



ONDOKUZ MAYIS ÜNİVERSİTESİ  
MÜHENDİSLİK FAKÜLTESİ  
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

# Makine Öğrenimi Ve Derin Öğrenme İle Duygu Analizi

## PROJE RAPORU

ALEYNA KAHRAMAN

DİLARA BIYIKLI

FEYZANUR AYTEKİN

HASAN BASRİ DARGA

Danışman

Dr. Öğr. Üyesi DURMUŞ ÖZKAN ŞAHİN

06, 2024

---

## TEŞEKKÜR

---

Proje boyunca yardımlarını esirgemeyen ve her konuda bilgisini paylaşan danışman hocamız Doktor Öğretim Üyesi Durmuş Özkan ŞAHİN'e teşekkür ederiz.

---

## Ö Z E T

---

Bu çalışmada, iki farklı veri seti üzerinde çeşitli metin temsil yöntemleri ile çeşitli makine öğrenimi ve derin öğrenme algoritmaları kullanılarak ikili ve üçlü sınıflandırma gerçekleştirilmiştir. Metin temsil yöntemleri olarak TF-IDF, GloVe, Word2Vec, FastText ve Bag of Words kullanılmıştır. Makine öğrenimi algoritmalarından Naive Bayes, Lojistik Regresyon, Destek Vektör Makineleri, Rastgele Orman, Yapay Sinir Ağı, En Yakın Komşu Algoritması, Karar Ağacı, XGBoost ve LightGBM uygulanmıştır. Derin öğrenme algoritmaları olarak ise Evrişimli Sinir Ağı, Tekrarlayan Sinir Ağı ve Uzun Kısa Süreli Bellek kullanılmıştır. Elde edilen sonuçlarla, kullanılan metin temsil yöntemleri ve algoritmaların performansları karşılaştırılmıştır. Amazon veri setinde, makine öğrenimi yöntemleri arasında en yüksek doğruluk oranı LightGBM algoritması, derin öğrenme yöntemleri arasında ise TF-IDF ve FastText kullanan LSTM algoritması tarafından elde edilmiştir. IMDb veri setinde, makine öğrenimi yöntemleri arasında en yüksek doğruluk oranı Logistic Regression algoritması, derin öğrenme yöntemleri arasında ise FastText kullanan LSTM algoritması tarafından elde edilmiştir.

**Anahtar Kelimeler:** Makine Öğrenimi, Derin Öğrenme, Duygu Analizi, Metin Sınıflandırma, Metin Temsil Yöntemleri

---

## İÇİNDEKİLER

---

### I GİRİŞ

|                   |   |
|-------------------|---|
| 1 Amaç            | 2 |
| 2 LİTERATÜR ÖZETİ | 3 |

### II MATERYAL

|                       |    |
|-----------------------|----|
| 3 Motivasyon ve Katkı | 35 |
|-----------------------|----|

### III Yöntem

|  |    |
|--|----|
| 4 Deneysel ayarlamalar                         | 38 |
| 4.1 Programlama Dili Ve Kütüphaneler . . . . . | 38 |
| 4.2 Veri Kümeleri . . . . .                    | 38 |
| 4.3 Veri Ön İşleme . . . . .                   | 39 |
| 4.4 Özellik Çıkarım Teknikleri . . . . .       | 39 |
| 4.5 Sınıflandırma Algoritmaları . . . . .      | 40 |
| 4.6 Performans Metrikleri . . . . .            | 40 |
| 5 Elde Edilen Sonuçlar                         | 42 |

### IV SONUÇ

|                       |    |
|-----------------------|----|
| 6 Genel Değerlendirme | 51 |
| KAYNAKÇA              | 53 |

---

## ŞEKİLLER LİSTESİ

---

|   |                                   |    |
|---|-----------------------------------|----|
| 1 | Veri Ön İşleme Adımları . . . . . | 39 |
|---|-----------------------------------|----|

---

## ÇİZELGELER LİSTESİ

---

|   |  |    |
|---|--|----|
| 1 | Çalışmaların Karşılaştırılması . . . . .             | 27 |
| 2 | Amazon Veri Seti Makine Öğrenimi Sonuçları . . . . . | 42 |
| 3 | Amazon Veri Seti Derin Öğrenme Sonuçları . . . . .   | 45 |
| 4 | IMDb Veri Seti Derin Öğrenme Sonuçları . . . . .     | 46 |
| 5 | IMDb Veri Seti Makine Öğrenimi Sonuçları . . . . .   | 47 |

