



#### **Sahibi / Owner**

Doç. Dr. M. Hanefi CALP

#### **Baş Editör / Editor in Chief**

Doç. Dr. M. Hanefi CALP

#### **Yardımcı Editörler / Co-Editors**

Doç. Dr. Ahmet DOĞAN

Doç. Dr. Serkan SAVAŞ

Dr. Öğr. Üyesi Ömer Çağrı YAVUZ

#### **Alan Editörleri / Field Editors**

Prof. Dr. Alptekin ERKOLLAR

Prof. Dr. Türksel BENSGHİR

Prof. Dr. Tülay İLHAN NAS

Prof. Dr. Üstün ÖZEN

#### **Yayın Kurulu / Editorial Board**

Prof. Dr. Abdulkadir PEHLİVAN, Karadeniz Teknik Üniversitesi, Yönetim Bilişim Sistemleri

Prof. Dr. Ali HALICI, Başkent Üniversitesi, Yönetim Bilişim Sistemleri

Prof. Dr. Aslıhan TÜFEKÇİ, Gazi Üniversitesi, Yönetim Bilişim Sistemleri

Prof. Dr. Bilal GÜNEŞ, Gazi Üniversitesi, Fizik Eğitimi

Prof. Dr. Bilal TOKLU, Gazi Üniversitesi, Endüstri Mühendisliği

Prof. Dr. Birgül Kutlu BAYRAKTAR, Boğaziçi Üniversitesi, Yönetim Bilişim Sistemleri

Prof. Dr. Bogdan PATRUT, Alexandru Ioan Cuza Üniversitesi, Matematik ve Bilgisayar Bilimleri

Prof. Dr. Bünyamin ER, Karadeniz Teknik Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Cevriye Gencer, Gazi Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Cihan TANRIÖVEN, Ankara Hacı Bayram Veli Üniversitesi, İşletme  
Prof. Dr. Efendi NASİBOĞLU, Dokuz Eylül Üniversitesi, Bilgisayar Bilimleri  
Prof. Dr. Erdoğan DOĞDU, Çankaya Üniversitesi, Bilgisayar Mühendisliği  
Prof. Dr. Erman COŞKUN, Bakırçay Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Hadi GÖKÇEN, Gazi Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Halil İbrahim OKUMUŞ, Karadeniz Teknik Üniversitesi, Elektrik-Elektronik Mühendisliği  
Prof. Dr. Hamdi Tolga KAHRAMAN, Karadeniz Teknik Üniversitesi, Yazılım Mühendisliği  
Prof. Dr. Hasan Erdinç KOÇER, Selçuk Üniversitesi, Elektrik-Elektronik Mühendisliği  
Prof. Dr. İlya LEVIN, Tel Aviv Üniversitesi, Bilim ve Teknoloji Eğitimi  
Prof. Dr. İsmail SARITAŞ, Selçuk Üniversitesi, Elektrik-Elektronik Mühendisliği  
Prof. Dr. İsmail ŞAHİN, Gazi Üniversitesi, Endüstriyel Tasarım Mühendisliği  
Prof. Dr. Kürşad ZORLU, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Latif ÖZTÜRK, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. M. Ali AKCAYOL, Gazi Üniversitesi, Bilgisayar Mühendisliği  
Prof. Dr. M. Nihat SOLAKOĞLU, Çankaya Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Mehmet AKTAN, Necmettin Erbakan Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Mehmet BAŞ, Ankara Hacı Bayram Veli Üniversitesi, İşletme  
Prof. Dr. Meltem ÖZTURAN, Boğaziçi Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Metehan TOLON, Ankara Hacı Bayram Veli Üniversitesi, İşletme  
Prof. Dr. Murat Paşa UYSAL, Başkent Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Nicu BIZON, Pitesti Üniversitesi, Elektronik, İletişim ve Bilgisayar Bilimleri  
Prof. Dr. Nursal ARICI, Gazi Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Oğuz KAYNAR, Sivas Cumhuriyet Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Rahmi CANAL, İnönü Üniversitesi, Biyomedikal Mühendisliği  
Prof. Dr. Sabri KOÇER, Necmettin Erbakan Üniversitesi, Bilgisayar Mühendisliği  
Prof. Dr. Selçuk KARAMAN, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Selçuk Kürşat İŞLEYEN, Gazi Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Sevinç GÜLSEÇEN, İstanbul Üniversitesi, Enformatik  
Prof. Dr. Shadi A. ALJAWARNEH, Jordan Üniversitesi, Bilim ve Teknoloji  
Prof. Dr. Suat ÖZDEMİR, Hacettepe Üniversitesi, Bilgisayar Mühendisliği  
Prof. Dr. Süleyman ERSÖZ, Kırıkkale Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Şükrü ÖZŞAHİN, Karadeniz Teknik Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Talip KELLEGÖZ, Gazi Üniversitesi, Endüstri Mühendisliği  
Prof. Dr. Tülay İlhan NAS, Karadeniz Teknik Üniversitesi, İşletme  
Prof. Dr. Türksel KAYA BENSGHIR, Ankara Hacı Bayram Veli Üniversitesi, İşletme  
Prof. Dr. Uğur YAVUZ, Atatürk Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Üstün ÖZEN, Atatürk Üniversitesi, Yönetim Bilişim Sistemleri  
Prof. Dr. Yılmaz GÖKŞEN, Dokuz Eylül Üniversitesi, Yönetim Bilişim Sistemleri  
Doç. Dr. Gürcan ÇETİN, Muğla Sıtkı Koçma Üniversitesi, Bilişim Sistemleri Mühendisliği  
Doç. Dr. Ekrem BAHÇEKAPILI, Karadeniz Teknik Üniversitesi, Yönetim Bilişim Sistemleri  
Doç. Dr. Hakan ÖZKÖSE, Bartın Üniversitesi, Yönetim Bilişim Sistemleri  
Doç. Dr. Muhammet BERİGEL, Karadeniz Teknik Üniversitesi, Yönetim Bilişim Sistemleri  
Doç. Dr. Murat DENER, Gazi Üniversitesi, Bilgisayar Mühendisliği  
Doç. Dr. Murat DÖRTERLER, Gazi Üniversitesi, Bilgisayar Mühendisliği  
Doç. Dr. Osman ÖZKARACA, Muğla Sıtkı Koçma Üniversitesi, Bilişim Sistemleri Mühendisliği  
Doç. Dr. Paolo TORRONI, Bologna Üniversitesi, Bilgisayar Bilimleri ve Mühendisliği  
Doç. Dr. Utku KÖSE, Süleyman Demirel Üniversitesi, Bilgisayar Mühendisliği

Doç. Dr. Ümit ATİLA, Gazi Üniversitesi, Bilgisayar Mühendisliği  
Doç. Dr. Ahmet DOĞAN, Osmaniye Korkut Ata Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Öğr. Üyesi Bilgehan İMAMOĞLU, Karadeniz Teknik Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Öğr. Üyesi Emin Sertaç ARI, Osmaniye Korkut Ata Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Öğr. Üyesi Güler KARAMAN, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Öğr. Üyesi Mevlüt UYSAL, Gazi Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Öğr. Üyesi Mustafa TANRIVERDİ, Gazi Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Öğr. Üyesi Ömer Çağrı YAVUZ, Trabzon Üniversitesi, Yönetim Bilişim Sistemleri  
Dr. Iulian FURDU, Vasile Alecsandri Üniversitesi, Bilişim ve Eğitim Bilimleri  
Dr. Pandian VASANT, Teknoloji Petronas Üniversitesi, Bilişim Sistemleri  
Dr. Tomayess ISSA, Curtin Üniversitesi, Bilişim Sistemleri  
Arş. Gör. Berat TAHTABİÇEN, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri  
Arş. Gör. Nadide Gizem GÜRSON DOLAR, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri  
Arş. Gör. Mahmud Zahid MUTLU, Ankara Hacı Bayram Veli Üniversitesi, Yönetim Bilişim Sistemleri

#### **Teknik Koordinatör / Technical Coordinator**

Arş. Gör. Mahmud Zahid MUTLU

#### **Sekreterlik / Secretarial**

[mahmud.mutlu@hbv.edu.tr](mailto:mahmud.mutlu@hbv.edu.tr), [hanefi.calp@hbv.edu.tr](mailto:hanefi.calp@hbv.edu.tr)

Ankara Hacı Bayram Veli Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü,  
Emniyet Mahallesi, Muammer Bostancı Caddesi, No:4, 06500 Beşevler/Ankara, Türkiye

#### **İÇİNDEKİLER (Cilt: 5 / Sayı: 2 - Aralık 2023)**

Gerçek Zamanlı Elektrik Yük Tahmini İçin Bir Derin Öğrenme Yaklaşımı <i>Alaa Harith Mohammed AL-HAMID, Serkan SAVAŞ</i>	01-09
Yapay Zeka Algoritmaları Kullanılarak Öğrencilerin Akademik Başarısı ile Stres İlişkisinin Keşifsel Bir Analizi <i>Hakan YÜKSEL</i>	10-20
Yapay Sinir Ağları ile Mısır Yapraklarında Hastalık Tespiti <i>Resul BÜTÜNER, Muzaffer ŞİMŞEK, Gökhan AKSÖZ, Sadi ŞAN</i>	21-31



## Gerçek Zamanlı Elektrik Yük Tahmini İçin Bir Derin Öğrenme Yaklaşımı

Alaa Harith Mohammed AL-HAMID <sup>\*,a</sup>, Serkan SAVAŞ <sup>b</sup>

<sup>a</sup>\* Çankırı Karatekin Üniversitesi, Mühendislik Fakültesi, Elektronik ve Bilgisayar Mühendisliği Anabilim Dalı, ÇANKIRI, 18100, TÜRKİYE

<sup>b</sup> Kırıkkale Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, KIRIKKALE, 71450, TÜRKİYE

### MAKALE BİLGİSİ

Alınma: 09.09.2023  
Kabul: 13.12.2023

#### Anahtar Kelimeler:

Dizi planı, Güç akışı,  
Derin öğrenme, Yük  
talebi tahmini, Yığın  
LSTM, Akıllı şebekeler

#### \*Sorumlu Yazar

e-posta:  
alaaahareth@gmail.com

### ÖZET

Doğru ve gerçek zamanlı elektrik talebi tahmininin artan önemi ışığında, bu araştırma, tahmin doğruluğunu önemli ölçüde artırmak amacıyla bir derin öğrenme modeli sunmaktadır. Doğrusal regresyon gibi geleneksel tahmin yöntemleri, elektrik kullanımıyla ilgili verilerde yer alan karmaşık kalıpları yakalamakta zorlanmaktadır. Standart makine öğrenimi yöntemlerinin, önerilen derin Uzun Kısa Vadeli Bellek (Long Short-Term Memory-LSTM) modeliyle karşılaştırıldığında yetersiz kaldığı görülmüştür. Ortalama Mutlak Hata (MAE) 5.454 ve Ortalama Karesel Hata (MSE) 18.243, derin LSTM modelinin bu sorunun üstesinden gelmedeki yeterliliğini göstermektedir. Doğrusal regresyon ise 47.352 MAE değeri ve 65.606 MSE değeri ile önerilen modelden daha düşük başarı sonucu elde etmiştir. Daha yüksek tahmin hassasiyeti ve güvenilirliği nedeniyle, derin LSTM modeli elektrik talebinin doğru, gerçek zamanlı tahmini için uygun bir seçenektir.

DOI: 10.59940/jismar.1357804

## A Deep Learning Approach to Real-Time Electricity Load Forecasting

### ARTICLE INFO

Received: 09.09.2023  
Accepted: 13.12.2023

#### Keywords:

Plan of array, Power  
flow, Deep learning,  
Load demand forecast,  
Stacked LSTM, Smart  
grids

#### \*Corresponding Authors

e-mail:  
alaaahareth@gmail.com

### ABSTRACT

In light of the increasing importance of accurate and real-time electrical demand forecasting, this research presents a deep learning model with the goal of dramatically improving predictive accuracy. Conventional methods of forecasting, such as linear regression, have trouble capturing the complex patterns included in data about electricity usage. Standard machine learning methods are shown to be wanting when compared to the suggested deep Long Short-Term Memory (LSTM) model. Mean Absolute Error (MAE) of 5.454 and Mean Squared Error (MSE) of 18.243 demonstrate the deep LSTM model's proficiency in tackling this problem. The linear regression, on the other hand, achieved a MAE of 47.352 and an MSE of 65.606, which is lower than the proposed model. Because of its greater predictive precision and reliability, the deep LSTM model is a viable option for accurate, real-time prediction of electricity demand.

DOI: 10.59940/jismar.1357804

### 1. INTRODUCTION (GİRİŞ)

The issues related with global warming and the energy resources scarcity are soaring. It therefore calls for strong Energy Management System (EMS) at this moment. Industries are on the look out for improved

EMS in order to realize its potential to revolutionize energy monitoring and budgeting. Smart meters represent very sophisticated tools for measuring power consumption at home or office levels. This makes them an extremely strong partner on the way

toward efficient and reduced costs when it comes to energy planning [1].

The smart meter's data becomes one of the valued resources that provides a huge pool of data for analytics-based management. Smart microgrid is an integral component in the tapestry of EMS. Examples include interpreting consumption trend, demand forecast, and optimal energy exchange to decipher energy load [2]. Within the realm of load forecasting, varying prediction horizons come to the fore: short term, midterm and longterm forecast . However, decisions concerning prosumers within the smart energy grids will require a short timeline, namely, minutes to days. In addition, medium-term forecasts, weeks to months ahead, are crucial in power systems scheduling [3-4], while long term forecasts, monthly/yearly predictions, support grid maintenance planning.

In essence, load forecasting seeks to predict electricity demand from end-use consumers in advance. This can be done using various methods, out of which machine learning (ML) strategies are becoming increasingly popular for having a more usable approach. To make informed decision making in energy management, accurate forecasts of future energy demands should be provided to enable proper planning and scheduling . Still, load forecasting is also not a walk in the park considering that energy consumption patterns change greatly over time. Dynamism results into concept drifting, which makes traditional ML approaches obsolete [5].

There are numerous reasons why people's energy consumptions behavior could change. These reasons include increasing or decreasing prices for fuel; temporal consideration relating to date or season. As an example, changes in pricing lead to the demand that the customers should react on the cost price increase. However, traditional ML approaches fail and deteriorate due to recognising dynamic nature of energy load demand [6].

Current research focuses on deep learning (DL) based approaches to interval load forecasting using the power of Long-Short Term Memory (LSTM) networks. It should however be noted that LSTM has shown very good results for load forecasting. An adaptive mechanism that grapples with new load consumption patterns due to concept drift in order to improve the efficiency of DL models -the model is updated automatically according to new energy usage patterns that signal changes. Nonetheless, active/passive tracking of concept drift is prone to various problems- particularly defining a magnitude threshold that would ensure overall good predictions [7-8].

This paper offers an interval-load forecast learning model called hybrid LSTM as a solution for these problems. A complete solution encompasses both pass and active drift adaptation. A hybrid LSTM network is developed to be able rapidly learn changing load consumption scenarios having captured the historical consumption patterns. There is a detailed comparison with baseline models in the paper and hence the effectiveness of the proposed hybrid LSTM model. In addition, a trade-off analysis of various adaptation strategies takes into account the predictive performance as well as computation costs that guides appropriate choice of adaptation.

In the subsequent sections, we consider related works in section 2. Section 3 proposes a new hybrid LSTM solution. Section 4 evaluates it experimentally and section 5 concludes the results.

## 2. RELATED WORKS (İLGİLİ ÇALIŞMALAR)

Researchers have extensively employed ML algorithms and DL to develop criteria for load-balanced forecasting across a plethora of fields. Using deep learning and the current spatio-temporal correlation in appliance load data, [9] develop a short-term home load forecasting method. Electricity consumption behaviours and their internal spatio-temporal relationship are studied using several time series in the framework. The proposed forecasting method also makes use of a deep neural network and an iterative process. The results demonstrate that both iterative ResBlocks and load data from appliances contribute to better predicting results. The proposed method reduces Root Mean Squared Error by 3.89 percentage points to 20.00 percentage points, Mean Absolute Error by 2.18% to 22.58%, and Mean Absolute Percentage Error by 0.69 percentage points to 32.78 percent. The suggested method is further tested with further trials to examine the effects of incorporating load data from appliances, iterative ResBlocks, and other parameters.

In order to effectively deploy demand response strategies in manufacturing facilities, [10] want to create a system to predict the electrical energy demand of metal cutting machine tools. Based on the findings of the previous research, the effectiveness of LSTM and convolutional neural networks (CNNs) in predicting the electric load of a machine tool for a 100-second time horizon is compared and contrasted. The results show that specifically the combination of CNN and LSTM in a DL strategy delivers accurate and robust time series forecasts with reduced feature preparation work. Different network topologies, such as an attention mechanism for the LSTMs, and other hyperparameter

combinations are assessed to see whether they can further enhance the predicting accuracy.

Using a combination of the factored conditional restricted boltzmann machine (FCRBM) and the conditional restricted boltzmann machine (CRBM), [11] provide a rapid and accurate short-term load forecasting system. A rectified linear unit (ReLU) and a sigmoid function are used in training the stacked FCRBM and CRBM. Utilities in the US have offline demand side load statistics that have been put via the suggested framework. Based on predicted demand, power plants may ramp up or down their output, add more generators, or trade energy with nearby grids. Their suggested approach is verified using three performance metrics: the mean absolute percentage error (MAPE), the normalised root mean square error (NRMSE), and the correlation coefficient. When compared to artificial neural networks (ANNs) and CNNs, the results demonstrate the accuracy and robustness of stacked FCRBM and CRBM.

Optimisation of an ANN model is presented by [12] using an embedded multi-population Differential Evolution (DE) micro-Genetic Algorithm (mGA). To begin, a method for optimising and balancing both global and local search—the mGA embedded multi-population DE—is proposed. The projected DE is then used to fine-tune the ANN's training-time weights. Four cutting-edge ML methods are used to compare the entire model's performance against the publicly available Panama electrical load dataset. Compared to the other chosen machine learning methods, the suggested DE based model is shown to have superior prediction accuracy in the evaluation results.

To enhance prediction accuracy, [13] suggested a self-adaptive DL model. Meanwhile, RSPSO is employed to determine the network's optimum architecture, which involves discrete variables (i.e. the amount of neurons in each layer and the quantity of hidden layers) and categorical variables (i.e. activation function in each layer and learning approach). Additionally, the architecture and structure of the dynamic DL model are updated using the moving horizon approach, allowing it to capture the most recent highlighting patterns in the building's electrical load. The electricity consumption of a school building and the local weather profile are used to evaluate the suggested load prediction model. The best model for predicting energy consumption beyond the next horizon is found to be the self-adaptive load prediction model, but this model's prediction performance degrades with increasing horizon length. Prediction accuracy and repeatability are shown by the

suggested prediction model's mean squared error, mean absolute error, and coefficient of determination all falling between 4.48 kW and 11.23 kW, 1.28 kW and 2.31 kW, and 97.52% and 98.92%, respectively. Adding Gaussian white noise to meteorological data results in an increase in mean absolute error between 2.08% and 15.33%, showcasing the reliability of the proposed prediction model in dealing with weather forecast uncertainty. Therefore, the suggested accurate, resilient, repeatable and self-adaptive load forecast model can be anchored in practical energy management systems thus facilitate building operation and system control.

