



Araştırma Makalesi - Research Article

Topluluk Öğrenmesi Algoritmaları Kullanarak Amazon Yemek Yorumları Üzerine Duygu Analizi

Sentiment Analysis on Amazon Food Reviews Using Ensemble Learning Algorithms

Adem Hilmi Bozkurt¹, Nesibe Yalçın^{2*}

Geliş / Received: 23/05/2023

Revize / Revised: 22/07/2023

Kabul / Accepted: 29/08/2023

ÖZ

Günümüzde kullanıcılar satın aldıkları ürün ve/veya hizmetler ile ilgili görüş, öneri ve şikayetlerini e-ticaret sitelerinde paylaşmayı daha çok tercih etmekte veya diğer kullanıcıların yazdıkları yorumları dikkate alarak satın alma işlemine karar verebilmektedir. Duygu analizi, kullanıcıların ürünlere karşı tutumlarını belirlemede oldukça yararlıdır. Bu çalışmada, topluluk öğrenmesi tabanlı duygu analizi yapılması amaçlanmıştır. Bu amaçla, ilk olarak Amazon yemek yorum verileri üzerinde veri ön işleme adımları gerçekleştirilmiş, metin ön işleme yapılmış ve öznelik mühendisliği ile veriler öğrenme modelleri için istenilen formata getirilmiştir. Daha sonra Rastgele Orman, CatBoost ve XGBoost topluluk öğrenmesi modelleri kullanılarak olumlu, olumsuz ve tarafsız (nötr) duygular şeklinde sınıflandırma yapılmıştır. Eğitim ve test doğruluğu, kesinlik, duyarlılık ve F1-skor metrikleri açısından sonuçlar değerlendirilmiştir. En yüksek %90,22 test doğruluk değerine Rasgele Orman ve CountVectorizer tekniğinin kullanıldığı sınıflandırmada ulaşılmıştır. Ayrıca web kazıma ile oluşturulan yeni bir veri seti üzerinde modeller test edilmiştir. Elde edilen sonuçlar işletmenin ürün ve hizmetleri ile ilgili bir öngörü sunmaktadır.

Anahtar Kelimeler- Topluluk Öğrenmesi, Doğal Dil İşleme, Duygu Analizi, Öznelik Mühendisliği, Web Kazıma

ABSTRACT

Nowadays, users prefer to share their opinions, suggestions, and complaints about the products and/or services they buy more on e-commerce sites, or they can decide to buy based on the comments written by other users. Sentiment analysis is very useful in determining their attitudes towards them. This study aims to perform ensemble learning-based sentiment analysis. For this purpose, firstly, data pre-processing was implemented on Amazon meal review data, the text was preprocessed, and the data was brought into a format suitable for learning models with feature engineering. Then the text has been classified according to whether it contained positive, negative, or neutral sentiments using Random Forest (RF), CatBoost, and XGBoost ensemble learning models. The obtained results have been evaluated in terms of training and test accuracy, precision, recall, and F1-score metrics. The highest test accuracy of 90.22% has been achieved in the classification using RF and CountVectorizer technique.

¹İletişim: 1030520343@erciyes.edu.tr (<https://orcid.org/0009-0004-1330-3569>)

Bilgisayar Mühendisliği Bölümü, Erciyes Üniversitesi, Kayseri, Türkiye

^{2*}Sorumlu yazar iletişimi: nesibeyalcin@erciyes.edu.tr (<https://orcid.org/0000-0003-0324-9111>)

Bilgisayar Mühendisliği Bölümü, Erciyes Üniversitesi, Kayseri, Türkiye

In addition, the models have been tested on a new dataset created with web scraping. The results obtained provide insight into the company's products and services.

Keywords- Ensemble Learning, Natural Language Processing, Sentiment Analysis, Feature Engineering, Web Scraping

I. GİRİŞ

Birçok ticari web sitesi kullanıcılarını satın aldıkları ürün ve hizmetler ile ilgili görüşlerini bildirmeye, aldıkları ürün ve hizmet hakkında yorum yapmaya ve geri bildirimde bulunmaya teşvik etmektedir. Bu geri bildirimler, müşterilerin ürün ve hizmetten memnun olup olmadığını ortaya koyan bir veri kaynağı olarak, diğer müşterilerin ürünü satın al(ma)ma hakkında daha kolay karar verebilmesinde, beklentilerini karşılayacak olan doğru ürünü seçebilmesinde, satıcının veya hizmet verenin tercih edilip edilmemesinde ve müşteri/kullanıcı beklentileri doğrultusunda satıcının yönlendirilmesinde (hizmet iyileştirme, ürün geliştirme, pazar araştırması yapma, reklam stratejisi belirleme, marka itibarı artırma gibi) oldukça önemli hale gelmiştir. Kullanıcı sayısı, yorum sayısı, ürün/hizmet sayısı ve çeşitliliği göz önüne alındığında yorumların analizinin hızlı ve otomatik yapılması ihtiyacı ise her geçen gün artmaktadır.

Doğal Dil İşleme (Natural Language Processing, NLP), bilgisayarların insan dilini okumasını, yorumlamasını, anlam çıkarmasını ve yine insanların anladığı dilde iletmesini sağlayan yapay zeka alanıdır [1]. Veri setinden bilgisayarların anlayabileceği anlamların çıkarılmasına imkan tanır. Böylece, bilgisayarlardan faydalanarak daha hızlı ve daha doğru analizler gerçekleştirilebilmektedir [2]. Bir NLP yöntemi olan duyu analizi veya fikir madenciliği, bir yazar tarafından yazılmış bir metni ya da belgeyi analiz ederek o metinde gömülü olan duyu veya hissin belirlenmesini, duyu kategorisinin ve polaritesinin tanımlanmasını sağlar [3-5]. Her dilin dil bilgisi (yapısı) belli başlı kurallar içerir. Bu kurallar doğrultusunda isim, nesne, sıfat, isim fiil vs. gibi cümlenin öğelerinin her birinin cümleye kattığı duygular vardır. Duyu analizi ile bu duyguların olumluluk, olumsuzluk ve tarafsızlık puanları belirlenerek cümleye skor tayin edilmektedir. Duyu analizine yönelik ilk çalışmalar, metinleri olumlu veya olumsuz duyu içermesi şeklinde ayırma üzerine olmuştur. Günümüzde dil biliminde, pazar araştırmasında, sosyal medya izlemelerinde, kamuoyu yoklamalarında, ürün ve hizmet deneyiminde, müşteri sesi (Voice of the Customer, VoC) ve çalışan sesi (Voice of the Employee, VoE) analizinde, kullanıcı eğilimlerini anlamada, reklam ve önerileri özelleştirmede fayda sağlamaktadır. Bununla birlikte ürün yorumları, müşteri sesi, sosyal medya gönderileri, blog yazıları, anket verileri ve kullanıcılar tarafından oluşturulan diğer çevrimiçi içerikler duyu analizi için veri kaynaklarıdır.