---

## KISALTMALAR

---

| <b>Türkçe</b>                      | <b>İngilizce</b>                          | <b>Kısaltma</b> |
|------------------------------------|---|-----------------|
| Doğal Dil İşleme                   | Natural Language Processing               | NLP             |
| Doğal Dil Araçları Kiti            | Natural Language Toolkit                  | NLTK            |
| Kelime Sıklığı-Ters Belge Frekansı | Term Frequency-Inverse Document Frequency | TF-IDF          |
| Kelimelerin Çantası                | Bag of Words                              | BOW             |
| Naive Bayes                        | Naive Bayes                               | NB              |
| Destek Vektör Makinesi             | Support Vector Machine                    | SVM             |
| Destek Vektör Sınıflandırma        | Support Vector Classification             | SVC             |
| Lojistik Regresyon                 | Logistic Regression                       | LR              |
| Karar Ağacı                        | Decision Tree                             | DT              |
| Rastgele Orman                     | Random Forest                             | RF              |
| Tekrarlayan Sinir Ağı              | Recurrent Neural Network                  | RNN             |
| Yapay Sinir Ağı                    | Artificial Neural Network                 | ANN             |
| Evrişimli Sinir Ağı                | Convolutional Neural Network              | CNN             |
| K-En Yakın Komşu                   | K-Nearest Neighbors                       | kNN             |

## BÖLÜM: İ

### GİRİŞ



---

## A M A Ç

---

Bu çalışmanın temel amacı, sosyal medya verilerinden duygu analizi yapmak için makine öğrenimi ve derin öğrenme tekniklerini kullanarak yüksek doğruluk oranları elde etmek ve farklı algoritmalar ile veri temsili yöntemlerini karşılaştırarak en etkili yaklaşımları belirlemektir. Bu amaç doğrultusunda, önceki araştırmaların genellikle sınırlı veri kümeleri ve belirli algoritma seçimleri üzerine odaklanmış olmasının aksine, bu çalışma daha kapsamlı bir yaklaşım benimseyerek çeşitli veri kümeleri ve geniş bir algoritma yelpazesi kullanarak duygu analizinin performansını çeşitli açılardan değerlendirmeyi hedeflemektedir.

Çalışmada öne çıkan amaçlar arasında farklı veri kaynaklarından elde edilen sonuçların karşılaştırılması, iki sınıflı ve üç sınıflı sınıflandırmalarla farklı sınıflandırma düzeylerinde duygu analizinin incelenmesi, beş farklı metin temsil yöntemi ile metinlerin duygu analizi için nasıl temsil edildiğinin araştırılması ve dokuz makine öğrenimi ile üç derin öğrenme algoritması kullanılarak farklı algoritmaların performansının değerlendirilmesi bulunmaktadır.

Bu hedefler doğrultusunda elde edilen bulguların, duygu analizi alanında mevcut bilgiyi genişleterek ve gelecekte yapılacak çalışmalara rehberlik ederek önemli bir katkı sağlaması beklenmektedir.

---

## LİTERATÜR ÖZETİ

---

Rawat ve arkadaşları yapmış oldukları çalışmada, aşıyla ilgili atılmış tweetleri kullanarak kamuoyunun aşıyla ilgili duyarlılığını değerlendirmek için duygu analizi yöntemini tercih etmişler [28]. Aşıyla ilgili tweetlerin veri seti kaggle platformundan alınmıştır. Özel karakterler, hashtag'ler, noktalama işaretleri, bağlantılar ve satır sonları gibi metnin ilgisiz kısımlarını kaldırmak için Spacy kütüphanesini kullanılmıştır. İlgisiz tüm kelime veya sembollerin kaldırılması için Nexttext kütüphanesi kullanılmıştır. TF-IDF yöntemini kullanarak işlenmiş tweetler vektörlere dönüştürülmüştür. sınıflandırma işlemi için Bernoulli Naïve Bayes (BNB), Destek Vektör Sınıflandırıcı (SVC) ve Lojistik Regresyon (LR) modelleri kullanılmıştır. Veri setinde 78.771 adet olumlu tweet, 23.993 adet olumsuz tweet bulunmuş ve 125.443 tweet'in çoğunun nötr olduğu tespit edilmiştir. SVC modeli %97 ile en yüksek doğruluğu sağlamıştır. BNB modelinin doğruluğu %81, LR modelinin doğruluğu ise %95 elde edilmiştir.