The dynamic drift-adaptive Long Short-Term Memory (DA-LSTM) architecture proposed by [14] can enhance the performance of load forecasting models without the need for a drift threshold. They incorporate a number of active and passive adaption mechanisms into the framework. They provide a comprehensive analysis of the proposed framework and apply it to a real-world problem in a cloud context in order to evaluate DA-LSTM in a realistic situation. Each method's efficiency is measured by how well it can make predictions and how much computing time it takes. The experimental findings reveal that, compared to the literature's baseline methodologies, our framework outperforms them across a variety of evaluation metrics.

### 3. MATERIALS AND METHODS (MATERİYALLER VE YÖNTEMLER)

Meticulous design is paramount given its prospective significant impact in real-world applications. The proposed approach harnesses DL techniques for forecasting power consumption. Notably, when regression is performed on the accumulated data, a model predicated on the LSTM algorithm is employed. This data encapsulates various factors influencing energy consumption patterns, such as weekly, monthly, or annual electricity usage.

Upon segregating the data into bifurcated sets, the methodology proceeds in two primary phases: data normalization and inference derivation. The Max-Min normalization technique is utilized to standardize the data, ensuring each constituent datum is normalized prior to the regression phase. The analytical phase is executed using two distinct strategies. Initially, a deep hybrid-based model is proposed, followed by the application of a machine learning linear regression method. Figure 1 delineates the sequential execution of these stages.

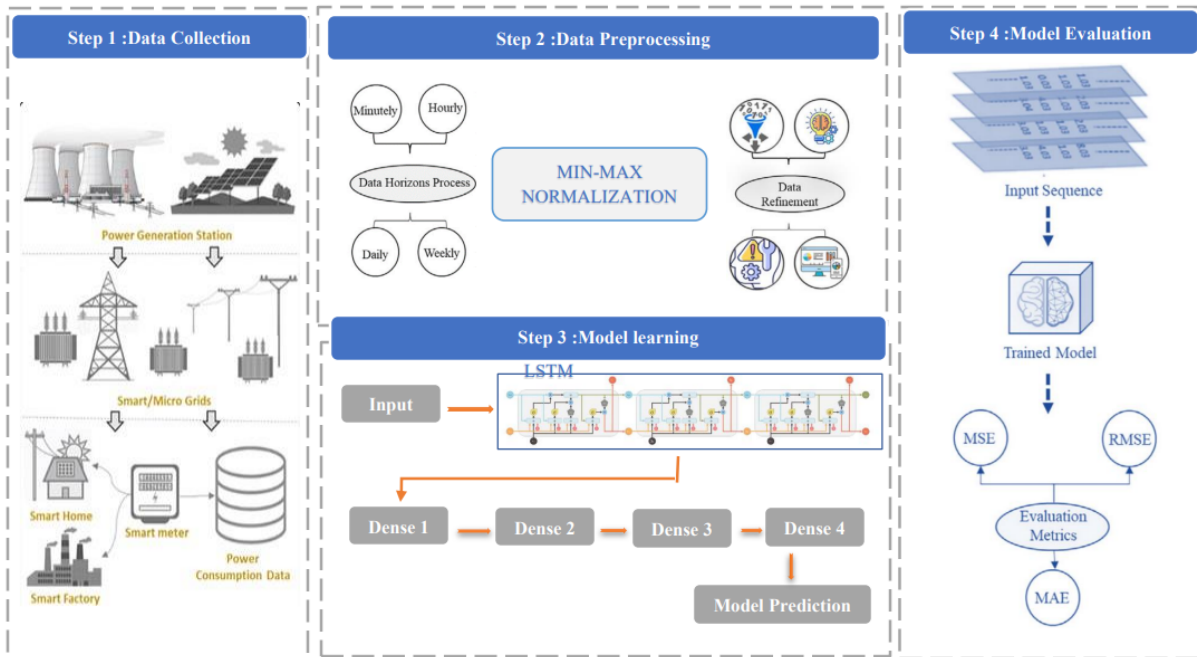


Figure 1. The proposed model

### 3.1 Dataset (Veriseti)

The Short-term Electricity Load Forecasting (Panama) dataset predicts Panamanian post-transmission electricity loads [15]. Weekly load projections are broken down hourly from daily and weekly pre-transmission power generation records in this dataset. Historical electrical load statistics, weekly projections, and calendar features like school sessions and vacations are included in its comprehensiveness. Notably, the information includes temperature, humidity, rainfall, and wind velocity for three major Panamanian cities.

### 3.2 Pre-processing (Ön İşleme)

For neural networks, pre-processing, especially

Algorithm 2 Linear Regression Computation

```

1: procedure LINEARREGRESSION( $n$ )
2:   Read Number of Data ( $n$ )
3:   for  $i = 1$  to  $n$  do
4:     Read  $x_i$  and  $y_i$ 
5:   end for
6:   Initialize:
7:    $sum\_x = 0$ 
8:    $sum\_x2 = 0$ 
9:    $sum\_y = 0$ 
10:   $sum\_xy = 0$ 
11:  for  $i = 1$  to  $n$  do
12:     $sum\_x = sum\_x + x_i$ 
13:     $sum\_x2 = sum\_x2 + x_i^2$ 
14:     $sum\_y = sum\_y + y_i$ 
15:     $sum\_xy = sum\_xy + x_i \cdot y_i$ 
16:  end for
17:  Compute the slope  $K$  and intercept  $J$ :
18:   $K = \frac{n \cdot sum\_xy - sum\_x \cdot sum\_y}{n \cdot sum\_x2 - (sum\_x)^2}$ 
19:   $J = \frac{sum\_y - K \cdot sum\_x}{n}$ 
20:  Compute the dependent variable:
21:   $y = J + K \cdot x$ 
22: end procedure
    
```

normalization, significantly impacts the effectiveness of the training phase. By normalizing raw inputs, the data becomes more conducive to training. Absence of normalization can decelerate the neural network training process, given that normalization's primary role is to ensure uniform scaling of data. A Min-Max normalization process applied to an electricity forecasting dataset [16]. For each feature  $x$  in  $S$ , the normalization as in Algorithm (1).

It takes as input a vector  $S$  consisting of features related to electricity forecasting. The algorithm initializes by determining the minimum  $X_{min}$  and maximum  $X_{max}$  values within  $S$ .

### 3.3 Linear Regression (Doğrusal Regresyon)

Linear regression stands as one of the most prevalent and intuitively comprehensible ML techniques [17], rooted deeply in statistical analysis. Its primary objective is to establish a linear relationship among multiple variables [18]. In the context of this paper, the linear regression methodology is employed to forecast electricity consumption, leveraging the normalized values from the dataset as illustrated in Algorithm (2).

### 3.4 Hybrid LSTM - Dense Model (Önerilen Hibrit LSTM – Dense Modeli)

This research aims to present an innovative model that integrates the capabilities of LSTM with a dense network. The ultimate goal of this fusion is to enhance data regression, with a specific emphasis on electricity forecasting. The proposed network model employs the renowned backpropagation technique. In this method, the weights affiliated with certain network connections undergo consistent adjustments. The primary objective behind this is to minimize the discrepancy between the original output vector, as generated by the designated network, and the anticipated output vector. Through this novel integration of LSTM and dense networks, we aspire to set a new benchmark in the realm of accurate electricity forecasting.

In a model, the output of each neuron in the dense layer is influenced by all of the neurons in the layer below it. Matrix-vector multiplication is performed by the neurons in the dense layer. This is where the row vector from the preceding layers is multiplied by the column vector from the dense layer. The primary rule for multiplying matrices and vectors is that the number of elements in the row vector must match the number of elements in the column vector. Because of this, the output of the dense layer will be an  $N$ -dimensional vector.

In general, the proposed hybrid deep model for load forecasting of electricity is made up of five layers as shown in Figure 2, and an optimizer as a final progress.

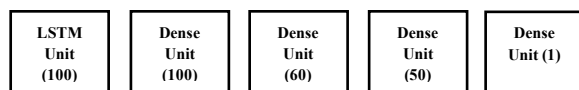


Figure 2. Proposed deep model layer structure

Table 1 outlines the architecture of a specific the proposed deep LSTM regression model, detailing the sequence of layers, the number of parameters for each layer, and the output shape. The model complexity is dictated by the number of parameters, where too many can cause overfitting and too few can result in underfitting.

Table 1. Proposed deep LSTM layers

Layer Type	Parameter	Output shape
LSTM	52400	(non, 100)
Dense	10100	(non, 100)
Dense	6060	(non, 60)
Dense	3050	(non, 50)
Dense	51	(non, 1)

The data's dimensionality at each processing stage is represented by the output shape of each layer. In this model, it begins with an LSTM layer with 52,400 parameters, generating 100 features per sample. Subsequently, a series of Dense layers follow, with varying parameters and output features: the first with 10,100 parameters and 100 features, the next with 6,060 parameters and 60 features, and another with 3,050 parameters and 50 features. The model concludes with a final Dense layer of 51 parameters, yielding a single output per sample, indicating its suitability for regression or binary classification tasks.

### 3.5 Performance Measurements (Performans Ölçümleri)

The performance of the model is evaluated based on specific metrics. The MAE (Mean Absolute Error) and MSE (Mean Squared Error) are two such measures used to assess the proposed model.

1- The Mean Absolute Error (MAE) measures the average discrepancy between two continuous variables [19],  $X$  and  $Y$ . These variables can represent paired observations of the same phenomenon. Comparisons such as predicted values vs. actual values, posterior time vs. beginning time, or a standard gauge method vs. an alternative measuring approach are examples where  $Y$  vs.  $X$  might be used. When there are  $n$  data points on a scatter plot, each with coordinates  $(x_i, y_i)$ , the MAE measures the average vertical distance of each point from the  $Y=X$  line. Contrary to the mention, MAE does not stand for the average horizontal distance from the  $Y=X$  line; it always denotes vertical distances. The method to compute this distance is presented in Eq. (2).

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (2)$$

Considering that there are  $n$  samples of model errors that are calculated as  $(e_i, i = 1, 2 \dots n)$ .



2- The Mean Squared Error (MSE) is the average of the squared differences between predicted and actual values. Because errors are squared before being averaged, the MSE will always yield non-negative values. Values closer to zero indicate better model performance. Graphically, the MSE can be thought of as the second moment of the error distribution. This characteristic allows it to capture both the variance of the estimator (the extent to which estimates differ from one data sample to another) and its bias. The calculation of MSE is detailed in Eq. (3), as described by [20].

Where  $X_i$  denote to the original sample data, and  $Y_i$  referred to the processed one.

3- Root Mean Squared Error (RMSE)

The RMSE, or root mean squared error [21], is calculated by taking the square root of the MSE. It's another popular metric for estimating the typical size of the mistake, and it uses the same measurement scale as the source data.

$$RMSE = \sqrt{MSE} \tag{4}$$

These metrics are critical for gauging the model's effectiveness in predicting future electricity demand. Better model performance is indicated by smaller values of MAE, MSE, and RMSE, as this indicates that the projected values are closer to the actual values.

**4. RESULTS AND DISCUSSION (SONUÇLAR VE TARTIŞMA)**

Comparing the best attained result from the deep model within the linear regression algorithm, there is a big difference, and much better regression and prediction performance when utilizing the proposed deep model rather than the machine learning linear algorithm as shown in Table 2.

Table 2. Comparing of two model performance

	MAE	MSE
Linear regression	47.352	65.606
Deep LSTM (Proposed study)	5.454	18.243

Also, Figure 2, and Figure 3, illustrated the MAE, and MSE for the best results of the proposed deep model and linear regression algorithm.

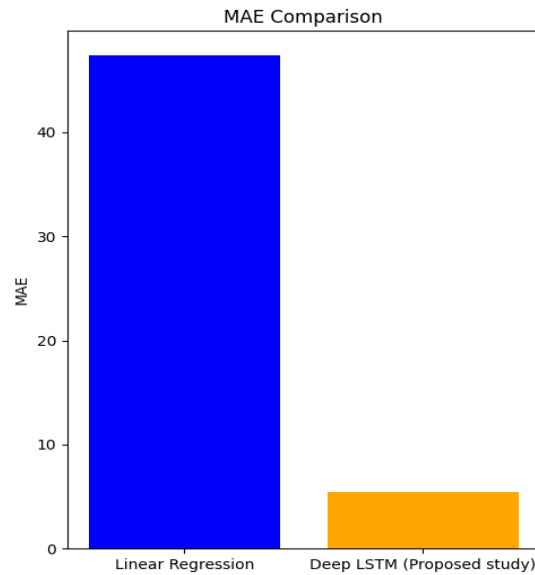


Figure 2. MAE for the proposed deep model and linear regression algorithm.

Visualising model performance shows how much better the proposed Deep LSTM model is than the more traditional Linear Regression method. The enormous gap between the two models is clearly demonstrated by the bar plot when looking at MAE.

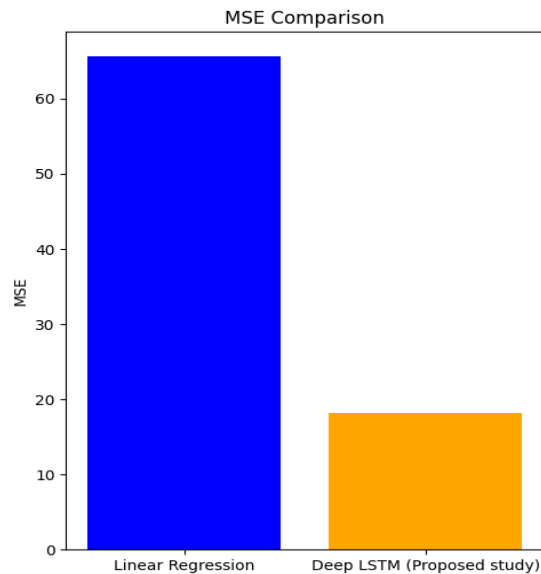


Figure 3. MSE for the proposed deep model and linear regression algorithm.

Linear Regression performs poorly with an enormous MAE of 47.352, however the proposed Deep LSTM model performs impressively better with an MAE of only 5.454. Compared to the more basic linear method of standard regression, the proposed Deep LSTM

model's superior performance is clear. This visual representation is a demonstration of the game-changing potential of deep learning for time-series forecasting.

Figure 4 displays the MAE, MSE, and RMSE at three different epochs (100, 200, and 300) for a variety of look-back periods (10, 20, 30, 40, and 50).

As the look-back period lengthens, the MAE, MSE, and RMSE for epoch 100 all tend to decrease. While both MSE and RMSE display variability in epoch 200, MAE increases, especially when considering a look-back of 30.

The MAE maintains its downward trend by epoch 300, and the MSE and RMSE exhibit stable behaviour. The choice of the look-back time effects the model's prediction performance, with a trade-off between short-term accuracy and capturing long-term relationships. These visualisations help assist the selection of an ideal look-back time based on the trade-off between bias and variance in the model.

Interesting insights on the model's generalisation performance can be gleaned from the outcomes of the

testing phase as shown in Figure 5. The MAE and RMSE values in epoch 100 are both low, indicating that the model has successfully learned patterns from the training data. However, both MAE and RMSE increase noticeably when we move on to epoch 200. This increase could indicate that the model is having trouble generalising to new data, or that the training approach needs to be tweaked to improve generalisation.

There is no clear pattern by epoch 300. It appears that the MAE is improving while the RMSE is still somewhat volatile. This trend emphasises the need for careful evaluation of model performance throughout numerous epochs, so that overfitting and underfitting can be detected.

These findings highlight the iterative nature of model learning and testing and suggest taking a systemic view when analysing the data. Insightful guidance for improving models for robust real-world applications can be gleaned from the interplay between MAE and RMSE.

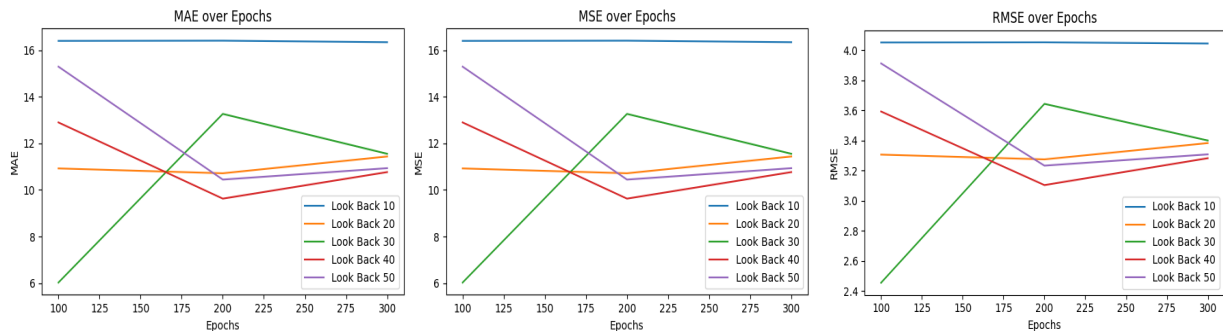


Figure 4 . Results from the training phase.

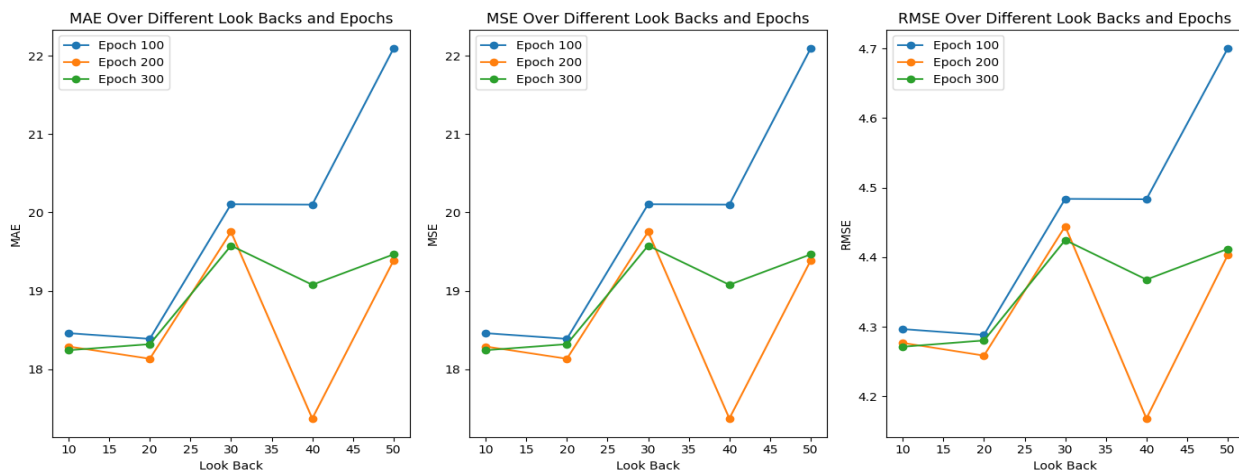


Figure 5 . Results from the testing phase.

