platformu olan .NET 6 ortamında yapılmış bir ORM performans çalışmasına rastlanılmamıştır. Bu çalışmada, belirtilen eksikliği gidermek ve yazılım geliştiricilere yol göstermek için .NET 6 ortamında belirli ORM araçlarının performans analizi gerçekleştirilmiştir. Çalışmada, işlem süresi ve RAM kullanım bilgilerinin yanında CPU kullanım bilgileri de yer almaktadır. Bu çalışmada başlıca ORM araçlarından Dapper, NHibernate ve Entity Framework Core (EF Core) ile kayıt okuma, ekleme, güncelleme, silme, arama ve sıralama işlemleri gerçekleştirilerek, işlem süresi, RAM ve CPU kullanımının ölçülebileceği bir yazılım geliştirilmiştir. Yapılan ölçümler sonucunda işlem süresi açısından; okuma, silme, arama ve sıralama işlemleri için Dapper; ekleme ve güncelleme işlemleri için EF Core en iyi sonuçları vermiştir. Kaynak kullanımını açısından Dapper’ın en iyi performansa sahip olduğu, EF Core ile NHibernate araçlarının sıralamalarının ise kayıt sayısı ve işlem türüne göre kendi aralarında değiştiği sonucuna varılmıştır.

**Anahtar Kelimeler**— orm, .net 6, dapper, nhibernate, entity framework core

## 1. INTRODUCTION

Object-oriented programming paradigm is frequently used for developing software. When using this paradigm, Object Relational Mapping (ORM) tools are indispensable for converting structures in relational model to object-oriented model. There are many ORM tools developed for different platforms, and each offers different advantages and disadvantages.

The performance of ORM tools is the most important criteria for developers when they need to select one of these tools. There are several studies ([1-3]) which analyze the performance of ORM tools on different software platforms and databases. These studies make comparisons based on operation processing times. Zmaranda et. al [4], measures the RAM usage besides processing times. A study by Balci [5], which is not performance comparison research on ORM tools but examines the performance analysis of the Entity Framework ORM tool on different databases, included all of the processing time, RAM and CPU usage information. There is a need to make performance comparison research on ORM tools by including not only processing times and RAM usage, but also the CPU usage.

Moreover, existing studies such as [6,7] perform comparisons on previous versions of .NET Framework by using Microsoft SQL Server database. Microsoft has a new .NET 6 environment which is a new generation, open-source, and cross-platform software development framework. Besides the framework, PostgreSQL database has gaining popularity and it has not been used in performance studies. These discussions show that, there is also a need for ORM tools comparison study on .NET 6 environment with PostgreSQL database.

In this study, we try to find the best performant ORM tools in .NET 6 environment with PostgreSQL database. We develop a software for measuring processing times, RAM usage, and CPU usage. We measure these for read, insert, update, delete (CRUD) [8], search and sort operations. We use Dapper, NHibernate and EF Core ORM tools in comparisons. These measurements will be analyzed to show the performance statistics of ORM tools in .NET 6 environment. The results may offer a guideline to developers for selecting the best ORM tools suitable for different operations.

## 2. CONCEPTUAL FRAMEWORK

There are two major methods for accessing to the databases from different programming languages or environments. The first one, which is the traditional one, is to use database providers' libraries to connect and execute operations. This method is fast since the libraries are generally optimized by the database providers to their databases. However, the application is bound to a specific database, and it is hard to switch between different databases. Moreover, the code for converting relational model to object-oriented model or vice versa must be written manually. This may need a lot

of work. The second method is to use an ORM tool for enabling modularity and decreasing the workload. Using an ORM tool may decrease the performance of the application. Joshi and Kukreti [9] compare ORM tools and traditional library access methods and they found that the complex code produced by ORM tools decreases the performance. They also indicate that when the advantages of using ORM tools are considered, the performance loss can be negligible.

Since ORM tools decrease the total performance of the applications, the importance of the performance of the ORM tools is paramount. There are some performance comparison studies in the literature. In a study implemented by Cvetkovic and Jankovic [7], the two ORM tools, NHibernate and Entity Framework are compared. Zmaranda et. al. [4] compare Dapper, EF Core and NHibernate tools. These studies shed the light on the performances of different ORM tools. However, these studies use Microsoft SQL Server as database, and they do not measure the CPU performance. Additionally, these studies don't analyze search and sort operations directly without the effects of other database structures. In another study [10], authors compare eight ORM frameworks with four different programming languages. Yousaf [11] evaluates the performance of Java-based ORM tools (Hibernate, EclipseLink, OpenJPA and Ebean) and his own GlycoVault lightweight persistence tool. In these studies, authors only compare read operations and Dapper is not included in ORM lists. For these reasons, we make a performance comparison of popular ORM tools which are Dapper, NHibernate and Entity Framework in .NET 6 environment by using a popular database which is PostgreSQL. We also compare read, insert, update, delete, sort and search operations. This study does not include a performance comparison of ORM tools and traditional library access methods. Colley et al. [12] has such a study which compares the effects of Entity Framework with SQL Server 2014 database and lists the negative behaviors of ORM tools.

Before describing the performance comparison methodology, we give general information about ORM technique and specific information about the ORM tools we use in this study.

### 2.1. Object Relational Mapping

An ORM tool is a bridge between a relational database and object-oriented programming language. It allows developers to work directly on the object-oriented programming concepts without thinking the details of the conversion of the components to relational tables or constraints. ORM tools have some advantages:

- They allow developing applications without being tied to a specific database. Different databases can be used with the same source code.
- Developers may execute database operations without writing SQL statements.

- Developers can easily concentrate to the OOP concepts.
- They decrease the time for writing database code.
- They increase the code readability.

Besides these advantages ORM tools have also some disadvantages:

- Writing direct SQL statements allows better performance.
- It is hard to write complex queries with ORM. Writing them with SQL may be easier.

## 2.2. ORM Tools

There are many ORM tools for object-oriented programming: Hibernate, TopLink and OpenJPA are used with Java; Django, Peewee, and SQLAlchemy are used for Python; and RedBeanPHP, Doctrine, and Propel are used for PHP. In .NET environment, Dapper, NHibernate and EF Core are highly used and, in this study, performances of these ORM tools are compared.

**Dapper:** Dapper is an open-source micro ORM tool developed for .NET environment. The main aim of Dapper is to provide performance to applications and to allow developers to decrease the effort of mapping operations.

**Entity Framework Core (EF Core):** EF Core is an open source and cross-platform ORM tool for ADO.NET data access library. It is a new version of Entity Framework ORM tool which has been distributed within .NET Framework. Starting from the Entity Framework version 6, Microsoft decided to deliver EF Core separately [13]. Because of this, EF Core, a more modern and sustainable ORM tool, is used in this study.

**NHibernate:** NHibernate is a .NET version of Hibernate which is frequently used ORM tool in Java environments. NHibernate is an open-source tool and includes almost all features of current Hibernate.

ORM tools are classified as full-featured ORM tools and micro ORM tools based on the features that they support. Micro ORM tools have limited capabilities according to the full-featured ones but they perform faster. A micro ORM tool may not support some caching capabilities, e.g., second level cache. Moreover, when a query is executed and an object is loaded from the database, other objects which are in relationship with this object are not automatically loaded. The programmer has to write special queries to load related objects. Besides these, micro ORM tools generally do not have graphical modelers and automatic database object creation capabilities [14].

ORM tools apply caching techniques for repetitive database operations. These techniques provide performance gains. EF Core has three types of caching: object caching, query plan caching and metadata caching.

Object caching is known as first level caching and it stores objects retrieved from database to memory. Query plan caching is used for storing queries executed more than once. This allows skipping the parsing and compiling operations of the query for later executions. EF Core supports metadata caching which is used for different connections to share the type and mapping information. NHibernate also supports first level caching to maintain objects in memory when they loaded first time. NHibernate has a second level cache for storing query plans and query results. EF Core and NHibernate, since they are full-features ORM tools, provide first level cache by default. However, Dapper only caches information for queries to materialize objects and process parameters quickly [15].

Another performance concern with ORM tools is loading related data with queries. This concern generally known as related data loading or fetching. Different ORM tools have different default characteristics for related data loading. Some ORM tools such as EF Core support eager loading by default [16]. Eager loading allows loading all the required entities with one query. Objects in relation with the parent object are also automatically loaded. Some ORM tools, such as NHibernate use lazy loading by default [17]. In this method, related objects are not loaded unless they are really needed. Dapper uses a multi mapping technique which is almost similar with eager loading [18]. However, because Dapper is micro ORM tool, third party libraries are needed for adding lazy loading property.

ORM tools use different mapping techniques between objects and tables and fields. Using an XML file, inserting annotations to source code or writing code to generate mappings are examples of metadata mapping. EF core and Hibernate support variety of these methods as shown in Table 1. In Dapper queries, we execute SQL statements by passing parameters. Beside the metadata mapping, ORM tools have capabilities for reflecting structure changes in object models to databases. If object model frequently changes, then these changes can be reflected to the database by executing a loading procedure. Executing a loading procedure frequently can be tedious. Because of this, some ORM tools use a reflection technique for the objects at runtime to reflect the changes. In this technique, the mapping between the object and table is stored in cache and upcoming calls use this mapping. Change reflection is only applied on first call [19].

Table 1. Properties of ORM Tools

	EF Core	NHibernate	Dapper
Mapping for Metadata	Code based, Attributes based	.HBM, .XML, Code based, Attributes based	SQL Statement
API	ADO.NET	ADO.NET	ADO.NET
Model Change Reflection Type	Automatic	With 3 <sup>rd</sup> party tools	-

Another concern with the ORM tools is transaction support. In EF Core, all the changes are tracked and handled in memory and when the “SaveChanges” method of “DbContext.Database” class is called the changes are applied to the database. This operation is atomic and all the changes are either committed or rolled back. We call this method after inserting, updating and deleting all the records. NHibernate has the similar methods. We use the “Save” method of “ISessionFactory” class in NHibernate API to save the changes permanently. However, Dapper has a different technique. Because Dapper, does not use object caching, it directly applies SQL statements through the classes of the related database. EF Core and NHibernate also have support for locking mechanisms for concurrent operations since they have object caches. However, our evaluation does not include concurrent access of the data and we have not utilized locking mechanisms.

### 2.3. .NET 6

.NET Framework is a software development environment produced by Microsoft and supports many languages like C#, Visual Basic, and F#. .NET 6, the latest version of this framework, is a platform for unifying web, mobile, desktop, games and IoT applications under a single framework. .NET 6 is released in November 2021 and targets cross-platforms from iOS, Mac OS, Windows, WatchOS, Android, tvOS etc. Programs written with different .NET compatible languages are compiled to platform-neutral Common Intermediate Language (CIL). Common Language Runtime (CLR), a platform specific runtime environment for .NET, compiles CIL to machine readable code.

### 2.4. PostgreSQL

PostgreSQL is an object-relational database management system developed by the University of California at Berkeley [20]. It is assumed that, PostgreSQL is the most advanced open-source relational database system [21]. This claim is supported by statistics of usage of PostgreSQL in high-level projects implemented by public and private organizations.

### 2.5. BenchmarkDotNet

BenchmarkDotNet is an open-source performance measurement library supported by .NET Foundation. BenchmarkDotNet, creates an isolated project for each method which are to be measured and executes them without other side-effects. By this way, processing time and resource consumption of each method can be measured precisely within their private processes [22].

### 2.6. Chinook Database

Chinook database is a sample database that can be created by a sample script file. It can be used by different databases such as, PostgreSQL, Oracle, SQL Server, and MySQL. The Chinook database has a data model which includes a digital media store, including tables for artists, albums, media tracks, invoices, and customers. In this study, we use the “Track” table from this model in performance measurements since it includes almost 3,500 records. We also use “Album” table for join operations.

## 3. METHODOLOGY

In this study, we get measures for read, insert, update, delete, sort and search operations. We test a one table select statement and a join statement for read measurements. We measure the processing times, RAM usage, and CPU usage of these operations by different ORM tools in .NET 6 environment. Dapper 2.0.123, EF Core 6.0.6 and NHibernate 5.3.12 versions are used for performance comparisons. Version 14.1 of PostgreSQL database is used. The records of the “Track” table of Chinook database are used in measurements. Additionally, we join “Album” table to “Track” table for measuring read operations of joined tables.

### 3.1. Architecture

We develop a software, ORMPY, for measuring processing times, RAM, and CPU usages of ORM tools in .NET 6 environment. We make the software open-source and publish it in GitHub<sup>1</sup>. The software is developed by a layered architecture including a model layer (Entity Layer), a persistence layer (Data Access Layer) and two application layers. Business layer is integrated within the application layer for executing the operations in isolation to get most accurate results.

There are two application layers in ORMPY. The first application layer measures processing times and RAM usage by BenchmarkDotNet library. In this application, each method is executed 100 times iteratively by BenchmarkDotNet, and averages are calculated. The second application layer measures CPU usage by Microsoft Diagnostics library [23]. Even though the methods are run in isolation, CPU usage is very fluctuating due to operating system processes. We execute each method 500 times and calculate the averages to normalize the CPU usage times. Moreover, we take the test computer in airplane mode and close the internet connections and all other applications. Measurements are implemented by using the computer with hardware properties given in Table 2. The detailed architecture of ORMPY software is given in Figure 1.

<sup>1</sup> <https://github.com/egvrcn/ORMPY>

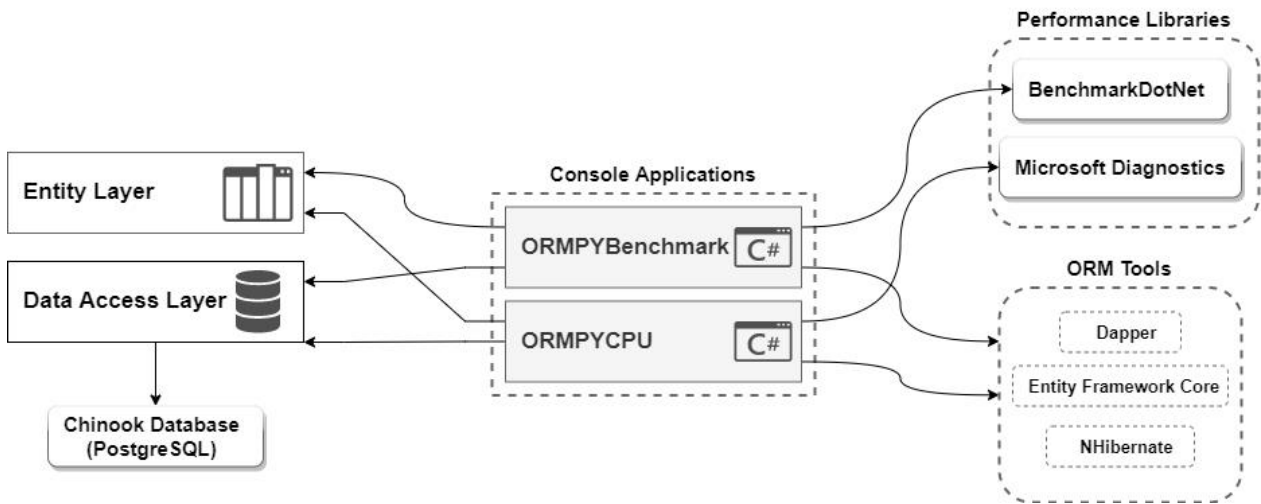


Figure 1. ORMPY Architecture

Table 2. Hardware/software information of test computer

Hardware/Software	Property
CPU	Intel Core i5-7300HQ CPU 2.50GHz (Kaby Lake), 4 logical and 4 physical cores
RAM	16 GB
Disk	Samsung SSD 860 Evo 250GB
Operating System	Windows 10 Pro

3.2. Data Collection

ORMPY is designed to execute read, insert, update, delete, search and sort operations. These operations are applied on the “Track” and “Album” tables of Chinook database for three ORM tools. The total number of records used for a process can be different depending on the process. The following list shows the total number of records used in different operations.

- Read operation is executed by 10,000, 50,000, and 100,000 records.
- Insert operation is executed by 1,000, 10,000, and 25,000 records.
- Update operation is executed by 1,000, 10,000, and 25,000 records.
- Delete operation is executed by 1,000, 10,000, and 25,000 records.
- Sort operation is executed by 10,000, 50,000, and 100,000 records.
- Search operation is executed by 10,000, 50,000, and 100,000 records.
- Read operation from joined tables is executed by 10,000, 50,000, and 100,000 records.

We store the results of processing times, RAM, and CPU usage data for each ORM tool into files after executing each operation. In addition, we extract the execution plans of the queries for understanding the background database operations triggered by ORM tools.

4. RESULTS

We implement 7 operations and measure processing times, RAM usage, and CPU usage. Totally, we collect data from 21 test scenarios. Each scenario includes measurements for three ORM tools and three different record count groups. We measure processing times through calculating the seconds needed for completing the tasks. RAM usage is measured by total MBs or KBs consumed by ORM tools. CPU usage is measured by getting the percentage of the total CPU usage throughout the process.

4.1. Read Operation

Read operations are executed by reading 10,000, 50,000 and 100,000 records on Dapper, EF Core and NHibernate ORM tools. The results of these operations are given in Table 3.

Table 3. Results of “Read” operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (MB)	CPU Usage (%)
Dapper	10,000	0.063	4	1.34
EF Core		0.091	12	2.09
NHibernate		0.097	12	3.14
Dapper	50,000	0.143	19	4.76
EF Core		0.265	60	7.79
NHibernate		0.400	64	11.64
Dapper	100,000	0.245	37	8.21
EF Core		0.465	121	11.87
NHibernate		0.803	128	16.6

Figure 2.a shows the processing time results of 10,000, 50,000, and 100,000 record reading operations of different

ORM tools. This figure shows that, Dapper is the fastest ORM tool for reading operations, whereas NHibernate is the slowest. These results are similar for 10,000, 50,000, and 100,000 record reading operations.

Figure 2.b shows that Dapper uses the least amount of memory, whereas NHibernate uses the highest. The difference between NHibernate and EF Core is small

according to the difference between them in processing times.

Similarly, Figure 2.c shows that Dapper uses the least CPU percentage, whereas NHibernate uses the highest. The difference between ORM tools for CPU usage is more significant according to the difference between them in RAM usage.

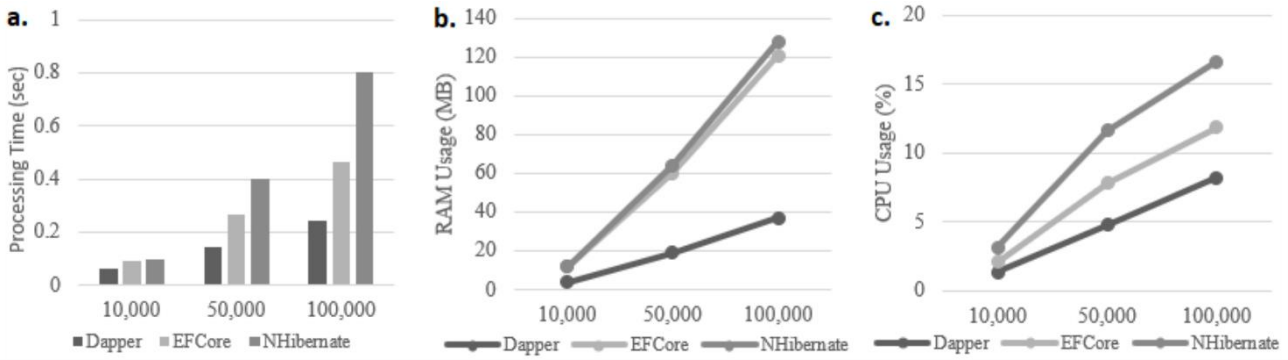


Figure 2. Results of "Read" operations

4.2. Insert Operation

Insert operations are executed by inserting 1,000, 10,000 and 25,000 records on Dapper, EF Core and NHibernate ORM tools. The results of these operations are given in Table 4.