Demir ve Bilgin tarafından yapılan çalışmada, Türkçe haber metinlerinden elde edilen veriler üzerinde BERT tabanlı dil modelleri ve makine öğrenmesi algoritmaları kullanılarak duygu analizi gerçekleştirilmiştir [10]. BERT tabanlı dil modelleri olarak ALBERT, DistilBERT ve RoBERTa, makine öğrenme algoritmaları olarak ise Naive Bayes, Destek Vektör Makinesi ve Rastgele Orman yöntemleri kullanılmıştır. Veri seti 5000 adet olumlu ve olumsuz cümle içermektedir ve verilerin %90'ı eğitim, %10'u test için kullanılmıştır. Çalışmaların

sonuçları incelendiğinde, dil modelleri ile yapılan çalışmaların doğruluk değerlerinin, makine öğrenmesi algoritmalarına göre daha yüksek değerlere ulaştığı görülmüştür. Dil modellerinin başarı oranları DistilBERT, RoBERTa ve ALBERT olup, elde edilen sonuçlar sırasıyla %80, %80 ve %77 olarak kaydedilmiştir. Makine öğrenmesi algoritmalarının sıralaması ise Naive Bayes, Destek Vektör Makinesi ve Rastgele Orman olup, elde edilen sonuçlar sırasıyla %71, %68 ve %68 olarak tespit edilmiştir.

Öge ve Kayaalp tarafından yapılan çalışmada, IMDB internet sayfasında yer alan film yorumlarından oluşan bir veri seti kullanılmıştır [24]. Veri setinde 25.000 adet pozitif ve 25.000 adet negatif olarak etiketlenmiş yorum bulunmaktadır. Python programlama dilindeki kütüphaneler kullanılarak Naive Bayes, Logistic Regression, Support Vector Machine, Decision Tree, Random Forest ve Artificial Neural Network sınıflandırma algoritmalarının duygu analizi kapsamında performans karşılaştırması yapılmıştır. Bag of Words, TF-IDF, Word2Vec ve FastText metin temsil yöntemleri kullanılarak temizlenmiş veri seti sayısal olarak temsil edilmiştir. Veri setinin eğitimi ve test edilmesi aşamasında  $k=5$  olacak şekilde  $k$ -fold cross validation tekniği kullanılmıştır. 6 farklı sınıflandırma yöntemi için elde edilen sonuçlar doğruluk, kesinlik, geri çağırma ve F1 score hesaplanarak ayrıntılı bir karşılaştırma yapılmış ve sonuçlar kaydedilmiştir. Sonuçlar karşılaştırıldığında en iyi performans, 0.87 doğruluk, 0.87 kesinlik, 0.88 geri çağırma değerleriyle 9.7 saniyede Word2Vec ile Logistic Regression yöntemi kullanılarak elde edilmiştir.

Çalışma [30]'da, Rumeli ve arkadaşları tarafından Türkçe metinler için bir duygu analizi modeli geliştirilmiştir. Bu model geliştirilirken, sözlük tabanlı yaklaşımlar ile makine öğrenmesi algoritmaları birlikte kullanılmıştır. Duygu analizinin en temel yöntemi olarak, cümledeki her kelimenin sözlükteki puan değerine bakarak toplama dayalı bir model ile hesaplama yapılmıştır. Makine öğrenmesi modelleri, metinlerin polarite puanlarını baz