## 5. CONCLUSION (SONUÇ)

The proposed deep LSTM model provides a clear leap in predictive accuracy. Additionally, the MAEs and MSEs are greatly reduced. Particularly, the deep LSTM model achieves a good MAE of 5.454 and an MSE of 18.243 while the linear regression counterpart is slow with an MAE of 47.352 and an MSE of 65.606. Such a big distinction highlights the ability of the advanced model of providing markedly better and more reliable forecasts, rather than plain linear regressions.

Even though it would be wrong not to recognize the strengths of the deep LSTM model, any study comes with its own limitations. The modelling is quite useful for a short-term power prediction but could have been distorted by the dynamic unpredictable factors. Thirdly, the nature of input data as well their presence affects the model effectiveness. In that order, future research endeavours can focus on addressing these problems in order to improve the reliability and suitability of this deep LSTM model.

Moreover, the study paves way for further research. Further research can be pursued in tuning model hyperparameters, examining additional features for better prediction accuracy, and also applying DL approaches across other forecasting horizons. Further, can contribute to the general discussion of the applicability and limitations of DL models in regression problems.

## REFERENCES (KAYNAKLAR)

- [1] S. G. Patil and M. S. Ali, "Review on Analysis of Power Supply and Demand in Maharashtra State for Load Forecasting Using ANN," *Int J Sci Res Sci Technol*, vol. 9, no.1, pp. 341-347, 2022, Doi: 10.32628/ijrsrst229152.
- [2] B. U. Islam, M. Rasheed, and S. F. Ahmed, "Review of Short-Term Load Forecasting for Smart Grids Using Deep Neural Networks and Metaheuristic Methods," *Mathematical Problems in Engineering*, vol. 2022, 4049685, 2022. Doi: 10.1155/2022/4049685.
- [3] B. Yildiz, J. I. Bilbao, and A. B. Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 1104-1122, 2017. Doi: 10.1016/j.rser.2017.02.023.
- [4] A. Azeem, I. Ismail, S. M. Jameel, F. Romlie, K. U. Danyaro, and S. Shukla, "Deterioration of Electrical Load Forecasting Models in a Smart Grid Environment," *Sensors*, vol. 22, no. 12, 4363, 2022, Doi: 10.3390/s22124363.
- [5] A. Talupula, "Demand Forecasting of Outbound Logistics Using Machine learning," *Faculty of Computing, Blekinge Institute of Technology, Karlskrona, Sweden*, February, 2018.
- [6] I. Zuleta-Elles, A. Bautista-Lopez, M. J. Catano-Valderrama, L. G. Marin, G. Jimenez-Estevez, and P. Mendoza-Araya, "Load Forecasting for Different Prediction Horizons using ANN and ARIMA models," in *2021 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies, CHILECON 2021*, 2021. Doi: 10.1109/CHILECON54041.2021.9702913.
- [7] M. L. Abdulrahman *et al.*, "A Review on Deep Learning with Focus on Deep Recurrent Neural Network for Electricity Forecasting in Residential Building," in *Procedia Computer Science*, vol. 193, pp. 141-154, 2021. Doi: 10.1016/j.procs.2021.10.014.
- [8] G. Chitalia, M. Pipattanasomporn, V. Garg, and S. Rahman, "Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks," *Appl Energy*, vol. 278, 115410, 2020, Doi: 10.1016/j.apenergy.2020.115410.
- [9] Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, "A deep learning method for short-term residential load forecasting in smart grid," *IEEE Access*, vol. 8, pp. 55785-55797, 2020, Doi: 10.1109/ACCESS.2020.2981817.
- [10] B. Dietrich, J. Walther, Y. Chen, and M. Weigold, "A deep learning approach to electric load forecasting of machine tools," *MM Science Journal*, vol. 2021-November, 2021, Doi: 10.17973/MMSJ.2021\_11\_2021146.
- [11] G. Hafeez *et al.*, "Short term load forecasting based on deep learning for smart grid applications," in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2019, pp. 276-288. Doi: 10.1007/978-3-319-93554-6\_25.
- [12] C. P. Joy, G. Pillai, Y. Chen, and K. Mistry, "Micro-genetic algorithm embedded multi-population differential evolution based neural network for short-term load forecasting," in *2021 56th International Universities Power Engineering Conference: Powering Net Zero Emissions, UPEC 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., Aug. 2021. Doi: 10.1109/UPEC50034.2021.9548262.

- [13] X. Luo and L. O. Oyedele, "A self-adaptive deep learning model for building electricity load prediction with moving horizon," *Machine Learning with Applications*, vol. 7, p. 100257, Mar. 2022, Doi: 10.1016/j.mlwa.2022.100257.
- [14] F. Bayram, P. Aupke, B. S. Ahmed, A. Kassler, A. Theocharis, and J. Forsman, "DA-LSTM: A dynamic drift-adaptive learning framework for interval load forecasting with LSTM networks," *Eng Appl Artif Intell*, vol. 123, Aug. 2023, Doi: 10.1016/j.engappai.2023.106480.
- [15] Ernesto Aguilar Madrid, "Short-term electricity load forecasting (Panama)."
- [16] H. Henderi, "Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer," *IJIIS: International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 13–20, Mar. 2021, Doi: 10.47738/ijjis.v4i1.73.
- [17] C. Xiong, H. Sun, D. Pan, and Y. Li, "A personalized collaborative filtering recommendation algorithm based on linear regression," *Mathematical Modelling of Engineering Problems*, vol. 6, no. 3, 2019, Doi: 10.18280/mmep.060307.
- [18] Yılmaz, Y. Doğrusal Regresyon Modeli. Teori ve Uygulamada Makine Öğrenmesi, (21-36), Nobel Akademik Yayıncılık, Ankara, 2022.
- [19] A. de Myttenaere, B. Golden, B. Le Grand, and F. Rossi, "Mean Absolute Percentage Error for regression models," *Neurocomputing*, vol. 192, pp. 38-48, 2016, Doi: 10.1016/j.neucom.2015.12.114.
- [20] T. O. Hodson, T. M. Over, and S. S. Foks, "Mean Squared Error, Deconstructed," *J Adv Model Earth Syst*, vol. 13, no. 12, e2021MS002681, 2021, Doi: 10.1029/2021MS002681.
- [21] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geoscientific Model Development*, vol. 15, no. 14. 2022. Doi: 10.5194/gmd-15-5481-2022.



## Yapay Zeka Algoritmaları Kullanılarak Öğrencilerin Akademik Başarısı ile Stres İlişkisinin Keşifsel Bir Analizi

Hakan YÜKSEL<sup>\*,a</sup>,

<sup>a,\*</sup> Isparta Uygulamalı Bilimler Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Isparta, 32100, Türkiye

### MAKALE BİLGİSİ

Alınma: 13.12.2023  
Kabul: 25.12.2023

#### **Anahtar Kelimeler:**

Yapay zeka, keşifsel veri analizi, stres, akademik başarı, makine öğrenmesi

#### **\*Sorumlu Yazar**

e-posta:  
hakanyuksel@isparta.edu.tr

### ÖZET

Günümüzde stres, insanların yaşamını ve yaşam kalitelerini önemli ölçüde etkileyen bir unsur olarak karşımıza çıkmaktadır. Öğrenciler ise günlük yaşam koşullarının yanı sıra akademik başarı olarak da stres düzeylerini dengelemek durumundadır. Bu noktada literatür incelendiğinde akademik çalışmalarda stres öğrenciler için psikolojik, fizyolojik, çevresel, akademik ve sosyal faktörler kapsamında değerlendirilmektedir. Bu faktörlerin alt düzey durumları da bulunmaktadır. Öğrencilerin stres düzeylerinin bulunduğu veri setinde bu faktörlere ait 20 alt başlık bulunmaktadır. Öğrenci stres veri seti üzerinde keşifsel bir veri analizi gerçekleştirilmiştir. Keşifsel veri analizi (EDA- Exploratory Data Analysis), verilerde gizlenmiş olan bilgilere ulaşmak için verileri keşfetme işlemidir. Bu çalışmada öğrenci stres seviyelerinin, akademik başarı ile olan ilişkisi keşifsel bir veri analizi ile gerçekleştirilmiştir. Elde edilen değerler görsel grafiklerle desteklenmiştir. Ayrıca veri seti üzerinde makine öğrenme algoritmalarından; Karar Ağaçları, Rassel Orman, K-En yakın komşu ve Gaussian Naive Bayes algoritmalarının karşılaştırmalı analizi yapılmıştır. En iyi öğrenim gösteren algoritma ise Gaussian Naive Bayes algoritması olmuştur.

DOI: 10.59940/jismar.1404452

## An Exploratory Analysis of the Relationship Between Students' Academic Success and Stress Using Artificial Intelligence Algorithms

### ARTICLE INFO

Received: 13.12.2023  
Accepted: 25.12.2023

#### **Keywords:**

Artificial intelligence, exploratory data analysis, stress, academic success, machine learning

#### **\*Corresponding Authors**

e-mail:  
hakanyuksel@isparta.edu.tr

### ABSTRACT

Today, stress emerges as an element that significantly affects people's lives and quality of life. Students, on the other hand, have to balance their stress levels in terms of academic success as well as daily living conditions. At this point, when the literature is examined, stress in academic studies is evaluated within the scope of psychological, physiological, environmental, academic and social factors for students. These factors also have lower-level states. There are 20 subheadings related to these factors in the data set containing students' stress levels. An exploratory data analysis was conducted on the student stress dataset. Exploratory data analysis (EDA) is the process of exploring data to find information hidden in the data. In this study, the relationship between student stress levels and academic success was conducted through an exploratory data analysis. The obtained values are supported by visual graphics. In addition, from machine learning algorithms on the data set; A comparative analysis of Decision Trees, Random Forest, K-Nearest Neighbor and Gaussian Naive Bayes algorithms was made. The best learning algorithm was the Gaussian Naive Bayes algorithm.

DOI: 10.59940/jismar.1404452

## 1. GİRİŞ (INTRODUCTION)

İnsanlar, varlığından bu yana hayatta kalmak için sürekli mücadele etmektedir. Tarih öncesi çağlarda avlanma, barınma ve iklim değişikliğine uyum sağlamakla başlayıp günümüzde ise yaşamlarının büyük çoğunluğunu, sınırlarını ve yeteneklerini zorlayarak geçirmektedir. 20. yüzyıldan itibaren gündelik hayata eklenen hız ve hareketlilik, insanların değişim ve rekabet halinde çalışmalarını da beraberinde getirmiştir. Dolayısıyla birçok durum, insanlık için temel stresleyici hale gelmiştir. Akademik olarak incelendiğinde; çalışma ortamı, rakipler, organizasyon yapısı, rol talepleri ve görev talepleri stresin temel nedenleri oluşturmaktadır. Günümüz gerçekleri göz önüne alındığında kişilerin stresle karşılaşması kaçınılmaz olmuştur [1].

Stres, günlük yaşamın akademik kısmında yaygın olarak kullanılmaktadır. Latince “stringere” kelimesinden türeyen ve sıkı, sıkıştırmak anlamına gelmektedir. Etimolojik olarak “stres” kelimesinin, İngilizcenin, Norman Fransız ve Anglo-Sakson karışımı ile geliştirildiği ve kökeninin 14. yüzyıla uzandığı bilinmektedir [2]. Başka bir görüşe göre stres, Latince “stringere” kelimesinden değil, “estricia” kelimesinden türemiştir. “Estricia” ise güç vermek, önem vermek, baskı yapmak demektir [3]. Stresi bir terim olarak tanımlamak, birçok disiplinin ilgisidir. Bu nedenle literatürde farklı tanımlar görülmektedir. Stres tanımları incelendiğinde üç kategoriye ayrıldığı gözlenmiştir. İlk kategori, stresi durumsal veya çevresel uyarıcı olarak gören, uyarıcı temelli ifade etmektedir. İkinci kategori, stresin bireyin çevresel/durumsal güçlere karşı psikolojik ya da fizyolojik tepkisi olduğunu tanımlamaktadır. Üçüncü kategori ise, stresi hem uyarıcı hem de yanıt olarak tanımlayan stresör (stres tepkisinin gelişmesine neden olan çevresel faktörler) olarak bilinmektedir [4]. Stres dolaylı unsurlar aracılığıyla oluşmaktadır. Bu unsurların ortaya çıkmasını “stresörler” adı verilen durumlar tetiklemektedir. Stresin ortaya çıkması için iki temel koşul sağlanmalıdır. Bunlardan ilki sonucun belirsiz olması, ikincisi de sonucun birey için önem arz etmesidir [5]. Dewe, O’Driscoll ve Cooper (2012), bir stresör, tepki ya da sonucun belirsiz veya önemli olması arasındaki etkileşimle ilgili stres tanımlarını sınıflandırmıştır [6].

Stres, ilk defa fizikçi Robert Hook tarafından 17. yüzyılda esneyen nesne ve ona uygulanan dış kuvvet arasındaki durumu açıklamak için kullanılmıştır. Thomas Young ise 18. yüzyılda stresi, dışarıdan uygulanan bir kuvveti dengeleyen elastik beden içindeki kuvvet olarak açıklamıştır [7]. 19. yüzyılda

stres, fizikte zorlanma veya deformasyona neden olan nesneye uygulanan basıncı tanımlamak için kullanılan bilimsel bir terim haline gelmiştir [8]. Ayrıca, 19. yüzyılın sonunda, Avrupalı fizyologlar (Bernard, Pflüger ve Frederic vd.) stresi, canlı organizmaya sunulan dinamik zorluklarla benzer nitelikte olarak ifade etmiştir [7]. 1900’lerin başlarında fizyolog Cannon stresi, zihinsel baskı veya bozulma olarak kullanmıştır. Cannon, “homeostaz” kelimesini, vücudun sıcaklığını nasıl koruduğunu, diğer anahtar değişkenleri referans olarak belirtmiştir [4, 8]. Selye ise, stresi vücudun herhangi bir değişim talebine karşı kapsamlı bir cevap olarak nitelendirmiştir. Selye, Cannon’ın çalışmalarını ilerleterek, stres yanıtının aşamalarını açıklamıştır. Bunlar; alarm aşaması, uyarlanabilir direnç aşaması ve bitkinlik aşamasıdır [8]. De Frank ve arkadaşlarına göre stres, birey ile çevreleri arasındaki etkileşimdir. Bu ortam, bireyin fiziksel ve zihinsel durumunu etkileyen duygusal baskı ile karakterizedir [9]. Cranwell (1998) stresi, bir insanın başa çıkma kapasitesinin ötesinde algılanan bir tehdide veya istenmeyen bir duruma karşı bilinçli veya bilinçaltı olarak hem fizyolojik hem de psikolojik bir tepki olarak nitelendirmiştir [10].

Genel bir anlam olarak stres, strese neden olan farklı iç ve dış faktörlerle duygusal ve fizyolojik bir tepki ve çevreden gelen baskıdır. Bu baskı ortamı harici veya dâhili olabilmektedir [11]. Ayrıca, bireyin yeteneği, strese göre istenmeyen durumlarla mücadele edemediğinde ortaya çıkmaktadır. Bir kişinin strese tepkisi ya onunla mücadele etmek ya da ondan kaçmak olarak belirtilmektedir. İnsan vücudu bir durumu tehdit olarak algıladığında, fiziksel ve psikolojik değişiklikler başlamaktadır. Birey bu duruma karşı koyamazsa, hastalıklar strese maruz kalmanın bir sonucu olarak ortaya çıkmaktadır [12].

Başarı terimi kişinin hedeflerine ulaşması olarak ifade edilmektedir. Akademik başarı ise öğrencilerin sınırlarını zorlamaları ve kendilerine verilen doğru eğitim ve öğretimle başarıya ulaşmaları anlamına gelmektedir [13]. Bu tanımdan yola çıkarak eğitim ve öğretimin amacı, günümüzün kültürel çağına ve toplumuna uyum sağlayabilen, insanların ihtiyaçlarına çözüm bulabilen bireyler yetiştirmektir. Buna göre eğitim kurumlarında öğrencilerin hedefi, hızla değişen ve karmaşık bir toplumun sorun ve sorunlarıyla yüzleşmek için teorik bilgileri uygulayabilen ve bu ortama zorluk çekmeden uyum sağlayabilen bireyler yetiştirmektir [14, 15]. Geliştirilecek eğitim-öğretim programı, öğrencilerin öğretim programının öğrenimine aktif katılımını ve bilgiye ulaşabilen, eriştiği bilgiyi kullanabilen ve paylaşabilenlerin yetiştirilmesini esas almalıdır [16].

Bu gelişim süreci aslında insanların topluma öncelikle “Akademik Başarı” kazandırılması gerektiğini tespit eden bir durumdur. Bireyin topluma uyum sağlayabilmesini, toplumsal sorunlara zorlanmadan çözüm bulabilmesini ve bu yolda engelleri kolaylıkla aşabilmesini sağlayan önemli faktörlerden biri de iyi bir akademik başarının gerekliliğidir. İyi bir akademik başarıya ulaşmak, iyi bir eğitmen, iyi bir bilgi birikiminden ve bu bilginin iyi bir şekilde uygulamaya konulmasından geçmektedir [17].

Akademik başarı, üniversite öğrencilerinin tüm eğitim kariyerinde önemli bir rol oynamaktadır. Bunun nedeni üniversite öğrencilerinin eğitim hayatları boyunca aldıkları dersler, öğrendikleri bilgiler ve sınavlardaki başarılarıyla en iyi olmayı ve akademik olarak başarılı olmayı hedeflemeleridir. Üniversite öğrencilerinin hayatlarında akademik başarının önemli bir rol oynamasının bir diğer nedeni de iş dünyasındaki gelecekteki kariyerlere yön verirken onların akademik başarılarının da dikkate alınmasıdır [18, 19].

Bu çalışmada üniversite öğrencilerinden oluşan bir grubun “akademik başarının” stres üzerindeki etkisi yapay zeka yöntemleriyle incelenmiş ve ortaya stresin akademik başarıyı nasıl etkilediği ile ilgili sonuçlara ulaşılmıştır. Çalışma aynı zamanda özellikle pandemi sonrası oluşan çevrimi eğitim süreçlerinde öğrencilerin nasıl etkilendiği ve bu süreçten sonra nasıl bir yol izlenmesi gerektiği ile ilgili stresin nasıl bir gösterge oluşturduğunu ifade etmektedir.

## 2. LİTERATÜR ÖZETİ (LITERATURE VIEW)

Günümüzde, öğrenciler günlük yaşam becerilerinde karşılaştıkları zorlukların yanı sıra ayrıca akademik isteklerden kaynaklı strese de maruz kalmaktadır. “Akademik Başarı” stresi, öğrencilerin eğitim-öğretim süreçlerinde yaşadığı, endişe, yoğun baskı ve kaygı durumu olarak belirtilmektedir [20]. Stresli ve karmaşık durumlar içerisinde akademik bilgiyi hatırlama, doğru kararlar alabilme ve uygun/verimli problem çözme becerisi kazandırmak oldukça önemlidir [21-23]. Öğrencilerin akademik başarıları üzerinde stres psikolojik, fizyolojik, çevresel ve sosyal faktörlerin etkisi açısından incelenmiştir [24-26]. Akademik başarı stresinin belirtileri incelendiğinde sürekli konsantrasyon güçlüğü, endişe ve kaygı duyma, motivasyon kaybı, özgüven yetersizliği, uykusuzluk gibi fiziksel semptomlar, duygusal değişimler ve sosyal geri çekilme yer alabilir [27, 28].