Bu çalışmada, Amazon yemek yorumları, yorumun olumlu, olumsuz veya tarafsız (nötr) duyu içerip içermediğini belirlemek için analiz edilmiştir. Analiz aşaması veri temizleme ile başlamış ve temizlenen veri seti, makine öğrenmesi algoritmaları için NLP yaklaşımları ile uygulanabilir bir formata (vektörlere) dönüştürülmüştür. Rastgele Orman (Random Forest, RF), Ekstrem Gradyan Yükseltme (eXtreme Gradient Boosting, XGBoost) ve Kategorik Yükseltme (Category Boosting, CatBoost) topluluk makine öğrenmesi algoritmaları ile sınıflandırma işlemi gerçekleştirilmiş ve farklı vektörleştirme tekniklerinde başarıyı karşılaştırılmıştır. RF ile CountVectorizer (sayı vektörleştirici) tekniğinde en başarılı sınıflandırma sonuçları elde edilmiştir. Çalışma kapsamında ayrıca Amazon web sitesinden web kazıma ile veri toplanmış ve yeni bir veri seti oluşturulmuştur.

II. LİTERATÜR ARAŞTIRMASI

Duyu analizi, birçok sektörde çeşitli ürünlere, hizmetlere, olaylara ve durumlara ilişkin mevcut yorumlara (görüşlere) göre yorumların etkinliğini ölçmek, verilen puanlarla yapılan yorumlar arasındaki ilişkiyi anlamak, yorumlardan ihtiyaç duyulan bilgiyi çıkarmak gibi çok çeşitli kullanım amaçlarına sahiptir. Makine öğrenme algoritmaları ile duyu analizi, ilk olarak Pang vd. tarafından [2, 6] sinema film yorumları üzerinde olumlu ve olumsuz sınıflandırma şeklinde gerçekleştirilmiştir. Sonrasında sosyal medya, pazarlama, politika, piyasalar, müşteri hizmetleri gibi birçok alanda uygulama alanı bulmuştur. Örneğin, sosyal medya akışından duyu analizi konulu bir çalışmada [7], Twitter web sitesinden politik bir olaya ilişkin alınan veriler kullanılmış, veri setine normalleştirme, köklendirme ve engellenmiş sözcükleri kaldırma ön işleme adımları uygulanmış, NLP yaklaşımı ile veriler uygulanabilir bir formata dönüştürülmüştür. Kelime torbası (Bag of Words, BoW), Terim Frekansı-Ters Doküman Frekansı (Term Frequency - Inverse Document Frequency, TF-IDF), Word2Vec metin özellikleri kullanılarak veri seti vektörleştirilmiştir. Lojistik Regresyon (Logistic Regression, LR), Destek Vektör Makinesi (Support Vector Machine, SVM), RF makine öğrenmesi ve derin öğrenme algoritmaları kullanılarak sosyal medya verileri olumlu veya olumsuz olarak etiketlenilmiştir. Derin öğrenme algoritmaları ile daha yüksek doğruluk (accuracy) elde edildiği gözlenmiştir. Twitter duyu analizine odaklanan bir başka çalışmada [8], Multinomial Naïve Bayes (NB), SVM, RF ve LR öğrenme algoritmaları ile birlikte Değerlik Farkındalıklı Sözlük ve Duygu Akıl Yürütücü (Valence Aware Dictionary and sEntiment Reasoner, VADER) aracı kullanılmıştır. Çalışma kapsamında ayrıca Etki Alanı Tabanlı Sözlük (Domain Based Lexicon) yaklaşımı geliştirilmiştir.

Geliştirilen yaklaşım ile olumlu ve olumsuz duygu şeklinde yapılan sınıflandırma sonucunda en yüksek %89,43 doğruluk elde edilmiştir. Ayrıca yapılan çalışmalarda duygular sadece olumlu ve/veya olumsuz olarak sınıflandırılmakla kalmayıp mutluluk, şaşkınlık, kızgınlık gibi metinde ifade edilmek istenen duygu belirlenmeye çalışılmıştır. Örneğin Topçuoğlu vd., TREMO veri seti kullanılarak yaptıkları Türkçe metinlerde duygu analizi çalışmasında [9], 6 duygu kategorisi (mutluluk, üzüntü, korku, şaşırma, öfke ve tikslenme) üzerine odaklanmış ve çeşitli makine öğrenmesi algoritmaları ile sınıflandırma gerçekleştirmiştir. Sınıflandırmada Yapay Sinir Ağı (Artificial Neural Network, ANN) en başarılı sonuçları vermiştir.

Müşteri hizmetleri yönetimi, şirketlerin müşterilerin ihtiyaçlarına odaklanmasını ve müşteri geri bildirimlerine göre daha hızlı aksiyon almasını sağlar. Duygu analizi, çevrimiçi müşteri geri bildirimlerini daha iyi analiz edebilmek için büyük bir avantajdır. Örneğin, Seyfioğlu ve Demirezen tarafından özel bir havayolu şirketine ilişkin müşteri değerlendirme verileri ele alınmış ve bu verilere yönelik yapılan bir duygu analizi ve kategorizasyonu çalışması [10] yapılmıştır. Çalışma kapsamında ilk olarak, veri karmaşıklığını azaltmak için veri ön işleme yapılmış, daha sonra Word2Vec ve TF-IDF gibi teknikler ile özellik çıkarımı gerçekleştirilmiş ve XGBoost sınıflandırıcısı ile eğitilmiştir. Olumlu ve olumsuz olmak üzere ikili sınıflandırma yapılmış ve daha sonra olumsuz duyguların daha fazla kategorize edilmesi için Doc2Vec yaklaşımını kullanan bir sınıflandırma modeli önerilmiştir. Çalışma sonucunda %71,16 doğruluk elde edilmiştir. Restoranlar için önerilen özellik tabanlı bir duygu analizi çalışmasında [3], Yelp yorumları kullanılmış ve ambiyans, fiyat, hizmet ve yemek kalitesi özelliklerine göre sınıflandırma gerçekleştirilmiştir. Sınıflandırma sonucunda en yüksek doğruluk %79,12 ile NB ile elde edilmiştir. Restoran (Yelp), film (IMDB) ve ürün (Amazon) yorumları içeren farklı veri setleri ve bu veri setlerinin birleştirilmiş versiyonu üzerinde yapılan bir duygu analizi çalışmasında [2], sınıflandırma için K-En Yakın Komşu (K-Nearest Neighbor, KNN), NB, SVM, Karar Ağacı (Decision Tree, DT) ve RF algoritmaları kullanılmıştır. Sınıflandırma sonucunda SVM algoritmasının diğer algoritmalara kıyasla daha iyi sonuçlar verdiği görülmüştür.