Figure 3.a shows that EF Core is the fastest ORM tool in all the record count groups for insert operation. However, Dapper is the slowest one. According to Figure 3.b, ORM tools show quite the opposite performance in terms of RAM usage. Dapper uses the smallest amount of RAM whereas EF Core uses the largest amount of RAM. CPU usage performance of ORM tools in Figure 3.c, shows interesting results. EF Core performs better than NHibernate for 1,000 records. However, for 10,000 and 25,000 records NHibernate performs better than EF Core. Dapper has the best performance for CPU usage in all record counts.

Table 4. Results of "Insert" operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (MB)	CPU Usage (%)
Dapper	1,000	0.207	2	1.75
EF Core		0.110	15	2.75
NHibernate		0.174	9	3.89
Dapper	10,000	2.023	20	4.35
EF Core		1.012	146	9.35
NHibernate		1.642	94	8.84
Dapper	25,000	5.249	49	4.43
EF Core		2.483	362	11.23
NHibernate		4.283	236	9.93

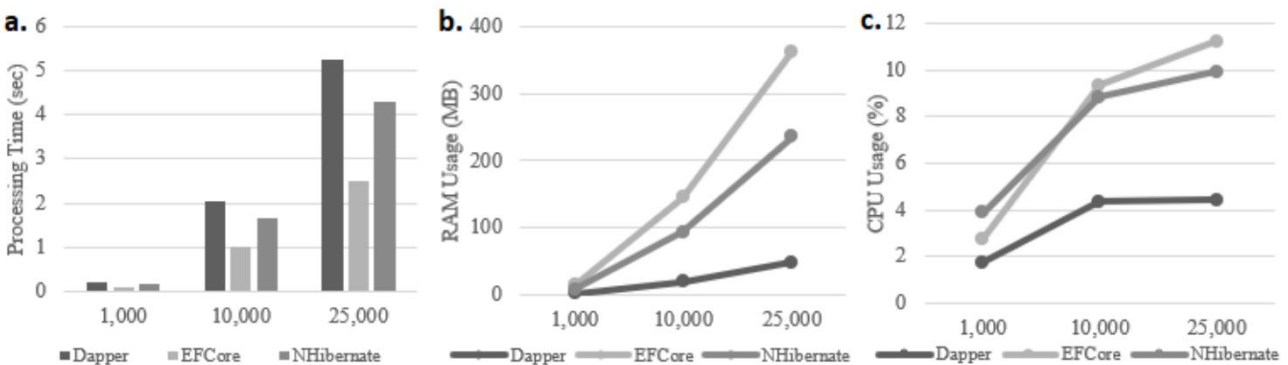


Figure 3. Results of "Insert" operations

### 4.3. Update Operation

We test Dapper, EF Core and NHibernate ORM tools for update operation by updating 1,000, 10,000, and 25,000 records. The results are shown in Table 5.

According to Figure 4.a, EF Core is the fastest ORM tool in all record count groups for update operations. Dapper shows the worst performance in terms of processing times. This shows that, while Dapper has a very good performance for reading operation, it doesn't have good performance for insert and update operations. However, Dapper is the best tool for RAM usage of update operations. NHibernate also consumes almost similar amount of RAM with Dapper. On the contrary to the performance on processing time, EF Core uses much memory, especially for 10,000 and 25,000 records (Figure 4.b). EF Core shows the worst performance in terms of CPU usage. Dapper is very efficient in CPU usage for all record count groups (Figure 4.c).

Table 5. Results of "Update" operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (MB)	CPU Usage (%)
Dapper	1,000	0.244	2	1.50
EF Core		0.126	14	2.49
NHibernate		0.175	3	2.45
Dapper	10,000	2.398	22	2.98
EF Core		1.310	127	7.44
NHibernate		1.726	25	6.56
Dapper	25,000	6.186	54	3.21
EF Core		3.130	313	8.56
NHibernate		4.514	64	7.4

### 4.4. Delete Operation

Dapper, EF Core and NHibernate ORM tools are used for deleting 1,000, 10,000 and 25,000 records. Table 6 shows the results of these delete operations.

According to the results in Figure 5.a Dapper is very fast in all record count groups. Even though EF Core is the slowest for 1,000 records, NHibernate performs worst with 10,000 and 25,000 records. The processing time of NHibernate for 25,000 record deletion is almost five times higher than deleting 10,000 records. Dapper is also the best ORM tool in terms of RAM usage for 10,000 and 25,000 records. Interestingly, for 1,000 records NHibernate uses less RAM than Dapper and EF Core (Figure 5.b). It is beyond any doubt that Dapper is the best tool in terms of CPU usage (Figure 5.c). However, EF Core and NHibernate tools produce different results for RAM usage and CPU usage. While NHibernate is better than EF Core for RAM usage, EF Core performs better in terms of CPU usage. NHibernate shows a drastic change in CPU usage when record count increase from 1,000 to 10,000.

Table 6. Results of "Delete" operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (KB)	CPU Usage (%)
Dapper	1,000	0.005	3.0	0.33
EF Core		0.091	6.5	0.99
NHibernate		0.030	1.9	1.29
Dapper	10,000	0.040	3.0	0.33
EF Core		0.789	61.4	3.92
NHibernate		1.841	19.1	19.31
Dapper	25,000	0.112	3.0	0.35
EF Core		1.820	150.8	5.2
NHibernate		11.567	47.6	23.58

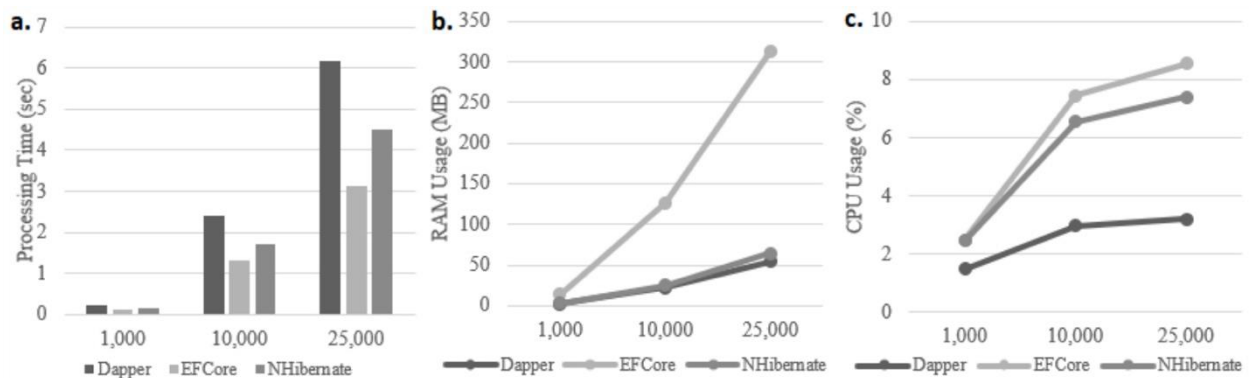


Figure 4. Results of "Update" operations



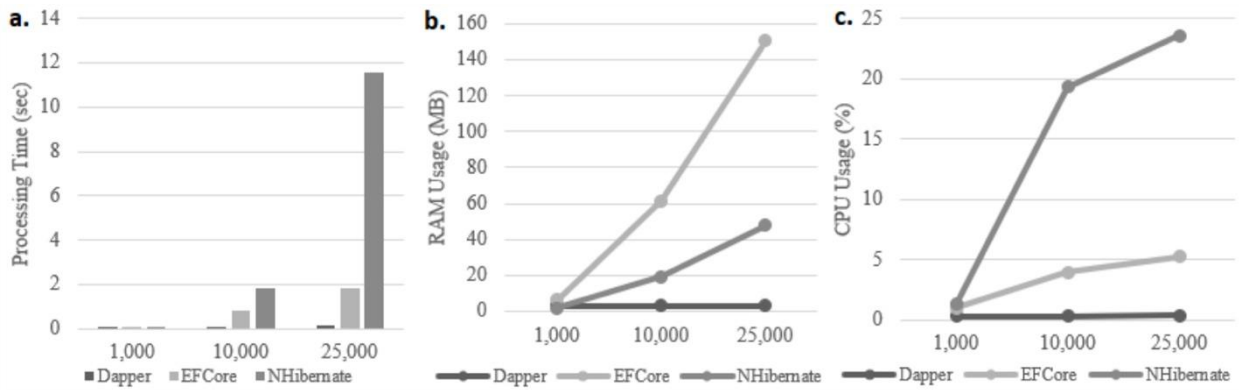


Figure 5. Results of "Delete" operations

4.5. Search Operation

We measure search operations with Dapper, EF Core and NHibernate ORM tools by using 10,000, 50,000, and 100,000 records. We search "Song" word in the "Track" table. The results are shown in Table 7.

The results in Figure 6.a shows that Dapper implements the fastest search in all record count groups. The slowest ORM tool is EF Core. Dapper is also better in RAM usage (Figure 6.b). However, for RAM usage, EF Core and NHibernate are close to each other. EF Core and NHibernate tools use RAM almost 3-4 times higher than Dapper depending on the record counts. On the contrary to processing time and RAM usage, Dapper, even though it is still the best, uses CPU time almost similar with the EF Core and NHibernate. EF Core is slightly better than NHibernate in terms of CPU usage (Figure 6.c).

Table 7. Results of "Search" operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (KB)	CPU Usage (%)
Dapper	10,000	0.003	34	0.36
EF Core		0.013	136	0.38
NHibernate		0.005	115	0.40
Dapper	50,000	0.009	154	0.38
EF Core		0.021	492	0.44
NHibernate		0.011	493	0.46
Dapper	100,000	0.016	307	0.44
EF Core		0.032	952	0.52
NHibernate		0.019	979	0.52

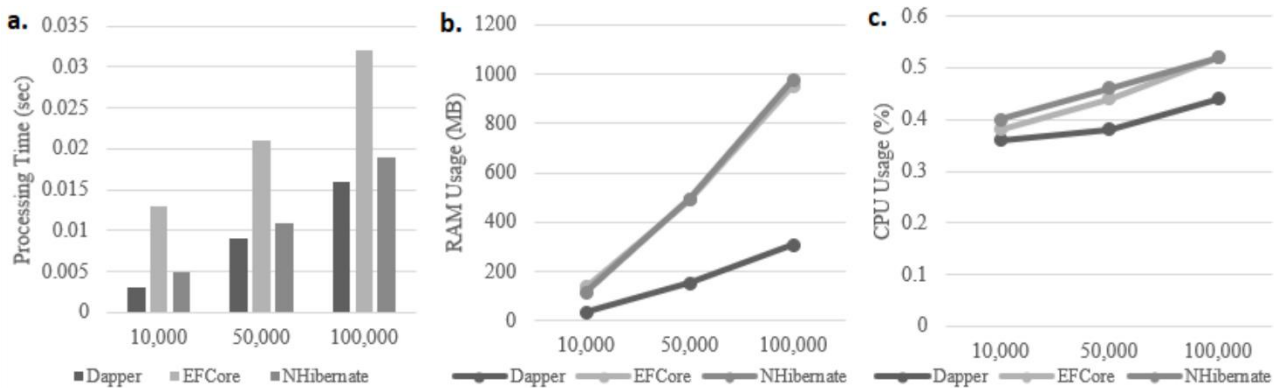


Figure 6. Results of "Search" operations

4.6. Sort operation

We test Dapper, EF Core and NHibernate ORM tools for sorting 10,000, 50,000 and 100,000 records. A descending (from Z to A) sort operation is applied to "Name" field of the "Track" table. The results are shown in Table 8.

Figure 7.a shows that the fastest ORM tool is Dapper. NHibernate and EF Core show almost similar processing

time performance for 10,000 records. However, for 50,000 and 100,000 records EF Core is faster than NHibernate. Dapper is still better for RAM usage (Figure 7.b) and CPU usage (Figure 7.c) according to EF Core and NHibernate. Dapper's RAM usage 3-4 times lower than EF Core and NHibernate. EF Core and NHibernate show similar performances for RAM usage. This is not the case for CPU usage, because EF Core has significantly lower CPU usage than NHibernate.

Table 8. Results of "Sort" operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (MB)	CPU Usage (%)
Dapper	10,000	0.115	4	1.20
EF Core		0.144	12	2.06
NHibernate		0.146	12	2.83
Dapper	50,000	0.401	19	3.14
EF Core		0.539	60	5.65
NHibernate		0.638	64	8.96
Dapper	100,000	0.884	37	4.62
EF Core		1.097	121	7.50
NHibernate		1.407	128	11.76

Table 9. Results of "Join" operations

ORM Tool	Record Count	Processing Time (sec)	RAM Usage (MB)	CPU Usage (%)
Dapper	10,000	1.449	5	0.42
EF Core		1.456	5	0.34
NHibernate		1.511	14	1.1
Dapper	50,000	1.478	25	2.04
EF Core		1.535	23	1.18
NHibernate		1.618	68	5.88
Dapper	100,000	1.594	50	3.88
EF Core		1.610	46	2.25
NHibernate		1.945	137	10.27

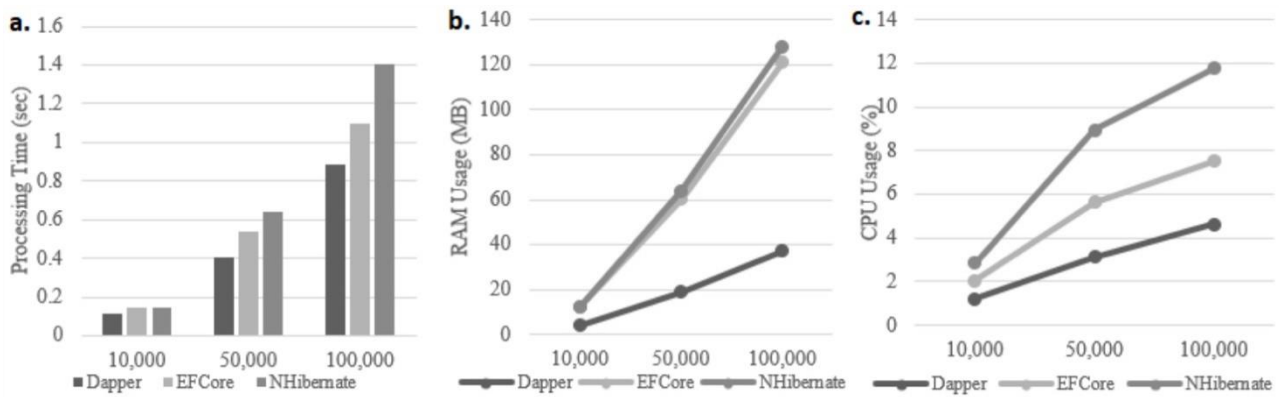


Figure 7. Results of "Sort" operations

4.7. Read operation from joined tables

We test Dapper, EF Core and NHibernate ORM tools for a join operation by reading 10,000, 50,000 and 100,000 records. We execute an inner join between "Track" and "Album" tables based on the album identifier attribute. The results are shown in Table 9

10,000 and 50,000 record readings. However, NHibernate reading operation takes longer than others for 100,000 records. Dapper and EF Core use almost similar RAM resources as shown in Figure 8.b. NHibernate consumes much memory especially for 50,000 and 100,000 records. CPU usage results are almost similar to the RAM usage. While NHibernate has the worst performance, EF Core is a little bit better than Dapper which is shown in Figure 8.c.

According to the results shown in Figure 8.a all ORM tools show similar performances for the processing time of

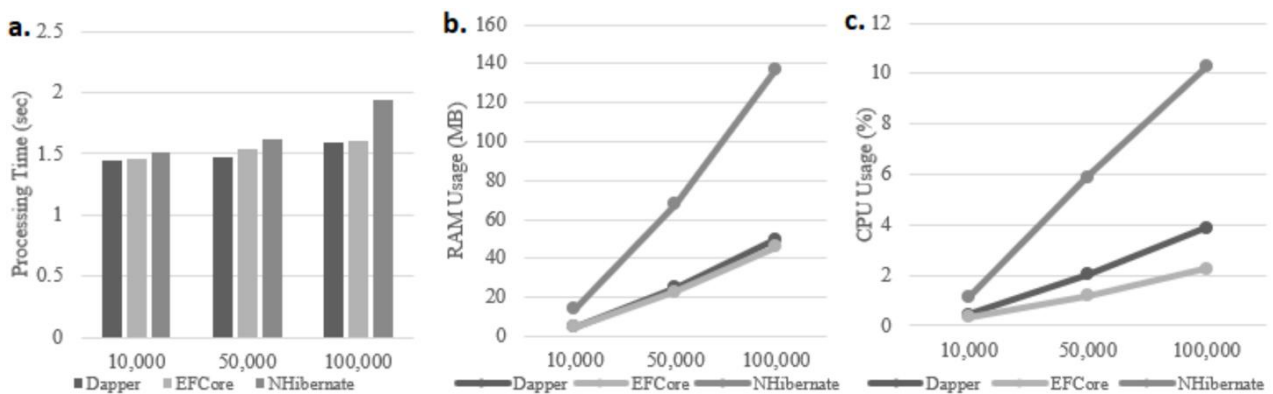


Figure 8. Results of "Join" operation

#### 4.8. Background operations

We examine the execution plans of the queries for understanding the background database operations triggered by ORM tools. We extract these execution plans by using the Session Manager of Dbeaver and Server Monitor of Navicat (see Table 10). We successfully access the execution plans for read, sort, search and reading from joined tables operations. However, for insert, update and delete operations we are able to access only to the execution plans of Dapper. For EF Core and NHibernate, we are able to catch only commit/rollback operations. In reading from one table operation, all ORM tools apply sequential scans. Similarly, for search operation, all ORM tools apply sequential scan by adding conditions. In sorting operation, all ORM tools first apply sorting and then apply sequential scan. For reading from joined tables operation, all ORM tools apply a nested loop first. Then, a sequential scan is applied on "Track" table. After then, "Momoize" operation, which is an optimization technique for making efficient computation, is applied. The plan is completed by applying index scan on "Album" table. The analysis of execution plans shows that the difference between ORM tools is not due to the operations at the database side. The techniques applied by the ORM tools such as mapping, caching etc. make the difference.

### 5. CONCLUSION AND DISCUSSION

In this study, we implement a performance analysis of specific ORM tools in .NET 6 environment. We measure the processing time, RAM usage and CPU usage for 7 operations including read, insert, update, delete, search, sort and a read from joined tables. We develop a software, ORMPY, for implementing these measurements. We give results of 21 different scenarios implemented with Dapper, EF Core, and NHibernate ORM tools and discuss the results with graphics.

According to performance analysis results, Dapper shows the best performance in terms of processing times for read, delete, search and sort operations. EF Core performs best for insert and update operations in terms of processing times. The worst performance for processing times is handled by NHibernate on read and sort operations and Dapper on insert and update operations. The ORM tools that complete the delete operation in the longest time differ according to the number of records. EF Core performs the worst for 1,000 records, and NHibernate for 10,000 and 100,000 records based on delete operation processing time. For reading from joined tables operation, Dapper, again, has the best results while EF Core and NHibernate come later. However, the gap between ORM tools for reading from joined tables operation is very small according to reading from one table operation. According to the RAM usage results of ORM tools, Dapper performed the best by consuming the lowest RAM resources in all operations except reading from joined tables operation. NHibernate showed almost similar performance with Dapper in terms of RAM usage in update operations. The worst

performances of RAM usage are handled by NHibernate for read, search, sort and reading from joined tables operations and by EF Core for insert, update, and delete operations. In RAM usage, Dapper has not been the worst performing ORM tool for any operation. NHibernate uses less RAM than Dapper for only deleting 1,000 records. Interestingly, EF Core outperforms other tools for reading from joined tables operation. This means that, if the query becomes more complicated, EF Core starts performing better than others in terms of RAM usage. This may occur due to advanced coding techniques of the EF-Core. We examined the logs produced by BenchmarkDotNet for the Garbage Collector (GC) operations. GC executes to release the memory for objects that are no longer used by the applications. If the memory is not used efficiently, GC performs frequently. Logs show that, for EF Core operations GC performs less than other ORM operations. According to the CPU usage results of ORM tools, Dapper performs best for all operations except the reading from joined tables operation. The worst performances of CPU usage are handled by NHibernate for read, delete, and sort operations and by EF Core for update operation. For insert operation, while EF Core is better than NHibernate for 1,000 records, NHibernate is better than EF Core for other record counts. For search operation, while EF Core is slightly better than NHibernate for 10,000 and 50,000 records, NHibernate and EF Core show the same performance for 100,000 records. EF Core uses less CPU for reading from joined tables operation. Again, this means that, if the query becomes more complicated, EF Core starts performing better than others in terms of CPU usage. This result may also be related with the GC operations. When GC starts, it suspends the application, releases the unnecessary objects and resumes the application. Less GC execution means less suspension and faster execution times for EF Core.