Doğan (2020), Covid 19 süresince öğrencilerin akademik kaygılarının ölçümlenmiştir. 18-25 yaş

aralığındaki öğrencilerden rastgele belirlenmiş 100 adet yazılan mektubun nitel araştırma yöntemleriyle değerlendirmişlerdir. Bu değerlendirme süreci öğrencilerin Covid 19 süresince akademik kaygı düzeylerinin yüksek olduğu ve bu durum sonucunda korku ve yalnızlık duygusu yaşadığı ortaya konmuştur [29].

Öğrenciler, kendilerini sürekli olarak değerlendirmek zorunda kalabilecekleri yoğun bir rekabet ortamında büyük bir başarı baskısı hissedebilirler. Bu durum, sınav kaygısı olarak ifade edilebilecek endişeyle birlikte gelir [28]. Uzun süre ve kontrolsüzce devam eden akademik stres, öğrencilerin genel yaşam kalitesini olumsuz etkileyebilir. Motivasyon kaybı, tükenmişlik, depresyon ve kaygı bozukluğu gibi daha ciddi sorunlar ortaya çıkabilir [30]. Bu nedenle, akademik stresin önlenmesi veya yönetilmesi önemlidir; böylece öğrenciler sağlıklı bir öğrenme deneyimi yaşayabilir ve başarılarını sürdürebilirler. Bu bağlamda, okula yönelik baskı ve stresle mücadele etmek; özellikle yoğun çalışma temposu, dersler, sınavlar ve diğer okul sorumlulukları nedeniyle enerjisi tamamen tükenmiş olan öğrenciler için önemlidir.

## 3. MATERYAL ve YÖNTEM (MATERIAL and METHOD)

### 3.1. Veri Seti (Dataset)

Çalışma kapsamında Kaggle veri tabanında yer alan ve öğrenci stresi olarak geçen veriseti ele alınmıştır [31]. Stres öğrenciler için büyük bir sorun teşkil etmektedir. Bu yüzden bu veri setinde öğrencileri strese sokabilecek 20 değer ile stres durumunun ortaya konulması beklenmektedir. Veri seti 1100 satır (öğrenci) ve 21 sütundan oluşmaktadır.

Bu veriseti, öğrencinin stres üzerinde en fazla etki oluşturan yaklaşık 20 çeşit özellik barındırmaktadır (Tablo 1). Özellikler 5 ana faktör dikkate alınarak belirlenmiştir. Bunlar psikolojik, fizyolojik, sosyal, çevresel ve akademik faktörlerdir. Bu faktörlerde kendi içerisinde bazı alt başlıklara ayrılmıştır. Verisetinin 21. Sütunu ise stres seviyesini belirtmektedir. İlk 20 sütun verisetinin özellikleri olarak ifade edilirken, 21. sütundaki stres seviyesi ise etiket olarak yer almaktadır. Veri seti üzerinde 20 farklı stres durumunu ortaya koymaktadır. Bu özellikler şu şekildedir;

- *Psikolojik Faktörler*; ‘kaygı düzeyi’, ‘benlik saygısı’, ‘zihinsel sağlık geçmişi’, ‘depresyon’,
- *Fizyolojik Faktörler*; ‘baş ağrısı’, ‘kan basıncı’, ‘uyku kalitesi’, ‘solunum problemi
- *Çevresel Faktörler*; ‘gürültü düzeyi’, ‘yaşam koşulları’, ‘güvenlik’, ‘temel ihtiyaçlar’,

- *Akademik Faktörler*; ‘akademik performans’, ‘çalışma yükü’, ‘öğretmen öğrenci ilişkisi’, ‘gelecek kariyer kaygıları’,
- *Sosyal Faktör*; ‘sosyal destek’, ‘akran baskısı’, ‘müfredat dışı aktiviteler’, ‘zorbalık’.

Tablo 1. Öğrenci stres seviyesi veriseti örnekleme  
(Student stress level dataset sample)

number	anxiety_level	self_esteem	mental_health_history	depression	headache	blood_pressure	sleep_quality	breathing_problem	noise_level	living_conditions	security	basic_needs	academic_performance	study_load	teacher_student_relationship	future_career_concerns	social_support	peer_pressure	extracurricular_activities	bullying	stress_level
0	14	20	0	11	2	1	2	4	2	3	2	2	3	2	3	3	2	3	3	2	1
1	15	8	1	15	5	3	1	4	3	1	3	2	1	4	1	5	1	4	5	5	2
2	12	18	1	14	2	1	2	2	2	2	1	2	2	3	3	2	2	3	2	2	1
3	16	12	1	15	4	3	1	3	4	2	3	2	2	4	1	4	1	4	4	5	2
4	16	28	0	7	2	3	5	1	3	2	2	3	4	3	1	2	1	5	0	5	1

### 3.2. Veri Ön İşleme (Data Preprocessing)

Bu çalışmada kullanılan, 1100 satır ve 21 sütundan oluşan veriseti hem hatalı hem de eksik değer hususunda gözden geçirilmiştir. Bu kapsamda pandas kütüphanesine ait duplicate() ve isna() fonksiyonları kullanılmıştır.

Pandas kütüphanesinde yer alan duplicate() fonksiyonu, yinelenen satırları bulmak, çıkarmak, saymak ve kaldırmak için kullanılmıştır. Diğer fonksiyon isna() ise eksik değerlerin tespitinde kullanılmıştır. DataFrame'deki her öge için, bir ögenin NA değeri olup olmadığını belirten bool değerlerinin maskesini vermektedir. İlgili kod blokları Şekil 1'de gösterilmiştir.

```
In [3]: # Checking for duplicates
ds.duplicated().sum()

In [4]: # Checking if we have missing values
ds.isna().sum()
```

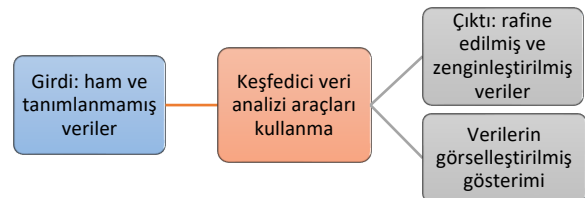
Şekil 1. Veri ön işlem komutları  
(Data preprocessing commands)

Veri kümesinde hem eksik değer hem de tekrarlanan değere rastlanılmamıştır. Dolayısıyla veri seti üzerinde farklı bir ön işlem aşaması bulunmamaktadır.

### 3.3. Keşifsel Veri Analizi (Exploratory Data Analysis)

Keşifsel veri analizi (EDA-Exploratory Data Analysis), bir kullanıcının tablo halindeki bir veriye karşı bir analiz eylemi yayınladığı, bazı yanıtlar aldığı, tablo halindeki verilerin alt bölümünü aldığı etkileşimli ve sıralı bir veri anlama ve içgörü oluşturma sürecidir [32]. Bu süreç sonunda bazı görselleştirmeler yapar ve ardından verilerin gizli özelliklerini ve ilişkili içgörülerini daha iyi anlamak için hangi sorguların yayınlanacağına karar verir [33, 34].

Keşifsel veri analizi, verilerde gizlenmiş olan bilgilere ulaşmak için verileri keşfetme sürecidir [35-38]. Bu sürecin algoritması Şekil 2'de gösterilmektedir [39].



Şekil 2. Keşifsel veri analizinin algoritması  
(Algorithm of exploratory data analysis)

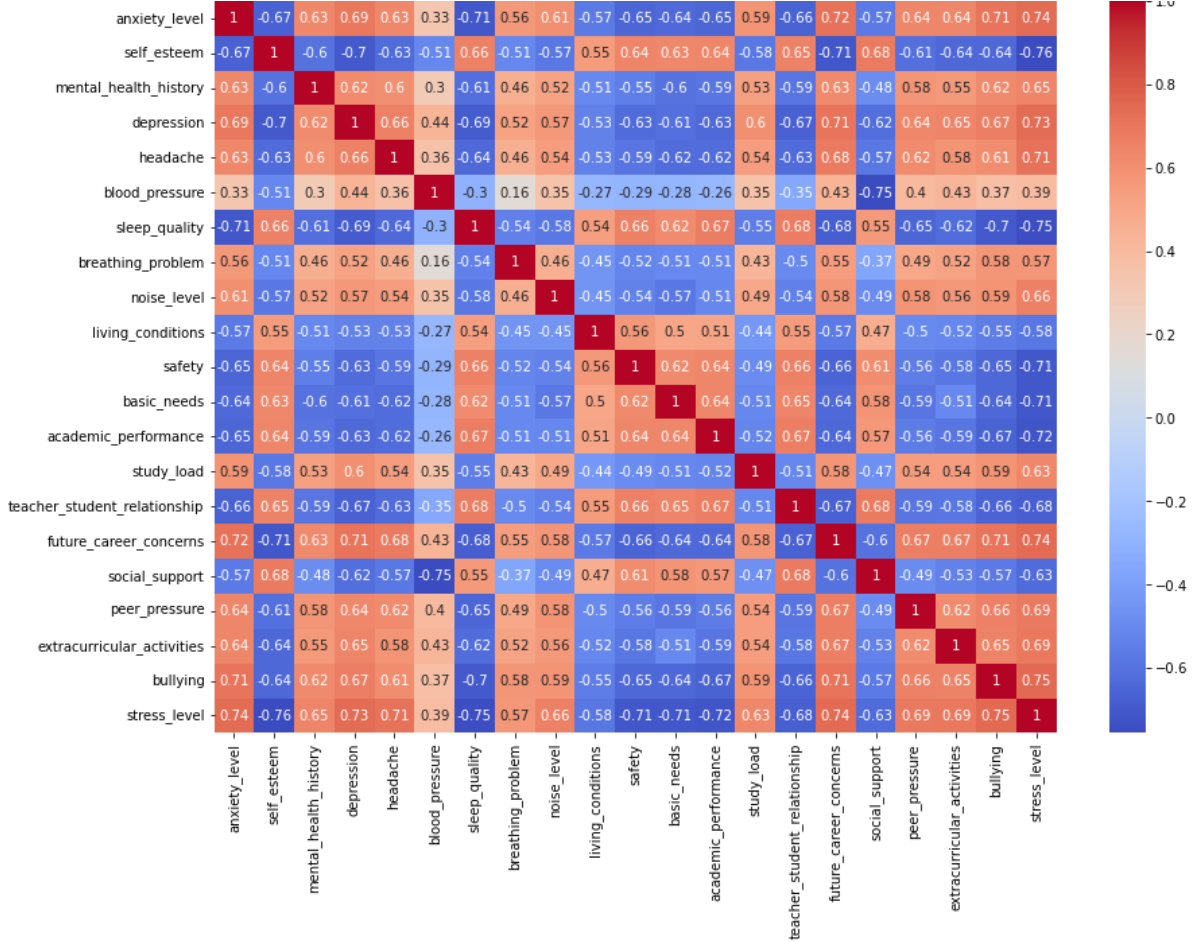


### 3.3.1. Öğrenci stres düzeyi karmaşıklık matrisi

(Student stress level confusion matrix)

Veriseti üzerinde, öğrenci stresine neden olan durum Şekil 3'te gösterilen karmaşıklık matrisinde gösterilmiştir. Karmaşıklık matrisi incelendiğinde, “kaygı, depresyon, gelecek kariyer kaygıları, zorbalık ve baş ağrıları” durumlarının artan stres seviyeleriyle yüksek oranda ilişkili olduğu gözlenmektedir.

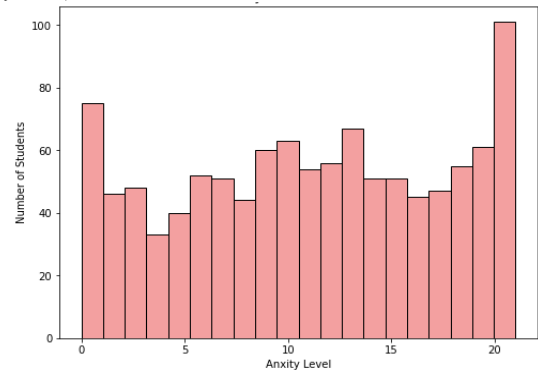
Elde edilen ısı haritası, stres üzerinde azaltıcı etkiye sahip görünen faktörler için de farklı bir model ortaya koymaktadır. “İyi bir öz saygıya, güçlü sosyal destek ağlarına, olumlu öğretmen-öğrenci ilişkilerine ve daha iyi uyku kalitesine” sahip olan öğrenciler, daha düşük stres seviyelerini yansıtmaya eğiliminde olduğu ortaya çıkmıştır.



Şekil 3. Stres faktörlerinin karmaşıklık matrisi  
(confusion matrix of stress)

### 3.3.2. Öğrenci stresini üzerinde kaygı düzeyi ve depresyon etkisinin incelenmesi (Examining the effect of anxiety level and depression on student stress)

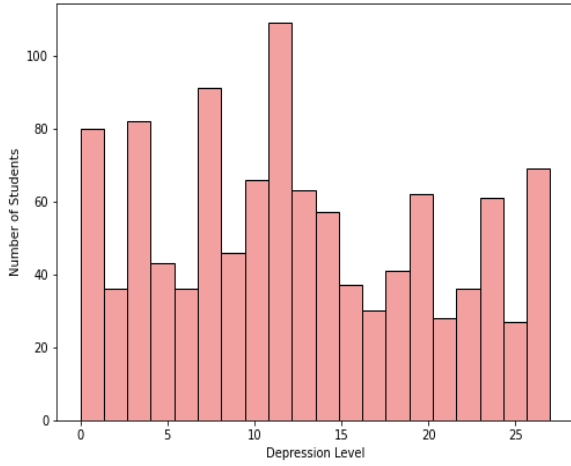
Veri setindeki öğrencilerin ortalama kaygı düzeyi, verilerden de anlaşılacağı üzere çoğunluğunun bir dereceye kadar kaygı yaşadığı görülmektedir. Şekil 4'te de görüleceği üzere, öğrencilerin önemli bir kısmının yüksek düzeyde kaygıyla karşı karşıya kalmasıdır. Veri setindeki öğrencilerin ortalama kaygısı 11.06 olarak tespit edilmiştir.



Şekil 4. Öğrenci stresinin kaygı düzeyi verisi  
(Anxiety level data of student stress)

Yine veri seti incelendiğinde öğrencileri stres düzeyinde, depresyon etkisinin yoğun bir görüntü sunduğu gözlenmektedir. Öğrencilerin birçoğunun yüksek düzeyde depresyona yakın olduğu görülmektedir (Şekil 5).

Toplam 1.056 kişiden oluşan öğrencilerin yaklaşık %96'sı hayatlarının bir noktasında depresyonla karşı karşıya kalmıştır. Bu durum, birçok öğrencinin günlük olarak karşılaştığı önemli sorunun altını çizmektedir.

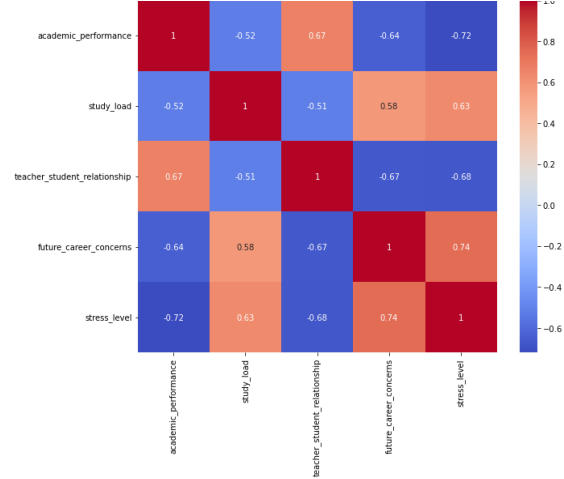


Şekil 6. Öğrenci stresinin depresyon düzeyi seviyesi  
(Depression level level of student stress)

### 3.3.3. Öğrenci stres düzeyine etki eden akademik faktörler (Academic factors affecting student stress level)

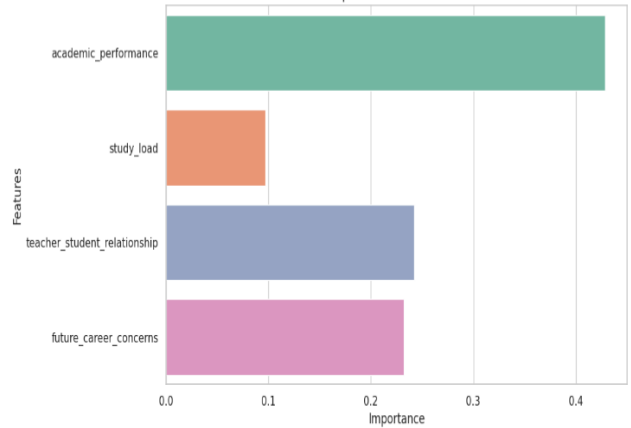
Akademik faktörler arasında akademik performans stres düzeyleriyle en güçlü korelasyonu sergilemektedir (Şekil 7). Bu durum aslında akademik performansı dikkate alan öğrencilerin yüksek stres düzeyi bildirme olasılıklarının daha yüksek olduğunu göstermektedir. Çalışma yükü ve gelecek kariyer kaygıları stresle pozitif ilişkiler göstermesi, bu akademik faktörlerin öğrencilerde artan stres düzeylerine de katkıda bulunabileceğini düşündürmektedir.

Aynı zamanda, elde edilen veriler potansiyel olarak güven verici bir eğilimi ortaya koymaktadır. Öğretmen öğrenci ilişki durumu daha iyi olan öğrenciler daha düşük stres seviyeleri sergileme eğilimindedir (Şekil 8). Bu ilişki, öğrencilerin daha iyi öğretmen öğrenci ilişki durumunda sahip olanların stres seviyelerini azaltmaya katkı da bulunabileceğini ima etmektedir.



Şekil 7. Akademik faktörlerin stres seviyesi kapsamında karmaşıklık matrisi  
(Complexity matrix within the context of stress level of academic factors)

- Akademik performansı ortalamasının altında olan öğrenci sayısı: 561,
- Öğrenciler tarafından bildirilen ortalama çalışma yükü: 2,62,
- Gelecekteki kariyerleriyle ilgili kaygısı olan öğrenci sayısı: 327.