Amazon yemek yorumları üzerine yapılan analiz çalışmasında [11], veri üzerinde ilk olarak lematizasyon (lemmatization, anlamsal köke inme) ve kök bulma (stemming) belirli ön işleme adımları uygulanmış ve bazı metin kodlama (text encoding) tekniklerinden yararlanılmıştır. LR ve SVM algoritmaları kullanılarak ikili sınıflandırma gerçekleştirilmiş ve sırasıyla %91,42 ve %93,11 test doğruluğu elde edilmiştir. Ayrıca ortalama Word2Vec metin kodlama tekniğinin daha yararlı olduğu tespit edilmiştir. Amazon yemek yorumları üzerine yapılan bir başka çalışmada [12], BoW ve TF-IDF metotları kullanılmış. Yorumlar, Multinomial NB algoritması ile pozitif ve negatif şeklinde ikili sınıflandırılmıştır. TF-IDF ile %88,65 ve BoW ile %85,75 doğruluk oranı elde edilmiştir. [13], kullanıcıların yemekten memnun kalıp kalmadığını doğru bir şekilde tahmin edebilmek için Evrişimli Sinir Ağı (Convolutional Neural Network, CNN)'na dayalı bir sınıflandırma modeli önermiştir. Önerilen model ile %86 üzerinde kesinlik (precision), duyarlılık (recall) ve F1-skor (F1-score) değerleri elde edilmiştir. [4] ise SVM, RF, NB ve KNN gibi makine öğrenmesi algoritmaları kullanarak yorumları, olumlu veya olumsuz olarak sınıflandırmıştır. Tüm algoritmalar arasında yaklaşık %94 doğruluk ile en iyi sonuçları SVM vermiştir. Kazmaier ve Van Vuuren tarafından duygu analizinde topluluk öğrenmesinin performans üzerindeki etkisinin araştırıldığı çalışmada [1], bireysel modellere göre yaklaşık %5,53 ortalama performans iyileştirmeleri elde edilmiştir. Uzavcı yüksek lisans tezinde [14], Twitter sosyal medya verileri üzerinde makine öğrenmesi ve topluluk öğrenmesi algoritmalarını kullanarak duygu analizi gerçekleştirmiştir. Ayrıca, uzman olmayan kişilerin de analiz yapabilmesine imkan tanıyan bir sanal asistan tasarlamıştır. Sanal asistan ile kullanıcıya, Twitter verisinin toplanması, işlenmesi ve analiz edilmesi aşamalarının öğretilmesi amaçlanmıştır.

Topluluk öğrenmesi algoritmaları, çoklu öğrenme algoritmalarının tahminlerini birleştirerek DT, SVM ve LR gibi temel öğrenme algoritmalarından daha yüksek sınıflandırma/tahmin performansı elde etmeyi amaçlar. Ayrıca yeterli veri olmaması veya sınırlı veri olması durumlarında umut verici sonuçlar verebilmektedir [1, 15-17]. Genelleştirme yeteneği daha yüksek olan modeller üretmeyi amaçlayan topluluk öğrenmesi algoritmalarına artan bir ilgi vardır, ancak duygu analizinde kullanımı hala sınırlıdır [1]. Bu çalışmada, duygu analizini daha yüksek başarımla gerçekleştirmek için topluluk öğrenme algoritmalarından RF, XGBoost ve CatBoost kullanılmıştır.

III. MATERYAL VE METOT

A. Veri Seti

Büyük ölçekli şirketler, milyonlarca kullanıcısı olan e-ticaret siteleri müşterilerinden/kullanıcılarından ürün ve hizmetleri ile ilgili deneyimlerine ilişkin günlük on binleri aşan geri bildirim (yorum, puan vb.) almaktadır. Bu web sitelerinden biri de Amazon'dur. Bu çalışmada, Kaggle'daki Amazon'un yemek hizmetine yapılan yorumların yer aldığı bir veri seti [18] kullanılmıştır. Veri seti, Ekim 1999 ve Ekim 2012 tarihlerine ait 256,059 kullanıcı tarafından 74,258 ürün hakkında yapılan toplamda 568,454 yorum içermektedir. Ayrıca 260 kullanıcıya ilişkin 50'den fazla yorum yer almaktadır. Veri setinin kısa bir örneği Şekil 1'de verilmiştir.

algoritmaları, RF'nin temelini oluşturan bir topluluk öğrenmesi yöntemidir. Orijinal eğitim veri setinden yinelemeli olarak tekrar yerine koyacak şekilde rastgele alt kümeler oluşturur (önyükleme). Her bir alt kümedeki örnek sayısı, veri setindeki örnek sayısına eşittir. Bazı örnekler alt kümelere yer almaz, bazıları ise iki veya daha fazla yer alır ve böylece orijinal veri setindeki örnek sayısına ulaşılmış olur. Daha sonra alt kümeler ile bireysel/temel öğrenme algoritması eğitilir. Oluşturulan her model birbirinden bağımsızdır, paralel olarak çalışır ve daha sonra modellerin çıktıları aritmetik ortalama veya çoğunluk oylama ile birleştirilir. Torbalamada aynı öğrenme algoritması kullanılır, bu yüzden homojen topluluk olarak kabul edilir. Yükseltme algoritmalarında, torbalama yönteminde olduğu gibi farklı dağıtılmış alt kümelere göre eğitim gerçekleşir, ancak süreç, sıralı ve bağımlıdır [1]. Orijinal veri setinden oluşturulan ve eşit sınıf dağılımına (ağırlığa) sahip alt küme ile birinci bireysel öğrenme algoritması eğitilir, model oluşturulur. Yanlış sınıflandırılan sınıflara daha fazla ağırlık verilecek şekilde veri dağılımı ayarlanır. İkinci model bu dağılıma göre eğitilir ve bir önceki modelin zayıflıkları ele alınarak tahmin hataları giderilmeye çalışılır. Süreç belirli sayıda tekrarlanarak tek bir güçlü öğrenme modeline ulaşmak amaçlanır. Bu çalışmada, RF (torbalama), XGBoost ve CatBoost (yükseltme) topluluk öğrenmesi algoritmaları kullanılmıştır.

1) Rastgele Orman (Random Forest, RF)

RF algoritması, sınıflandırma ve regresyon amaçlı kullanılabilir [8, 24]. Veri setinden seçtiği rastgele örneklerle dayalı olarak çok sayıda DT oluşturur, kök düğümü bulma ve düğümleri bölme işlemlerini rastgele gerçekleştirir. Model doğruluğunu artırmak için her bir DT tarafından döndürülen tahmin sonuçlarına göre ağırlıklandırma yapar. Sınıflandırma probleminde ağırlıklandırma yaklaşımı, her bir tahmin sonucu için oylama (çoğunluk/ağırlıklı) yapılarak, regresyon probleminde ise sonuçların ortalaması (aritmetik/ağırlıklı) alınarak gerçekleşir. Elde edilen sonuç, nihai tahmin sonucu olarak sunulur ve böylece en yüksek doğrulukla tahmin amaçlanır. RF içinde oluşturulacak ağaç sayısı ($n_{estimators}$), ağaç oluşturulurken dikkate alınacak maksimum özellik sayısı ($max_features$), ağacın maksimum derinliği (max_depth) ve bir düğümü bölmek için gereken minimum örnek sayısı gibi ($min_samples_split$) gibi hiperparametreleri için en uygun değerler seçilerek başarılı sonuçlara ulaşılmaya çalışılır.