It is very hard to compare the results of our study with other studies. The test environments, databases that are used and operations are not standard within these studies. For example, in [4], one of the insert scenarios includes three insert operations to the tables in one-to-one relationship. In this scenario NHibernate performs better in terms of processing time with all record count groups including 500, 1,000, 2,000, 5,000, and 10,000. On the contrary to our insert operation results, the results of [4] may show that for complex scenarios, NHibernate may have better performances. Similarly with processing times, memory usage statistics in [4] are not coherent with our study. While Dapper is the best tool in terms of memory usage in our study, it is not the case in [4], especially for complex operations. We can also compare our joined reading operation with the get operation of [4]. In our one-to-many joined select statement, Dapper is the fastest tool while NHibernate is the fastest in [4]. Because of these test environment and operation structure differences, we mostly use simple queries and don't apply any configuration parameter changes to ORM tools and PostgreSQL database to get results without any side effects.

Table 10. Execution plans of the operations

Oper.	ORM	Execution Plan					
		Node Type	Entity	Cost	Rows	Time (ms)	Condition
Read	Dapper	Seq Scan	track	0.00 - 271836.60	100000	1.505.325	[NULL]
	EF Core	Seq Scan	track	0.00 - 271836.60	100000	1.495.283	[NULL]
	NHibernate	Seq Scan	track	0.00 - 271836.60	100000	1.472.832	[NULL]
Insert	Dapper	ModifyTable	track	0.00 - 0.01	0	0.030	[NULL]
		Result	[NULL]	0.00 - 0.01	1	0.006	[NULL]
No accessible data for EF Core and NHibernate							
Update	Dapper	ModifyTable	track	8.43 - 12.44	0	0.033	[NULL]
		Bitmap Heap Scan	track	8.43 - 12.44	1	0.016	[NULL]
		Bitmap Index Scan	PK_Track	0.00 - 8.43	1	0.009	(track_id = 1)
No accessible data for EF Core and NHibernate							
Delete	Dapper	ModifyTable	track	11.51 - 15.53	0	0.074	[NULL]
		Nested Loop	[NULL]	11.51 - 15.53	1	0.058	[NULL]
		Aggregate	[NULL]	3.08 - 3.09	1	0.038	[NULL]
		Subquery Scan	[NULL]	0.00 - 3.08	1	0.034	[NULL]
		Limit	[NULL]	0.00 - 3.07	1	0.029	[NULL]
		Seq Scan	track	0.00 - 271836.60	1	0.028	[NULL]
		Bitmap Heap Scan	track	8.43 - 12.44	1	0.010	[NULL]
		Bitmap Index Scan	PK_Track	0.00 - 8.43	1	0.007	(track_id = "ANY_subquery". track_id)
No accessible data for EF Core and NHibernate							
Search	Dapper	Seq Scan	track	0.00 - 272058.00	742	1.507.293	((name)::text ~ '%Song%':::text)
	EF Core	Seq Scan	track	0.00 - 272279.40	742	1.549.108	(strpos((name)::text, 'Song':::text) > 0)
	NHibernate	Seq Scan	track	0.00 - 272058.00	742	1.533.420	((name)::text ~ '%Song%':::text)
Sort	Dapper	Sort	[NULL]	283051.24 - 283272.64	100000	2.381.069	[NULL]
		Seq Scan	track	0.00 - 271836.60	100000	1.610.319	[NULL]
	EF Core	Sort	[NULL]	283051.24 - 283272.64	100000	2.212.061	[NULL]
		Seq Scan	track	0.00 - 271836.60	100000	1.486.284	[NULL]
	NHibernate	Sort	[NULL]	283051.24 - 283272.64	100000	2.188.023	[NULL]
		Seq Scan	track	0.00 - 271836.60	100000	1.471.869	[NULL]
Join	Dapper	Nested Loop	[NULL]	0.16 - 274098.82	100000	1.497.240	[NULL]
		Seq Scan	track	0.00 - 271836.60	100000	1.453.257	[NULL]
		Memoize	[NULL]	0.16 - 0.18	1	0.000	[NULL]
		Index Scan	album	0.15 - 0.17	1	0.001	(album_id = t.album_id)
	EF Core	Nested Loop	[NULL]	0.16 - 274098.82	100000	1.619.155	[NULL]
		Seq Scan	track	0.00 - 271836.60	100000	1.577.679	[NULL]
		Memoize	[NULL]	0.16 - 0.18	1	0.000	[NULL]
		Index Scan	album	0.15 - 0.17	1	0.001	(album_id = t.album_id)
	NHibernate	Nested Loop	[NULL]	0.16 - 274098.82	100000	1.537.926	[NULL]
		Seq Scan	track	0.00 - 271836.60	100000	1.496.500	[NULL]
Memoize		[NULL]	0.16 - 0.18	1	0.000	[NULL]	
Index Scan		album	0.15 - 0.17	1	0.001	(album_id = track0_.album_id)	

In terms of processing time, Dapper gives the best results in 5 of 7 operations including read, delete, search sort and reading from joined tables. However, Dapper could not win the feature of being the best ORM tool in processing times due to the worst results in insert and update operations. EF Core gives the best result in all record count groups to the nearest ORM tool, with a margin of about 70% for inserts and about 40% for updates. In terms of transaction times,

the NHibernate ORM tool doesn't have the best performance for any operation. As a result, Dapper ORM tool should be used to get the fastest results in applications where operations such as reading, searching, and sorting will be done intensively. In applications which adding and updating operations will be carried out intensively, the EF Core ORM tool should be used to get the fastest results.

In terms of RAM usage, Dapper gives the best results in 6 out of 7 operations while EF Core performs better only in the reading from joined tables operation. However, in reading from joined tables operation, there are very close consumption results between EF Core and Dapper in terms of RAM usage.

Dapper gives the best results in terms of CPU usage in all operations except reading from joined tables operation. Moreover, it performs 2 times or better than other ORM tools in many operations. However, from these results, we can infer that when queries become more complicated, EF Core outperforms others in terms of CPU and RAM usage.

If we evaluate the RAM usage and CPU usage results together, Dapper is still the ORM tool that gives the best results in terms of resource usage. Dapper ORM tool should be preferred in software projects developed with .NET 6 where resource consumption is higher.

As a result of our research, Dapper provides the best performance within the ORM tools in .NET 6 environment, considering the processing time, RAM usage, and CPU usage. When choosing Dapper, it should be considered that some of its features are clipped as a micro ORM tool. If a full-featured ORM tool needs to be used, the ORM tool that we can get the best performance is EF Core.

According to the execution plan results, the differences between ORM tools are not due to the operations at the database side. Databases apply almost similar execution plans for simple operations. The main reason for the differences is the techniques applied by the ORM tools such as mapping, caching etc.

Different conclusions about different ORM tools can be obtained by doing similar research to ours. The following list can be given as research recommendations in this regard:

- Analysis studies can be done with different or new ORM tools.
- Performance analyzes can be made with different databases other than PostgreSQL.
- ORM tools performance analysis can be performed for different software development environments.
- A benchmark suite can be developed by standardizing operations, software configuration parameters and hardware characteristics of the test computers.

## REFERENCES

- [1] A. Gruca, P. Podsiadlo, "Performance Analysis of .NET Based Object-Relational Mapping Frameworks", **10th International Conference Beyond Databases, Architectures, and Structures (BDAS 2014)**, Ustron, Poland, 40-49, May 27-30, 2014.
- [2] S. M. Bhatti, Z. H. Abro, F. R. Abro, "Performance Evaluation of Java Based Object Relational Mapping Tool", *Mehran University Research Journal of Engineering and Technology*, 32(2), 159-166, 2013.
- [3] M. Kopteff, "The Usage and Performance of Object Databases compared with ORM tools in a Java environment", **1st International Conference on Objects and Databases (ICOODB'08)**, Berlin, Germany, 199-218, 13-14 March, 2008.
- [4] D. Zmaranda, L.-L. Pop-Fele, C. Györödi, R. Györödi, G. Pecherle, "Performance comparison of crud methods using net object relational mappers: A case study", *International Journal of Advanced Computer Science and Applications*, 11(1), 55-65, 2020.
- [5] T. Balcı, **Entity framework'ün farklı veri tabanlarındaki performans analizi**, Yüksek Lisans Tezi, Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü, 2018.
- [6] B. Irakli, A. Bardavelidze, K. Bardavelidze, "Study And Analysis Of The .Net Platform-Based Technologies For Working with the Databases", **33rd International Conference on Information Technologies**, Bulgaria, 1-8, 19-20 September, 2019.
- [7] S. Cvetkovic, D. S. Janković, "A comparative study of the features and performance of orm tools in a .net environment.", **3rd International Conference on Object and Databases (ICOODB'10)**, Frankfurt am Main, Germany, 147-158, 28-30 September, 2010.
- [8] J. Martin, **Managing the Database Environment**, Prentice Hall, New Jersey, 1983.
- [9] A. Joshi, S. Kukreti, "Object Relational Mapping in Comparison to Traditional Data Access Techniques", *International Journal of Scientific & Engineering Research*, 5(6), 540-543, 2014.
- [10] V. Sivakumar, T. Balachander, Logu, R. Jannali, "Object Relational Mapping Framework Performance Impact", *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(7), 2516-2519 2021.
- [11] H. Yousaf, **Performance evaluation of java object-relational mapping tools**, Yüksek Lisans Tezi, The University of Georgia Graduate Faculty, 2012.
- [12] D. Colley, C. Stanier, and M. Asaduzzaman, "The Impact of Object-Relational Mapping Frameworks on Relational Query Performance", **International Conference on Computing, Electronics & Communications Engineering (icCECE)**, Southend, UK, 47-52, August 2018.
- [13] Internet: Microsoft technical documentation, Compare EF Core & EF6, <https://docs.microsoft.com/tr-tr/ef/efcore-and-ef6/>, 25.01.2021.
- [14] Internet: A. Shapovalov, Micro ORM vs ORM, <https://www.yaplex.com/blog/micro-orm-vs-orm/>, 24.06.2022.
- [15] Internet: Dapper - a simple object mapper for .Net, <https://github.com/DapperLib/Dapper>, 24.06.2022.
- [16] Internet: Eager Loading of Related Data, <https://docs.microsoft.com/en-us/ef/core/querying/related-data/eager>, 24.06.2022.
- [17] Internet: NHibernate Reference Documentation, Chapter 21 Improving performance, <https://nhibernate.info/doc/nhibernate-reference/performance.html#performance-fetching>, 24.06.2022.
- [18] Internet: D. Paquette, Loading Related Entities: Many-to-One, <https://www.davepaquette.com/archive/2018/02/06/loading-related-entities-many-to-one.aspxreference/performance.html#performance-fetching>, 24.06.2022.

- [19] Internet: A. Chan, High Performance Reflection ORM Layer, <https://www.codeproject.com/Articles/22188/High-Performance-Reflection-ORM-Layer>, 24.06.2022.
- [20] Internet: PostgreSQL 14.1 Documentation, What is PostgreSQL? <https://www.postgresql.org/docs/current/intro-what-is.html>, 25.01.2021.
- [21] Internet: The World's Most Advanced Open-Source Relational Database, <https://www.postgresql.org/>, 25.01.2021.
- [22] Internet: BenchmarkDotNet Powerful .NET library for benchmarking, <https://benchmarkdotnet.org/index.html>, 25.01.2021.
- [23] Internet: .NET documentation, Diagnostics client library, <https://docs.microsoft.com/en-us/dotnet/core/diagnostics/diagnostics-client-library>, 25.01.2021.

# Examination of Risks in Management of Software Licenses in Enterprises and Cost-Benefit Analysis

*Araştırma Makalesi/Research Article*

 Ali DURDU<sup>1</sup>,  Abdullah TUTĞAÇ<sup>2\*</sup>

<sup>1</sup>Yönetim Bilişim Sistemleri, Ankara Sosyal Bilimler Üniversitesi, Ankara, Türkiye

<sup>2</sup>Denetim ve Risk Yönetimi, Ankara Sosyal Bilimler Üniversitesi, Ankara, Türkiye

[ali.durdu@asbu.edu.tr](mailto:ali.durdu@asbu.edu.tr), [abdullahtutgac@gmail.com](mailto:abdullahtutgac@gmail.com)

(Geliş/Received:07.11.2021; Kabul/Accepted:03.10.2022)

DOI: 10.17671/gazibtd.1019869

**Abstract**— In this study, a proposal on how software license assets should be managed with an enterprise risk management approach has been put forward. The study examined licenses, which mean software usage rights, and various license models with a complex structure. Within the scope of the study, the risks that might arise concerning software licenses were determined and these risks were categorized. Considerable risks prioritized by senior management in businesses were identified and responses and strategies were set out for each risk factor that needs to be kept under control. The method suggested in the study was applied in a large-scale enterprise acting in the telecommunications sector. As part of this application, the enterprise's current situation in software asset management was revealed. The outcomes of the application are expected to be beneficial in determining the risk models of the enterprises. The study presents the study's management of software assets, determination of risks, and solutions to be applied to risks.

**Keywords**—audit, enterprise, risk, risk management, software, license

## İşletmelerde Yazılım Lisanslarının Yönetiminde Risklerin İncelenmesi ve Fayda-Maliyet Analizi

**Özet**— Bu çalışmada, yazılım lisans varlıklarının kurumsal risk yönetimi yaklaşımıyla nasıl yönetilmesi gerektiğine dair bir öneri ortaya konmuştur. Çalışmada yazılım kullanım hakları anlamına gelen lisanslar ve karmaşık bir yapıya sahip çeşitli lisans modelleri incelenmiştir. Çalışma kapsamında yazılım lisanslarına ilişkin oluşabilecek riskler belirlenmiş ve bu riskler kategorize edilmiştir. İşletmelerde üst yönetimin önceliklendirdiği önemli riskler belirlenmiş ve kontrol altında tutulması gereken her bir risk faktörü için yanıtlar ve stratejiler oluşturulmuştur. Çalışmada önerilen yöntem telekomünikasyon sektöründe faaliyet gösteren büyük ölçekli bir işletmede uygulanmıştır. Bu uygulama kapsamında işletmenin yazılım varlık yönetimindeki mevcut durumu ortaya çıktı. Uygulama sonuçlarının işletmelerin risk modellerinin belirlenmesinde faydalı olması beklenmektedir. Çalışma, çalışmanın yazılım varlıklarının yönetimini, risklerin belirlenmesini ve risklere uygulanacak çözümleri sunmaktadır.

**Anahtar Kelimeler**— denetim, kurumsal, risk, risk yönetimi, yazılım, lisans

### 1. INTRODUCTION

Risk management has become one of the areas that gain importance day by day and requires focus. As a result of the shift from traditional risk management to enterprise risk management, the risks are no longer limited to financial risks but have expanded to address the strategic,

human resources, operational, and particularly information technology risks of the organizations holistically. With the development of technology and the rapid increase in internet access, the expectations of customers have also changed, and companies have increased their investments in information technologies to keep pace with these expectations. A considerable part of

the investments made by companies undergoing digital transformation in this context consists of software. Therefore, institutions now need to pay more attention to the management of software assets, whose share is increasing compared to other budget items.

Today, we hear the concepts of digitalization and digital transformation frequently. We often witness that it is not only heard but also it is in our daily lives. So what does this concept mean? Let's try to explain the difference between the two concepts thanks to a few examples. When you enter the office restroom in your workplace, the lighting is an example of digitalization, which detects you through sensors and turns on automatically. The lighting will automatically turn off after a certain period when you stand still and have not come out of yet. However, if you have not left yet, you will be in the dark and you will have to move somehow to make the lights come on. This is an indicator of the fact that digital transformation has not started yet. Whenever the aforesaid sensors recognize you and estimate the duration of the lights according to your age and gender and turn them off accordingly, then we can say that digital transformation has begun. To set another example of digitalization, being able to control an air conditioner running in your home through remote control and stabilizing the ambient temperature can be cited. When your air conditioner automatically informs you or the technical service by sending a text message when the gas is running out of, or when it calculates the energy consumption of the surrounding air conditioners, considers that it consumes more energy and informs you that there must be a problem in the engine, we can say that the digital transformation process has started. Thus, digital transformation is a process that starts with the customer (user) [1].

Institutions are increasing their technology investments day by day to adapt to the digital transformation process and not to stay behind it. Particularly, largescale companies have already begun to position themselves as technology companies. For instance, the CEO of Goldman Sachs, a multinational investment bank, describes his company as a technology company. In the headline of an article published in MIT Technology Review, expressions are drawing our attention that the Walmart Company has turned into a technology company [1]. In advertisements for the bank industry after 2015 in Turkey, now, we see that emphasis is placed on customers making their conversations with robot assistants.

Digital Transformation Office established under the Presidency of the Republic of Turkey in 2018, enables our country to follow its adaptation to this transformation process in a more institutionalized structure. In the Information and Communication Security Guide released by the Office on 24 July 2020, the obligations are included, which makes software asset management mandatory for public institutions and private companies providing critical infrastructure. A 2-year schedule has

been planned for the institutions to adapt to this transition [2]. Durdu and Eren worked on integrating information security with information systems. The software has been proposed so that institutions can carry out and maintain their continuity by transferring all the processes required by ISO 27001, which is the information security management system standard, to the digital environment [3].

It is software, which is one of the most fundamental elements that constitute the infrastructure of products produced by companies for digital transformation in terms of information technologies. Bill Gates highlights that the software development business is the closest profession to magic in the world [1]. Companies take advantage of software-based platforms while introducing their products to their customers in different media (websites, social media, mobile applications, etc.). These software are procured either by in-house coding with the help of their own human resources or by acquisition from outside. Just as an artist demanding a certain fee for the lyrics s/he writes or a poet demanding fee for the poetry s/he composes, as it is subject to copyright, individuals and institutions that generate a work (software) by coding also demand a certain fee and copyright. The fee required is called the copyright license [4].

Table 1. Worldwide information technology spending forecast (S:Spending, G:Growth %)[5].

	2019 S	2019 G	2020 S	2020 G	2021 S	2021 G
Data center systems	210,053	0.6	188,365	-10.3	200,094	6.2
Enterprise software	476,687	11.7	449,506	-5.7	482,666	7.4
Devices	711,525	-0.3	596,914	-16.1	611,303	2.4
IT services	1,040,263	4.8	969,438	-6.8	1,023,179	5.5
Commun. services	1,372,236	-0.6	1,326,492	-3.3	1,366,419	3.0
<b>Overall IT</b>	<b>3,810,764</b>	<b>2.3</b>	<b>3,530,714</b>	<b>-7.3</b>	<b>3,683,661</b>	<b>4.3</b>

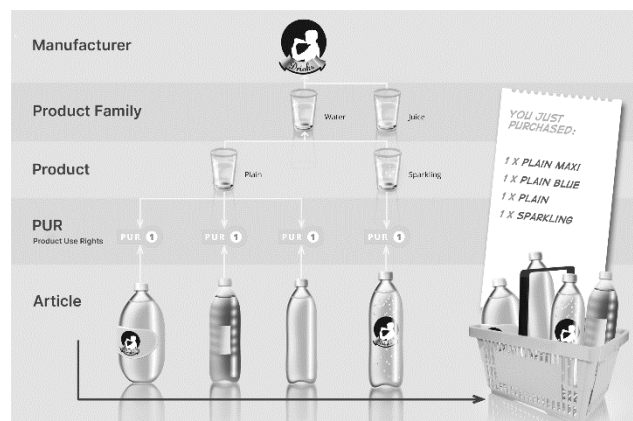


Figure 1. An example of product use rights [6].