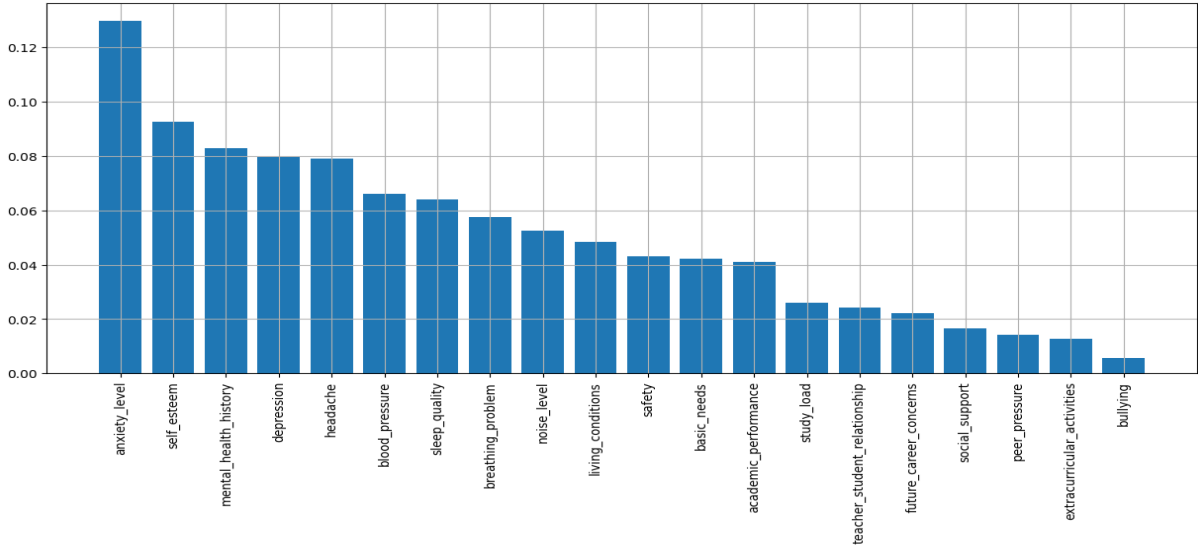
Şekil 8. Öğrenci stresinin akademik faktörlerinin özellik önemi  
(Feature importance of academic factors of student stress.)

### 3.3.4. Öğrenci stres üzerindeki en önemli etkenler (The most important factors on student stress)

Bu noktada öğrenci stres veri seti üzerinde özellik önem analizi yapılmıştır. Bu analiz, her bir faktördeki hangi özelliğin en önemli etkiye sahip olduğunu belirlemek için kullanılmıştır (Şekil 9). Sonuçlar Tablo 2'de gösterilmiştir.

Tablo 2. Öğrenci stres seviyesi üzerinde önem değerleri  
(Importance values on student stress level)

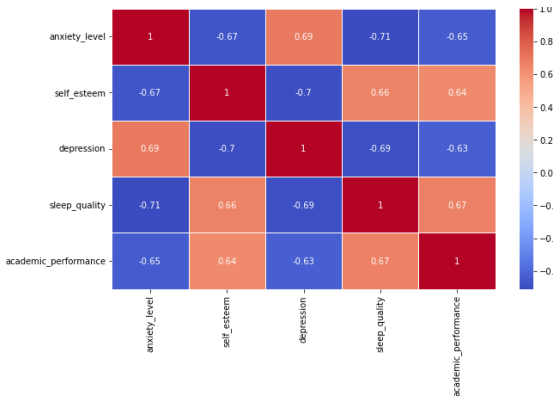
Önem Değeri														
Psikolojik	anxiety_level	0.26	Fizyolojik	headache	0.40	Çevresel	noise_level	0.15	Akademik	academic_performance	0.43	Sosyal	social_support	0.09
	self_esteem	0.50		blood_pressure	0.05		living_condition	0.12		study_load	0.10		peer_pressure	0.16
	mental_health_history	0.02		sleep_quality	0.47		safety	0.31		teacher_student_relationship	0.24		extracurricular_activities	0.25
	depression	0.22		breathing_problem	0.07		basic_needs	0.41		future_career_concerns	0.23		bullying	0.50



Şekil 9. Öğrenci stres seviyesi üzerinde önem değerleri  
(Importance values on student's stress level)

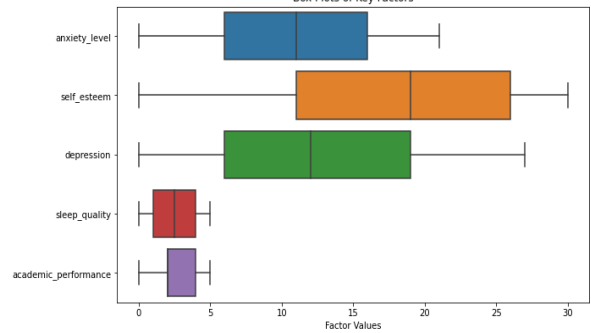
### 3.3.5. Öğrenci stres düzeyinin temel faktörlerin korelasyonu (Correlation of student stress level with basic factors)

Temel faktörler arasındaki korelasyonları görselleştirmek için bir ısı haritası oluşturun. Bu, güçlü olumlu veya olumsuz ilişkileri tanımlamanıza yardımcı olmaktadır (Şekil 10).



Şekil 10. Öğrenci stres seviyelerinin temel durumları  
(Baselines of student stress levels)

Faktörlerdeki eğilimleri ve farklılıkları belirlemek için kutu grafiklerini kullanın. Bu, dağılımları karşılaştırmak için kullanışlıdır (Şekil 11).



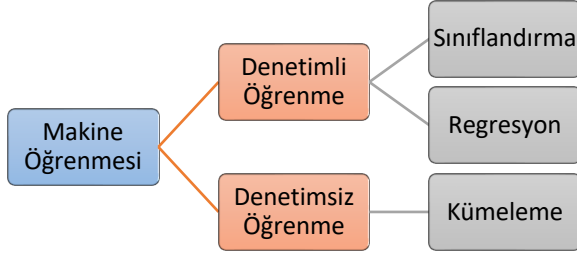
Şekil 11. Öğrenci stres seviyelerinin temel durumları kutu grafiği  
(Key cases boxplot of student stress levels)

### 3.5. Makine Öğrenmesi (Machine Learning)

İnsanlar, bilgi birikimlerinin bazılarını deneyim yoluyla algılamaktadır. Makine öğrenimi de robotlara aynısını yapmayı öğretmektedir. Makine öğrenimi

algoritmaları, model olarak bir denkleme güvenmek yerine, bilgiyi doğrudan verilerden "öğrenmek" için bilgisayar tekniklerini kullanmaktadır.. Algoritmalar, öğrenme için mevcut örneklerin sayısı kadar performanslarını dinamik olarak artırabilir [40].

Makine öğreniminde, Şekil 12’de gösterildiği gibi iki teknikten oluşmaktadır. Denetimsiz öğrenme, girdi verilerindeki gizli kalıpları veya içsel yapıları aramaktadır. Buna karşılık, denetimli öğrenme, gelecekteki çıktıları tahmin etmek için bilinen girdi ve çıktılara sahip verileri kullanan bir modelin eğitilmesini içermektedir [41].



Şekil 12. Makine Öğrenmesi Kategorileri  
(Machine Learning Categories)

Veri seti üzerinde, veri ön işleme işlemleri tamamlandıktan sonra veri setinin kullanılacak makine öğrenmesi algoritmaları ile uyumu da göz önüne alınarak, Karar Ağaçları, Rassal Orman, K-En yakın komşu ve Gaussian Naive Bayes öğrenme algoritmaları tercih edilmiştir [42-46].

Veriseti üzerinde sırasıyla %80 eğitim- %20 test, %70 eğitim- %30 test, %60 eğitim- %40 test değerlendirmeleri yapılmıştır. Veri seti üzerinde ilgili makine öğrenmesi algoritmalarının başarı durumları her bir eğitim ve test oranı doğrultusunda tahminlenmiştir. Değişen test verisi oranlarında en düşük başarı oranının, %20 ve %40 test verisinin makine öğrenmesi algoritmalarında uygulanması sonucu elde edildiği gözlemlenmiştir. Test verisinin %30 olarak belirlenmesi sonucu elde edilen doğruluk değerleri ise Tablo 3’ te verilmiştir. Değerler incelendiğinde ise makine öğrenmesi algoritma seçimini etkileyecek çok büyük farkların olmadığı gözlemlenmiştir. Bu algoritmalar arasında Gaussian Naive Bayes algoritması ise diğer algoritmalara göre daha iyi sonuçlar ortaya koymuştur.

Tablo 3. Makine öğrenmesi algoritmalarının karşılaştırılması  
(Comparison of machine learning algorithms)

	Eğitim Doğruluk	Test Doğruluk
Karar Ağaçları	1.0	0.893
Rassal Orman	1.0	0.9
K-En yakın komşu	0.876	0.887
Gaussian Naive Bayes	0.875	0.903

#### 4. TARTIŞMA ve SONUÇ ((DISCUSSION and CONCLUSION)

Öğrencilerin, eğitim-öğretim hayatındaki akademik başarıları stres faktörleri tarafından etkilenmektedir. Bu çalışmada öğrencilerinin akademik başarılarının stres faktörleri arasındaki kişisel bir veri analizi gerçekleştirilmiştir. Veri seti incelendiğinde birçok öğrencinin stresten karşı karşıya olduğu açıktır. Zorbalıktan, gelecekteki kariyerleri hakkında endişelenmekten veya zihinsel sağlık sorunlarıyla uğraşmaktan kaynaklanan stres, yaşamlarını etkilemektedir. Stres, öğrencilerin öğrenmesini ve gelişim göstermesini engelleyebilmektedir [47].

Stresin, nicel olarak ölçülebilmesi tam anlamıyla mümkün olamamaktadır. Stres seviyesinin belirlenmesinde çeşitli anket çalışmaları ya da algılayıcılar tarafından ölçümler yapılmaktadır. Bu algılayıcılar arasında EEG (Elektroensefalografi), GSR (elektrodermal aktivite), EMG (Elektromiyografi) ve SpO2 (Solunum aktivasyonu) en çok kullanılanlarıdır. Makine öğrenmesi algoritmalarıyla öğrenci stresinin tespitinde benzer çalışmalara rastlanılmıştır [48-51]. 200’den fazla öğrencinin verisinden oluşan ve Naive Baye’s, Linear Regression, Çok Katmanlı Algılayıcılar, Bayes Net, J48 ve Rassal Orman gibi farklı sınıflandırma algoritmalarının kullanıldığı çalışmada Rassal Orman algoritması %94,73 oranında doğruluk vermiştir [52]. Benzer bir çalışma 650 katılımcıdan 5’li Likert ölçeği kullanılarak toplanmıştır. Farklı veri görselleştirme teknikleri ve Rassal Orman algoritması uygulanarak, 25 özellikten oluşan listeden 15 önemli katkıda bulunan faktör belirlenmiş ve stres seviyesinin tahmini yapılmıştır. Bu stres faktörlerinden en önemlisi ise akademik başarı olarak ifade edilmiştir [53].

Üniversite öğrencilerinin stres düzeyi ile akademik başarı ilişkisinin incelendiği bu çalışmada hem elde edilen kişisel veri analizleri kapsamında hem de literatürdeki yer alan benzer çalışmaların [54-57] da ortaya koyduğu üzere stresin akademik başarıyı etkilediği görülmüştür [58].

Bu noktada okulların öğrencilerin stresleriyle baş etmelerine yardımcı olması önem arz etmektedir. Bu gelişimi, gelecek kariyer tavsiyeleri, çalışma yükü ve öğretmen öğrenci ilişkisini destekleyerek yapabilirler. Bu durumların pozitif yönde gelişim göstermesi öğrencileri daha mutlu yapabilir. Ayrıca onlara daha iyi öğrenme ortamı ve gelecek kariyer kaygılarını en aza indirecek etkiyi de dolaylı olarak sağlayabilir.

**KAYNAKLAR (REFERENCES)**

- [1] Ertekin, Y., (1993). Stres ve Yönetim. Ankara: Türkiye ve Orta Doğu Amme İdaresi Enstitüsü Yayınları No:253. Devlet İstatistik Enstitüsü Matbaası.
- [2] Graham, H. (1999). Stresi kendi yararınıza kullanın. (Çev. M. Sağlam ve T. Tezcan). İstanbul: Alfa Yayınları.
- [3] Köknel, Ö. (1987). Zorlanan İnsan, İstanbul: Altın Kitaplar Yayınları.
- [4] De Vito, N. (2009). *The relationship between teacher burnout and emotional intelligence: A pilot study* (Doctoral dissertation, Fairleigh Dickinson University).
- [5] Stora, J.B. (1994). Stres (2.Baskı.). (K. Ayşen, Çev.) İstanbul: İletişim.
- [6] Dewe, P. J., O'Driscoll, M. P., & Cooper, C. L. (2012). Theories of psychological stress at work. *Handbook of occupational health and wellness*, 23-38.
- [7] Brantley, P. J., & Thomason, B. T. (1995). Stress and stress management. In *Handbook of health and rehabilitation psychology* (pp. 275-289). Boston, MA: Springer US.
- [8] Lehrer, P. M., Woolfolk, R. L., & Sime, W. E. (Eds.). (2007). *Principles and practice of stress management* (3rd ed.). The Guilford Press.
- [9] DeFrank, R. S., & Ivancevich, J. M. (1998). Stress on the job: An executive update. *Academy of Management Perspectives*, 12(3), 55-66.
- [10] Cranwell-Ward, J., & Abbey, A. (2005). *Organizational stress*. Springer.
- [11] Gökler, R. (2012). Modern Çağın Hastalığı; Stres Ve Etkileri/The Disease of Modern Era; Stress and its Effects. *Journal of history culture and art research*, 1(3), 154-168.
- [12] Erdoğan, T., Ünsar, Y. S., & Süt, D. (2009). Stresin Çalışanlar Üzerindeki Etkileri: Bir Araştırma. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 14(2), 447-461.
- [13] Muslu, M. (2023). Öğrencilerde akademik başarının artırılmasında masanın dört ayağı; beslenme, uyku, fiziksel aktivite ve sosyal yaşam. *Current Perspectives on Health Sciences*, 4(1), 17-25.
- [14] Karaman, G. (2020). *Üniversite öğrencilerinin akademik başarısı ile stres ve kaygının ilişkisi* (Master's thesis, İstanbul Gelişim Üniversitesi Lisansüstü Eğitim Enstitüsü).
- [15] Aydın, A., Üçüncü, K., & Taşdemir, T. (2011). Akademik performansı etkileyen stres kaynaklarının belirlenmesine yönelik bir alan çalışması. *Uluslararası İnsan Bilimleri Dergisi*, 8(2), 387-399.
- [16] Durmaz, M., Hüseyinli, T., & Güçlü, C. (2016). Zaman yönetimi becerileri ile akademik başarı arasındaki ilişki. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 5(7), 2291-2303
- [17] Hayri, S., Mahmut, Ö., Sadri, Ş., Bekir, G., Selahattin, G., & Petek, A. (2021). Türkiye'de akademik başarının belirleyicileri. *Journal of Economy Culture and Society*, (64), 143-162.
- [18] ILTER, İ. (2021). Lisans öğrencilerinin akademik başarıları ve kariyer kararı öz-yeterliklerinin lisansüstü eğitim niyetlerine etkisi. *Yükseköğretim ve Bilim Dergisi*, 11(1), 1-13.
- [19] Uyar, M. Y., & Doğanay, A. (2018). Öğrenci merkezli strateji, yöntem ve tekniklerin akademik başarıya etkisi: bir meta-analiz çalışması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 14(1), 186-209.
- [20] Maykel, C., deLeyer-Tiarks, J., & Bray, M. A. (2018). Academic stress: What is the problem and what can educators and parents do to help?. *Positive schooling and child development: International perspectives*, 27-40.
- [21] Van der Wath, A. E., & Du Toit, P. H. (2015). Learning end-of-life care within a constructivist model: Undergraduate nursing students' experiences. *Curationis*, 38(2), 1-9.
- [22] Aliakbari, F., Parvin, N., Heidari, M., & Haghani, F. (2015). Learning theories application in nursing education. *Journal of education and health promotion*, 4.
- [23] Goldfarb, E. V., Froböse, M. I., Cools, R., & Phelps, E. A. (2017). Stress and cognitive flexibility: cortisol increases are associated with enhanced updating but impaired switching. *Journal of cognitive neuroscience*, 29(1), 14-24.
- [24] Bedel, A., & Ulubey, E. (2015). Ergenlerde Başa Çıkma Stratejilerini Açıklamada Bilişsel Esnekliğin Rolü. *Elektronik Sosyal Bilimler Dergisi*, 14(55), 291-300.

- [25] Laçın, B. G. D., & Yalçın, İ. (2018). Üniversite öğrencilerinde öz-yeterlilik ve stresle başa çıkma stratejilerinin bilişsel esnekliği yordama düzeyleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(2), 358-371.
- [26] Karaca, A., Yıldırım, N., Ankaralı, H., Açıkgöz, F., & Akkuş, D. (2015). Hemşirelik Öğrencileri İçin Algılanan Stres, Biyo-psiko-sosyal Cevap ve Stresle Başetme Davranışları Ölçeklerinin Türkçe'ye Uyarlanması. *Journal of Psychiatric Nursing/Psikiyatri Hemşireleri Derneği*, 6(1).
- [27] Erten, S. B. (2020). *Lise son sınıf öğrencilerinin sınav kaygısı ile algılanan stres düzeyleri arasındaki ilişkinin incelenmesi* (Master's thesis, İnönü Üniversitesi Sağlık Bilimleri Enstitüsü).
- [28] Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International journal of adolescence and youth*, 25(1), 104-112.
- [29] Doğan, S. (2020). Üniversite Öğrencilerinin Covid-19'a Yazdıkları 100 Mektubun İncelenmesi. *Türk Dünyası Araştırmaları*, 126(248), 25-40.
- [30] Ribeiro, Í. J., Pereira, R., Freire, I. V., de Oliveira, B. G., Casotti, C. A., & Boery, E. N. (2018). Stress and quality of life among university students: A systematic literature review. *Health Professions Education*, 4(2), 70-77.
- [31] Kaggle, (2023). Veri seti. <https://www.kaggle.com/datasets/rxnach/student-stress-factors-a-comprehensive-analysis/> (Erişim Tarihi: 01.09.2023)
- [32] Ma, P.; Ding, R.; Han, S.; and Zhang, D. 2021a. MetaInsight: Automatic Discovery of Structured Knowledge for Exploratory Data Analysis. In SIGMOD
- [33] Bar El, O.; Milo, T.; Somech, A.; Bar El, O.; Milo, T.; and Somech, A. 2020. Automatically Generating Data Exploration Sessions Using Deep Reinforcement Learning. In SIGMOD.
- [34] Milo, T.; Somech, A.; Milo, T.; and Somech, A. 2018a. Next-Step Suggestions for Modern Interactive Data Analysis Platforms. In KDD.
- [35] Tukey, J. W., & Mosteller, F. (1991). *Data analysis*. Princeton University.
- [36] Behrens, J. T. (1997). Principles and procedures of exploratory data analysis. *Psychological methods*, 2(2), 131.
- [37] Akbaş, A. (2023). Doktora Öğrencilerinin Akademik bir Formasyon Dersini Değerlendirmesi: Keşifsel bir Analiz Örneği. *Social Sciences Studies Journal (SSSJurnal)*, 9(116), 9027-9042.
- [38] Beştaş, M. (2023). Keşifçi Veri Analizi ile Eczane Satış Analizi ve Satış Tahmini. *Third Sector Social Economic Review*, 58(1), 765-782.
- [39] Khan, S. A., & Velan, S. S. (2020, June). Application of exploratory data analysis to generate inferences on the occurrence of breast cancer using a sample dataset. In *2020 International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 449-454). IEEE.
- [40] Mahesh, B. (2020). Machine learning algorithms- a review. *International Journal of Science and Research (IJSR)*. [Internet], 9(1), 381-386.
- [41] Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, 4, 51-62.
- [42] Sinap, V. (2023). Makine Öğrenmesi Teknikleri ile Counter-Strike: Global Offensive Raunt Sonuçlarının Tahminlenmesi. *Journal of Intelligent Systems: Theory and Applications*, 6(2), 119-129.
- [43] Çakar, Ö. K., & Esen, A. (2023). Yerel yönetimlerde stratejik planlama uygulamalarının kurum performansına etkisi. *Business & Management Studies: An International Journal*, 11(3), 1010-1025.
- [44] San Millan-Castillo, R., Morgado, E., & Goya-Esteban, R. (2019). On the use of decision tree regression for predicting vibration frequency response of handheld probes. *IEEE Sensors Journal*, 20(8), 4120-4130.
- [45] Doğan, S., Büyükkör, Y., & Atan, M. (2022). A comparative study of corporate credit ratings prediction with machine learning. *Operations Research and Decisions*, 32(1), 25-47.
- [46] Liang, L., Di, L., Huang, T., Wang, J., Lin, L., Wang, L., & Yang, M. (2018). Estimation of leaf nitrogen content in wheat using new hyperspectral indices and a random forest regression algorithm. *Remote Sensing*, 10(12), 1940.
- [47] Mukhiya, S. K., & Ahmed, U. (2020). *Hands-On Exploratory Data Analysis with Python: Perform EDA techniques to understand, summarize, and investigate your data*. Packt Publishing Ltd.
- [48] Sulaiman, N. (2016). Determination and classification of human stress index using nonparametric analysis of EEG signals.