2) XGBoost

Gradyan Yükseltme ve DT'ye dayalı bir topluluk öğrenmesi yöntemidir. Kısa zamanda çok fazla veriyi analiz etme, aşırı öğrenmeyi ve tahmin yanlılığını azaltma, yüksek doğrulukla tahmin, daha hızlı sınıflandırma gibi avantajlar sunabilmektedir [25]. Paralel çalışır ve ağaç oluştururken max_depth değerini kullanır. RF'deki ağaç oluşturma parametrelerine ek olarak aşırı öğrenmeyi önlemek için 0 ile 1 aralığında değer alan öğrenme oranı ($learning_rate$) ve eğitim örneklerinin alt örneklem oranı ($subsample$) gibi parametreleri de ayarlanabilmektedir. Dengesiz sınıflandırma problemleri için $scale_pos_weight$ (genellikle, negatif örnek sayısının pozitif örnek sayısına oranı) parametre ayarı yapılabilir [26].

3) CatBoost

Daha iyi tahmin çıktıları üretmek için önceki hatalarından ders alan gradyan yükseltme tabanlı bir sıralı yükseltme topluluk öğrenmesi algoritmasıdır. Regresyon ve sınıflandırma problemlerine yüksek doğrulukla çözüm sunmakla birlikte çok fazla hiperparametre ayarına ihtiyaç duymamaktadır. CatBoost algoritması ayrıca kategorik değişken desteği sunmaktadır. Başka bir ifade ile kategorik değişkenler üzerinde ek bir ön işlem gerektirmeden çalışabilme özelliğine sahiptir. Veri setindeki kategorik değişkenleri, çeşitli istatistiksel bilgileri kullanarak sayısal dönüştürür ve başarılı bir şekilde işler [25, 27].

D. Başarım Kriterleri

Topluluk öğrenmesi algoritmalarının başarımını değerlendirmek için Tablo 1'de detayları verilen doğruluk, kesinlik, duyarlılık ve F1-skor kriterleri [28] kullanılmıştır. Her bir kriter, Doğru Pozitif (True Positive, TP), Doğru Negatif (True Negative, TN), Yanlış Pozitif (False Positive, FP) ve Yanlış Negatif (False Negative, FN) terimleri kullanılarak hesaplanır. Bu terimler, veri setindeki her bir Pozitif/Negatif etiketli örneğin doğru veya yanlış sınıflandırılmasını/tahmin edilmesini ifade eder.

Tablo 1. Başarım kriterleri.

Kriter	Açıklama	Formül
Doğruluk	Doğru tahminlerin tüm örneklere oranı	$\frac{TP + TN}{TP + FP + TN + FN}$
Kesinlik	Doğru pozitif tahminlerin tüm pozitif tahminlere oranı	$\frac{TP}{TP + FP}$
Duyarlılık	Doğru pozitif tahminlerin toplam pozitif örneklere oranı	$\frac{TP}{TP + FN}$
F1-skor	Kesinlik ve duyarlılık değerlerinin harmonik ortalaması	$2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık}$

E. Kütüphaneler

Yemek yorumları üzerinde duygu analizi için Jupyter Lab, Chrome, Splash, Anaconda Environment ve Windows Command Line Interface platformlarından yararlanılmıştır. Uygulama Python 3.8 ve Markdown Language programlama dilleri ile gerçekleştirilmiş ve Tablo 2’de sunulan kütüphaneler kullanılmıştır.

Tablo 2. Kullanılan kütüphaneler.

Kullanım amacı	Kütüphane
Veri Ön İşleme	numpy pandas
Metin Ön İşleme	re textblob nltk nltk.corpus/stopwords nltk.stem.porter nltk.chunk/ne_chunk, word_tokenize nltk/pos_tag wordcloud/WordCloud
Öznetelik Mühendisliği	nltk.sentiment.vader/SentimentIntensityAnalyzer sklearn.feature_extraction.text/TfidfVectorizer, CountVectorizer
Topluluk Öğrenmesi	sklearn.preprocessing sklearn.ensemble sklearn.model_selection sklearn.model_selection/train_test_split sklearn.model_selection/GridSearchCV, cross_val_score xgboost/XGBClassifier catboost/CatBoostClassifier sklearn.ensemble/RandomForestClassifier sklearn.ensemble/BaggingClassifier sklearn.metrics/accuracy_score
Web Kazıma	bs4/BeautifulSoup requests
Görselleştirme	seaborn matplotlib.pyplot

IV. UYGULAMA

Çalışmada NLP’nin bir parçası olarak çeşitli ön işleme yöntemleri benimsenmiştir. Topluluk öğrenmesi algoritmaları ile daha iyi sonuç elde edebilmek için tokenizasyon (tokenization), kök bulma ve lematizasyon metin ön işleme yapılmış ve veri setindeki etkisiz kelimeler (stopwords), HTML etiketleri (tags) gibi gürültülü veriler kaldırılmıştır. Veri setinin sadeleştirilmesinden sonra sayısal olarak temsil edilebilmesi için TF-IDF, CountVectorizer ve n-gram yöntemleri kullanılmıştır. Vektörleştirme işlemlerinin detayları açıklanmış ve uygulama adımları tek tek gösterilmiştir. Her bir vektörleştirme elemanı (count, n-gram, word, char) ayrı ayrı değerlendirilmiş ve topluluk öğrenmesi algoritmaları tarafından koşulmuştur.

A. Veri Ön İşleme

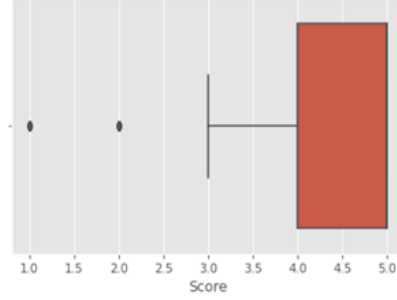
Veri ön işleme, veri setindeki aykırı ve eksik gözlemlerin tespit edilmesi, izlenmesi, görselleştirilmesi ve bir hareket konusunda yargıya varılması işlemidir. Birçok makine öğrenmesi algoritması eksik veriyi desteklememektedir. Bu nedenle, eğitim öncesi bu eksik değerlere sahip verilerin işlenmesi gerekmektedir. Topluluk öğrenme modellerinin geliştirilmesi için kullanılan veri seti, 568.454 satır ve 10 sütundan oluşmaktadır. Veri tipleri incelenmiş, eksik gözlem analizi yapılmış ve Şekil 3’te gösterildiği “Profile Name” ve “Summary” sütunlarında sırasıyla 16 ve 27 eksik gözlem olduğu tespit edilmiştir. Ancak çalışmada bu eksik veriler, uygulanacak yöntemleri etkilememektedir.

```
df.isnull().any()
Id                False
ProductId        False
UserId           False
ProfileName      True
HelpfulnessNumerator  False
HelpfulnessDenominator False
Score            False
Time             False
Summary         True
Text             False
dtype: bool

df.isnull().sum()
Id                0
ProductId        0
UserId           0
ProfileName      16
HelpfulnessNumerator  0
HelpfulnessDenominator  0
Score            0
Time             0
Summary         27
Text             0
dtype: int64
totally 43 NaN values
```

Şekil 4. Eksik gözlem analizi sonuçları.