Information as to how much companies invest in information technology across the world and how much these investments are expected to increase in the long view is available in Table 1. This table illustrates the trend of worldwide information technology expenditures



between the years 2019-2021, which was published by Gartner in July 2020. Realizing in particular that corporate software expenditures will increase at a greater rate in 2021 compared to other investment items reveals the fact that it is a demanding area for companies.

The concept of license defines the complexity and limits of the authorization to use the software owned by the licensee. When you buy software, you have a license that means the right to use it. With exceptions, the software itself always belongs to the software vendor. No price is paid for free software, open-source software or software for a definite period of use. Our research will not cover this type of software. There may be differences in licensing models of software purchased on the basis of a certain contract for a fee. In the research we performed in a telecommunication company, more than 1000 different software purchased from more than 100 different companies were encountered. Considering the licensing models of this software, it is seen that there are more than 100 different license models. Some of these licenses are user-based, earning-based, per central processing unit (CPU), per seat, usage-based etc. These models are changed by software vendors at certain periods and differ in terms of the infrastructures they are installed on. User licenses are commonly installed on client hardware. To set an example, laptop or desktop computers, mobile phones, tablets, etc. devices can be specified. Moreover, there is hardware such as servers or storage devices in data centers. Many different operating systems and software such as virtualization run on this hardware. A large-scale telecommunications company has more than 20,000 clients, 10,000 servers, and devices such as storage. Considering the software traces coming from these devices, it was observed that over 1 million data were received.

Apart from the variety and different software license models, it is also significant for companies to clearly define the right to use the purchased software in the contract. An example of product usage rights is given in Figure 1. In the example in Figure 1, a water order is received from a company that produces water and fruit juice. Some of the water is sparkling water and some is normal water. But some water is packaged in 1-liter, some in 0.5-liter bottles. Some of them are packaged in 6 pieces and they contain 3 normal and 3 sparkling water bottles, each of which is 0.5 liters. Even though it seems that only water has been ordered, the fee and usage rights paid for each product are different

The points that demonstrate how important it is for organizations to manage assets effectively are the following: the product variety in the above-mentioned software licenses, the differences in licensing models, the change of usage rights according to the contract, and the fact that it constitutes an important expenditure item in the budget. The underlying idea of this study is what kind of risks organizations will face in case they cannot

effectively manage these complex assets in a manner and how these risks should be managed.

In addition to the lack of a mere explanation agreed on the concept of risk, it is considered as a threat or opportunity that may have an impact on the achievement of the company or institutions' goals [7]. This concept was first used in the field of insurance in the 1950s. While the traditional understanding of risk management focused more on financial risks in the 1970s and 1980s, we can observe that organizations also focused on operational risks as of the early 1990s and switched to an enterprise risk management approach. After 2004, the risk was not only deemed as a threat but also as an opportunity that could be turned into an advantage [8].

In traditional risk management, silo-based risk management was adopted. Every single unit within the company focused on its own risks. Financial risks were mostly prioritized, and hedge transactions were performed using derivative instruments in order to avoid these risks. While shifting to enterprise risk management, a holistic approach to risks was adopted. Not only financial risks but also operational and strategic risks came to the fore. In addition to financial risks, the following risks also started to be managed in coordination: supply chains, distribution systems, corporate management, information technologies, and human resources [9].

Especially with the emergence of the enterprise risk management approach, that is, after the 1990s, the number of studies and publications on risk management has increased significantly. As can be seen in Figure 2, we can say that since the second half of the '90s, the rate of increase of publications on enterprise risk management has increased significantly compared to the increased rate of publications on internal control and internal audit.

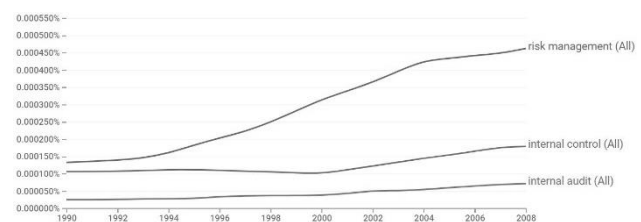


Figure 2. Publications on risk management, internal control, and internal audit [10].

The outbreak of the Enron crisis in 2001 and the Worldcom scandal at about the same time heralded a new era in financial circles and particularly in the enterprise risk management discipline. In the Enron crisis, thanks to the manipulation of the company's accounting records, a company that was indeed at loss was depicted as profitable for years. Following this crisis, the Public Company Accounting Reform and Investor Protection Act, known as the Sarbanes-Oxley Act, was enacted in the United States in 2002 [11]. With this act, it was obligatory for public

companies to adopt the risk management principles, and to document and evaluate the controls related to the identified risks. Moreover, the major responsibility for these processes is given to top executives [11].

In 2004, the COSO (Committee of Sponsoring Organizations of the Treadway Commission) put forth the Enterprise Risk Management Framework. Normally, in 1992 COSO introduced the Internal Control Framework, but especially following the Sarbanes-Oxley Act, it tried to clarify which methods and how should be employed in risk management and internal control issues during the implementation phase of this act [11].

Along with the COSO framework, many different organizations issued guidelines on risk management [12]. In 2009, ISO (International Standards Organization) published a guide that includes the application of risk management principles well-accepted in international standards known as ISO 31000 [11].

The proliferation of risk management in the course of time, its being subject to laws and its continuous improvement are the indicators demonstrating the fact that it has become more important for organizations. It has been criticized in the literature that there are few empirical studies evaluating the contributions that enterprise risk management (ERM) makes to organizations [11]. In general, there are studies on the following subjects: practices that measure the maturity levels of companies in different sectors in enterprise risk management, the effects of enterprise risk management on company performance, and its interaction with internal audit and internal control. For instance; to measure the impact of ERM on the financial performance of companies, samples were taken from publicly traded companies in the Istanbul Stock Exchange, which is among the largest 500 industrial companies published by the Istanbul Chamber of Industry (ISO) in 2014 [8]. In a study performed on 62 companies selected from the first 250 companies in Bursa, it was observed that the level of ERM implementation was positively affected by the increase in the turnover and degree of centralization of the companies [13]. 25 companies operating in the energy sector, which are among the ISO 500 companies in Turkey, were selected as samples, and with the surveys carried out, it was tried to reveal the maturity level and profile of enterprise risk management practices in our country [14].

Apart from the above-mentioned examples, many studies were also discussed in terms of managing software development project risks. A bank ATM project was taken as an example and an approach was put forward on how to identify and analyze risks in software development projects, plan risk response strategies, and design control activities [15]. There are also studies in which smart methods are employed in order to manage costs and risks in software projects. The software data between 2007 and 2014 of a company acting in the telecommunications sector were compiled. Thanks to these data, an approach was suggested that can predict how a software project will

result when it starts using the fuzzy logic method [16]. In parallel, artificial intelligence-based risk management studies became apparent. In the study conducted on 467 software project data collected from 774 different software companies, a web-based model that performs risk management based on artificial intelligence was produced. Using this model, the project budget, time plan, and deviations from achievements can be estimated as per the risk levels [17]. Fuzzy methods were used in the selection and evaluation of enterprise resource planning [18].

On the management of software assets, a thesis study was conducted in 2018 within the University of Twente, School of Management and Governance MBA (Business Administration) program in the Netherlands. Within the scope of this thesis, Thales Huizen, a large technology company, was chosen as a case study. In the study, it was mentioned that the purchasing costs cover 50 to 80 percent of the operations of the companies and the place of software costs in this cost was emphasized. Suggestions were put forward on how to improve the software acquisition and asset management process and what should be done in order to minimize the risks. The study is based on three main questions. Answers to the following questions were sought: Is software asset management applied in Thales company and similar companies acting in the same industry? What are the risks and problems in the current software sourcing and management process? Finally, what are the requirements for developed software procurement and management process in order to minimize risks and problems? A descriptive qualitative research method was adopted by using process modeling and semi-structured interviews [19]. In terms of software asset management, it can be argued that it is the most comprehensive academic study we encountered.

In a thesis study performed within the MBA program of Helsinki Metropolia University in Finland, the current situation of software asset management of a large telecom operator in Finland was analyzed. Using the results of the current situation analysis, improving the company's software asset management was set as the objective. As in the example of Thales, in terms of its method, it was qualitative research where interview workshops were performed. Participants representing the entire company were carefully selected in these interviews. In the study, it was revealed that there were positive developments in software asset management within the company, but the process remained isolated and the whole company should become organized. It was particularly emphasized that a tool (system) should be used in order to manage software asset management effectively and a program should be initiated within the company to improve the competencies [20].

Academic studies and private companies' surveys contain important examples regarding the prevalence of enterprise risk management in the world and the level of

maturity achieved by the institutions in Turkey [19]. Also, rapidly developing and transforming technology tools have a serious impact on institutions. One of the key components of these technology tools, the software licenses, are complicated and expensive, thus it's vital to manage them systematically.

Licenses for software assets are very complex for businesses and useless licenses are created if not managed effectively. In the report published by Gartner in July 2020, it was emphasized that it should be realized that especially enterprise software expenditures in enterprises will increase more than other investment items in 2021. Therefore, it has been clearly stated that the management of software licenses is an important risk for businesses to be managed. With the holistic perspective of enterprise risk management, companies need to focus on financial and technological risks.

In this study, it has been tried to start an application for the effective management of software licenses used by companies by adopting the enterprise risk management approach and to analyze the profit-benefit analysis provided to the institutions by discussing the post-implementation results. In the second part of the study, information is given about the methods to be used within the scope of the study. The methods used in the study were applied to a private company operating in the telecommunications sector and the details of the risks identified after the application were given. In the third part, current risks are given. In the fourth chapter, current risks are answered and suggestions are made. In the conclusion part, information was given to the institutions for the purpose of advice.

## 2. METHODS USED IN THIS STUDY

During the implementation phase, a company was selected as a pilot company acting in the telecommunications sector in Turkey. Having over 30 thousand employees and serving over 40 million customers, the company is among the leading companies in terms of its size in its sector. In this respect, it is expected that the application made specific to the selected company provides an advantage in terms of being inclusive for other companies implementing software license management.

Two different methods were employed in order to identify risks in information technology (IT) software license management. One-on-one interviews and focus group meetings were held. Moreover, in order to take advantage of the enterprise risk management approach, risk classes mentioned in the Public Internal Audit Manual were utilized [21].

### 2.1. Current situation (One-to-one interviews)

In order to determine the maturity level of software license management in the pilot company, the existing inventory was revealed. 166 units were contacted one-on-

one and a set of 25 questions was delivered. These questions were: the name of the software, the number of licenses, license metric, contact information, validity periods, whether the measurement was made or not (Annex-1).

As a result of the study, it was revealed that there are more than 1600 different types of software licenses in the inventory of the company. These licenses were observed to be obtained from more than 100 different companies/vendors. Also, it was seen during the examinations that there were 132 different license models (Annex-2). The sub-category licenses given in Annex-2 are the licenses under the main category licenses. For example, "User" based license is determined as the main category, but there are "User" licenses in many different subcategories such as "concurrent" or "named user". It was stated that the licenses in the inventory run by around 15,000 servers and approximately 21,000 client devices. According to the declaration of the relevant teams, it was understood that 15% of the software licenses in their inventory were not managed effectively. 65% of the software in the inventory was observed to be in the IT category and 35% in the Network/Access category. When the contract prices are considered, it was seen that an annual total investment and maintenance budget of more than 100 million TL was allocated.

### 2.2. Focus group meetings

The main purpose of this study is to determine the risks with one-to-one and group meetings in order to manage software licenses and make profit-benefit analysis in businesses. As a result of one-on-one interviews, the current inventory was revealed. Data revealing the complexity of license management was obtained. As the second stage, focus group meetings were held. Five competent experts in their fields, who engage in contract and license management, attended these meetings.

At the meetings held, the current software inventory was examined and the possible risks that might be encountered by the company in this regard were determined. The probability of occurrence and possible effects of the identified risks scored between 1-5. Risk scores were determined by multiplying the obtained impact-probability scores and prioritized considering the risk appetite of the company. The strategies to be adopted by the company against these risks, the responses to the risks and the controls to be applied were determined.

## 3. RISKS

A focus group can be broadly defined as a type of group discussion on a topic under the guidance of a trained group moderator [22]. As stated in the literature, one of the main purposes of the research is to manage software licenses and to identify risks with group meetings in order to make profit-benefit analysis in businesses. To summarize the picture that emerged as a result of the interviews; keeping the asset inventory up-to-date, which

includes software running on thousands of servers and clients in physical and virtual environments and with hundreds of complex license models, is a necessity and is of great importance. It is believed that the evaluation of this area, for which a remarkable amount of budget has been allocated, from an enterprise risk management perspective will be beneficial for the members of the board of directors and shareholders of the companies. Details of the 12 risks identified in focus group meetings are illustrated in Table 2.

Table 2. Risks identified in focus group meetings.

No	Risks	Risk causes
1	Inactive (Idle) software cost risks	Risk of financial loss that may occur due to the failure in pursuing the issues regarding the existing software: necessity, benefits and use cases
2	Duplicate procurement risks	Risk of duplicate procurements and possible financial loss due to the lack of existing inventory control in new software requests.
3	External audit risks	Loss of reputation and financial risks that may arise as a result of the audits performed by the software suppliers
4	Failure in being up to date with the versions	Security and business continuity interruption risk that may occur due to the lack of support by the relevant vendors for the old version software used
5	Ineffective inventory control	Workforce loss risk due to repeating of inventory work during contract renewal periods
6	Deficiency in alternative product analysis	Risk of financial loss due to not using more cost-effective or open source products having the same functions
7	Long procurement processes	The risk of loss of workforce and bargaining power due to prolonged procurement processes caused by not keeping track of the software inventory up to date.
8	Inadequate contract control	The risk of financial and business loss due to software contracts not being concluded in favor of the enterprise
9	Ineffective contract tracking	Business continuity interruption risk that may occur due to the inefficient tracking of software license contract periods
10	Internal audit risks	Possible risk of loss of workforce in the internal audit process due to inefficient software inventory management
11	Inefficiency in following the license model	Financial loss risk that may arise due to delays in adapting to licensing models that have changed as a result of the strategies of the vendors
12	Risks from incomplete feasibility	The risk of financial and business loss that may arise in case the software requirement request is not analyzed properly during the planning phase.

After the risks were determined, the impacts of these risks on the company were evaluated. While performing this assessment, the risk analysis model (risk map) was utilized. In the light of the scores obtained as a result of

the focus group meetings, the dangers that will be encountered by the company were evaluated through the risk analysis model.

The information on the impact range of the scored risks as a result of the creation of the risk analysis model is available in Table 3. According to the scores, it can be seen that 3 risks marked with dark grey are determined as high level, 5 risks marked with grey as medium level, and 4 risks marked with light grey as low-level risks. How the Company's top management plans to respond to these risks considering the risk appetite is also illustrated in the relevant table.

Table 3. Effects of risks and responses (I: Impact, P: Possibility, S: Score, R: Responses, C: Controls, CT: Control Type).

No	Risks	I	P	S	R	C	CT
1	Inactive (Idle) software cost risks	4	4	16	Check it	Preventive	Semi-automatic
2	Duplicate procurement risks	4	4	16	Check it	Preventive	Semi-automatic
3	External audit risks	5	2	10	Check it	Detective	Semi-automatic / Strict
4	Failure in being up to date with the versions	5	2	10	Check it	Preventive	Semi-automatic
5	Ineffective inventory control	1	4	4	Admit it	-	-
6	Deficiency in alternative product analysis	3	2	6	Admit it	-	-
7	Long procurement processes	2	3	6	Admit it	-	-
8	Inadequate contract control	4	2	8	Check it	Directing	Manual
9	Ineffective contract tracking	5	3	15	Check it	Preventive	Automatic
10	Internal audit risks	2	2	4	Admit it	-	-
11	Inefficiency in following the license model	3	3	9	Check it	Corrective	Semi-automatic
12	Risks from incomplete feasibility	4	3	12	Check it	Directing	Manual

#### 4. RESPONSES AND PROPOSALS TO RISKS

It was determined that the 3 risks, which are considered as higher level among the 12 risks illustrated in Table 3, are the ones that primarily need to be taken precautions. How the company will respond to these risks (risks 1, 2 and 9) and what kind of controls should be applied are clarified. These 3 risks were evaluated collectively.

Three different proposals were put forward in order to keep under control the High-Level Risks (1,2,9), which are Inactive (Idle) Software Cost Risks, Duplicate Procurement Risks and Ineffective Contract Tracking risks, and to minimize their impacts.

##### 4.1. 1st proposal; team building (SAM team)

To build a license inventory management-focused team, referred to as the Software Asset Management (SAM) Team. This is the team that centrally creates and manages the software license inventory. It can obtain information from relevant users, request contract details from purchasing teams, and follow up updates. The team to be built is positioned as the internal control team. It is of importance to be in a strong position within the technology organization, with equal measure to each unit. Otherwise, it will have problems in reaching the necessary information quickly and its power of sanction will be limited. Considering the size of company, where the pilot application is performed, the team is recommended to consist of at least 5 people. The team can be expanded according to the number of software vendor companies to be managed. In order to ensure an effective inventory management in light of the best practices, 1 expert should be allocated per 10 vendors.

##### 4.2. 2nd proposal; SAM system setup

It is recommended to provide a license inventory management system (Software Asset Management Tool) in order to manage license management with actual data. The following information about the software procured by the company should be kept regularly: what the license models are, number, right of use, duration of use, etc. What's more, actual usage and installation data are of vital importance. Within the scope of usage data, it is necessary to access real data through automation by scanning servers and clients rather than statements of individuals. Due to the size of the organization, different units can perform installations unaware of each other within the company. By ensuring a central control, possible under licensing or over licensing information can be monitored in an effective manner. In the light of this information, duplicate procurements can be prevented thanks to analyzing new license requests by checking current uses. By monitoring less frequent usage, savings can be achieved by excluding the related products from the scope of the contract during maintenance times. It is possible to be reminded about the renewal of license contracts by automatically delivering the usage periods to the teams responsible for the system.

##### 4.3. 3rd proposal; software life cycle process

During the phases of team building and establishment of license inventory management system, it is of vital importance to write the software life cycle process, which determines roles and responsibilities, and to be implemented after approval by top managers. If an approved process is not initialized, the responsibility areas of the SAM team will not be clear. Therefore, the established system will become dysfunctional in the course of time.

If these three proposals are put into practice effectively in the organization of the company, the impacts of risks 1, 2 and 9 will be minimized as far as possible. For risks 1 and 2, automatic and manual controls will be provided with the help of both system and process support. Automatic controls will be ensured for the number 9 risk. As a result of the market researches, it was estimated that the cost of the first investment (CAPEX) that the company has to bear for these controls will be around 2 and 3 million TL. The annual operational (OPEX) cost for the team (5 people) was calculated to be 0.75 million TL. It can be suggested that this cost is acceptable in terms of effective management of assets for a company that spends more than 100 million TL on license annually.

In an analysis released by Gartner in 2016, it was stated that companies that built the SAM system could save up to 30% in license expenditures [23]. In another analysis of Flexera company released in 2011, it was stated that companies that built the SAM system would make annual savings of up to 10% in license expenditures [24].

##### 4.4. External audit risks

In software license contracts notably owned by largescale international corporate companies, the relevant vendor's right to audit is included as standard. The products of 5 big companies, referred to as "Top 5 Vendors", are commonly used across the world. These big 5 companies are known as Oracle, Microsoft, IBM, SAP and Micro Focus (formerly HPE). As a matter of course, there are other companies that include the audit article in the contracts they made. However, these 5 companies have higher audit risks. Because the business volume with these companies is higher, the products used are directly related to critical infrastructure and business processes. As a priority, it is recommended to follow the licenses of these companies. No matter how many precautions are taken, the contractual request of the vendor for an external audit will not be prevented legally. However, once the risk occurs, some actions can be taken in order to minimize the possible impact. Especially, thanks to the building of the SAM system, license usage can be tracked automatically. Furthermore, resolutions should be released and it should be ensured that the software licenses, which are used within the company with the ethical license model, are not used over the amount of product (under licensing) purchased. If the problems that may arise after the audit are known, negotiations can be

made with the relevant vendor accordingly and the audited company can overcome this process with the least damage.