olarak doğru sonuçlar bulan duygu analizi yapmak için eğitilmiştir. Projede NB, RF, SVM ve k-NN algoritmaları olmak üzere dört farklı makine öğrenmesi yöntemi kullanılmıştır. NB ve SVM için R dilindeki 'e1071' kütüphanesi kullanılmıştır. Bu paket, SVM algoritması için maliyet çekirdeği ve gama değerlerini değiştirmeye olanak tanımıştır. Projede radyal, lineer ve polinom çekirdekler değiştirilerek model hayata geçirilmiş ve RF için "randomForest" isimli paket kullanılarak 'tree' parametresi 100, 425 ve 750 olarak değiştirilerek model eğitilmiştir. kNN algoritması için 'class' kütüphanesi kullanılmış ve k değeri sadece tek sayılar kontrol edilerek 3'ten 71'e çıkarılmıştır. Son olarak doğrulama setindeki performansı yüksek olduğundan k değeri 33 olarak belirlenmiştir. Çapraz doğrulamanın, makine öğrenimi yöntemlerinin değerlendirilmesi üzerinde önemli bir etkisi olduğu belirtilmiştir. Bu nedenle, 100 defa çapraz doğrulama yapılmıştır. Veri setinin %80'i eğitim, %20'si test verisi olarak bölünmüştür. Tüm modellerin ortalama doğruluğunun 0,73 civarında değiştiği görülmüştür. Bu da modellerin doğrulukları arasında çok büyük bir fark olmadığını göstermektedir. NB ve kNN'nin 1 dakikadan az sürede işlemi tamamlarken, SVM ve RF'nin daha uzun süreye ihtiyaç duyduğu gözlemlenmiştir.

Poornima ve Priya tarafından yapılan çalışmada, Twitter verileri kullanılarak cümle sınıflandırmasında SVM, Multinomial Naive Bayes ve Logistic Regression algoritmalarının performans karşılaştırması yapılmıştır [26]. Twitter'dan alınan ham tweetler genellikle gürültülü bir veri kümesine neden olmaktadır. Çeşitli yöntemler kullanılarak veri seti temizlenmiştir. Bigramlar kullanılarak veri kümesinden özellikler çıkarılmıştır. Bu özellikler, sınıflandırmada olumsuzluğu modellemek için etkili bir yol sağlamıştır. Lojistik regresyonun doğruluğunun %86,23, destek vektör makinesinin doğruluğunun %85,69 ve Multinomial Naive Bayes'in doğruluğunun %83,54 olduğu gözlemlenmiştir.

Sharma ve arkadaşlarının yapmış oldukları çalışmada, çok terimli Naive Bayes, Destek Vektör Makineleri ve Lojistik Regresyon'dan yararlanarak duygu analizi çalışmalarında kullanılan tüm makine öğrenimi yaklaşımları için araştırma yapılmış, karşılaştırma analizi yapılmış ve performans değerlendirme verileri sunulmuştur [35]. CNN ve LSTM gibi derin öğrenme modellerinin, fikir madenciliği ve duygu sınıflandırma görevlerini gerçekleştirmek için kullanıldığında diğer modellerden daha iyi performans gösterdiği görülmüştür. Makine öğrenimi modeli olan Rastgele Orman (RF) ve derin öğrenme modeli olan RNN'in performansının benzer olabileceği, ancak sınıflandırıcıların veri kümelerine göre değişen eğilimlerinin olduğu tespit edilmiştir. Sonuçlara göre en iyi doğruluk oranı %93,5 ile CNN kullanılarak elde edilmiştir. Aynı veri kümesinde yapılan çalışmada, LSTM'nin %77,75 doğruluğa sahip olduğu gözlemlenmiştir.

Joshi ve arkadaşları tarafından yapılan çalışmada, duygusal analiz ve kategorizasyon için Twitter verileri kullanılmıştır [18]. Duygu analizinin uygun veri görselleştirmesi ve ön işleminin gerekli olduğu belirtilmiştir. Makine öğrenimi algoritması kullanılarak olumlu veya olumsuz duyguların analizi yapılmıştır. Naive Bayes sınıflandırma algoritması kullanılmıştır. Verilerin kesinlik değerlerinin random forest yönteminde %75, karar ağacı yönteminde %75,36, eğitimde %84,93 ve testte %78,58 olduğu sonucuna ulaşılmıştır.