Eksik gözlem analizinden sonra aykırı (uç) ve çoklu gözlem analizi yapılmıştır. Böylece yanıltıcı yorumlara neden olabilecek verilerin temizlenmesi amaçlanmıştır. Aykırı gözlemler, model normaline uymayan ve modeli bozan; çoklu gözlemler ise tekrar eden, dolayısıyla yararsız olan, verilerdir (aynı değerlere sahip satırlardır). Şekil 5'te görüldüğü üzere puanların çoğunluğu 4-5'tir ve 1-2 puanları ise azınlıktadır. Puan ortalaması 4,2'dir ve puanlamada aykırılık söz konusu değildir.



Şekil 5. Aykırı ve çoklu gözlemler.

Web kazıma ile oluşturulan veri seti ise 5.248 satır ve 2 sütundan oluşmaktadır. Eksik ve çoklu gözlem analizi sonucu sadece puan içeren (boş yorum) 22 adet ve tekrarlı yorum bulunan 506 adet satır silinmiştir. Ön işleme sonucu veri setinde toplamda 4,720 adet yorum ve puan bilgisi yer almaktadır. Puan ortalaması 3,95 ve standart sapması 1,52'dir.

B. Metin Önışleme

Verilerden anlam çıkarmak amacıyla verilerin makine öğrenmesi tarafından işlenebilir hale getirilmesi gerekmektedir. Bu amaçla, yorumlar küçük harflere çevrilmiş ve yorumlarda yer alan HTML etiketleri, noktalama işaretleri ve sayısal veriler silinmiştir. Anlamda değişiklik yapmayan başka bir ifadeyle duygu skoruna katkısı olmayan kelimeler (stopwords: i, you, the, ...) yorumdan çıkartılmıştır. Ayrıca diğer yorumlarda ortak olmayan ve dilde az geçen (nadir) kelimeler de temizlenmiştir. Son olarak lematizasyon (kelimenin çekimli biçimlerini yalın haline dönüştürme) işlemi [29] yapılmış ve yorumların son haline ilişkin bir kesit Şekil 6'da verilmiştir.

```
lemmatized = reviews["Text"].apply(lambda x: " ".join([word(i).lemmatize() for i in x.split()]))
lemmatized.head()
0    bought several vitality canned dog food produc...
1    product arrived labeled jumbo salted peanutsth...
2    confection around century light pillowy citrus...
3    looking secret ingredient robittussin believe f...
4    great taffy great price wide assortment yummy ...
Name: Text, dtype: object
```

Şekil 6. Lematizasyon.

C. Öznitelik Mühendisliği

Bir öğrenme modelinin performansı yalnızca modelin yapısına ve hiperparametrelerine değil, aynı zamanda farklı türdeki nitelikleri nasıl işlediğine ve modele nasıl beslendiğine de bağlıdır. Çoğu makine öğrenmesi algoritması, yalnızca sayısal değişkenleri kabul etmektedir, bu nedenle kategorik değişkenlerin ön işlenmesi gerekli bir adım haline gelir. Modelin değerli bilgileri anlayabilmesi ve anlam çıkarabilmesi için bu kategorik değişkenler sayısallaştırılarak temsil edilmelidir. Öznitelik mühendisliği ile model için gerekli olan en önemli öznitelikler seçilmektedir. Bu aşamada, kelimelerin kökleri ve cümlenin öğeleri analiz edilmiştir. Yorum verileri sayısallaştırılmış ve n-gram (n-kelime sayısı), kelime (word) ve karakter (char) seviyesinde vektörleştirme yapılmıştır. Böylece topluluk öğrenmesi için giriş verileri oluşturulmuştur.

1) VADER Duygu Skorlama

VADER, metinde ifade edilen duyguları tahmin etmek için özel olarak geliştirilmiş, sözlük ve kural tabanlı bir duyarlılık analizi aracıdır [8]. Metnin ne kadar olumlu veya olumsuz duyguya sahip olduğunu (duygu yoğunluğunu) ölçmek için kullanılır. Veri setindeki örneklere ilişkin skorların VADER kullanılarak elde edilme işlemi Şekil 7'de verilmiştir. Şekilde görüldüğü gibi veri setindeki ilk yorum cümlesi; %48,3 olumlu (pos), %51,7 tarafsız (neu) ve %0 olumsuz (neg) olarak derecelendirilmiştir ve birleşik (compound) puanı 0,914'tür. Duygu sınıflarına göre skor dağılımları ise Şekil 8'de sunulmuştur. VADER ile üretilen 5 duygu ölçümünden, olumlu örneklerin çoğunlukla 5 ve olumsuz örneklerin de yoğun olarak 1'de toplandığı görülmüştür.

```
vader = reviews["Text"].apply(lambda x: sia.polarity_scores(x))

vader.head()

0   {'neg': 0.0, 'neu': 0.517, 'pos': 0.483, 'comp...
1   {'neg': 0.129, 'neu': 0.762, 'pos': 0.11, 'com...
2   {'neg': 0.13, 'neu': 0.584, 'pos': 0.286, 'com...
3   {'neg': 0.0, 'neu': 0.868, 'pos': 0.132, 'comp...
4   {'neg': 0.0, 'neu': 0.369, 'pos': 0.631, 'comp...
Name: Text, dtype: object

vaderDf = pd.DataFrame()
vaderDf["neg"] = vader.apply(lambda x: x.get("neg"))
vaderDf["neu"] = vader.apply(lambda x: x.get("neu"))
vaderDf["pos"] = vader.apply(lambda x: x.get("pos"))
vaderDf["compound"] = vader.apply(lambda x: x.get("compound"))

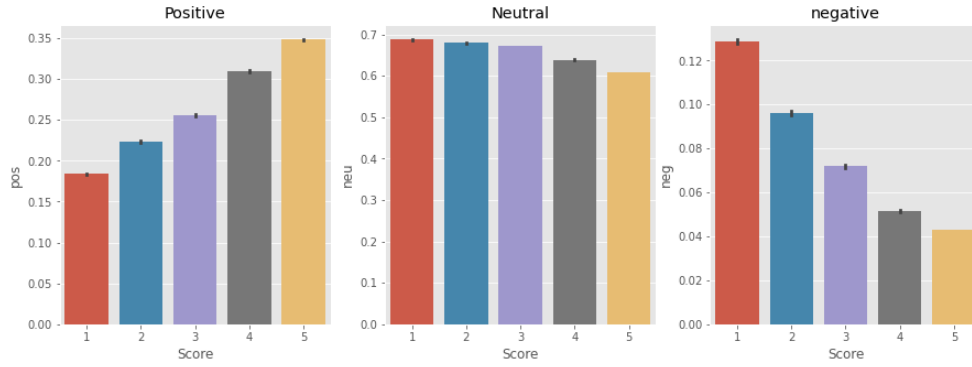
vaderDf.head()

   neg  neu  pos  compound
0  0.000  0.517  0.483   0.9413
1  0.129  0.762  0.110  -0.1027
2  0.130  0.584  0.286   0.8624
3  0.000  0.868  0.132   0.4404
4  0.000  0.369  0.631   0.9468

vaderDf.reset_index()

***
```

Şekil 7. VADER skorlamasının tüm veri setine uygulanması.