#### 4.5. Failure in being up to date with the versions

SAM system is thought to be used in order to control the license versions of software products. With the automatic discovery, software licenses coming from data sources to the SAM system can come together with version information. However, it will not be enough to get this information alone. It should be compared with an information repository that checks the up-to-dateness of the incoming version information and whether it is supported by the vendor. This information repository is formed by the shares of vendors who openly share license information with the public. When a catalog, referred to as the "Master Catalog", is included in the SAM system, automatic checks can take place. Products referred to as "End of Support" or "End of Life" can be reported to responsible users at certain periods.

Version information of companies whose license information is not openly disclosed can be added to the system manually and reported. Reports must also be delivered simultaneously to cyber security teams. Through implementing these controls, it can be targeted to minimize the risk impact by preventing possible security gaps and interruptions in business continuity.

#### 4.6. Inadequate contract control

In the technical specifications (RFX) and administrative contracts to be issued when purchasing software products, arrangements regarding license models or source code ownership of the software must be made in favor of the company. Especially in purchases by tender, in projects for which the solution is not yet clear, the license model may be requested to be perpetual. Making the preference not dependent on a variable metric such as the number of users, the number of processes and subscribers is important in terms of preventing additional costs that may arise after the project. Also, if possible, requesting the source code of the software to be procured will reduce the dependency of the company. In order to take these measures, the claimant technical units, procurement units and legal teams must be informed in writing in advance. In addition, obtaining consultancy from the SAM team during the technical specification preparation phase by the claimant teams and checking the documents by the SAM team are directing manual controls that will minimize the risk.

#### 4.7. Inefficiency in following the license model

It is known that software license models are changed over time by the vendors. They make this change sometimes to simplify complex license models, sometimes to increase their revenues, and sometimes due to the developments and changes in technology. Many manufacturing companies announce these changes to their customers and

the public in advance. Companies, on the other hand, need to take necessary precautions beforehand for these changes. With respect to these changes, the "Master Catalog" information repository working with the aforementioned SAM system can be utilized automatically, and it will be possible for the SAM team to follow up these changes manually. SAM team can analyze the impacts of license model changes that may be encountered in the future on the company and report them to the relevant parties. In this way, the company will be least affected by this risk. One point that should be kept in mind is that in order for the license model and metric change to occur, this change must be registered with a new contract.

#### 4.8. Risks from incomplete feasibility

The needs analysis should be performed accurately for the software that are planned to be procured at the stage of making a new investment decision. If the need is not determined accurately, there may be a risk of higher costs than the initial investment for the company. In order to minimize these risks, the SAM team can be included in the feasibility processes. Predictions and recommendations can be added to the checklist as to what the software model should be, which will be created without feasibility approval.

The responses to be given by the company for 8 risks out of the 12 identified risks and proposals regarding the controls are summarized above. For the remaining 4 risks, the company is not considering any control; these risks are (risks 5, 6, 7 and 10): Ineffective Inventory Control, Deficiency in Alternative Product Analysis, Long Purchase Processes, Internal Audit Risks. Due to the fact that the probability and possible impacts of the risks will be low, these risks are among the acceptable risks. It is expressed as the residual risk in the literature.

## 5. CONCLUSION AND DISCUSSION

We can also consider the old quotation of "if you can't measure it, you can't manage it" attributed to Peter Drucker as a principle for license inventory management [25]. It is the ability to measure, the precondition for the discipline of software products tracking purchased by companies for millions of liras. Only then can forwardlooking strategic decisions be made easily. Otherwise, it will continue to be the bleeding wound of companies as an uncertainty.

The aim of this study is to reveal the risks or opportunities that companies face while managing their IT software license assets and the measures to be taken against these risks with an enterprise risk management approach. While revealing these risks, one of the largest companies in terms of sample size was chosen as a pilot. The results are expected to be a guide for companies using software. The emergent results as a recipe are particularly advisory for companies managing complex and diverse software assets in their inventory. Moreover, it is thought to make a

contribution to the literature in terms of demonstrating whether the application of enterprise risk management adds value to companies.

Among the possible gains to be achieved with the study, the following can be listed: prevention of duplicate software procurement, optimization opportunities that may arise as a result of reducing maintenance costs by tracking the usage of purchased software.

As proposals, the following can be stated: to be prepared against vendor audits (external audit) while managing software assets, to protect against both reputational loss and business continuity risks thanks to the detection of old version software that may cause security vulnerabilities.

Companies need to create their own software inventories as a priority step. When starting the inventory conduct, taking a certain starting date as a baseline will make the process run more effectively. Otherwise, searching for very old-dated contracts will lead to both a waste of time and will not make any additional contribution. Also, it may not always be possible to access very old-dated contracts and information. There may be no detailed information in the contract documents, the contracts may be missing in the archives, the responsible persons may have quitted the job or the owner of the vendor may have changed.

Following a software inventory conduct based on a certain date, priority should be given to the software licenses of the most used and risky vendors. Building a SAM system according to the size and complexity of the companies is recommended.

In future academic studies, new risks can be added to the aforementioned risks, and the responses to be given according to the risk appetite, strategies and controls of the companies can be changed. Moreover, the maturity levels of companies regarding their information technology units in our country in their license inventory management can be questioned. The following issues can be examined: How many companies have set up and operated the SAM system and how many companies have not centrally managed any software license inventory.

## REFERENCES

- [1] H. Aksu, Dijitopya Dijital Dönüşüm Yolculuk Rehberi, Pusula 20 Teknoloji ve Yayıncılık, İstanbul, 2018.
- [2] Internet: Presidency of the Republic of Turkey Digital Transformation Office, Information and Communication Security Guide, <https://cbddo.gov.tr/bgrehber>
- [3] A. Durdu, A. Eren, (2021) "ISO 27001 Bilgi Güvenliği Yönetim Sistemi Yazılım Tasarımı", *Bilişim Teknolojileri Dergisi*, 14 (3), 255-266, 2021.
- [4] Internet: Wikipedia. Software Licence, [https://en.wikipedia.org/wiki/Software\\_license](https://en.wikipedia.org/wiki/Software_license)

- [5] Internet: Gartner, Newsroom, <https://www.gartner.com/en/newsroom/press-releases/2020-07-13-gartner-says-worldwide-it-spending-to-decline-7-point-3-percent-in-2020>
- [6] Internet: Aspera Training Catalog, Compiled from Aspera Company's End User Training Catalog, <https://learning-aspera.usu.group>
- [7] Ş. Güneş, **Enterprise Risk Management and A Survey Study Related ERM Awareness In Turkey**, Master Thesis, İstanbul Technical University, Institute of Science, 2009.
- [8] Z. Şenol, Z. **The effect of enterprise risk management on firm performance: A Case Study on Turkey**, Ph. D. Thesis, Gaziosmanpaşa University, Social Science Institution, 2016.
- [9] M. K. McShane, A. Nair, E. Rustambekov, "Does Enterprise Risk Management Increase Firm Value?", *Journal of Accounting, Auditing & Finance*, 26(4), 641-658, 2011.
- [10] Internet: Google, Google Books Ngram Viewer. [https://books.google.com/ngrams/graph?content=risk+management%2C+internal+audit%2C+internal+control&year\\_start=1920&year\\_end=2000&corpus=15&smoothing=3&share=&direct\\_url=t1%3B%2Crisk%20management%3B%2Cc0%3B.t1%3B%2Cinternal%20audit%3B%2Cc0%3B.t1%3B%2Cinternal%20control%3B%2Cc0](https://books.google.com/ngrams/graph?content=risk+management%2C+internal+audit%2C+internal+control&year_start=1920&year_end=2000&corpus=15&smoothing=3&share=&direct_url=t1%3B%2Crisk%20management%3B%2Cc0%3B.t1%3B%2Cinternal%20audit%3B%2Cc0%3B.t1%3B%2Cinternal%20control%3B%2Cc0)
- [11] H. Kırıl, "Risks of Enterprise Risk Management", *Denetim*, (18), 5-14, 2018.
- [12] H. Kırıl, İç Denetimin Kurumsal Risk Yönetimindeki Rolü, H. Kırıl, (Ed.), İç Denetim "Yönetime Değer Katmak", Ankara: İç Denetim Koordinasyon Kurulu Yayınları No:1, 317-332, 2014.
- [13] M. M. Şener, **A Research On The Determinants of Implementation Level of Enterprise Risk Management In Organizations**, Master Thesis, Bursa Uludağ University, Social Science Institution, 2019.
- [14] Ş. Güneş, S. Teker, "Enterprise Risk Management Awareness In Turkish Energy Sector", *Doğuş University Journal*, 11(1), 64-76, 2010.
- [15] H. R. Yazgan, P. Sönmez, "Risk Management of Software Development Project: An Example of A Bank ATM Project", *Ege Academic Review*, 15(1), 111-125, 2015.
- [16] A. B. Olcaysoy, **Usage of Intelligent Methods In Cost and Risk Management Software Projects**, Ph. D. Thesis, Yıldız Technical University, Institute of Science, 2016.
- [17] M. H. Calp, **Artificial Intelligence Based Risk Management For Software Projects**, Ph. D. Thesis, Gazi University. Informatics Institute, 2017.
- [18] Y. D. Ö. Özen, A. Koçak, "Selection and Evaluation of Enterprise Resource Planning Software by Using Fuzzy Analytical Hierarchy Process and Fuzzy Dematel", *Journal of Management & Economics*, 24(3), 929-957, 2017.
- [19] J. B. Gugler, Procurement and Asset Management of Commercial-Off-The-Shelf Software, Master Thesis, University of Twente School of Management and Governance Chair of Technology Management, 2018.
- [20] V. P. Peltonen, **Software Asset Management, Current State and Use Cases**, Master Thesis, Helsinki Metropolia University of Applied Sciences, Master of Business Administration, 2015.

- [21] Internet: Minister of Treasury and Finance, Public Internal Audit Manual, <https://ms.hmb.gov.tr/uploads/2019/06/4046elkitabipdf.pdf>
- [22] J. Sim, J. Waterfield, “Focus group methodology: some ethical challenges”, *Qual Quant* 53, 3003–3022, 2019.
- [23] Internet: H. Marquis, G. Spivak, B. Victoria, Cut Software Spending Safely with SAM, <http://gartner.com/home> ID: G00301780.
- [24] Flexera, Software Asset, and Licence Management Best Practice, Guidelines for IT Management, Number 340.
- [25] Ö, Yanar, “Entelektüel Sermaye ve Örgütlerin Performansı Üzerine Etkisi”, Maltepe University Faculty of Economics and Administrative Sciences, *Journal of Economic, Social and Political Analysis*, 2013(2), 79-96, 2013.



## Annex 1. Questions

1	Licenses
2	Product/license description
3	Application/system name where the license is used
4	License model (category)
5	License model (sub-category)
6	Declared license model
7	Current quantity
8	Activation date
9	Expiration date
10	Contact person
11	Department name
12	Directorate name
13	Cost center
14	Supplier name
15	Manufacturer name
16	Purchase order number
17	Price per unit
18	Renewal price
19	Remarks
20	The licenses are managed?
21	Is there any unlicensed position?
22	The license can be replaced?
23	What are the alternatives?
24	Is the meeting completed?
25	Operation contact

Annex 2. License models

Category	Category Description	Sub Category
<b>User</b>	Licensing is based on users directly using the application	User
		Non Employee User
		UVU (User Value Unit)
		VUH (Virtual User Hour)
		Active User
		Concurrent User
		Active Concurrent User
		Authorized User
		Profile-based User
		Named User
		Virtual User
		Floating User
		Admin User
		GUI User
		Retail User
		Professional User
		Limited Professional User
		Employee Self-Service User
		Manager Self-Service User
		Employee User
		Developer User
		Expert User
		Business Information User
		Process Control User
		Partner User
		MSS User
		ESS User
		Agent
		Concurrent Agent
		Seat
		Site
		<b>CPU</b>
Processor		
Socket		
Core		
PVU (Processor Value Unit)		

		RVU (Resource Value Unit)
<b>Device</b>	Licensing is based on using all kinds of device and device concepts (port, IP, instance)	Device
		Device (Concurrent)
		Network Device
		Mobile Device
		Agent
		Disk Drive
		Probe
		Point
		Instance
		Operating System Instance
		Home Gateway
		Rack
		Cabin
		Station (Concurrent)
		Phone
		Mobile Phone
		Server
		Application Server
		Exchange Server
		Virtual Server
		Database Server
		Client
		PC
		IP
		Port
		Node
		Platform
		Host
		Hardware
		RAM
		Named Host
		Environment
Terminal		
Firewall		
Access Point		
Active Concurrent Access Point		
Router		
Switch		
Installation		
<b>Transaction / Traffic</b>	Licensing is based on using metrics	TPS (Transaction Per Second)

	containing transaction/traffic information of the service provided by the application	MBPS (Mega Bit Per Second)		
		GBPS (Gigabit Per Second)		
		Gigabyte Per Day		
		(Inter)Connection Per Day		
		Call		
		Conference		
		Concurrent Call		
		Concurrent Conference		
		Call per Month		
		Call per Second		
		EPS (Event Per Second)		
		MSU (Message Signal Unit)		
		MPS (Message Per Second)		
		Query per Month		
		MMS per Day		
		CDR per Day		
		Concurrent Session		
		Log per Day		
		CAL (Client Access License)		
		<b>Capacity / Activity</b>	Licensing is based on using metrics containing transaction/traffic information of the service provided by the application	Subscribers
				Mobile Subscribers
				Prepaid Subscribers
				Gigabyte
				Concurrent Video Stream
				VoD (Video on Demand) Content
				Concurrent VoD (Video on Demand)
Connection				
Concurrent Connection				
Process				
Scenario				
Channel				
Object				
Data Objects				
Master Data Objects				
HAVE				
Payroll				
Engine				
Document				

<b>Indirect</b>	Licensing based on indicators not related to the service provided by the application	Personnel
		Mailbox
		Subscription
		Application
		Customers
		Mobile Customers
		All Subscribers
		All Mobile Subscribers
		Budget
		Employee
		Full-Time Equivalent
		Spend Volume
		Flat Fee
		Revenue
<b>Freeware</b>	Situations that do not require any payment	Freeware
		Inhouse
		Bundle

# Yapay Zekâ Tabanlı Sistemlerin Kişisel Veri Mahremiyeti Üzerine Etkisi: Sohbet Robotları Üzerine İnceleme

*Araştırma Makalesi/Research Article*

 Fatma BAŞKAYA\*,  Hacer KARACAN

Bilgisayar Mühendisliği, Gazi Üniversitesi, Ankara, Türkiye  
[fatma.baskaya@gazi.edu.tr](mailto:fatma.baskaya@gazi.edu.tr), [hkaracan@gazi.edu.tr](mailto:hkaracan@gazi.edu.tr)  
(Geliş/Received:07.01.2022; Kabul/Accepted:24.10.2022)  
DOI: 10.17671/gazibtd.1053803

**Özet**—Günümüz gelişen teknolojiyle birlikte yapay zekâ kullanımı artmıştır. Bu durumun çoğunlukla faydaları üzerine konuşulsa da kişisel veri mahremiyetine ve güvenliğine olumsuz etkisi göz ardı edilmemelidir. Çünkü yapay zekâ için ham madde olan veriler, kişilerin mahremiyet haklarına aykırı durumlara neden olabilmektedir. Bu doğrultuda ilgili çalışmada; genel olarak yapay zekânın kişisel veri mahremiyeti açısından oluşturduğu risklere, bu risklerin giderilmesi için önerilen bazı yöntemlere ve bu yöntemlerden biri olan federe öğrenme metodu ile ilgili detaylara yer verilmiştir. Ayrıca günümüzde çeşitli sektörlerde ait uygulamalarda sıklıkla kullanılan yapay zekâ ürünü, sohbet robotları (chatbot) ile ilgili bir inceleme ve anket çalışması yapılmıştır. Bu anlamda; sohbet robotlarının kullanımı ve yapısının kişisel veri güvenliğine uygunluğu ile ilgili detaylara ve önerilere yer verilirken anket çalışmasında ise kişilere yöneltilen dört adet soru ile kişilerin, sohbet robotu kullanma boyutunu ve sohbet robotu kullanımının kişisel verilere etkisine yönelik farkındalığını ölçmek amaçlanmıştır. Anket sonuçları, sohbet robotlarının en fazla bankacılık işlemleri için kullanıldığını ve bireylerin %62,1'inin ise bu servislerin kişisel verileri kaydedip işleyebileceğini bilmediğini göstermiştir.

**Anahtar Kelimeler**— yapay zekâ, kişisel veri güvenliği, KVKK, GDPR, federe öğrenme, sohbet robotları.

## The Impact of Artificial Intelligence-Based Systems on Personal Data Privacy: A Study on Chatbots

**Abstract**— Use of artificial intelligence has increased with the developments in technology. Although the benefits of this situation have mostly been discussed, its negative impacts on personal data privacy and security shouldn't be overlooked. Artificial intelligence-based systems can cause antithetical situations in terms of the privacy rights of individuals because of the fact that data is the raw material for artificial intelligence. Accordingly, the risks posed by artificial intelligence, some methods recommended to overcome these risks, and one of these methods, federated learning, were covered in the current study. Moreover, a study and survey were conducted on chatbots which are artificial intelligence products and are frequently used in applications applied on various sectors. In this sense, some details and recommendations regarding the use of chatbots and the compatibility of their structure with personal data security were included. In the survey with four questions, it was aimed to understand the extent of using chatbots and the awareness on the effects of chatbot using personal data. The survey results showed that chatbots are mostly used for banking transactions and 62.1% of individuals do not know that these services can save and process personal data.

**Keywords**— artificial intelligence, data privacy, KVKK, GDPR, federated learning, chatbots.

### 1. GİRİŞ (INTRODUCTION)

Günümüzün teknolojik gelişimiyle birçok alanda yer bulan yapay zekânın en önemli faydalarından birkaçı; verileri,

insan gücünden daha kısa sürede daha hızlı analiz etmeleri ve bu verileri işlerken öğrenme sürecine girmeleridir. Tüm faydaları düşünüldüğünde; son yıllarda yapay zekâ ve yapay zekânın alt dallarına ait kollarda, literatüre yeni

algoritmalar ve yöntemler kazandırılmış ve kazandırılmaya devam etmektedir.

Bu gelişmeler yapay zekâ dünyası için yeni öngörülerin ve yeni kapıların açılmasına olanak sağlamaktadır. Yapay zekâ uygulamalarına; tahmin, ses ve/veya görüntü tanıma, öneri ve tespit sistemleri, sohbet robotları, vb. gibi birçok uygulama ve sistem örnek olarak verilebilir [1].

Yapay zekânın hammaddesini oluşturan kaynak büyük veridir ve yapay zekâ teknolojisini kullanan sistemlerde, toplanan büyük miktarda verinin işlenmesi durumu söz konusudur [2]. Bu anlamda, kişisel veriler ve yapay zekâ iki taraflı ve birbirini besleyen bir ilişki içindedir. Bir yandan kişisel veriler yapay zekâ teknolojilerini beslerken öte yandan yapay zekâ teknolojileri de çıkarımlar yaparak daha fazla kişisel veri üretir [3].

Büyük veri ihtiyacı karşılanırken toplanan verilerin mahremiyeti, dikkat edilmesi gereken önemli konulardan biridir [1]. Bu durum, yapay zekânın sahip olduğu faydalar yanında çeşitli riskleri de barındırdığı anlamına gelmektedir. Örneğin; hastalık teşhisi için kullanılan yapay zekâ sistemlerinde, veri olarak kullanılan kan testlerinin işlenmesi kişisel veri mahremiyeti ilkesine aykırı durumlara sebebiyet verebilmektedir. Öte yandan, doktorların elinde bulunacak büyük miktardaki örnek test, derin öğrenme yöntemleri ile hastalık teşhisini kolaylaştırabilmektedir [1]. Anlaşıldığı üzere yapay zekâ, bahsedilen riskler düşünüldüğünde vazgeçilen bir alan olmamalı; tam aksine gerekli ve doğru bir şekilde kullanılarak faydasından en doğru şekilde istifade edilen bir alan olmalıdır.