- [49] Xu, Q., Nwe, T. L., & Guan, C. (2014). Cluster-based analysis for personalized stress evaluation using physiological signals. *IEEE journal of biomedical and health informatics*, 19(1), 275-281.
- [50] Liu, D., & Ulrich, M. (2014). Listen to your heart: Stress prediction using consumer heart rate sensors. *Final. Proj. Stanford CS*, 229, 229-233.
- [51] Jung, Y., & Yoon, Y. I. (2017). Multi-level assessment model for wellness service based on human mental stress level. *Multimedia Tools and Applications*, 76, 11305-11317.
- [52] Sharma, D., Kapoor, N., & Kang, S. S. (2020). Stress prediction of students using machine learning. *International Journal of Mechanical and Production Engineering Research and Development*, 10(3).
- [53] Pabreja, K., Singh, A., Singh, R., Agnihotri, R., Kaushik, S., & Malhotra, T. (2021). Stress prediction model using machine learning. In *Proceedings of International Conference on Artificial Intelligence and Applications: ICAIA 2020* (pp. 57-68). Springer Singapore.
- [54] Ranasinghe, P., Wathurapatha, W. S., Mathangasinghe, Y., & Ponnampereuma, G. (2017). Emotional intelligence, perceived stress and academic performance of Sri Lankan medical undergraduates. *BMC medical education*, 17(1), 1-7.
- [55] Joseph, N., Joseph, N., Panicker, V., Nelliyanil, M., Jindal, A., & Viveki, R. (2015). Assessment and determinants of emotional intelligence and perceived stress among students of a medical college in south India. *Indian journal of public health*, 59(4), 310-313.
- [56] Duari, P. (2012). Importance of self-esteem among students in academic performance and coping with stress. *Indian Journal of Positive Psychology*, 3(4), 474.
- [57] Khanekhesi, A. L. I. (2011). The Relationship of Academic Stress with Aggression, Depression and Academic Performance of College Students in Iran. *Journal on Educational Psychology*, 5(1), 24-31.
- [58] Brailovskaia, J., Schillack, H., & Margraf, J. (2020). Tell me why are you using social media (SM)! Relationship between reasons for use of SM, SM flow, daily stress, depression, anxiety, and addictive SM use—An exploratory investigation of young adults in Germany. *Computers in human behavior*, 113, 106511.



## Yapay Sinir Ağları ile Mısır Yapraklarında Hastalık Tespiti

Resul BÜTÜNER<sup>\*,a</sup>, Muzaffer ŞİMŞEK<sup>b</sup>, Gökhan AKSÖZ<sup>c</sup>,

Sadi ŞAN<sup>d</sup>

<sup>a\*, b, c, d</sup> Beypazarı Fatih M.T.A.L., Ankara, 06730 TÜRKİYE

### MAKALE BİLGİSİ

Alınma: 02.11.2023  
Kabul : 26.12.2023

#### Anahtar Kelimeler:

Yapay zekâ, Yapay sinir ağı, Mısır, Mısır hastalığı, Tespit etme, Mısır yaprağı  
**\*Sorumlu Yazar**  
e-posta:  
rbutuner@gmail.com

### ÖZET

Mısır, tarihte yaklaşık 10.000 yıl önce güney Meksika'daki yerli halklar tarafından evcilleştirilen ve günümüzde dünya genelinde popüler bir gıda maddesi olarak kabul edilen bir tahıl çeşididir. Dünya nüfusunun artmasıyla birlikte, diğer birçok besin maddesinde olduğu gibi mısır tüketiminde de belirgin bir artış gözlenmektedir. Ancak, bu önemli ürünün üretiminde meydana gelen hastalıklar, özellikle yoğun tüketim yaşanan ülkeler için ciddi bir sorun oluşturabilir. Mısır üreticileri, olası hastalıkları belirlemek ve izlemek konusunda sürekli bir takip içerisindeyler. Ne var ki, modern teknoloji mısır hastalıklarının tespiti ve izlenmesi için tam anlamıyla kullanılamamaktadır. Geleneksel tespit yöntemlerinin yanı sıra Yapay Zekâ, Görüntü İşleme, Yapay Sinir Ağları, ve Derin Öğrenme gibi yeni teknolojiler, hastalıkların tespiti için kullanılmaktadır. Bu yöntemler ve teknikler, hastalıkların etkili ve hızlı bir şekilde teşhis edilmesine ve gereken müdahalelerin teknolojik çözümlerle yapılmasına olanak tanımaktadır. Bu araştırmada, dört farklı mısır hastalığı ve sağlıklı mısır görüntüleri toplanmış ve bu görüntülerin analizinde K-En Yakın Komşu Algoritması, Rastgele Orman Algoritması ve Yapay Sinir Ağları kullanılmıştır. Bu algoritmalar, yapay zekâyı eğitmiş ve öğrenmiş modeller ortaya koymuştur. Eğitim süreci sonucunda model oluşturulmuş ve sağlıklı mısır, mısır pası, yaprak lekesi ve yaprak yanığı görüntüleriyle karşılaştırılmıştır. Mısır yaprağı hastalığı tespitinde elde edilen başarı oranları şu şekildedir: K-En Yakın Komşu Algoritması %97, Rastgele Orman Algoritması %97,6 ve Yapay Sinir Ağları %97,7. Bu sonuçlar doğrultusunda, oluşturulan Yapay Sinir Ağı modelinin yüksek doğruluk oranıyla mısır hastalıklarının tespiti için kullanılabileceği ve gelecekte yapılacak çalışmalara rehberlik edebileceği öngörülmektedir.

DOI: 10.59940/jismar.1384930

## Disease Detection in Corn Leaves with Artificial Neural Networks

### ARTICLE INFO

Received: 02.11.2023  
Accepted: 26.12.2023

#### Keywords:

Artificial intelligence, Artificial neural network, Corn, Corn disease, Detection, Corn leaf  
**\*Corresponding Authors**  
e-mail:  
rbutuner@gmail.com

### ABSTRACT

Corn, a grain variety domesticated by indigenous peoples in southern Mexico around 10,000 years ago, is considered a popular food globally today. With the increase in the world's population, there has been a significant rise in corn consumption, as observed in many other food items. However, diseases affecting the production of this crucial crop could pose a serious problem, especially in countries with high consumption rates. Corn producers remain continuously vigilant in identifying and monitoring potential diseases. However, current modern technology cannot fully cater to the detection and monitoring of corn diseases. Besides traditional detection methods, new technologies such as Artificial Intelligence, Image Processing, Artificial Neural Networks, and Deep Learning are being utilized for disease detection. These methods and techniques enable effective and swift diagnosis of diseases, allowing for technological interventions when necessary. In this study, images of four different corn diseases and healthy corn were gathered and analyzed using the K-Nearest Neighbors Algorithm, Random Forest Algorithm, and Artificial Neural Networks. These algorithms trained and produced learned models in artificial intelligence. Following the training process, a model was created and compared with images of healthy corn, corn rust, leaf spots, and leaf blights. The success rates in detecting corn leaf diseases are as follows: K-Nearest Neighbors Algorithm at 97%, Random Forest Algorithm at 97.6%, and Artificial Neural Networks at 97.7%. Based on these results, it is anticipated that the created Artificial Neural Network model, with its high accuracy rate, could be used for identifying corn diseases and could serve as a guide for future research endeavors.

DOI: 10.59940/jismar.1384930



## 1. GİRİŞ (INTRODUCTION)

Tarım ve tarımsal sektörün ilerlemesi, küresel anlamda yoksulluğun azaltılmasına dair en etkili hususlardan biridir. Küresel kalkınma hedeflerine ulaşma yolunda sağlam, sürdürülebilir ve kapsayıcı gıda sistemleri kritik bir öneme sahiptir. Ortak refahı artırmak ve 2050 yılına kadar tahmini 9,7 milyar insanı beslemek adına tarımsal kalkınma, en etkin stratejilerden biri olarak öne çıkmaktadır.

Tarım, ekonomik büyüme için çok önemlidir, çünkü dünyanın GSYİH(Gayri Safi Yurt İçi Hasıla)'sinin %4'üne ve bazı az gelişmiş ülkelerde bu oran GSYİH'nın %25'ine kadar çıkabilmektedir [1].

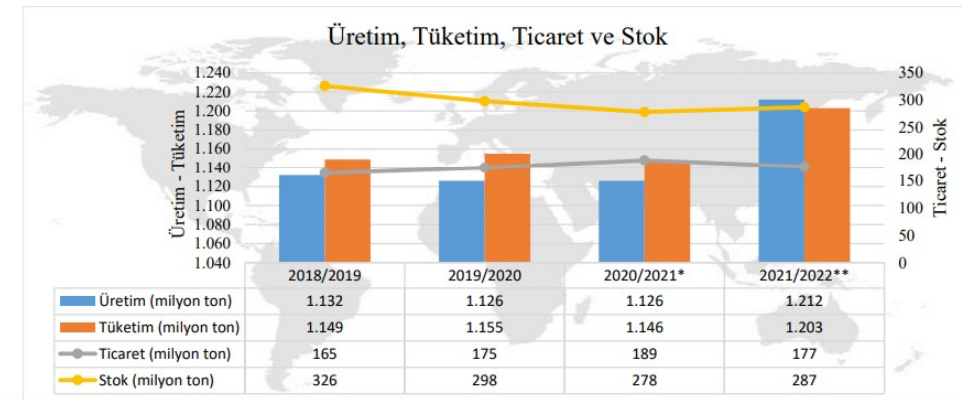
Tarımsal alanda, özellikle sebze üretiminde, hastalıkların erken teşhisi ve sınıflandırılması, üreticilerin hastalıkları önlemek için alacakları tedbirlerin yanı sıra ekonomik kayıpların ve üretim kayıplarının azaltılmasına sebep olmaktadır. Uzmanlar uzun süredir sebze ve meyve hastalıklarını çiplak gözle izleyerek tespit etmektedirler. Ancak, bu yöntemler hem zaman alıcıdır hem de bazı hastalıklar için yetersiz veya yanıltıcı olabilmektedir. Sebze ve meyvelerdeki hastalıklar, belirli görünür belirtiler

biraktığından, yapay zeka teknikleri sebzeledeki bu kusurların görüntü analizini yapma imkanı sunmaktadır.

Mısır, tek yıllık bir tahıl bitkisi olup özellikle sıcak iklimlerde yetişen Buğdaygiller ailesine ait bir bitkidir [2]. Mısır, tarımının binlerce yıldır yapıldığı ve uzun bir tarihe sahip nadir bitkilerden biridir. İklim çeşitliliği göz önüne alındığında, mısır dünya genelinde ılıman, subtropikal ve tropikal iklim kuşaklarında yetiştirilebilme özelliğine sahiptir.

Dünya genelinde, buğday ve pirinçten sonra en fazla tarımı yapılan tahıl bitkisi olarak mısır öne çıkmaktadır. Yüzde seksen oranında gelişmekte olan ülkelerde bulunan yaklaşık 70 milyon çiftçi ailesinin genellikle mısır tarımıyla meşgul olduğu bilinmektedir.

Uluslararası Hububat Konseyi'nin verilerine göre, 2021/2022 pazarlama döneminde küresel mısır üretiminin 1,21 milyar ton, tüketimin ise 1,20 milyar ton olması öngörülmektedir. Küresel mısır ticaretinin ise 177 milyon ton, stok miktarının ise 287 milyon ton olacağı tahmin edilmektedir. [3].



Kaynak: IGC, \*Tahmin, \*\*Öngörü, Erişim Tarihi: 03.01.2022

Şekil 1. Dünya Mısır Üretim, Tüketim Ve Stok Miktarları (World Corn Production, Consumption and Inventory Quantities)

Şekil 1'e göre dünya mısır üretim, tüketim ve stok miktarı verilmiş olup, Dünya genelinde, ekiliş alanı bakımından buğday ve çeltikten sonra üçüncü sırada yer alan mısır, üretim miktarı açısından önde gelen bir konumda bulunmaktadır. ABD, Çin, Brezilya, Arjantin, Meksika, Hindistan, Ukrayna ve Endonezya gibi ülkeler en yüksek mısır üretimine sahipken Türkiye, üretim sıralamasında 23. sırada yer almaktadır. Son on yılda, dünya çapında mısır ekim alanları %24 oranında genişlerken üretimde %42,3 oranında bir artış yaşanmıştır. Bu durum, mısır tarımında tekrar edilen ekimler ve pestisit kullanımı gibi unsurların ürün verimliliği ve kalitesini olumsuz etkilemesiyle ilişkilendirilirken,

bu etkenler zamanla daha belirgin hale gelmektedir [2].

Mısır bitkisi, gelişme süreci boyunca patojenlerin saldırısına maruz kalmasıyla çeşitli hastalıklara yol açabilir (Genç Kesimci, T. 2022). Ancak, mısır bu patojenlere oldukça hassas bir tepki göstermektedir. Hastalıklar genellikle mısır bitkisinin büyümesini bozarak üretim miktarında azalmaya yol açarlar [4]. Hastalığın etki derecesi, mısırın hastalık seviyesine bağlıdır. Hastalığın etkilediği bitkilerin genellikle büyümesi bozulur ve bitki dokuları ile hücrelerinin normal aktivitesi bozulur. Hastalıktan etkilenen mısır bitkileri kuruyabilir, boy atmada kısıtlanabilir, renk değişiklikleri gösterebilir ve yapraklar sararır

solabilir veya kuruyabilir [5].

Fungal hastalıklar, temel etkenler arasında yer alır. Mısırdaki yaprak yanıklığı gibi fungal hastalıklar, başlıca ürün kayıplarına neden olan hastalıklardan biridir. Bu hastalıklar, dünya genelinde neredeyse her mısır yetiştirilen bölgede görülen ve uygun iklim koşullarında %20 ile %70 arasında değişen verim kayıplarına sebep olabilen fungal kökenli rahatsızlıklardır.

Ülkemizde de mısır tarımının yapıldığı bölgelerde, uygun iklim koşulları sıklıkla hastalığın ortaya çıkmasına yol açar; bu durum, dünya genelinde benzer bir eğilimi yansıtmaktadır. Mısır hastalıklarının teşhis edilmesi için bitki hastalıklarının otomatik tespiti için çeşitli uzman sistemler geliştirilmiştir [6]. İnternetin hızlı, kolay ve ücretsiz erişimi sayesinde birçok yapay zekâ uygulaması geliştirilebilmektedir. Yapay zekâ terimi, makine öğrenimi ve derin öğrenme gibi konuları içerir (LeCun ve ekibi, 2015). Derin öğrenmedeki ilerlemeler, yapay zekânın görüntü işlemede kullanımına olan ilgiyi artıran ana etkindir [7]. Son yıllarda bilgisayarlı görü alanında yapılan çalışmalar, derin öğrenme ve makine öğrenmesi tekniklerinin kullanılarak başarılı yapay zekâ modellerinin oluşturulduğunu göstermektedir. Örneğin, insanlığın yaşamını sürdürebilmesi ve sağlıklı beslenebilmesi için gereken tarım ürünlerinin yetiştirilmesinde bu uygulamalar önemli bir rol oynamaktadır. Mısır üretiminde verimliliği azaltan ve yapraklarda görülen hastalıkların tespiti de bu yöntemlerle başarıyla gerçekleştirilmektedir. Bu alanda yapılmış birçok araştırma literatürde bulunmaktadır [8].

### 1.1. Literatür Taraması (Literature Review)

Literatürdeki bir çalışmada, mısırın hastalıklarını tespit etmek için farklı görüntü işleme tekniklerini kullanarak bir sistem geliştirmeyi amaçlamışlar. RGB (Red Green Blue) renklerin tanınmasını sağlamak için çeşitli özellik çıkartma yöntemleri kullanmışlardır. Destek vektör makineleri, karar ağaçları, rastgele orman ve naive bayes algoritmaları gibi çeşitli algoritmaları denemişler ve çoğu algoritma başarılı sonuçlar vermiştir [4]. Başka bir çalışmada, mısır hastalıklarının sınıflandırılmasında destek vektör makineleri (SVM) yerine genetik algoritma destek vektör makineleri (GA-SVM) adlı bir algoritmanın kullanılarak SVG algoritmasında başarılı sonuçlar elde edilebileceği gözlemlenmiştir [9]. Panigrahi vd., parametrelerin ayarlanmasıyla geliştirilmiş bir CNN modeli kullanarak %98.78'lik bir ortalama doğruluk oranına erişmişlerdir. Bu geliştirme, düzeltilmiş doğrusal birim aktivasyon fonksiyonu ile bir Adam optimize edicisi kullanılarak gerçekleştirilmiştir. [10].

Mısır hastalıklarının sınıflandırılması amacıyla

yapılan başka bir çalışmada Evrişimli sinir ağları (CNN) algoritmasının eğitimi için bir akıllı telefon kamerası kullanılarak mısır yaprağı hastalıklarının görüntüleri toplanmış ve bu görüntüleri tanıyan ve sınıflandıran bir CNN algoritması ile %92.85 başarı elde edildiği görülmüştür [11].

Mısır hastalıklarını sınıflandırmak için LeNet temelli bir CNN kullanılarak geliştirilen sınıflandırma algoritmasının başarı oranının %97.89 olduğu tespit edilmiştir [12].