Şekil 8. VADER görselleştirme.

2) Kodlama (Encoding)

VADER skorlama ile elde edilen 5 dereceli duygu skorları, SScoring yardımı ile 0 (olumsuz), 1 (tarafsız) ve 2 (olumlu) olacak şekilde normalize edilmiştir. Şekil 9'da sunulduğu gibi birleşik puanı -0,5'den küçük olanlar olumsuz, -0,5 büyük eşit ve 0,5 küçük eşit olanlar tarafsız, 0,5'den büyük olanlar olumlu olarak etiketlenmiştir.

```
vaders["SScoring"] = vaders["compound"].apply(lambda x: 0 if x < -0.5 else 1 if (x >= -0.5 and x <= 0.5) else 2)
```

Şekil 9. Kodlama.

3) Vektörleştirme

Yorum verileri, CountVectorizer ve TfidfVectorizer fonksiyonları kullanılarak vektörlere dönüştürülmüştür. TF-IDF dönüşümü ile elde edilmiş veri setinde n-gram ile öznitelik seçimi yapılmıştır. N-gram modeli, kelime gruplarını öznitelik olarak kullanarak BoW modeline kıyasla bilgi kaybını azaltmaktadır [25]. Çalışmada, ikili ve üçlü kelime grupları içeren n-gram (Şekil 10) ile karakter ve kelime düzeyinde öznitelikler kullanılmıştır. Özellikler, kelime düzeyinde kelimeler ile karakter düzeyinde ise kelimeleri oluşturan karakterler ve karakter grupları ile temsil edilmektedir.


```

tf_idf_ngram_vectorizer = TfidfVectorizer(ngram_range = (2,3))
tf_idf_ngram_vectorizer.fit(x_train)

TfidfVectorizer
TfidfVectorizer(ngram_range=(2, 3))

tf_idf_ngram_vectorizer.get_feature_names()[10:15]

['_ instead water_',
 '_ this',
 '_ this bit',
 '_ variety',
 '_ variety calculated']

x_train_tf_idf_ngram = tf_idf_ngram_vectorizer.transform(x_train)
x_test_tf_idf_ngram = tf_idf_ngram_vectorizer.transform(x_test)

```

Şekil 10. İkili ve üçlü kelime düzeyinde öznitelikler.

V. SONUÇLAR

Veri seti, %75'i eğitim ve %25'i test amaçlı kullanılmak üzere iki gruba ayrılmıştır. Duygu analizi için üç farklı topluluk öğrenmesi algoritması (RF, XGBoost ve Catboost) kullanılmış ve olumlu, olumsuz ve tarafsız duygu sınıflandırması yapılmıştır. Yapılan analizler sonucunda algoritmalar için belirlenen en uygun hiperparametre değerleri Tablo 3'te sunulmuştur.

Tablo 3. Parametre optimizasyonu.

Algoritma	Hiperparametreler
RF	max_depth = 8 min_samples_split = 5 max_features = 8 n_estimators = 1000
XGBoost	min_samples_split = 2 subsample = 0,6 max_depth = 6 learning_rate = 0,001 n_estimators = 1000
CatBoost	max_depth = 5 learning_rate = 0,03

Her bir algoritma için 10 katlı çapraz-doğrulama (10 fold cross-validation) uygulanmıştır. Farklı seviyede kelime grupları kullanılarak analizler yapılmış ve analizler sonucunda Tablo 4'te verilen eğitim doğruluğu ile test verisine ilişkin elde edilen doğruluk, kesinlik, duyarlılık ve F1-skor değerlerine ulaşılmıştır.

Tablo 4. Sınıflandırma sonuçları.

Algoritma	Vektörleştirme Yöntemi	Eğitim Doğruluğu (%)	Test Doğruluğu (%)	Kesinlik	Duyarlılık	F1-skor	
RF	Count vector	99,99	90,22	0,9343	0,9990	0,9656	
	TF-IDF vector	n-gram	99,99	88,91	0,9592	0,9999	0,9791
		kelime düzeyi	99,99	89,79	0,9362	0,9991	0,9666
		karakter düzeyi	99,99	89,04	0,9795	0,9999	0,9896
XGBoost	Count vector	90,18	89,41	0,8437	0,9761	0,9051	
	TF-IDF vector	n-gram	90,53	89,57	0,8545	0,9733	0,9100
		kelime düzeyi	90,54	89,52	0,8529	0,9778	0,9111
		karakter düzeyi	90,02	88,29	0,8915	0,9824	0,9347
CatBoost	Count vector	89,76	89,43	0,8302	0,9745	0,8966	
	TF-IDF vector	n-gram	89,81	89,54	0,8408	0,9742	0,9026
		kelime düzeyi	90,05	89,63	0,8369	0,9778	0,9019
		karakter düzeyi	86,98	86,46	0,8539	0,9729	0,9095

Tabloda görüleceği üzere doğruluk değerleri açısından en yüksek performans RF algoritması ile elde edilmiştir. En yüksek doğruluk değerlerine ise RF ve CountVectorizer yöntemi birlikteliğinde ulaşılmıştır. RF yöntemi ile eğitim setinde %99,99 ve test setinde %88,91 ve üzeri doğrulukla sınıflandırma gerçekleştirilmiştir. XGBoost yöntemi ile %90 ve üzeri, CatBoost yöntemi ile yaklaşık %87 ve üzeri doğrulukla eğitimler gerçekleştirilmiştir. Yükseltme yöntemleri ile ulaşılan kesinlik değerleri, RF yöntemine nazaran daha düşüktür.

Farklı vektörleştirme işlemleri sonucu elde edilen duyarlılık değerlerinin öğrenme yöntemi bazında birbirlerine yakın olduğu gözlenmiştir. Karakter düzeyinde vektörleştirme yapıldığında elde edilen kesinlik ve F1-skor değerlerinin, diğer vektörleştirme yöntemleri kullanılarak elde edilenlere göre daha yüksek olduğu görülebilir. XGBoost ve CatBoost yöntemleri, RF yöntemine kıyasla daha düşük doğruluk oranlarında sınıflandırma gerçekleştirmiştir (Tablo 4). Buna gerekçe olarak veri setindeki duygu sınıflarının dengesiz dağılımı verilebilir. Yükseltme öğrenme modellerinin eğitiminde sınıf dengesizliği ile mücadele etmek için `scale_pos_weight` parametresi ayarlanabilir. Ayrıca sınıf bazında performans analizi yapılarak daha doğru sınıf tanımlama için daha gelişmiş tekniklere (SMOTE [30], MSMOTE gibi örnekleme yöntemlerine) başvurulabilir.