Tüm bu açıklamalardan yola çıkarak; bu çalışmada, genel olarak yapay zekâ tabanlı sistemlerin kişisel veri güvenliğine olan olumsuz etkileri, kullanıcıların karşı karşıya kalabileceği riskler ve bu risklerin nasıl aşılabileceğine yönelik öneriler anlatılmıştır.

Belirtilen amaç ile ilgili yapılan literatür taramasında, yapay zekâ tabanlı sistemlerin hangi önlemler alınarak tasarlanması gerektiği ve böylece kişisel veri güvenliğinin nasıl sağlanabileceği ile ilgili hem akademik hem de ulusal ve uluslararası otoriteler tarafından yapılan çalışmalar incelenmiştir. Bu bağlamda çalışmada; Genel Veri Koruma Yönetmeliği'ne (*General Data Protection Regulation - GDPR*) ve 6698 sayılı Kişisel Verileri Koruma Kanunu'na (KVKK) da yer verilmiştir.

İncelenen çalışmalar neticesinde, gizlilik prensibine uygun sistemler tasarlamaya imkân sağlayan *Federe Öğrenme* yönteminin son birkaç yıldır literatürde yer etmeye başladığı görülmüştür. Yapay zekâ tabanlı sistemlerde, federe öğrenme yöntemi kullanarak kişisel verilere saygılı tasarımlar yapmanın mümkün olabileceği anlatılmıştır. Ayrıca yine yapay zekâ tabanlı bir sistem olan ve günümüzde özellikle elektronik ticaret sektörü başta olmak üzere birçok sektörde sıklıkla kullanılan sohbet robotları (chatbot) hakkında inceleme çalışması yapılmış ve bu

çalışmaya ait detaylara yer verilmiştir. Kişisel verileri koruma hakkında yayımlanan kanun ve yönergelere uygun sohbet robotları tasarlamak için hangi şartların sağlanması gerektiği anlatılmıştır.

Çalışma kapsamında; literatür taramasına, öncelikli olarak insanların kişisel veri mahremiyeti farkındalığına ve bilincine ne düzeyde sahip olduğunu araştıran bir çalışma incelenerek başlanmıştır. Bu kapsamda; Eroğlu [4] tarafından yapılan çalışmada, günlük hayatlarında çoğunlukla dijital ortam ile etkileşim içinde olan öğrencilerin farkındalığı analiz edilmiştir. Bunun için 34 sorudan oluşan web tabanlı bir anket yapılarak 280 öğrenciye gönderilmiş ve ilgili anket 151 öğrenci tarafından yanıtlanmıştır.

Çalışma sonucunda ise öğrencilerin 72'si kişisel veriyi bildiğini, 19'u ise kişisel veri kavramını bilmediğini belirtmiştir. Ayrıca, öğrencilerin çoğunlukla kimlik numarasını kişisel veri olarak nitelendirdiği görülmüştür. Bunun dışında; kredi kartı bilgileri, ad, soyadı ve doğum tarihi verilerini içeren kimlik bilgileri ve ev adresi verilerinin de öğrenciler tarafından kişisel veri olarak nitelendirildiği görülmüştür.

Kişisel veri ve yapay zekâ birlikteliğine gelindiğinde ise Mitrout yaptığı çalışmada [3], yapay zekâ ve makine öğrenmesi kavramlarına yer vermiş ve bu kavramların kişisel veriler üstündeki etkilerine değinmiştir. Mitrout, kişisel veri ve yapay zekâ arasında çift taraflı bir yol olduğunu, kişisel verilerin yapay zekâyı beslediğini ve buna karşılık yapay zekânın da daha fazla kişisel veri ürettiğini belirtmiştir.

Kamarinou ve arkadaşları [5], GDPR'a göre veri sahiplerinin profillenmesi için makine öğrenimi algoritmalarının kullanılmasının etkilerinin bir analizini sunmuşlardır. Veri koruma hak ve yükümlülüklerini, makine öğrenmesi sistemlerinin geliştirilmesinin sonuçlarını ve kişisel verilerin nasıl toplandığını ve kullanıldığını araştırmışlardır.

Li ve arkadaşları [6], endüstriyel alanlar ve bilişim dünyası uygulamaları açısından federe öğrenmeyi incelemiş ve mobil cihazlar, sağlık uygulamaları, sektör uygulamaları alanında federe öğrenme yöntemini kullanan uygulama örnekleri ile ilgili bir araştırma çalışması yapmışlardır.

Chandiramani ve arkadaşları [7], 2016 yılında Google tarafından önerilen federe öğrenme modeli üzerine bir çalışma yapmıştır. Çalışmada; klasik makine öğrenmesi, dağınık öğrenme modeli ve federe öğrenme yaklaşımları üzerine analiz ve karşılaştırmalar yapılmıştır. Üç modelin doğruluk oranlarında önemli bir fark olmadığı ancak modellerin oluşturulması için geçen sürede fark olduğu görülmüştür. Temel makine öğrenimi sınıflandırıcısının eğitilmesi 25,4 saniye sürerken federe öğrenme sınıflandırıcısının 16,7 saniye; dağıtılmış makine öğrenimi sınıflandırıcısının ise 14,1 saniye sürdüğü gözlemlenmiştir. Böylece en kısa sürede federe öğrenme modeli ile erişildiği

gözlemlenmiş ve fazla veriyle çalışırken verileri dağıtmanın ve küçük görevlere ayırmanın performans artırımı için bir avantaj olduğu görülmüştür. Bu nedenle; üç model arasından federe öğrenmenin, modelleri eğitirken gizliliği korumak için en iyi model olduğu anlaşılmıştır.

Wang ve arkadaşları [8]; kullanıcı görüşlerinin, duygularının ve e-postaların yer aldığı üç farklı hazır veri seti üzerinde metin madenciliği yöntemini uygulamıştır. Model, federe öğrenme yöntemi ile birleştirilen Gizli Dirichlet Ayırımı (Latent Dirichlet Allocation) algoritması kullanılarak eğitilmiş ve kullanıcıların veri gizliliğinin korunduğu görülmüştür.

Ülkemizde de son yıllarda özellikle 6698 sayılı KVKK ile birlikte mahremiyet kavramına verilen önem artmış ve bu alanda yapılan çalışmalara büyük ilgi olduğu gözlemlenmiştir. Süzen ve Kayaalp [1] yaptıkları çalışmada; klasik makine öğrenmesi uygulamaları için federe öğrenme yöntemini incelemiş ve federe öğrenmenin gizlilik açısından faydalı olsa da problemlere uygulanması sırasında karşılaşılabilecek zorluklarının olduğunu belirtmişlerdir. Bununla birlikte yeni algoritmalar ve bu alandaki araştırma çalışmalarının artması ile bu zorlukların aşılabileceği ve gizlilik ve güvenlik ilkelerine uygun sistemler geliştirilebileceği belirtilmiştir.

Süzen ve Şimşek [9], sağlık alanında yapılan uygulamalar için makine öğrenmesi modellerinin uygulanışını incelemişlerdir. Sağlık verilerinin işlenmesinde, gizlilik ilkesine uyumu yakalamak adına model içine federe öğrenme yöntemlerinin entegre edilmesi gerektiğini belirtmişlerdir.

Sağiroğlu ve Canbay [10] yapmış oldukları çalışmada, büyük veride yapay zekâ teknolojilerini kullanırken mahremiyet ihlallerinin önüne geçmek amacıyla yeterli önlemlerin alınması gerektiğini belirtmişlerdir. Bunun için, KVKK ve GDPR gibi yasal mevzuatlara uygunluğun önemli olduğunu belirtmişler ve mahremiyeti yakalamak için kullanılan anonimleştirme konusunun üstünde durmuşlardır.

Tüm bu çalışmaların yanı sıra Türkiye’de yapay zekâ ve kişisel veri mahremiyeti konusunun kamu idareleri tarafından da önemli ölçüde dikkate alındığını gösteren örneklerle karşılaşmak mümkündür. Türkiye Cumhuriyeti Cumhurbaşkanlığı Dijital Dönüşüm Ofisi tarafından yapılan “AçıkVeri” projesinde [11], federe öğrenme yöntemi gibi yenilikçi yöntemler ile gizliliği korunan kişisel veriler konusuna vurgu yapılmıştır. Ayrıca, Kişisel Verileri Koruma Kurumu tarafından Eylül 2021 tarihinde “Yapay Zekâ Alanında Kişisel Verilerin Korunmasına Dair Tavsiyeler” [12] isimli rehber çalışması yayımlanmış ve bu çalışmada önemli tavsiyelere yer verilmiştir. Bu tavsiyeler; yürürlükte bulunan yasal mevzuatlar doğrultusunda, yapay zekâ sistemini geliştiren, üreten, bu sisteme servis sağlayan gerçek veya tüzel kişiler ile karar alıcıları kapsamıştır. Bu duruma yönelik benzeri önerilerin

bulunduğu çalışmalar her geçen gün artmaktadır. Hem akademik düzeyde hem de kamu eli ile desteklenen bu alan, yenilikçi yöntemlerin keşfi ile gelecek vadeden alanlardan biri haline gelmiştir.

## 2. YAPAY ZEKÂ VE KİŞİSEL VERİ MAHREMİYETİ (ARTIFICIAL INTELLIGENCE AND PERSONAL DATA PRIVACY)

Bu bölümde; kişisel veri kavramının ne olduğu ve hangi türlerden oluştuğu, yapay zekânın kişisel verilerle olan ilişkisi ve bu sistemlerin veri toplama ve işleme faaliyetlerinde kişisel veri güvenliği açısından ne tür riskler oluşturabileceği anlatılmıştır.

Kişisel veri kavramı, 6698 sayılı KVKK Madde-3/1’de “*Kimliği belirli veya belirlenebilir gerçek kişiye ilişkin her türlü bilgi...*” [13] olarak tanımlanmıştır. Bu tanımdan hareketle kişilere ait, ad, soyad, telefon numarası, adres, banka hesap numarası, resim, vb. gibi kişiyi tanımlayabilen ya da kişi ile ilişkilendirilebilen [14] tüm verilerin kişisel veri olduğunu söylemek mümkündür. Tüm bunların yanı sıra, KVKK’da “*Özel Nitelik Kişisel Veri*”; GDPR’de ise “*Hassas Veri*” olarak ifade edilmiş olan veriler bulunmaktadır. Bu veriler, bir başkası tarafından öğrenildiğinde kişinin mağduriyet yaşayabileceği veya ayrımcılığa maruz kalabileceği [14] durumlara neden olabilecek verilerden oluşmaktadır. KVKK Madde-6/1’de özel nitelikli kişisel veriler, “*Kişilerin ırkı, etnik kökeni, siyasi düşüncesi, felsefi inancı, dini, mezhebi veya diğer inançları, kılık ve kıyafeti, dernek, vakıf ya da sendika üyeliği, sağlığı, cinsel hayatı, ceza mahkûmiyeti ve güvenlik tedbirleriyle ilgili verileri ile biyometrik ve genetik verileri...*” olarak ifade edilmiş ve kanunda bu verilerin işleme şartları ayrı bir madde ile (Madde-6) ele alınmıştır.

Gelişen dünya ile yapay zekâ gibi veriye ihtiyaç duyan ve veri ile beslenen teknolojiler ortaya çıkmıştır. Yapay zekâ, makinelerle insansı özellikler kazandırarak karmaşık problemlere çözüm üretmeyi amaçlayan önemli konulardan biridir. Bir insanın karar verebilme gücünü taklit etmeyi amaçlayan bu sistemler, çeşitli algoritmalar aracılığı ile modellenir. Böylece, insan gücü ile yapılacak işler daha kısa sürede daha başarılı bir şekilde çözüme ulaşır [15].

Yapay zekâ teknolojisinin günümüz kullanım alanlarına bakıldığında; yapay zekânın sağlıktan elektronik ticarete, ulaşımdan pazarlamaya değin birçok alanda kullanıldığını görmek mümkündür. Hastalıkları, bilişim sistemlerine yapılan saldırıları [15], seri üretim bandı hatalarını, dolandırıcılığı tespit edebilen sistemler ile kişisel asistanlar, suç analizi, duygu analizi, fikir tespiti, vb. gibi çalışma alanları, yapay zekâ konusunun birer ürünüdür. Yapay zekâ teknolojisi ise tüm bu işlemler için yüksek miktarda veriyi işlemekte ve kendini eğitmektedir. Görüldüğü üzere yapay zekânın hemen her alanda hayatımızın içinde oluşu verilerin de aynı oranda yaygın bir şekilde işlenmesi sonucuna neden olmaktadır.

Veri ve yapay zekâ ise birbirinden beslenen kavramlardır. Bu durumu somutlaştırmak için gündelik hayatta sıklıkla karşılaşılan durumlardan örnek vermek gerekirse; arama motorunda daha önce aradığımız kelimelerle ilgili olan reklamların daha sonra karşımıza çıkarılması, aynı şekilde önceden izlediğimiz bir videoyla benzer konuları barındıran videoların öneri listemize gelmesi veya sosyal medya platformlarının arkadaş listemiz veya beğendiğimiz sayfalardan yola çıkarak kişi veya sayfa önermesi yapay zekânın verileri işleyerek ürettiği durumlardan sadece birkaçıdır [2]. Kısacası; çözmek istenilen probleme yönelik uygun veriye sahip olan sistemler, akıllı yaklaşımlar üretebilmektedir. Yapay zekâ için bir nevi besin olan bu veriler, her alanda her türlü sektör veya kurumdan elde edilebilmektedir.

Yapay zekânın birçok probleme başarılı çözümler üretmesinin yanı sıra tüm bu işlemleri gerçekleştirirken kaynağı olan veriyi ne şekilde aldığı, nasıl kullandığı ve koruduğu soruları da üzerinde düşünülmesi gereken önemli konulardandır. Tüm bunlar son yıllarda önem verilen konular olsa da maalesef yapay zekânın faydalarının yanında görece göz ardı edilmektedir. Bu durum ise zamanla yapay zekâyı; kişisel verileri, kanunlara aykırı bir şekilde işleme potansiyeli olan platformlar haline getirmektedir. Kişisel verileri kanunlara aykırı toplayan, işleyen ve saklayan uygulamalar; kötü niyetli kişilerce veri sahiplerini zor durumda bırakacak şekilde kullanılabilirlerdir.

Dolayısıyla bir yapay zekâ sisteminin veri işleme görevini yerine getirirken veri mahremiyeti ilkesini gözetmesi için bazı zorunluluklara uyması gerekmektedir. Aksi halde işlenen veriler; veri sahibi konumundaki kişi veya kurumları, kişisel verilerin ifşa edilmesi, yetkisiz üçüncü kişilere aktarılması, fide amacıyla kullanılması ve daha birçok şekilde zor duruma düşürebilmektedir.

### 3. ÖNERİLEN YÖNTEMLER (PROPOSED METHODS)

Yapay zekâ veriden bağımsız düşünülemez. Özellikle günümüz gelişen teknolojiyle adeta yenedünya düzeni oluşturan teknolojiler, hayatın her alanına ve her sektöre girmişken bundan kaçmak imkânsız ve yanlıştır. O halde, yapılması gereken yapay zekâdan kaçmak değil; yapay zekâyı en etkili şekilde kullanıp hizmetlerinden en maksimum faydayı elde etmek ve ancak bunu gerçekleştirirken veri koruma ilkesine sadık kalarak kişisel verilerin korunmasını sağlamaktır.

Kişisel verilerin otomatik işlenmesi ile ilgili çalışmalar yapan Avrupa Konseyi, 108 numaralı Kişisel Verilerin Otomatik İşleme Danışma Komitesi (*Consultative Committee of The Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*) tarafından Ocak 2019'da hazırlanan rehber çalışmasında [16]; yapay zekâ teknolojisine dayanan sistemlerde, kişisel veri işleme faaliyetinde bulunurken hangi hususların göz önünde bulundurulması gerektiği ile ilgili bilgilere yer verilmiş ve aşağıda belirtilen hususlara değinilmiştir:

- Verilerin işlenmesi ve teknolojinin gelişimi aşamasında veri koruma ilkeleri esas alınmalıdır. Ayrıca, ilgili kişi hakları gözetilmelidir.
- Veri işleme faaliyeti sırasında gerekli risk analizleri yapılmalı, bu analizler yapay zekâ teknolojisinin özellikleri göz önüne alınarak hassas bir şekilde yapılmalıdır.
- İlgili kişilere, açık, doğru ve anlaşılır bir şekilde gerekli bilgilendirmeler yapılmalıdır.
- Veri minimizasyonu ilkesi gözetilmeli, gereksiz ve aşırı veri işleme faaliyetinden kaçınılmalıdır.
- Amaç dışı veri işleme faaliyetinde bulunulmamalıdır.

Bahsi geçen bazı temel veri işleme politikaları GDPR Madde-5 ve 6698 sayılı KVKK Madde-4 ile düzenlenmiştir. KVKK Madde-4'te yer alan hükümler aşağıda bulunmaktadır [13]:

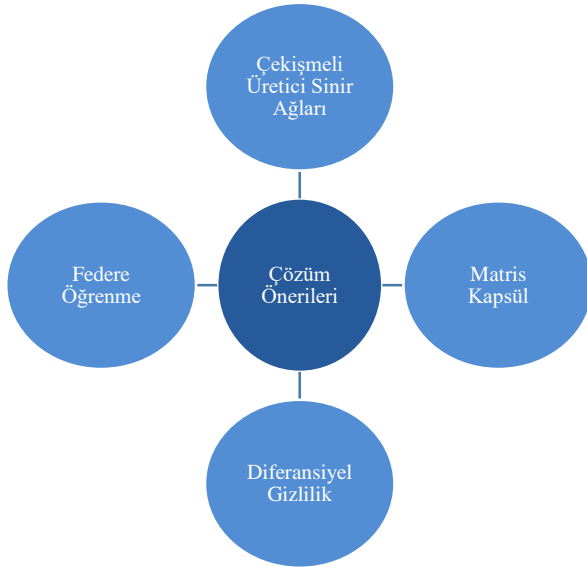
“...(1) Kişisel veriler, ancak bu Kanunda ve diğer kanunlarda öngörülen usul ve esaslara uygun olarak işlenebilir. (2) Kişisel verilerin işlenmesinde aşağıdaki ilkelere uyulması zorunludur:

- a) Hukuka ve dürüstlük kurallarına uygun olma.
- b) Doğru ve gerektiğinde güncel olma.
- c) Belirli, açık ve meşru amaçlar için işlenme.
- ç) İşlendikleri amaçla bağlantılı, sınırlı ve ölçülü olma.
- d) İlgili mevzuatta öngörülen veya işlendikleri amaç için gerekli olan süre kadar muhafaza edilme...”

Bu hükümler içinde belirtilen amaçla sınırlı olma ilkesi, kişisel verilerin belirli, açık ve meşru amaçlarla toplanması gerektiğini; veri minimizasyonu prensibi ise kişisel verilerin işleme için gerekli olduğu kadar, amaçla ilgili ve ölçülü bir biçimde işlenmesi gerektiğini ifade etmektedir [17]. Bu duruma örnek vermek gerekirse; seyahat için bilet alınmak istendiğinde, kişiden kan grubu bilgisinin istenmesi verinin amaçla ilgili olmadığı anlamına gelmektedir.

Norveç Veri Koruma Otoritesi tarafından Ocak 2018'de hazırlanan yapay zekâ ve gizlilik raporunda [18], GDPR'de yapay zekâ ile veri işlemenin hususlarına değinilmiştir. Temel olarak yasallık, şeffaflık, amaçla sınırlı kalma, veri minimizasyonu, amaca bağlılık ilkelerinden söz edilmiştir. Bunun yanı sıra veri sorumlularının, ilgili kişi konumundaki veri sahiplerine karşı hesap verilebilirlik ilkesi gereğince hareket etmeleri gerektiği belirtilmiştir. Bu bilgilendirmeler ışığında; bahsi geçen raporda ilgili otorite, kişisel veri mahremiyetine önem veren yapay zekâ tabanlı sistemlerin teknik alt yapıları hususunda bazı çözüm önerilerinden bahsetmiştir. Bu çözüm önerileri Şekil 1' de belirtilmiştir.