He ve arkadaşları tarafından yapılan çalışmada, çevrimiçi ürün deneyimi elde etmeyi amaçlayan ve metinsel analiz tekniklerini makine öğrenimi algoritmalarıyla birleştiren bir füzyon duyarlılığı analizi yöntemi geliştirilmiştir [15]. Yöntem temel olarak üç adımdan oluşmaktadır. İlk olarak, duygu sözlüğünün duygusal bilgiye duyarlılığından esinlenerek, sözlük duygu özellikleri çıkarılmıştır. Daha sonra, incelemelerin duyarlılık kutuplarını belirlemek için SVM algoritması kullanılmıştır. Buna dayanarak, duyarlılık konuları LDA modeli aracılığıyla incelemelerden çıkarılmıştır. Ayrıca, duygusal bilgilerin atlanmasını önlemek için

sözlük anlamsal benzerliğe dayalı olarak genişletilmiştir. Tüketicilerin Amazon'daki çevrimiçi kitap okuma deneyimlerinin araştırılması, yöntemin uygulanabilirliğini ve geçerliliğini doğrulamıştır. Sonuçlar, yöntemin incelemelerin duygusal eğilimlerini doğru bir şekilde belirlediğini ve incelemelerden okuma deneyimlerini etkileyen unsurları yakaladığını göstermektedir. Genel olarak, araştırmanın çevrimiçi ürün deneyimini araştırmak ve müşteri taleplerini takip etmek için etkili bir yol sağladığı, gelecekteki ürün iyileştirme ve pazarlama stratejisi optimizasyonunu güçlü bir şekilde desteklediği görülmüştür.

Duygu analizi, yaklaşık 20 yıl önce ortaya çıkan bir alan olarak önemli gelişmeler kaydetmiş ve ticari uygulamalarda geniş bir kullanım alanına sahip olmuştur. Ancak, Poria ve arkadaşlarının yapmış oldukları çalışmada da vurgulandığı gibi, bu alanın olgunlaştığı algısına rağmen hâlâ keşfedilmemiş ve yeterince araştırılmamış yönler bulunmaktadır [27]. Duygu analizi genellikle basit bir olumlu-olumsuz sınıflandırma görevi olarak algılsa da çalışma, bu görevin aslında karmaşık ve çok yönlü olduğunu, insan güdülerini, niyetleri ve bağlamsal nüansları içerdiğini vurgulamıştır. Proje örnekleri, duyarlılık analizi, sarkazm analizi ve çok modlu duyarlılık analizi gibi alt alanlardaki çalışmaların önemine işaret ederken, bu çalışmaların duygu analizi alanında hâlâ çözülmesi gereken sorunlara ışık tuttuğunu belirtmiştir. Bu nedenle çalışma, duygu analizi alanındaki keşfedilmemiş potansiyeli vurgulayarak bu alanda daha fazla araştırma yapma çağrısında bulunmaktadır.

Dhariwal ve arkadaşları tarafından yapılan çalışmada, sesli duygu analizi araçlarından transkripsiyon hizmetlerine ve doğal dil işleme tekniklerine kadar geniş bir yelpazede kullanılan hesaplamalı bir yaklaşım olan "çatallama ve karıştırma" yöntemini önermeyi amaçlamışlardır [11]. Metin duyarlılığı analizi sözlüklerini, AssemblyAI ve VADER gibi, ses verilerinden duyguları çıkarmak ve kategorilere ayırmak için kullanarak gerçek zamanlı duyarlılık analizi modeli oluşturmak hedeflenmiştir. Bu analizin sonuçları, tüm Hindistan

Radyosu'nun 'Akashwani' tarafından yapılan radyo yayınlarında ifade edilen duygulardaki kalıpları ve eğilimleri ortaya koyarak haber servislerinin duygu analizi konusundaki yöntemlerini geliştirmeyi amaçlamıştır. Metin duyarlılığı analizi ve sesli duyarlılık analizi arasındaki avantaj ve dezavantajları değerlendiren araştırma, bu iki yöntemin birleşimiyle daha kapsamlı ve etkili bir sonuç elde edilebileceğini ortaya koymuştur. Ayrıca, bu çalışmanın gelecekte afet yönetimi gibi gerçek dünya sorunlarına çözüm bulmak ve radyo sinyallerinin kullanıldığı iletişimdeki rolünü güçlendirmek adına önemli bir araştırma kaynağı olabileceği vurgulanmıştır. Ancak, analiz modellerini çalıştıran bağımsız bir sistem geliştirmenin zorlukları, modelin eğitildiği veri setinin kalitesi ve çeşitliliği, afet yönetimi durumlarında kullanılacak olan NVV veri setlerinin ihtiyaçları gibi konular da ele alınması gerektiği sonucuna varılmıştır.