Literatürde aynı veri seti üzerine yapılmış çalışmalar ile incelendiğinde (Tablo 5), “olumlu” veya “olumsuz” olmak üzere ikili sınıflandırma yapıldığı görülmüştür. LR ve SVM algoritmaları kullanılarak yapılan sınıflandırmada [11], sırasıyla %91,42 ve %93,11 test doğrulukları elde edilmiştir. Multinomial NB algoritması ile yapılan sınıflandırma çalışmasında [12], en yüksek %88,65 doğruluk oranına ulaşılmıştır. SVM, RF, NB ve KNN algoritmalarının kullanıldığı çalışmada [4] ise yaklaşık %94 sınıflandırma ile en iyi sonuçları SVM vermiştir. Bu çalışmada, topluluk öğrenme yöntemlerinin duygu analizinde etkinliği araştırılmıştır. Tablo 5 incelendiğinde literatürdeki çalışmalardan farklı olarak “olumlu”, “tarafsız” veya “olumsuz” şeklinde sınıflandırma gerçekleştirilmiştir. Çalışma kapsamında, %99,99 eğitim ve %90,22 test doğruluğu ile başarılı sonuçlar elde edilmiştir. TF-IDF vektörleştirme yöntemini kullanan ve ikili sınıflandırma gerçekleştiren [11] ve [12] çalışmalardan daha yüksek başarımlar elde edilmiştir. Ayrıca RF ve TF-IDF birlikteliği ile yaklaşık %89 oranında test doğruluğuna da ulaşılmıştır.

Tablo 5. Literatür karşılaştırması.

Referans	Açıklama	Yöntem	En Yüksek Başarım (%)
[4]	Olumlu ve olumsuz duygu sınıflandırma	SVM ve TF-IDF	~94
		RF ve TF-IDF	~92
		KNN ve Average Word2Vec	~88
[11]	Olumlu ve olumsuz duygu sınıflandırma	LR ve Average Word2Vec	91,42
		LR ve TF-IDF	89,15
		LR ve BoW	90,03
		SVM ve Average Word2Vec	93,11
		SVM ve TF-IDF	89,45
[12]	Olumlu ve olumsuz duygu sınıflandırma	SVM ve BoW	90,89
		Multinomial NB ve TF-IDF	88,65
Bu çalışma	Olumlu, olumsuz ve tarafsız duygu sınıflandırma	Multinomial NB ve BoW (Count vector)	85,75
		RF ve Count vector	90,22
		XGBoost ve TF-IDF	89,57
		CatBoost ve TF-IDF	89,54

Topluluk öğrenme modelleri ayrıca web kazıma yöntemi ile elde edilen özel (custom) bir veri seti üzerinde test edilmiştir ve performans sonuçları Tablo 6’da sunulmuştur. Bütün modeller ile kesinlik, duyarlılık ve F1-skor açısından başarılı sonuçlar elde edilmiştir. En başarılı sınıflandırma ise XGBoost ve CountVectorizer birlikteliği sonucu %86,10 doğruluk oranı ile gerçekleştirilmiştir.

Tablo 6. Web kazıma ile oluşturulan veri seti için modellerin performansı.

Algoritma	Vektörleştirme Yöntemi	Doğruluk (%)	Kesinlik	Duyarlılık	F1-skor	
RF	Count vector	83,81	0,9215	0,9950	0,9568	
	TF-IDF vector	n-gram	81,86	0,9105	0,9975	0,9520
		kelime düzeyi	83,30	0,9171	0,9975	0,9556
XGBoost	Count vector	karakter düzeyi	77,11	0,9199	1,0000	0,9583
		86,10	0,9196	0,9786	0,9482	
		TF-IDF vector	n-gram	83,81	0,9269	0,9878
CatBoost	Count vector	kelime düzeyi	84,40	0,9229	0,9850	0,9529
		karakter düzeyi	80,08	0,9136	0,9946	0,9524
		84,57	0,9081	0,9881	0,9464	
		TF-IDF vector	n-gram	83,98	0,9233	0,9808
		kelime düzeyi	84,15	0,9229	0,9929	0,9566
		karakter düzeyi	80,76	0,9144	0,9973	0,9541

VI. DEĞERLENDİRME

Günümüzde bir şeyler satın almak için insanlar e-ticaret sitelerini daha çok tercih etmektedir. Ürün satın almadan önce çevrimiçi yorumları gözden geçirme yaygınlaşmakta ve ürünlere ilişkin yorumların önemi daha da artmaktadır. Bu çalışmada, Amazon yemek yorumlarının olumlu, olumsuz veya tarafsız olup olmadığını analiz

etmek için RF, XGBoost ve CatBoost topluluk öğrenmesi algoritmaları kullanılmıştır. En yüksek doğruluk değerlerini elde etmek için çeşitli ön işleme adımları ve öz nitelik mühendisliği yöntemleri uygulanmıştır. En başarılı tahmin sonuçları (%99,99 eğitim ve %90,22 test doğrulukları) RF ile Count Vectorizer yönteminde elde edilmiştir. RF yönteminin dengesiz veri seti için daha sağlam olduğu görülmüştür. XGBoost ve CatBoost ile en yüksek test doğruluk değerleri kelime ve n-gram düzeyinde sunulmuştur ve bu iki yükseltme algoritmaları için elde edilen doğruluk değerleri ise sırasıyla %90,54 (kelime) - %90,53 (n-gram) ve %90,05 (kelime) - %89,81 (n-gram)'dir. Çalışmanın özellikle müşteri hizmetleri yönetimi, müşteri memnuniyeti, kullanıcı eğilimleri, ürün geliştirme, pazar araştırması gibi alanlarda e-ticaret şirketlerine ve hizmet sağlayıcılara katkı sunması beklenmektedir. Ayrıca gerçek zamanlı hedef bir ürün bazında duygu analizi yapılırsa ürüne ilişkin iyileştirme ve reklam stratejisini daha doğru belirleme hususlarında şirket yöneticilerini de destekleyebilir. Bir sonraki çalışmada, geliştirilen topluluk öğrenmesi modelleri Amazon e-ticaret sitesinden web kazıma ile yemeklere ilişkin elde edilen gerçek zamanlı veriler üzerinde test edilecektir.

Gelecek çalışmalarda, noktalama işaretleri veya etkisiz kelimeler nitelik olarak bırakılabilir ve performans üzerinde etkisi analiz edilebilir. Ayrıca öz niteliklerin belirlenmesinde istatistiksel yöntemler (bilgi kazancı, χ^2 gibi) tercih edilebilir, farklı skorlama (örneğin; Roberta, Bert) ve vektörleştirme yöntemleri (örneğin; Word2Vec, BoW) kullanılarak sınıflandırma gerçekleştirilebilir.