Belirtilen çözüm önerilerine kısaca değinmek gerekirse, Çekişmeli Üretici Sinir Ağları (Generative Adversarial Networks) modeli, 2014 yılında Google araştırmacısı *Ian Goodfellow* tarafından geliştirilmiştir. Bu modelde, üretici ağ ve ayırt edici ağ olmak üzere iki türlü ağ bulunmaktadır. Üretici ağ gerçeğe yakın veriler üretirken ayırt edici ağ, verinin gerçek mi sahte mi olduğunu anlamaktadır [18].



Şekil 1. Otoritenin çözüm önerileri [18]  
(Solutions of the authority)

Matris kapsül yöntemi ise sistemi eğitmek için kullanılan yapay sinir ağları temelli bir modeldir [18]. Diferansiyel Gizlilik (Differential Privacy) modelinde ise veriye gürültü ekleme işlemi yapılmaktadır [19, 20].

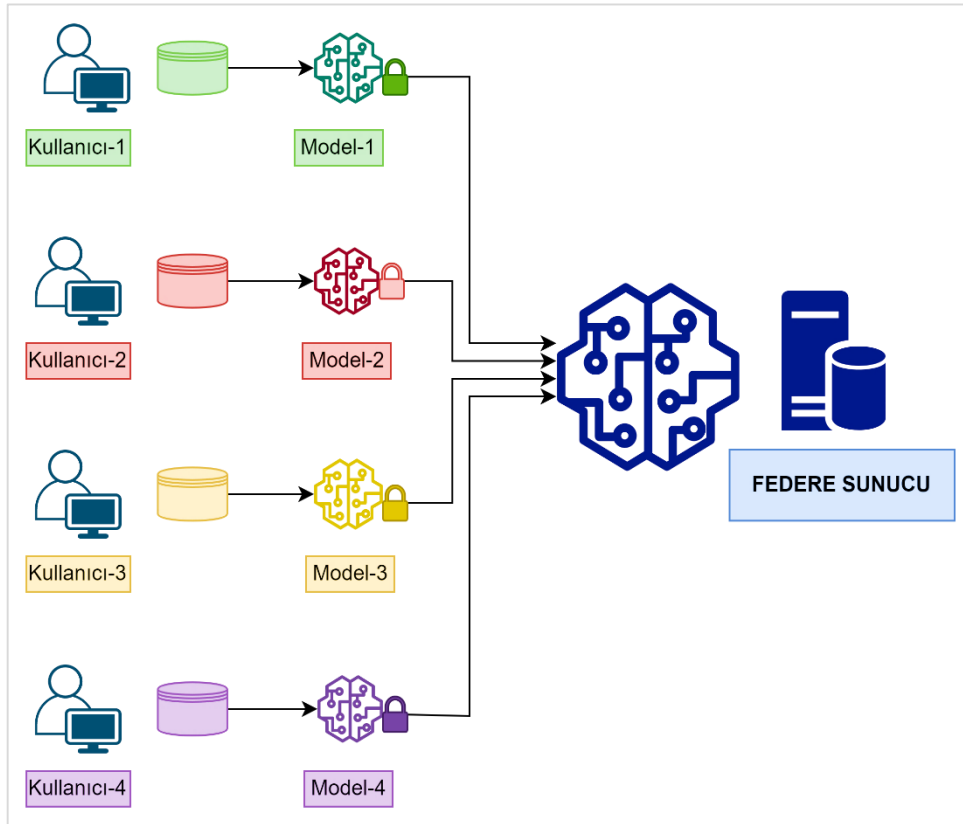
Bu çalışmada, önerilen yöntemlerden biri olan federe öğrenme konusu ile ilgili detaylara yer verilmiştir.

### 3.1. Federe Öğrenme (Federated Learning)

Federe öğrenme, kullanıcıların verileri paylaşmadan sistemi eğittiği bir makine öğrenmesi biçimidir. Bu öğrenme metodunda her kullanıcı kendi verileri ile yerel modelini oluşturur ve tüm model ağırlıkları birleştirilerek federe sunucuda bulunan ana modelin ağırlıkları oluşturulur [21]. Federe öğrenme, makine öğrenmesindeki eğitim sürecinin sistemdeki kullanıcılar arasında paylaşıldığı işbirlikçi bir yöntemdir [18].

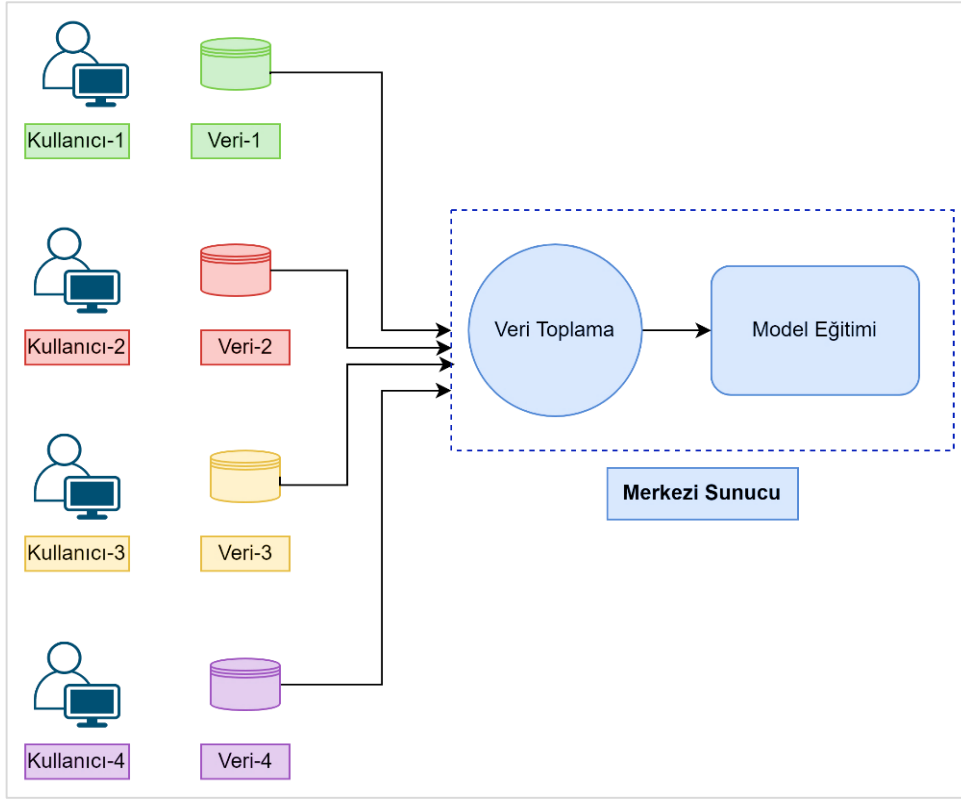
Federe öğrenmede, koordinatör rolüne sahip bir sunucu [22] ve sunucu tarafından belirlenen kullanıcılar bulunmaktadır. Koordinatör rolündeki federe sunucu bu kullanıcıları, kullanıcının erişilebilir olup olmadığı ve model eğitimine izin verip vermediği gibi kısıtları göz önünde bulundurarak belirlemektedir. Belirlenen kullanıcılar verilerini paylaşmadan kendi modellerini oluşturup ağırlıklarını federe sunucuya aktarırlar. Şekil 2'de [7, 18, 23] görüldüğü gibi federe sunucu koordinatör rolü ile tüm kullanıcılardan gelen ağırlıkları birleştirir ve nihai modeli oluşturur [21].

Federe öğrenmede mevcut durumda eğitilmiş bir model bulunuyorsa bu modelin ağırlıkları ve parametreleri, bahsi geçen kriterlere göre seçim yapılarak belirlenen kullanıcılar ile paylaşılır. Her kullanıcı, gelen modeli kendi verisi ile oluşturduğu yerel modelin ağırlıkları ile günceller ve güncellenmiş ağırlıkları federe sunucuya iletir.



Şekil 2. Federe öğrenme [7, 18, 23]  
(Federated learning)





Şekil 3. Klasik makine öğrenmesi modeli [7]  
(General machine learning model)

Klasik makine öğrenmesi modelinde Şekil 3'te görüldüğü üzere kullanıcıların verileri paylaşması söz konusudur. Merkezi sunucu rolündeki koordinatör, kullanıcılardan gelen verileri toplayarak modeli eğitir [7] ve sonrasında geleneksel makine öğrenmesi adımları izlenir. İlk aşamada veri, ön işlemeden geçirilerek yapısal formata dönüştürülür. İkinci aşamada, makine öğrenmesi modeli oluşturulur. Bir sonraki aşamalarda ise model eğitilir ve test edilir.

Federe öğrenmenin kullanıldığı çeşitli alanlara ait uygulama ve sistemler bulunmaktadır. Bunlardan biri olan Google Klavye (Gboard), kullanıcının yazdıklarına erişerek kullanıcı için otomatik metin düzeltme, bir sonraki kelimeyi tahmin etme ve kelimeyi tamamlama gibi özellikler içeren sanal bir klavyedir [24] ve federe öğrenmeyi kullanarak kullanıcı gizliliği önem vermiştir.

Çin'de Google'un yasaklanmasıyla birlikte Baidu şirketi, yapay zekâ ürünlerine dâhili olarak uygulamak amacıyla federe öğrenme tekniğini kullanarak PaddleFL platformunu geliştirmiştir [25].

Bir başka örnek olarak Owkin isimli bir Fransız tıp-biyoloji şirketi hastalık tanıları ile ilgili etkili bir çalışma yapmak amacıyla sağlık kurumları ve ilaç firmalarıyla ortaklık yürüterek federe öğrenmeye dayalı bir makine öğrenmesi modeli oluşturmuştur. Böylece, hasta mahremiyeti korunurken kuruluşlar ve araştırmacılar arasında veri paylaşımı yapılmıştır [1].

NVIDIA şirketi, GDPR'de hassas nitelikli kişisel veri olarak da ifade edilen (6698 sayılı Kişisel Verileri Koruma Kanunu'nda da özel nitelikli kişisel veri) sağlık verilerinin hastane içinde kalmasını sağlayan ClaraFL geliştirmiştir [1]. Bazı üst düzey sağlık hizmeti sunan kurumlarda bulunan radyologlar tarafından kullanılmaya başlanan ClaraFL, önceden eğitilmiş modelleri kullanmaktadır. Öğrenme tekniklerini aktarırken radyologlara etiketleme konusunda yardımcı olmaktadır.

Intel şirketi, Mayıs 2020'de Pensilvanya Üniversitesi ile yaptığı işbirliği neticesinde beyin tümörlerini tespit eden yapay zekâ teknolojisi üzerine çalışmaya başladıklarını açıklamıştır [26]. Geliştirilecek sistemde, 29 tıp merkezinden alınacak hasta verileri için kişisel veri gizliliğini sağlamak adına federe öğrenme yönteminin kullanılacağı belirtilmiştir.

Federe öğrenme, yapay zekâ sistemlerinde kişisel veri mahremiyetini korumak için uygun bir çözümdür ancak yanında getirdiği bazı teknik zorluklar ve riskler bulunmaktadır. İlki, tasarlanacak federe öğrenme sisteminin karmaşıklığıdır. Sistemi iyi tasarlamak ve karmaşıklığı azaltacak çözümler üzerine çalışmak gerekmektedir. Bir diğer zorluk, federe sistem koordinasyonundan sorumlu olan sunuculara iş yükü yüklemektir. Bu durumun sunucuları yoracağı düşünüldüğünden sistemin, bu dezavantaj düşünülerek inşa edilmesi gerekmektedir.

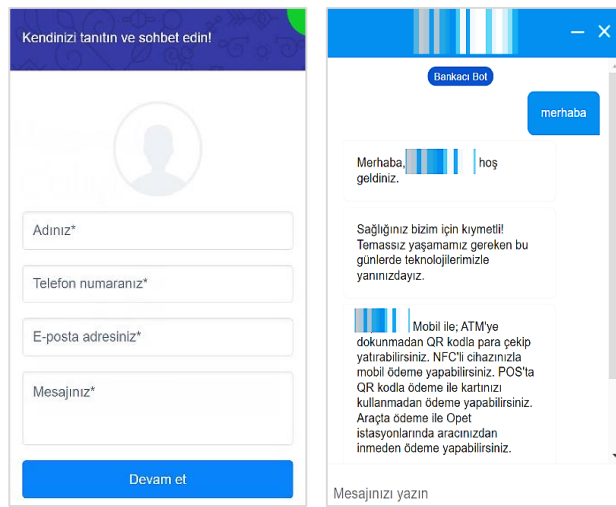
#### 4. SOHBET ROBOTLARI (CHATBOTS) ÜZERİNE İNCELEME ÇALIŞMASI (STUDY ON CHATBOTS)

Bu bölümde, yapay zekâ tabanlı uygulamaların ne sıklıkla ve çoğunlukla hangi işlemler için kullanıldığıyla ilgili genel yargı çıkarımı yapmak amacıyla yapay zekâ tabanlı uygulama olan sohbet robotları özelinde deneysel bir çalışma yapılmıştır. Yapay zekâ tabanlı uygulama olan sohbet robotlarının ne olduğu, nerelerde kullanıldığı, hangi tür verilere ihtiyaç duyduğu ve veri toplama faaliyetini yürütürken barındırdığı riskler anlatılmıştır. Kişisel veri mahremiyeti ve buna yönelik kanun, yönerge ve rehberlere uygun sohbet robotları tasarlamak için nelere dikkat edilmesi gerektiği konuları üstünde durulmuştur.

##### 4.1. Sohbet Robotu (Chatbot)

Chatbot kelimesi, sohbet anlamına gelen *Chat* ve robot anlamına gelen *Bot* sözcüklerinin birleşiminden meydana gelmiştir [27]. Sohbet robotu olarak bilinen bu uygulamalar, kullanıcı ile birebir etkileşimde bulunarak daha önceden belirlenmiş bazı durumlar üzerine kullanıcının isteğine göre cevaplar üreten ve bunu kullanıcıya ileten yapay zekâ ürünleridir [28]. Yapay zekâ teknolojisi ile geliştirildiklerinden dolayı kullanıcı ile yaptıkları her ikili konuşma neticesinde bir öğrenim sürecine girerler. Bu durum; sohbet robotlarına, karşılaştıkları yeni durumlar için kullanıcı isteğine cevap üretme imkânı sağlamıştır [29].

Sohbet robotlarının; kesintisiz erişebilir olmaları, basit sorunlar için kısa sürede çözümler oluşturması, sipariş işlemlerini yönetmesi, zaman ve maliyet anlamında etkin sonuçlar doğurması, vb. gibi birçok faydası bulunmaktadır [28]. Tüm bu faydaları ile her sektörün kullanmak istediği bu interaktif araç, müşteri deneyimine yeni bir bakış açısı kazandırmış; her an her yerde kullanıcı karşısına çıkmaya başlamıştır.



Şekil 4. Örnek sohbet robotu uygulamaları  
(Examples for the chatbot apps)

Telefonlarda kullanılan kişisel asistanlardan bankaların kişisel asistanlarına ve elektronik ticaret sitelerinin çevrimiçi destek servisine kadar her alanda hizmet verebilir niteliğe ulaşmıştır. Her sektörde boy gösteren yapay zekânın ve bunun bir yansıması olan sohbet robotlarının, kullanıcı deneyimi için artık bir devrim niteliğinde olduğu yadsınmaz bir gerçektir. Şekil 4'te ülkemizde, bankacılık ve elektronik ticaret sektöründe hizmet veren iki firmanın web sitelerinde kullandıkları sohbet robotlarına ait örnekler verilmiştir.

Sohbet oturumu sonrasında robot; elde ettiği verileri, CRM (*Customer Relationship Management*) adı verilen müşteri ilişkileri yönetimi veya benzer teknolojilerle bağlantılı duruma getirmek için saklamaktadır. Böylece bu durum; çıkarım yapmak, gelecek stratejileri belirlemek veya satış raporları oluşturmak için firma tarafından kullanılmaktadır [30].

Tüm bu özellikler ve toplanan veriler, sohbet robotlarının kullanım özeline göre (hangi uygulama ve hangi sektörde olduğu) değişiklik gösterse de temel olarak yukarıda anlatılan prensibin kullanıldığını söylemek mümkündür. Görüldüğü üzere bir sohbet robotu uygulaması, müşteri verilerini toplamakta büyük bir potansiyele sahiptir. Her ne kadar kullanıcı ile etkileşimli sistem oluşumu içinde onlara efektif bir hizmet sunmayı amaçlasa da amaç dışı sonuçlara da sebep olabilmektedir.

Kişiler; söz konusu bu uygulamadan faydalandıkları sırada paylaştıkları kişisel verilerin kaydedilmesi, işlenmesi ve aktarılması gibi sonuçlarla karşı karşıya kalabilmektedirler [31]. Öyleyse bu uygulamaların, bireyin kişilik haklarına aykırı bir durum teşkil etmeden tasarlanması gerekmektedir. Bunun yanı sıra veri sahibi konumundaki kişiler de yasal haklarını bilerek hareket etmelidirler.

##### 4.2. Deneysel Çalışma (Experimental Study)

Bu çalışma kapsamında, Gazi Üniversitesi Etik Komisyonunun 27.01.2022 tarih ve E.274489 sayılı yazısı ile etik kurul izni alınan bir anket çalışması yapılmıştır. Yapılan anket çalışmasında; kişilerin, sohbet robotu kullanma boyutunu ve sohbet robotu kullanımının kişisel verilere etkisine yönelik farkındalığını ölçmek amaçlanmıştır. Bu amaçla, *Google Formlar* [32] aracı ile kişilere Tablo 1'de belirtilen dört adet soru yöneltilmiştir.

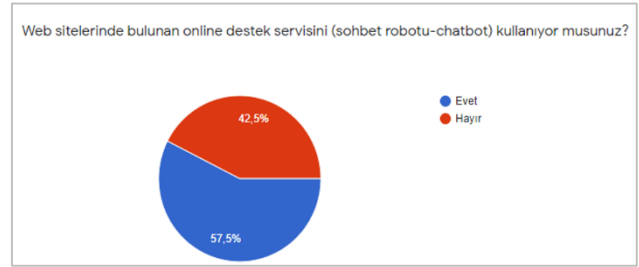
Anket soruları ve soru seçeneklerinin belirlenmesinde, kişilerin; sohbet robotlarını kullanımı ve kişisel veri farkındalığı bilincini ölçmek için temel olarak nitelendirilebilecek, cevaplaması fazla zaman almayan ve katılımcıları yormayan öz metinlerin belirlenmesi amaçlanmıştır.

Tablo 1. Katılımcılara yöneltilen anket soruları ve seçenekleri  
(Questions and options)

Sorular	Anket Soruları	Seçenekler
Soru-1	Web sitelerinde bulunan online destek servisini (sohbet robotu-chatbot) kullanıyor musunuz?	<ul style="list-style-type: none"> <li>• Evet</li> <li>• Hayır</li> </ul>
Soru-2	Kullanıyorsanız çoğunlukla hangi işlemlerinizi için kullanırsınız?	<ul style="list-style-type: none"> <li>• Bankacılık İşlemleri</li> <li>• E-ticaret (sipariş işlemleri)</li> <li>• Müşteri Hizmetleri/Destek</li> <li>• Kişisel Asistanlar</li> <li>• Kullanmıyorum</li> </ul>
Soru-3	Bu servislere güveniyor musunuz?	<ul style="list-style-type: none"> <li>• Evet, güveniyorum</li> <li>• Emin değilim</li> <li>• Hayır, asla güvenmiyorum</li> </ul>
Soru-4	Bu servislerin kişisel verilerinizi kaydedip işleyebileceğini biliyor musunuz?	<ul style="list-style-type: none"> <li>• Evet</li> <li>• Hayır</li> </ul>

Tablo 1’de belirtildiği üzere; Soru-1 ile sohbet robotlarının kullanılma düzeyi; Soru-2 ile hangi işlemler için kullanıldığı, Soru-3 ile kullanılan uygulamalara olan güven ve Soru-4 ile sohbet robotu uygulamasının kişisel veri işlediğinin ne düzeyde bilindiği ölçülmüştür. Soru-2’nin seçenekleri belirlenirken fazlaca kişisel veri işleyen uygulamaların seçeneklerde yer alması hususu göz önünde bulundurulmuştur.