Çiftçilerin meyve hastalıklarını tespit etmelerine yardımcı olmak amacıyla bir sistem üzerinden meyve fotoğraflarını yükleyerek bir teknik önerildi. Bu yöntemde, önceden öğrenilmiş bir veri kümesi olan nar tohumlarından elde edilen veriler kullanıldı. Kullanıcılar, çeşitli fotoğrafları sisteme yükleyerek analiz için katkıda bulundular. Bu fotoğraflar daha sonra eğitim veri setindeki görüntülerle karşılaştırılarak hastalık şiddeti belirlendi. Önerilen yöntemin deneysel sonuçlarına göre %82 doğruluk elde edildi [13].

Kapon, mısır yaprağı hastalıkları üzerine yaptığı çalışmada EfficientNet adlı CNN mimarisini tercih etmiştir. Yapılan yöntemde, dört farklı sınıf için sırasıyla %91,30, %97,71, %93,10 ve %100 doğruluk oranları elde edilmiştir [14].

Whxna, araştırmalarında SqueezeNet adlı bir CNN mimarisini tercih etmiş ve optimize edici olarak Adam yöntemini kullanmıştır [15].

Bilgisayarlı görü kullanılarak, mısır hastalıklarının 5 farklı kategorisine yönelik sınıflandırma işlemi %80'in üzerinde bir başarı elde etmiştir [16].

AlexNet'in kullanıldığı bir çalışmada, oluşturulan sınıflandırıcının %98.62'lik bir başarı oranına sahip olduğu gözlemlenmiştir [17].

Mısır kuraklığını belirleme ve kategorize etme amacıyla gerçekleştirilen bir çalışmada, derin konvolüsyonel sinir ağının kullanıldığı tespit edilmiştir. Elde edilen sonuçlar, bu önerilen sistemin tanımlama ve sınıflandırma doğruluğunun sırasıyla %98.14 ve %95.95 olduğunu göstermektedir [18].

Song ve ekibi, yaprak hastalıklarının tespiti ve sınıflandırılmasında sayısal görüntü işleme teknikleri, destek vektör makineleri, yapay sinir ağları ve diğer metodları kullanmışlardır. Farklı mısır yaprağı hastalıklarını %89.6 oranında doğrulukla tespit etmek için bir SVM yaklaşımı benimsemişlerdir. Bu SVM tabanlı sınıflandırma metodolojisi, en yüksek doğruluğu sunmasa da hem küçük veri setleri hem de büyük veri setleri için kullanışlıdır [19].

Rage, mısır yaprağı hastalıkları üzerine yaptığı çalışmada VGG16 adlı bir CNN mimarisini tercih etmiş ve optimize edici olarak Adam yöntemini kullanmıştır. Çalışmada mısır yaprakları için elde edilen doğruluk oranı %90,82 olarak kaydedilmiştir [20].

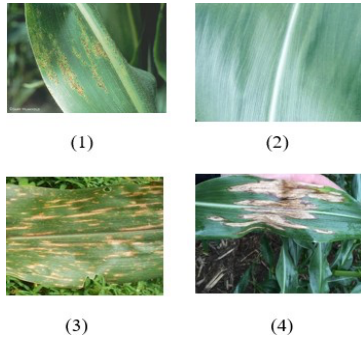
## 1.2. Veri Seti Özellikleri (Data Set Characteristics)

Bu çalışmada kullanılan veri seti kaggle.com platformundan alınmıştır (<https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset>).

Orange Data Mining uygulaması, veri analizi için veri birimine odaklanan bir platformdan yararlanılmıştır. Bu yazılımın temel amacı, veri analizi, veri madenciliği tekniklerinin kullanımı ve yeni modellerin oluşturulmasıyla bu modellerin değerlendirilmesidir.

Orange Data Mining platformu, açık kaynaklı bir araç setidir ve kullanıcılara veri hazırlama, veri analizi ve model oluşturma gibi olanaklar sunar. Veri madenciliği işlemleri, görsel programlama veya Python betikleme yöntemleriyle Orange Data Mining'de gerçekleştirilebilir.

Bu araştırmada, 4 ayrı veri grubu üzerinde çalışılmıştır. Bu gruplardan üçü hastalıklı verilerden oluşurken bir grup sağlıklı verilerden oluşmaktadır. Veri setindeki materyaller JPG formatındadır ve toplamda 1306 adet Mısır Pası, 574 adet Yaprak Lekesi, 1146 adet Yaprak Yanığı ve 1162 adet Sağlıklı mısır yaprağı olmak üzere 4188 resim bulunmaktadır.



Şekil 2. Veri Setinden Örnek Resimler (Sample Images from the Data Set)

Şekil 2'de sunulan veri setinden örnek resimler gösterilmiştir. Bu resimler, (1) numarada mısır pası örneği, (2) numarada sağlıklı veri örneği, (3) numarada yaprak lekeli örneği ve (4) numarada yaprak yanığı veri örneği olarak tanımlanmaktadır. Veri setinin dağılımı ise Tablo 1'de sunulmuştur.

Tablo 1. Çalışmada kullanılan veri sayıları ve yüzdeleri (The Numbers and Percentages of the Data Used in the Study)

Kullanım Alanı	Veri Sayısı	Yüzde(%)
Eğitim(training)	3351	%80
Test(testing)	837	%20
Toplam	4188	%100

Araştırmanın temel amacı, günümüzde sıkça karşılaşılan mısır yaprağı hastalıklarını hızlı ve kesin bir biçimde öngörebilen bir yapay zekâ modelinin oluşturulmasıdır. Bu çalışmanın hedefleri arasında

tarım sektöründe yapay zekânın benimsenmesini desteklemek ve bu yolla hastalık teşhisinde zaman ve çaba kaybını en aza indirgeyerek doğru teşhisi yapmak yer almaktadır. Ek olarak, bu çalışmanın bir diğer amacı, geliştirilen yapay zekâ modelinin çeşitli ve geniş veri setlerini analiz ederek kendini sürekli olarak eğitmesi ve teşhis doğruluğunu artırmasıdır. Bu sayede, üretim maliyetlerinde ve zaman kullanımında tasarruf sağlanması öngörülmektedir.

## 2. YÖNTEM (METHOD)

Bu bölümde, çalışma sürecinde önerilen model ve adımlar ayrıntılı bir şekilde sunulmuştur. İlk aşamada, kullanılan algoritmaların açıklaması yapılmış, daha sonra veri setinin edinilme ve hazırlanma süreci açıklanmış ve en nihayetinde modelin oluşturulma süreci detaylandırılmıştır. Model oluşturulurken tercih edilen araç, Orange veri analiz programıdır.

Bu çalışmada, mısır yapraklarındaki hastalıkların tespiti için görüntü işleme ve sınıflandırma araçlarını kullanan, yapay zekâ temelli bir yöntem ile Orange veri analiz platformu incelenmiştir. Bu sistem için, hastalıklı ve sağlıklı mısır yaprağı resimleri JPG formatında girdi olarak sunulmuş ve bu resimler üzerinde hastalık varlığıyla ilgili bir analiz görüntü işleme teknikleriyle gerçekleştirilmiştir.

### 2.1. Algoritma Seçimi (Algorithm Selection)

Orange veri analiz programı, modellerin oluşturulmasında tercih edilen bir araçtır. Bu araştırma, mısır yapraklarında hastalıkların teşhisini amaçlamaktadır. Bu hedef doğrultusunda, Rastgele Orman Algoritması, K-En Yakın Komşu algoritması ve Yapay Sinir Ağları gibi çeşitli algoritmalar kullanılmıştır. Çalışmada, üçten fazla makine öğrenmesi algoritmasıyla testler gerçekleştirilmiş ve en yüksek başarı oranını sunan KNN, Naïve Bayes ve Yapay Sinir Ağları tercih edilmiştir.

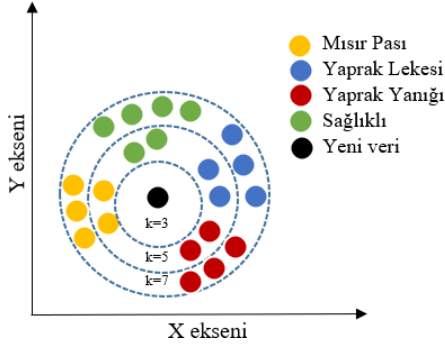
#### 2.1.1. K-En Yakın Komşu (KNN) Algoritması

(K-Nearest Neighbor (KNN) Algorithm)

Yeni verinin örnek veri setine katılması durumunda, mevcut verilere olan benzerliği hesaplanır ve k en yakın komşusuna bakılır. K-En Yakın Komşu (KNN) algoritması, yeni veriyi var olan verilerle karşılaştırarak uzaklığını hesaplayıp sınıflandıran bir yapıya sahiptir. KNN, temel bir makine öğrenimi algoritmasıdır. Yeni bir veri örneği eklendiğinde, bu örneğe en yakın olan k veri belirlenir ve yeni örneğin sınıfı bu k verinin sınıfına dayalı olarak belirlenir. [21].

K-En Yakın Komşu (KNN) algoritması, yeni veriyi alır ve ardından uzaklıkların hesaplanmasını gerçekleştirir.

Bu hesaplama sonrasında en yakın komşular belirlenir ve bu modele göre veri, mısır pası, yaprak lekesi, yaprak yanığı veya sağlıklı olup olmadığına göre sınıflandırılır. Şekil 3'te çıktı katmanlarına göre KNN modeli verilmiştir.



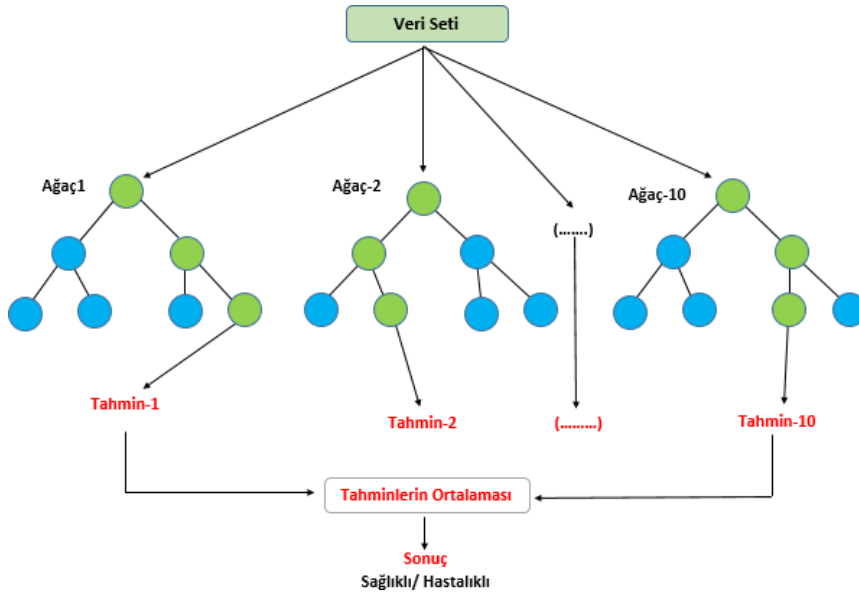
Şekil 3. KNN modeli (KNN model)

## 2.1.2. Rastgele Orman Algoritması

(Random Forest Algorithm)

Rastgele Orman, Leo Breiman ve Adele Cutler tarafından geliştirilmiş bir sınıflandırma yöntemidir. Bu yöntem, oylama tekniğini içerir ve birden fazla karar ağacının bir araya getirilmesiyle oluşturulur. Her bir ağaç, kendi sınıfını oylar ve kazanan sınıf bu oylarla belirlenir. Karar ağaçlarının oluşturulmasında, bağımsız bir şekilde veri setinden bootstrap tekniğiyle seçilen örnekler kullanılır. [22].

Çalışmada Rastgele Orman algoritması modeli Şekil 4'te kullanılmıştır. Bu modelde 10 ağaç yer almaktadır ve elde edilen sonuçlar için tahminlerin ortalaması alınmıştır.



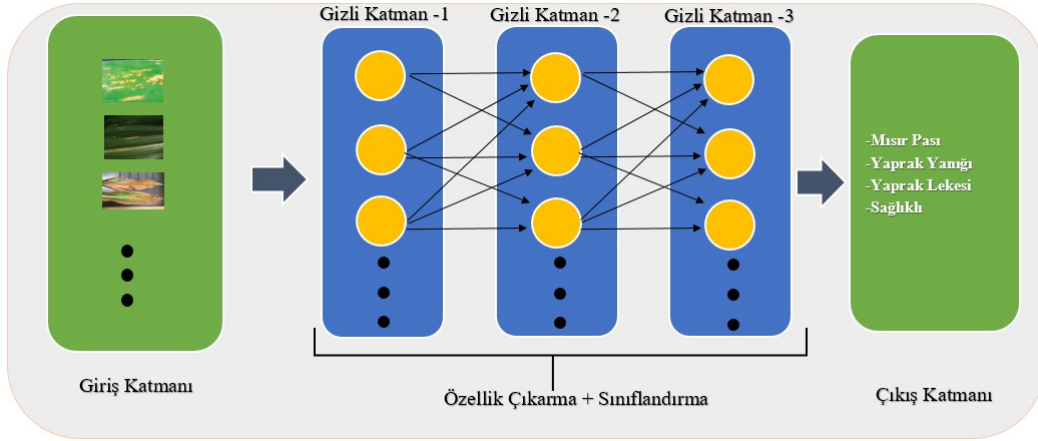
Şekil 4. Rastgele orman modeli (Random forest model)

## 2.1.3. Yapay Sinir Ağı Algoritması

(Artificial Neural Network Algorithm)

Yapay sinir ağları, insan beynindeki nöronlar gibi bağlantılı elemanlardan oluşur. Bu elemanlar bellek kapasitesine sahiptir ve bilgi işleme yeteneklerine sahiptir. Yapay Sinir Ağları, biyolojik sinir

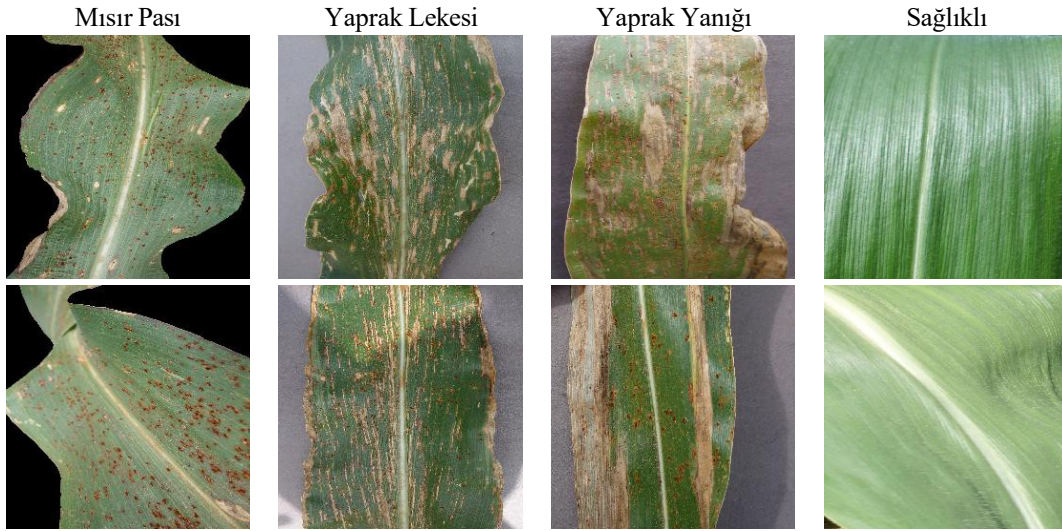
sisteminin yapı ve yeteneklerini taklit etmektedir. Ayrıca, kendi kendine öğrenme, bilgi depolama, yorumlama gibi yetenekleri olan bir algoritma yapısıdır [23]. Çalışmada kullanılan YSA modelinde 5 gizli katman ve 15 düğüm bulunmaktadır. Bu YSA algoritma modeli Şekil 5'te verilmiştir.



Şekil 5. Yapay sinir ağı modeli (Neural Network model)

### 3. BULGULAR (FINDINGS)

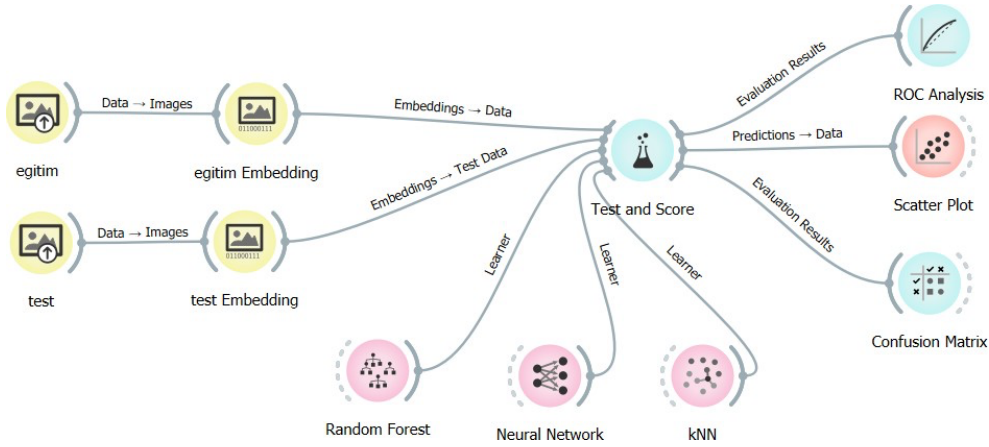
Modelin eğitimi için, mısır yaprağı hastalığına ait görüntüler ile sağlıklı mısır yaprağı görüntüleri kullanılmıştır. Bu nedenle modelin giriş değerleri belirlenmiştir. Bu giriş değerleri, mısır pası, yaprak lekesi, yaprak yanığı ve sağlıklı mısır yaprağı görüntülerine aittir. Şekil 6'da çalışmada kullanılan veri görüntülerinden örnekler verilmiştir.



Şekil 6. Veri setinden örnek veri görüntüleri (Sample data images from the data set)

Çalışmanın Orange veri analizi ekranı, Şekil 7'de sunulmuştur. Orange veri analizi programında, veri görüntüleri eğitim ve test verisi olarak iki farklı gruba ayrılır. Bu veriler Embedding bölümünde işlenir ve daha sonra belirlenen üç farklı algoritma kullanılarak modeller oluşturulur. Elde edilen bu modellerin sonuçları ve başarı değerleri Test and

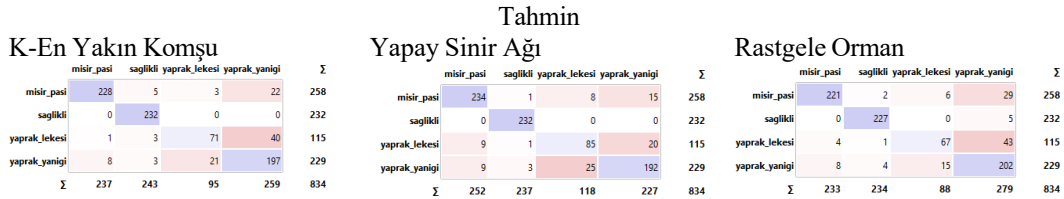
Score bölümünde tablo formatında sunulur. Test ve Skor bölümünden sonra, test verileri algoritmalarla ilişkilendirilerek görselleştirme yapılmıştır. Bu görseller, her bir algoritmanın yüzdelik doğruluk değerlerini, çıkış katmanına göre doğruluk ve kayıp grafiklerini içermektedir.