Li ve arkadaşları tarafından yapılan çalışmada, canlı yorumlama platformu Danmaku üzerine odaklanılmıştır [22]. Danmaku, kullanıcıların videoları izlerken gerçek zamanlı olarak yorum yapabildikleri bir sistemdir. Ancak, mevcut duygu sınıflandırma yöntemleri Danmaku verileri için uygun değildir. Bu sorunu çözmek için makalede bir duygu sözlüğü oluşturulmuş ve Naive Bayes tabanlı yeni bir yöntem önerilmiştir. Tencent Video'dan seçilen Danmaku verileri üzerinde yapılan testler, önerilen SDD-NB modelinin diğerlerine göre %14.8 daha yüksek doğruluk oranına sahip olduğunu göstermiştir.

Fu ve arkadaşları tarafından yapılan çalışmada, duygu analizi için kullanılan LSTM'lerin başarılı bir yöntem olarak vurgulandığı ancak yalnızca kelime gömülerinin yetersiz olduğu belirtilmiştir [12]. Bu eksikliği gidermek amacıyla LE-LSTM adlı bir sözlük ile geliştirilmiş LSTM modeli önerilmiştir. GloVe kullanılarak yapılan deneylerde, LE-LSTM'nin standart LSTM modeline kıyasla %2 daha yüksek doğruluk elde ettiği gözlemlenmiştir. Ayrıca, duygu kelimesinin dil bilimsel rolünü modellemek için kayıp fonksiyonuna düzenleyici eklenmesi ve farklı boyutlardaki CNN mimarisi de metin kategorizasyonu için önerilmiştir. Gerçekleştirilen

deneyler, LE-LSTM, Bi-LSTM ve Tree-LSTM modellerinin tüm veri kümelerinde standart LSTM'den daha iyi performans sergilediğini göstermiştir. Bu modellerin test edildiği beş veri kümesi (IMDB, Yelp2013, NB4000, MR ve Book4000) üzerinde elde edilen sonuçlar, özellikle LE-LSTM'nin diğer modellere göre daha etkili olduğunu ortaya koymuştur. Bu önerilen model, kelime gömülerinin anlamsal bilgi taşıma sorununu çözmek ve duygu analizi performansını artırmak için etkili bir yöntem sunmaktadır.

Wang ve arkadaşları tarafından yapılan çalışmada, geleneksel kelime gösterimine duygu bilgisi ekleyememe sorununu çözmek amacıyla Word2Vec ve GloVe temelli RGWE yöntemi önerilmektedir [39]. RGWE, duyarlılık, konum ve POS özelliklerini ekleyerek duygu analizinin doğruluğunu artırmayı hedeflemektedir. Yapılan deneylerde, RGWE'nin SSWE ve Seninfo+TF-IDF'ye göre daha iyi bir performans sergilediği görülmüştür. İkili sınıflandırma, üçlü sınıflandırma ve CNN için sırasıyla %89.86, %69.1 ve %48.85 ortalama F1 değerleri elde edilmiştir. Bu başarı, SemEval, SST1, SST2, IMDB, Amazon ve Yelp gibi altı klasik veri kümesi üzerinde gözlemlenmiştir. RGWE, duygusal içerikli metin analizi için güçlü bir kelime gömme teknolojisi olarak öne çıkmaktadır.

Zhang ve Zheng'in çalışmasında, Çince metinler üzerinde duygu analizi için TF-IDF ile ağırlıklı özellik hesaplaması kullanılmıştır [41]. SVM ve ELM (Extreme Learning Machine) çekirdekli modeli, eğitim örnek kümesini öğrenerek test kümesinde duygu sınıflandırması yapmıştır. Çalışmada, ELM'nin yüksek verimliliği ve uygulanabilirliği ile duygu analizi alanında önemli bir çözüm sunduğu vurgulanmıştır. Metin dosyaları kullanılarak oluşturulan veri setinde, gürültü temizleme, segmentasyon ve durdurma kelimesi filtreleme gibi adımlar gerçekleştirilmiştir. TF-IDF kullanılarak belgeler arasındaki fark belirlenmiş ve MATLAB ortamında simülasyonlar yapılmıştır. Deney sonuçları, çekirdekli ELM'nin Çince metinlerde duygu analizi uygulamalarında SVM'ye kıyasla daha etkili olduğunu göstermiştir. Hem