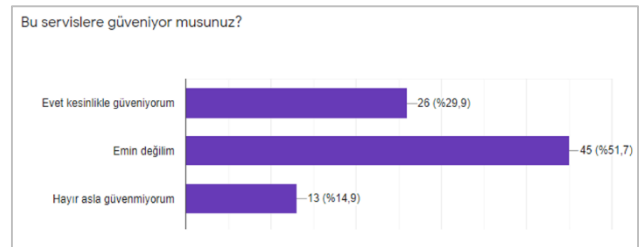
Yapılan anket çalışması 100 katılımcı tarafından yanıtlanmıştır. Şekil 5’te görüldüğü gibi Soru-1 için verilen cevaplarda, katılımcıların %57,5’i sohbet robotlarını kullandığını; %42,5’i ise kullanmadığını ifade etmiş ve sohbet robotlarının kullanım oranının azımsanmayacak boyutta olduğu sonucuna varılmıştır. Katılımcıların Soru-2’ye vermiş oldukları yanıtlarda, %21,4 oranı ile sohbet robotu uygulamalarının en fazla bankacılık işlemleri için kullanıldığı görülmüştür. Banka faaliyetlerinde bulunan, kimlik doğrulama, kişisel bankacılık hizmetleri, vb. gibi operasyonlarda sohbet robotlarının kullanım oranının fazla olması, bu uygulamaların kişisel veri güvenliği konusunda hassas olması gerektiği sonucunu doğurmuştur. Şekil 6 ile bankacılık işlemlerini sırasıyla, e-ticaret ve müşteri hizmetleri işlemlerinin takip ettiği görülmektedir. Anket sonuçlarına ait Şekil 7’de verilen grafiksel ögede ise Soru-3 için katılımcıların verdiği yanıtlara ait sonuçlar görülmektedir. Buna göre, katılımcıların %45’i bu uygulamaların güvenilirliği konusunda emin olmadığını düşünürken, %26’sı bu uygulamalara güvendiğini, %13’ü ise güvenmediğini belirtmiştir. Soru-3 için ulaşılan bu yanıtlar ile katılımcılar, çoğunlukla bankacılık işlemlerinde kullandıkları sohbet robotu uygulamalarına büyük oranda güvenmediklerini ifade etmişlerdir.



Şekil 5. Soru-1 için katılımcı yanıtları  
(Responses to Question-1)



Şekil 6. Soru-2 için katılımcı yanıtları  
(Responses to Question-2)



Şekil 7. Soru-3 için katılımcı yanıtları  
(Responses to Question-3)



Şekil 8. Soru-4 için katılımcı yanıtları  
(Responses to Question-4)

Şekil 8'e göre ise kullanıcıların %62,1'i bu servislerin kişisel verileri kaydedip işleyebileceğini bilmediğini; %37,9'luk kesim ise bildiğini belirtmiştir. Bu sonuçlardan anlaşıldığı üzere katılımcıların büyük bir kısmı sohbet robotu uygulamalarını bankacılık hizmetleri, vb. gibi veri güvenliği ve mahremiyetin önem arz ettiği operasyonlarda kullanırken bu uygulamalara güvenmediğini ifade etmiştir.

Yapılan anket çalışması, yapay zekâ tabanlı uygulamalardan sadece biri olan sohbet robotları temel alınarak gerçekleştirilmiştir ve anket sonuçlarından da anlaşılacağı üzere yapay zekâ tabanlı uygulamalar sıklıkla kullanılmaktadır. Hatta bu uygulamaların kişisel verileri işlediğini bilmeyen ve/veya bu uygulamaların güvenilirliğinden emin olmayan kişiler tarafından da kullanıldığı görülmektedir. Dolayısıyla yapay zekâ, insan hayatına değer katmayı amaçlasa da çıktı üretmesi için işlemesi gerekli olan veriyi en doğru şekilde ve kişisel veri mahremiyeti ve veri güvenliği ilkelerine bağlı kalarak kullanılmalıdır.

Federe öğrenme modeli, bir önceki bölümde detaylıca anlatıldığı gibi mahremiyet korumalı yapay zekâ uygulamaları geliştirmek için kullanılacak önemli yöntemlerden biridir. Bununla beraber bir sonraki başlıkta önerilen özelliklerin de yapay zekâ uygulamalarını veri mahremiyeti konusunda güçlü kılacağı düşünülmektedir.

#### 4.3. Öneriler (Recommendations)

Sohbet robotlarının, kişisel veri mahremiyeti ilkesine uygun geliştirilmesi, tasarlanması ve kullanılması için gerekli önlemlerin alınması gerekmektedir. Bununla ilgili olarak yasal mevzuatlarla ve tavsiye niteliğindeki bazı rehberler ile kişisel veri mahremiyetine uygunluk sağlanabileceği görülmüştür.

- *Privacy by Design* ilkesi ile sohbet robotu uygulamaları gizlilik prensipleri düşünülerek tasarlanmalıdır.
- *Amaca Bağlılık* ilkesi gereğince kişilere, verilerinin hangi amaçla alındığı kaydedildiği ve işlendiği açık ve doğru bir şekilde açıklanmalı ve amaçla bağdaşmayan türden veriler toplanılmamalıdır.

- *Veri Minimizasyonu* ilkesi ile kişisel bilgilerin toplanması, yasal olarak belirtilen amaçlar için gerekli olan kadarı ile sınırlı kalmalıdır. Kişisel olarak tanımlanabilir bilgilerin toplanması kesinlikle asgari düzeyde tutulmalıdır [33].
- *Silme hakkı* gereğince; veri sahipleri, işlenen verilerini silme talebinde bulunabilirler. Bu talepler karşılanmalı veya silinmesi talep edilen veriler, yasal dayanaklarda belirtilen uygun yöntemler ile anonim hale getirilmelidir.
- *Açık rıza* ilkesi ile sohbet robotunu kullanacak kullanıcının açık rızası alınmalıdır. Kullanıcıya, kişisel veri toplandığına, işlendiğine ve bunun hangi amaçla yapıldığına ilişkin bilgilendirmeler ışığında açık rıza isteği sorulmalıdır [13].
- *Aydınlatma metni* ilkesi gereğince veri sahipleri, hangi verilerinin hangi amaçla ve ne kadar süre işleneceğine dair bilgilendirilmelidir.

#### 4.4. Privacy by Design

*Privacy by Design* ilkesi, geliştirilmesi planlanan uygulamanın tasarım ve mimari aşaması dâhil olmak üzere ürünün tüm yaşam döngüsü boyunca veri koruma kurallarının dikkate alınmasıdır. Bu ilke kapsamında *Kanada Ontario Bilgi ve Gizlilik Kurulu Başkanı Ann Cavoukian* tarafından yedi temel prensip yayımlanmıştır [33]:

- Çalışmaların önleyiciliği ilkesi gereğince bir uygulamanın kullanımı esnasında çıkabilecek mimari veya teknik veri koruma sorunlarının önceden tespit edilmesi ve ortaya çıkmalarının engellenmesi hedeflenmelidir.
- *Privacy by Default* ilkesi gereğince uygulamaların standart varsayılan ayarları kullanıcı gizliliğini koruyacak şekilde yapılandırılmalıdır.
- *Gizlilik ilkesi* tasarıma entegre edilmelidir. Gizlilik prensipleri, uygulamanın tasarlanması aşamasında dikkate alınmalı ve uygulama tasarımı bu temel üzerine inşa edilmelidir.
- Kullanıcı, gizlilik prensibi çerçevesinde estetikten mahrum bırakılmamalıdır.
- Verinin, uygulama içinde toplandığı andan itibaren geçirdiği tüm yaşam döngüsü süresince güvenliği sağlanmalıdır.
- Uygulamada, veri işlemi sırasındaki her aşama hakkında şeffaf olunmalı ve kullanıcılar bilgilendirilmelidir.
- Kullanıcı çıkarları en üst seviyede tutulmalı ve kullanıcıya saygı ilkesi gereğince hareket edilmelidir.

## 5. SONUÇLAR (RESULTS)

Yapay zekâ ve veri birbirinden beslenen iki önemli kavramdır. Yapay zekâ için kaynak olan veriler, yapay zekâ teknolojilerine sahip sistemler için bir malzeme işlevi görür. Son yıllarda hem ülkemizde hem de dünyada kişisel veri mahremiyetine verilen önemin artmasıyla birlikte yapay zekâ konusu da kişisel verilerin mahremiyeti açısından önemli bir olgu haline gelmiştir. İnsanların kişisel veri mahremiyeti üstündeki bilinci ve farkındalık düzeyi artınca bu teknolojilerin mahremiyete etkisi üstünde daha kapsamlı çalışmaların karşımıza çıkacağı ise yadsınamaz bir gerçektir.

Bu çalışma kapsamında, genel olarak yapay zekâ tabanlı sistemlerin kişisel veri mahremiyeti üzerine etkisi anlatılmıştır. Yapay zekânın veri ile olan ilişkisine ve veri işlemenin kişileri mağdur edebilecek riskler taşıdığına değinilmiştir. Bu riskleri ortadan kaldırmak ve yapay zekâyı risklerinden arındırarak kişisel verilere saygılı sistemler haline dönüştürmek için ulusal ve uluslararası düzenlemelerde ve kanunlarda üstünde durulan ilkelere ve önerilere yer verilmiştir.

Yapılan literatür taramasında, bu alandaki çalışmaların büyük bir kısmının uygulama tabanlı çalışmalardan çok araştırma şeklinde olduğu görülse de son yıllarda uygulama çalışmalarına yönelik büyük bir artış yaşandığı da gözlemlenmiştir. Yapay zekânın kişisel verileri suiistimal etmesine engel olmak amacıyla özellikle veri koruma otoriteleri tarafından sunulan çözüm önerilerden en dikkat çeken ve üstünde durulan önerinin ise *Federe Öğrenme* yöntemi olduğu görülmüştür. Federe öğrenmenin karmaşık bir sisteme sahip olması, sunucuları yormaya yatkın olması gibi dezavantajları olsa da bu dezavantajlar, iyi tasarlanan sistemlerle ortadan kaldırılarak daha güvenli sistemler tasarlanabileceği görülmüştür.

Ayrıca çalışma kapsamında, bir yapay zekâ ürünü olan ve kullanımı her sektörde gittikçe artan sohbet robotları üzerine bir inceleme çalışması yapılmıştır. Sohbet robotlarının uygulamalarda kullanım yaygınlığı ve bir sohbet robotunun kullanıcı verilerini ne şekilde topladığı ve işlediği ile ilgili detaylara yer verilmiştir. Bu bağlamda; ülkemizde hem sohbet robotlarının kullanımını hem de kişilerin veri güvenliği ve mahremiyeti farkındalığını ölçmek için bir anket çalışması yapılmıştır. Yapılan anket çalışması 100 kişi tarafından yanıtlanmış ve bunların %57,5'i sohbet robotlarını kullandığını; %42,5'i ise kullanmadığını ifade etmiştir. Kullanıcıların sohbet robotlarını %21,4 oranı ile en fazla bankacılık işlemleri için kullandığı görülmüştür. Ayrıca, kullanıcıların %45'i bu uygulamaların güvenilirliği konusunda emin olmadığını düşünürken, %26'sı güvendiğini, %13'ü ise güvenmediğini belirtmiştir. Bunun yanı sıra kullanıcıların %62,1'i bu servislerin kişisel verileri kaydedip işleyebileceğini bilmediğini; %37,9'luk kesim ise bildiğini belirtmiştir. Bu durum bir bakıma, kişisel veri güvenliği farkındalığı konusunda toplumda eksikler olduğunu göstermektedir.

Tüm bunların yanında, sohbet robotlarını kişisel veri güvenliğini sağlayarak kanunlara uygun duruma getirmek için uygulanması gereken önemli ilkeler anlatılmış ve *Kanada Ontario Bilgi ve Gizlilik Kurulu Başkanı Ann Cavoukian* tarafından yayınlanan ve bugün birçok kesim tarafından kabul edilen *Privacy by Design* ilkesi ve buna ait yedi temel prensip anlatılmıştır.

Sonuç olarak, yapay zekâdan faydalanılması ve bu teknolojiler aracılığıyla çeşitli verilerin işlenmesi her ne kadar ticari ve sosyal açıdan önemli bir gelişme olarak kabul edilse de veri güvenliği açısından önemli riskleri içinde barındırmaktadır. Ancak, yapay zekâ riskleri düşünüldüğünde vazgeçilen bir alan olmamalı; tam aksine gerekli ve doğru bir şekilde korunmuş ve gizlenmiş verilerle beslenerek geliştirilmeli ve kullanılmalıdır.

## KAYNAKLAR (REFERENCES)

- [1] A. Süzen, K. Kayaalp, "Büyük Verilerde Gizlilik Tabanlı Yaklaşım: Federe Öğrenme", *International Journal of 3d Printing Technologies and Digital Industry*, 3(3), 297-304, 2019.
- [2] İnternet: M.V. Dülger, Yapay Zekâ Teknolojileri ve Veri Koruma Hukuku, <https://www.researchgate.net/publication/349552759>, 14.12.2021.
- [3] L. Mitrou, "Data Protection, Artificial Intelligence and Cognitive Services: Is The General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", *SSRN Electronic Journal (3386914)*, 2018.
- [4] Ş. Eroglu, "Dijital Yaşamda Mahremiyet (Gizlilik) Kavramı ve Kişisel Veriler: Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü Öğrencilerinin Mahremiyet ve Kişisel Veri Algılarının Analizi", *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, 35(2), 130-153, 2018.
- [5] D. Kamarinou, M. Christopher, S. Jatinder, "Machine Learning with Personal Data", *Queen Mary University of London, School of Law Legal Studies Research Paper 247/2016*, 2016.
- [6] L. Li, F. Yuxi, M. Tse, K. Lin, "A Review of Applications in Federated Learning", *Computers & Industrial Engineering*, 149(5), 2020.
- [7] K. Chandiramani, D. Garg, N. Mahesvari, "Performance Analysis of Distributed and Federated Learning Models on Private Data", *Procedia Computer Science*, 165, 349-355, 2019.
- [8] Y. Wang, Y. Tong, D. Shi, "Federated Latent Dirichlet Allocation: A Local Differential Privacy Based Framework", *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04), 6283-6290, 2020.
- [9] A. A. Süzen, M. A. Şimşek, "A Novel Approach to Machine Learning Application to Protection Privacy Data in Healthcare: Federated Learning" *Namik Kemal Tıp Dergisi*, 8(1), 22-30, 2020.
- [10] Y. Cambay, Y. Vural, Ş. Sağıroğlu, "Mahremiyet Korunmalı Büyük Veri Yayınlama İçin Kavramsal Model Önerileri", *Politeknik Dergisi*, 23(3), 785-798, 2020.
- [11] İnternet: Türkiye Cumhuriyeti Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, Açık Veri Projesi, <https://cbddo.gov.tr/projeler/acik-veri/>, 13.12.2021.

- [12] İnternet: Kişisel Verileri Koruma Kurumu, Yapay Zekâ Alanında Kişisel Verilerin Korunmasına Dair Tavsiyeler, <https://www.kvkk.gov.tr/Icerik/7048/Yapay-Zekâ-Alaninda-Kisisel-Verilerin-Korunmasına-Dair-Tavsiyeler>, 15.12.2021.
- [13] İnternet: Kişisel Verileri Koruma Kanunu, <https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=6698&MevzuatTur=1&MevzuatTertip=5>, 14.12.2021.
- [14] İnternet: Kişisel Verileri Koruma Kurumu, 100 Soruda Kişisel Verilerin Korunması Kanunu, <https://www.kvkk.gov.tr/SharedFolderServer/CMSFiles/185c2130-8070-4b2b-a91e-1d48322ca352.pdf>, 26.09.2022.
- [15] O. Sarı, “Yapay Zekânın Sebep Olduğu Zararlardan Doğan Sorumluluk”, *Türkiye Barolar Birliği Dergisi*, 147, 251-312, 2020.
- [16] Council of Europe, **Guidelines on Artificial Intelligence and Data Protection**, Directorate General of Human Rights and Rule of Law, 01, Fransa, 2019.
- [17] A. Akıncı, **Büyük Veri Uygulamalarında Kişisel Veri Mahremiyeti**, T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı Uzmanlık Tezi, 2019.
- [18] The Norwegian Data Protection Authority Committee, **Artificial Intelligence and Privacy Report**, Norwegian Data Protection Authority, Norwegian, 2018.
- [19] O. Choudhury, A. Gkoulalas-Divanis, T. Salonidis, I. Sylla, Y. Park, G. Hsu, A. Das, “Differential Privacy-enabled Federated Learning for Sensitive Health Data”, *arXiv:1910.02578*, 2020.
- [20] M. Pettai, P. Laud, “Combining Differential Privacy and Secure Multiparty Computation”, *Proceedings of the 31st Annual Computer Security Applications Conference*, 421-430, 2015.
- [21] İnternet: M. A. Sarıkaya, Yapay Zekâ ve Mahremiyet, [https://medium.com/@sarikayameh?source=post\\_page-45f8731736-](https://medium.com/@sarikayameh?source=post_page-45f8731736-), 11.12.2021.
- [22] K. Hu, Y. Li, M. Xia, J. Wu, M. Lu, S. Zhang, L. Weng, "Federated Learning: A Distributed Shared Machine Learning Method", *Complexity*, 2021(2), 1-20, 2021.
- [23] İnternet: R. Yeşil, Federated Learning, <https://medium.com/datarunner/federe-%C3%B6%C4%9Frenme-federated-learning-8ad87791c0b5>, 12.12.2021.
- [24] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, F. Beaufays, “Applied Federated Learning: Improving Google Keyboard Query Suggestions”, *arXiv:1812.02903*, 2018.
- [25] İnternet: F. Hartman, Federated Learning for Firefox, <https://florian.github.io/federated-learning-firefox/>, 14.12.2021.
- [26] İnternet: Intel Works with University of Pennsylvania in Using Privacy-Preserving AI to Identify Brain Tumors, <https://newsroom.intel.com/news/intel-works-university-pennsylvania-using-privacy-preserving-ai-identify-brain-tumors/#gs.oak7q9>, 01.02.2022.
- [27] J. Vogel, **Chatbots: Development and Applications**, Yüksek Lisans Tezi, HTW Berlin University of Applied Sciences, International Media and Computing Faculty, 2017.
- [28] Z. Seyitoğlu, **Türkiye’de Dijital Halkla İlişkilerde Değişen Müşteri Deneyimi: Chatbot Uygulamaları**, İstanbul Kültür Üniversitesi, Yüksek Lisans Tezi, Mayıs 2019.
- [29] İnternet: Türkiye Cumhuriyeti Ticaret Bakanlığı, Chatbot Kurgusunun Yapılması, <https://www.eticaret.gov.tr/cevrimiciegitim/chatbot-kurgusunun-yapilmasi-64>, 10.12.2021.
- [30] İnternet: H. Charatan, Chatbots vs. GDPR, <https://chatamo.com/chatbots-vs-gdpr-interact/>, 13.12.2021.
- [31] İnternet: Yapay Zekâ Reklamcılığı ve Kişisel Verilerin korunması, <https://cukurpartners.com/tr/yapay-zeka-reklamciligi-ve-kisisel-verilerin-korunmasi>, 26.10.2022.
- [32] İnternet: Google Formlar (Google Forms), [https://www.google.com/intl/tr\\_tr/forms/about/](https://www.google.com/intl/tr_tr/forms/about/), 31.01.2022.
- [33] A. Cavoukian, **Privacy by Design**, The 7 Foundational Principles Implementation and Mapping of Fair Information Practices, Information & Privacy Commissioner, Canada, 2010.