KAYNAKLAR

- [1] Kazmaier, J. & Van Vuuren, J. H. (2022). The Power of Ensemble Learning in Sentiment Analysis. *Expert Systems with Applications*, 187, 115819.
- [2] Çelik, E., Dal, D., & Aydın, T. (2021). Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, 27, 880-889.
- [3] Hemalatha S., & Ramathmika, R. (2019). Sentiment Analysis of Yelp Reviews by Machine Learning. *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2019)*, IEEE, Madurai, Hindistan, 700-704.
- [4] Yarkareddy, S., Sasikala, T., & Santhanalakshmi, S. (2022). Sentiment Analysis of Amazon Fine Food Reviews. *2022 IEEE 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 20-22 Ocak 2022, Tirunelveli, Hindistan, 1242-1247.
- [5] Tokcaer, S. (2021). Türkçe Metinlerde Duygu Analizi. *Yaşar Üniversitesi E-Dergisi*, 16 (63), 1514-1534.
- [6] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs Up? Sentiment Classification using Machine Learning Techniques. *Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, 6-7 Temmuz 2002, Philadelphia, ABD, 79-86.
- [7] Demirci, G. M., Keskin, Ş. R., & Doğan, G. (2019). Sentiment Analysis in Turkish with Deep Learning. *2019 IEEE International Conference on Big Data (Big Data)*, 09-12 Aralık 2019, Los Angeles, ABD, 2215-2221.
- [8] B. Engülü. 2018. *Twitter Sentiment Analysis*, Bahçeşehir Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 62s, İstanbul.
- [9] Toçoğlu, M. A., Çelikten, A., Aygün, İ., & Alpkoçak, A. 2019. Türkçe Metinlerde Duygu Analizi için Farklı Makine Öğrenmesi Yöntemlerinin Karşılaştırılması. *DEÜ Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 21 (63), 719-725.
- [10] Seyfioglu, M., & Demirezen, M. (2017). A Hierarchical Approach for Sentiment Analysis and Categorization of Turkish Written Customer Relationship Management Data. *2017 IEEE Federated Conference on Computer Science and Information Systems*, 03-06 Eylül 2017, Prag, Çek Cumhuriyeti, 361-365.
- [11] Das, M., Roy, S., & Saha, R. (2020). Sentiment Analysis on Amazon Fine Food Reviews by using Linear Machine Learning Models. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 8 (IX), 675-678.
- [12] Kumar, P. (2018). *Kaggle - Amazon Fine Food Reviews Analysis - Naive Bayes*, <https://www.kaggle.com/code/premvardhan/amazon-fine-food-reviews-analysis-naive-bayes> (14.01.2023).
- [13] Lasne, K. S., Nandrekar, S. S., Khan, A. A., & Ghorpade, T. (2021). Food Reviews Classification using Multi-Label Convolutional Neural Network Text Classifier. *ITM Web of Conferences (ICACC-2021)*, 14-15 Temmuz 2021, Navi Mumbai, Hindistan, 40, 01009 (1-5).
- [14] Uzavcı, M. (2022). *Türkçe Sosyal Medya İçeriklerinin Analizi İçin Sanal Asistan Tasarımı*, Sakarya Üniversitesi, İşletme Entitüsü, Yüksek Lisans Tezi, 85s, Sakarya.
- [15] Onan, A. (2018). An Ensemble Scheme Based on Language Function Analysis and Feature Engineering for Text Genre Classification. *Journal of Information Science*, 44(1), 28-47.
- [16] Kuş, İ., Bozkurt Keser, S., & Yolaçan, E. N. (2021). Saldırı Tespit Sistemlerinde Topluluk Öğrenme Yöntemlerinin Kıyaslanması. *Avrupa Bilim ve Teknoloji Dergisi*, (31), 725-734.
- [17] Akcan, F., & Sertbaş, A. (2021). Topluluk Öğrenmesi Yöntemleri ile Göğüs Kanseri Teşhisi. *Turkish Studies*, 16(2), 511-527.
- [18] Kaggle. (2017). *Amazon Fine Food Reviews*, <https://www.kaggle.com/datasets/snap/amazon-fine-food-reviews> (12.11.2022).

- [19] Castellano, R. (2016). *Exploratory Data visualization of Amazon fine food reviews*, <https://nycdatasience.com/blog/student-works/exploratory-data-visualization-of-amazon-fine-food-reviews/> (20.07.2023)
- [20] Ahmetoğlu, H., & Daş, R. (2020). Türkçe Otel Yorumlarıyla Eğitilen Kelime Vektörü Modellerinin Duygu Analizi ile İncelenmesi. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 24 (2), 455-463.
- [21] Bozkurt, A. H. (2023). *Kaggle – Amazon Snacks & Sweets Reviews*, <https://www.kaggle.com/datasets/ademhilmibozkurt/newdata?select=5reviews.csv> (20.07.2023)
- [22] Bozkurt, A. H. & Yalçın, N. (2023). Web Kazıma ve Makine Öğrenme Yöntemleri ile Duygu Analizi, *III. Uluslararası Yapay Zeka ve Veri Bilimi Kongresi (ICADA 2023)*, 28 Temmuz 2023, İzmir, Türkiye, 13.
- [23] Rokach, L. (2010). Pattern Classification Using Ensemble Methods. *World Scientific*, 75, 244s.
- [24] Ünalı, S., & Yalçın, N. (2022). Hava Kirliliğinin Makine Öğrenmesi Tabanlı Tahmini: Başakşehir Örneği. *Mühendislik Bilimleri ve Araştırmaları Dergisi*, 4 (1), 35-44.
- [25] Yıldırım, E. (2022). *Hızlandırılmış Makine Öğrenmesi Algoritmaları ile Türkçe Sahte Haber Tespiti*, Karabük Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yüksek Lisans Tezi, 69 s, Karabük.
- [26] dmlc XGBoost. (2022). *XGBoost Parameters*, <https://xgboost.readthedocs.io/en/stable/parameter.html> (20.07.2023)
- [27] Wattanakitrunroj, N., Pinpo, N., & Tongman, S. (2021). Sentiment Polarity Classification using Minimal Feature Vectors and Machine Learning Algorithms, *12th International Conference on Advances in Information Technology (IAIT2021)*, 29 Haziran - 1 Temmuz 2021, Bangkok, Thailand, 18, 1-8.
- [28] Yalçın, N., & Ünalı, S. (2022). Symptom Based COVID-19 Prediction Using Machine Learning and Deep Learning Algorithms. *Journal of Emerging Computer Technologies*, 2(1), 22-29.
- [29] Tahiroğlu, B. T. (2021). Lematizasyon ve Türkçe için bir Lematizasyon Uygulaması: elemanTR. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (24), 475-486.
- [30] Al-Azani, S., & El-Alfy, E. S. M. (2017). Using Word Embedding and Ensemble Learning for Highly Imbalanced Data Sentiment Analysis in Short Arabic Text. *Procedia Computer Science*, 109, 359-366.