Şekil 7. Çalışmanın orange veri analiz ekranı (Orange data analysis screen of the study)

Yapay sinir ağının işlediği veriler, giriş, gizli ve çıkış katmanlarından geçerek mısır yaprağı görüntülerinin hastalıklı olup olmadığını sınıflandırmak için kullanılmaktadır. Şekil 8'de, çalışmada kullanılan algoritmalar için Karışıklık Matrisi (Confusion Matrix) değerleri sunulmuştur. Bu algoritmaların

performansı değerlendirildiğinde, 834 test verisi içinde KNN'in 728, Yapay Sinir Ağı'nın 743 ve Rastgele Orman'ın 717 doğru tahminde bulunduğu gözlemlenmiştir. En yüksek başarıyı Yapay Sinir Ağı algoritması göstermiştir.



Şekil 8. Algoritma modelleri karışıklık matrisi sonuçları (Algorithm models results of the confusion matrix)

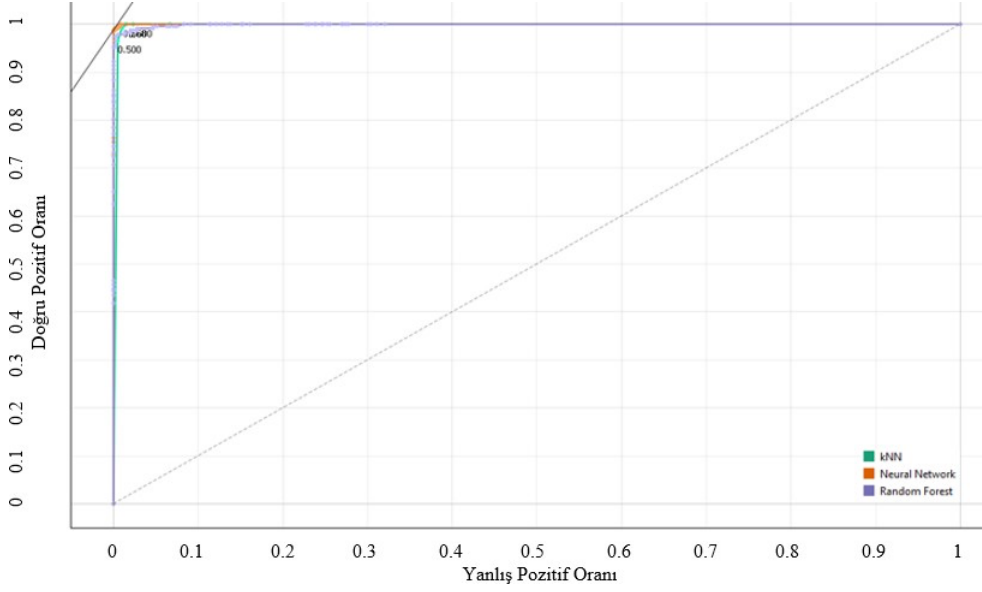
Tablo 2'de projede kullanılan algoritmaların eğitim ve test verilerine göre yüzdeler değeri sunulmuştur. Yapay Sinir Ağı, %97.0 eğitim ve %89.1 test doğruluğuyla en üstün sonucu göstermiştir.

Tablo 2. Modellerin karşılaştırılması (Comparison of models)

Model Adı	AUC	Sınıflandırma Doğruluğu	F1 Skoru	Kesinlik	Duyarlılık
Yapay Sinir Ağı	0.970	0.891	0.891	0.891	0.891
Rastgele Orman	0.967	0.877	0.859	0.867	0.860
kNN	0.961	0.873	0.872	0.875	0.873

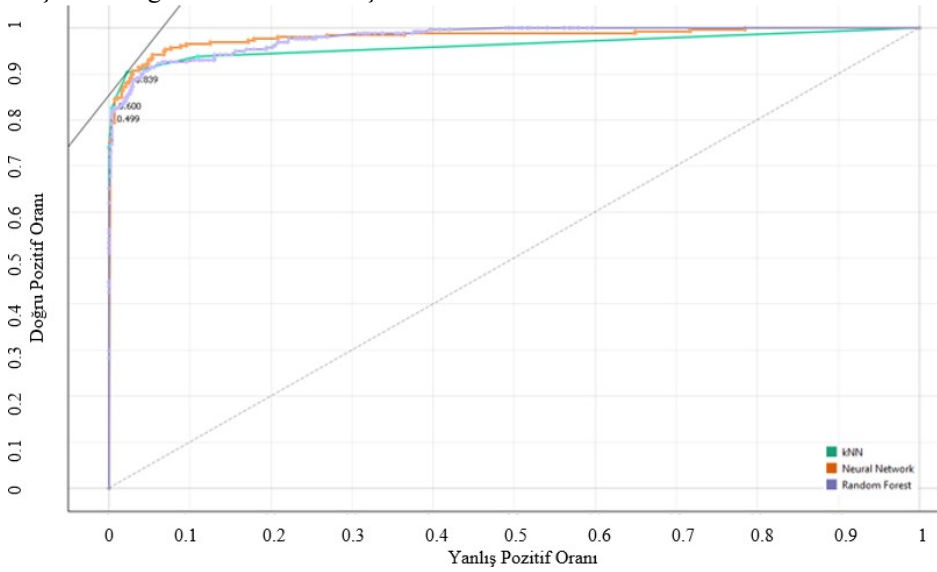
Roc grafiği, herhangi bir sınıflandırma modelinin performansını değerlendirmek için kullanılan temel metriklerden biridir. Özellikle dengesiz veri setleriyle çalışıldığında, makine öğrenimi algoritmalarının performansını değerlendirmek için en sık tercih edilen ölçümlerden biridir. Çalışmada sağlıklı mısır, mısır pası, yaprak lekesi ve yaprak yanığı için Roc doğruluk grafikleri şekil 9, şekil 10, şekil 11 ve şekil 12'de sunulmuştur. Grafiklerde,

kullanılan algoritmaların doğruluk oranları, doğru pozitif (TP) ve yanlış pozitif (FP) değerlerine göre belirlenmiştir. Şekil 9'da sağlıklı mısır verileri için kNN 0.989, yapay sinir ağı 0.994 ve rastgele orman algoritması için 0.987 doğruluk oranları gözlemlenmiştir. Grafik incelendiğinde, yapay sinir ağı algoritmasının diğerlerine göre daha başarılı bir doğruluk oranı sergilediği belirlenmektedir.



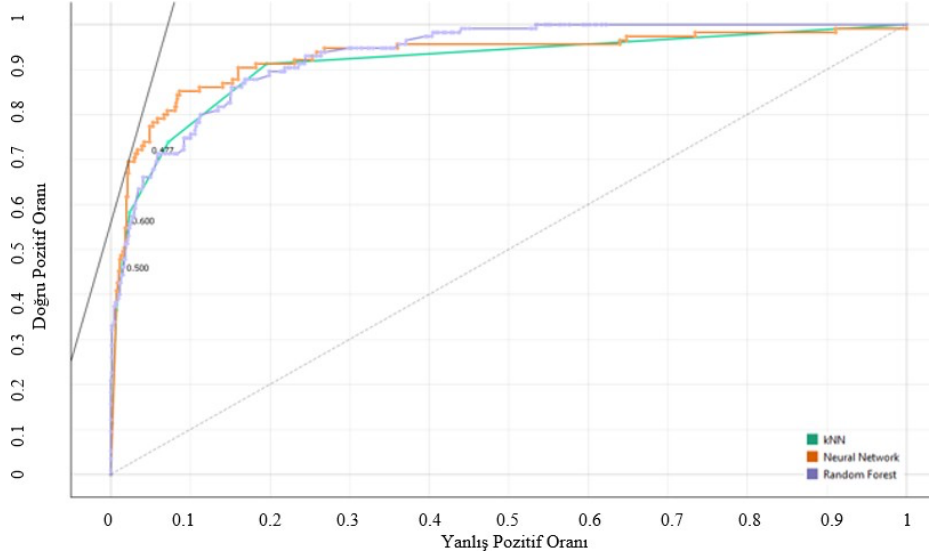
Şekil 9. Sağlıklı mısır verileri için roc analiz grafiği (Roc analysis graph for healthy corn data)

Şekil 10'da mısır pası hastalığı verileri için kNN 0.947, yapay sinir ağı 0.950 ve rastgele orman algoritması için 0.940 doğruluk oranları sonucu ortaya çıkmıştır. Grafiğe göre yapay sinir ağı algoritmasının doğruluk oranının diğerlerine göre başarılı olduğu sonucuna varılmıştır.



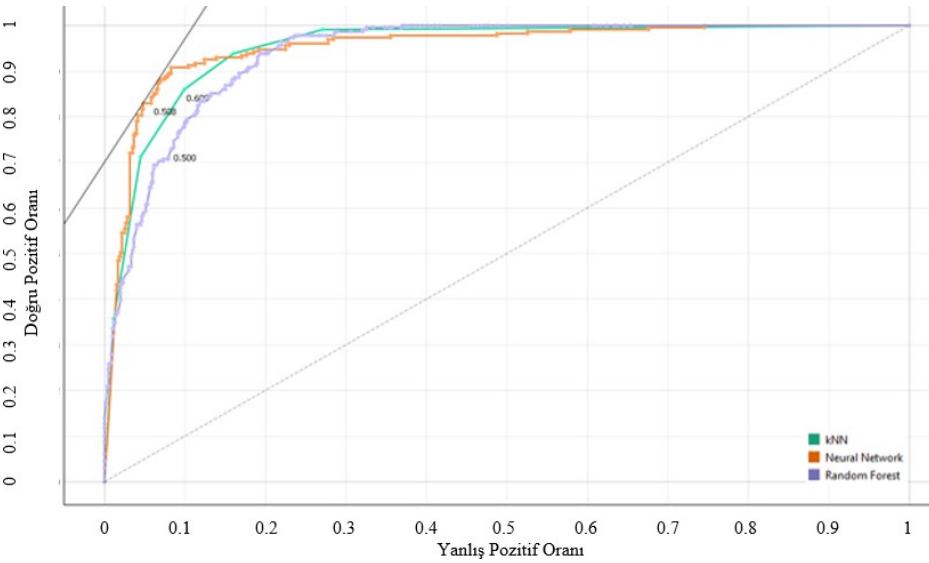
Şekil 10. Mısır pası hastalığı verileri için roc analiz grafiği (Roc analysis graph for maize rust disease data)

Şekil 11'de yaprak lekeli hastalığı verileri için kNN 0.922, yapay sinir ağı 0.924 ve rastgele orman algoritması için 0.911 doğruluk oranları sonucu ortaya çıkmıştır. Yapay sinir ağı algoritmasının doğruluk oranının diğerlerine göre daha başarılı olduğu görülmektedir.



Şekil 11. Yaprak lekesi hastalığı verileri için roc analiz grafiği (Roc analysis graph for leaf spot disease data)

Şekil 12’de yaprak yanığı hastalığı verileri için kNN 0.890, yapay sinir ağı 0.912 ve rastgele orman algoritması için 0.869 doğruluk oranları ortaya çıkararak, yapay sinir ağı algoritmasının doğruluk oranının diğerlerine göre daha başarılı olduğu sonucuna varılmıştır.



Şekil 12. Yaprak yanığı hastalığı verileri için roc analiz grafiği (Roc analysis graph for leaf burn disease data)

#### 4. SONUÇ, TARTIŞMA VE ÖNERİLER

(CONCLUSION, DISCUSSION AND SUGGESTIONS)

Elde edilen sonuçlar, oluşturulan yapay sinir ağı modelinin mısır yaprağı hastalığını tahmin etme yeteneğinin eğitimde %97 ve test aşamasında %89 başarı oranıyla olduğunu göstermektedir. Yapılan araştırmada kullanılan yapay zekâ tekniklerinin, geleneksel yöntemlere göre mısır yaprağı hastalıklarını tespit etme konusunda daha etkili olduğu belirtilmektedir. Bu önerilen sistem, mısır hastalıklarının tespitinde zaman ve maliyet açısından büyük bir tasarruf sağlamaktadır. Yapılan çalışma,

yapay sinir ağı modelinin sınıflandırma ve kalıp tanıma uygulamalarında oldukça etkili olduğu sonucuna varmaktadır. Bu sistem, hastalığın otomatik olarak teşhis edilmesine katkıda bulunmayı hedeflemektedir. Oluşturulan model, değerlendirme aşamasında sadece www.kaggle.com sitesinden alınan yaprak resimleriyle değerlendirme yapmaktadır. Farklı bölgelerden ve daha geniş bir veri setinden elde edilecek bilgiler, daha kapsamlı bir değerlendirme yapabilmek için daha doğru sonuçlara ulaşmayı sağlayabilir. Bu verilerin kullanımı, küresel, ulusal veya bölgesel düzeyde çeşitli analizlerin ve yeni bilimsel



araştırmaların yapılmasına büyük katkı sağlayabilir. Belki de bu tür araştırmalar, tarımsal alanlarda geniş bir gelişim sağlamak adına gerçekleştirilecektir. Sonuçlar, düşük hata oranı ve minimum maliyetle hastalık tespitinin yapay zekânın tarım ve bitki yetiştiriciliği alanında kullanılması gerekliliğini vurgulamaktadır. Mobil uygulama geliştirilerek çiftçilerin kullanımına sunulabilir ve bu sayede çekilen resimlerle veri setleri zenginleştirilerek tahmin hassasiyeti artırılabilir. Yapay sinir ağı algoritması, diğer sınıflandırıcılar gibi, mısır yaprağı hastalıklarının teşhisinde önemli bir tanı yöntemi olarak değerlendirilebilir.

#### KAYNAKLAR (REFERENCES)

- [1] Dünya Bankası (30 Eyl 2022), “Genel Bakış”. Erişim Adresi: <https://www.worldbank.org/en/topic/agriculture/overview>
- [2] Turgay, E. B., Büyük, O., Tunalı, B., Kurt, Ş., Akçalı, E., Baran, B., & Kansu, B. (2017). Türkiye’de önemli mısır (*Zea mays* L.) alanlarında Kuzey [Exserohilum turcicum (Pass.) KJ Leonard & Suggs] ve Güney [Bipolaris maydis (Y. Nisk. & C. Miyake) Shoemaker] mısır yaprak yanıklığı hastalıklarının yaygınlığı. Griffiths, M., (1999), “Internet addiction: Fact or fiction?” The Psychologist, c. 12, S. 5, s. 246–251.
- [3] Tarım ve Orman Bakanlığı (Ocak 2022, Sayı 19). Dünyada Mısır [Mısır bülteni]. Erişim adresi: <https://www.tarimorman.gov.tr/BUGEM/Belgeler/B%C3%BCltenler/OCAK%202022/M%C4%B1s%C4%B1r%20Ocak%20B%C3%BClteni.pdf>Karaman
- [4] Kusumo, B. S., Heryana, A., Mahendra, O., & Pardede, H. F. (2018, November). Machine learning-based for automatic detection of corn-plant diseases using image processing. In 2018 International conference on computer, control, informatics and its applications (IC3INA) (pp. 93-97). IEEE.
- [5] Khoirunnisak, K. M. (2020). Sistem Pakar Diagnosa Penyakit pada Tanaman Jagung dengan Metode Dempster Shafer. Publikasi Tugas Akhir S-1 PSTI FT-UNRAM.
- [6] Wiesner-Hanks, T., Stewart, E. L., Kaczmar, N., DeChant, C., Wu, H., Nelson, R. J., ... & Gore, M. A. (2018). Image set for deep learning: field images of maize annotated with disease symptoms. BMC research notes, 11(1), 1-3.
- [7] Le, E. P. V., Wang, Y., Huang, Y., Hickman, S., & Gilbert, F. J. (2019). Artificial intelligence in breast imaging. Clinical radiology, 74(5), 357-366.
- [8] Göksu, M., Sünnetçi, K. M., & Alkan, A. Derin öğrenme ağları kullanılarak mısır yapraklarında hastalık tespiti. Computer Science, (Special)(2021), 208-216.
- [9] Zhang, Z. Y., He, X. Y., Sun, X. H., Guo, L. M., Wang, J. H., & Wang, F. S. (2015). Image recognition of maize leaf disease based on GA-SVM. Chemical Engineering Transactions, 46, 199-204.
- [10] Panigrahi, K. P., Sahoo, A. K., & Das, H. (2020, June). A cnn approach for corn leaves disease detection to support digital agricultural system. In 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184) (pp. 678-683). IEEE.
- [11] Sibiya, M., & Sumbwanyambe, M. (2019). A computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks. AgriEngineering, 1(1), 119-131.
- [12] Ahila Priyadharshini, R., Arivazhagan, S., Arun, M., & Mirnalini, A. (2019). Maize leaf disease classification using deep convolutional neural networks. Neural Computing and Applications, 31(12), 8887-8895.
- [13] Bhangé, M., & Hingoliwala, H. A. (2015). Smart farming: Pomegranate disease detection using image processing. Procedia computer science, 58, 280-288.
- [14] Kapon, O. (2022, February). Kaggle: <https://www.kaggle.com/code/omreekapon/corn-and-maizediseases-classification/notebook>. Erişim: 04.10.2022
- [15] Whxna. (2022, August). Kaggle: <https://www.kaggle.com/code/jiaoowoguanren/corn-ormaize-leaf-dataset-tf-squeezenet/notebook>. Erişim: 04.10.2022
- [16] Zhao, Y.-X., Wang, K.-R., Bai, Z.-Y., Li, S.-K., Xie, R.-Z., & Gao, S.-J. (2009). Research of Maize Leaf Disease Identifying Models Based Image Recognition. Crop Modeling and Decision Support, 1(2004), 317–324. [https://doi.org/10.1007/978-3-642-01132-0\\_35](https://doi.org/10.1007/978-3-642-01132-0_35).
- [17] Lv, M., Zhou, G., He, M., Chen, A., Zhang, W., & Hu, Y. (2020). Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet. IEEE Access, 8, 57952–57966. <https://doi.org/10.1109/ACCESS.2020.2982443>.
- [18] An, J., Li, W., Li, M., Cui, S., & Yue, H. (2019). Identification and classification of maize drought stress using deep convolutional

- neural network. *Symmetry*, 11(2), 1–14. <https://doi.org/10.3390/sym11020256>.
- [19] Song, K., Sun, X. Y., & Ji, J. W. (2007). Corn leaf disease recognition based on support vector machine method. *Transactions of the CSAE*, 23(1), 155-157.
- [20] Rage, S. (2022). Kaggle: <https://www.kaggle.com/code/sailikhitara/m-aize-mutantclassification-using-vgg16/notebook>. Erişim: 04.10.2022
- [21] Kiliç, d., Borandağ, E., Yücalar, f., Tunali, v., Şimşek, M., & Özçift, A. (2016). Knn Algoritması Ve R Dili Ile Metin Madenciliği Kullanılarak Bilimsel Makale Tasnifi. *Marmara Fen Bilimleri Dergisi*, 28(3), 89-94.
- [22] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- [23] Elmas, Ç. (2016). *Yapay Zeka Uygulamaları 3*. Baskı. Ankara: Seçkin Yayıncılık.