doğruluk oranları hem de eğitim-test süreleri açısından çekirdekli ELM, avantajlı bir performans sergilemiştir. Bu çalışma, Çince metinlerde duygu analizinde yeni yaklaşımların ve çözümlerin geliştirilmesine katkı sağlamıştır.

Xu ve Song'un yapmış oldukları çalışmada, metin duygu sınıflandırması alanında geleneksel ve derin öğrenme modellerini aşan yeni bir yaklaşım sunulmuştur [40]. Çalışmada, BERT ile birleştirilmiş iki hibrit sinir ağı modeli tanıtılmış ve bu modeller Chnsenticrop ve Weibo veri kümeleri üzerinde eğitilmiştir. Veriler, temizlenmiş ve SST-2 formatına dönüştürülmüş olup, 4:1 oranında eğitim-test setlerine ayrılmıştır. Kelime tabanlı metinler, 512 boyutlu Word2Vec vektörleri kullanılarak temsil edilmiştir. Çeşitli yöntemler, örneğin CNN, LSTM, GRU, Bert+BiLSTM ve Bert+CNN karşılaştırılmış ve en etkili yaklaşım belirlenmiştir. Bert+CNN modeli, diğerlerine göre daha yüksek doğruluk, geri çağırma ve F1 skorlarına sahip olduğunu göstererek ön plana çıkmıştır. LSTM en yüksek geri çağırma oranını, GRU ise en yüksek F1 puanını elde etmiştir. Analiz sonuçları, metin duygu sınıflandırmasında Bert+CNN modelinin etkili bir performans sergilediğini ortaya koymaktadır.

Rizk ve Asal yapmış oldukları çalışmada, çeşitli makine öğrenimi, sinir ağları ve kelime gömme modellerini değerlendiren kapsamlı bir çalışmanın özetini sunmuşlardır [29]. Araştırma, BERT katıştırmaları ile LSM-çift yönlü kodlama modelinin en yüksek test doğruluğuna ulaştığını ve kelime gövdeleri ile derin öğrenme mimarilerinin geleneksel modellere göre üstün performans sergilediğini göstermiştir. Ayrıca, Word2Vec'in farklı modellerle birleştirilmesini inceleyen çalışma, bu hibrit yaklaşımın %89.2'lik bir doğruluk oranıyla diğer modellere göre daha etkili olduğunu ortaya koymuştur. DistilBERT adlı önceden eğitilmiş bir dil modelinin tanıtıldığı ve orijinal BERT'e kıyasla %40 daha küçük, %97 oranında bilgi koruyabilen ve %60 daha hızlı olduğu kanıtlanmıştır. Araştırma, IMDB film incelemeleri üzerinde yapılan denemelerde Transformatörden Çift Yönlü Kodlayıcı Temsilleri (BERT) gö-

mülü Uzun Kısa Süreli Bellek (LSTM) modelinin %93'lük doğrulukla en iyi performansı sergilediği sonucuna ulaşmıştır. Bu bağlamda, DistilBERT'in daha küçük boyutta ve hızlı çalıştığı vurgulanmıştır.

Hammad ve Anwar yapmış oldukları çalışmada, Sindhi dilinde duygu analizi gerçekleştirmek için denetimli makine öğrenimi tekniklerini kullanmışlardır [13]. Sindhi tweetlerini içeren bir veri kümesi, önceden tanımlanmış sözlüksel veri tabanı ve Google Translate API yardımıyla etiketlenmiştir. Veri kümesi, temizlenmiş ve tokenizasyon işleminden geçirilmiştir. Duygu analizi için SVM, NB, DT ve k-NN algoritmaları kullanılmıştır. Elde edilen sonuçlar, Sindhi tweet veri setinde DT ve k-NN algoritmalarının iyi bir performans sergilediğini, SVM'nin bunları takip ettiğini göstermektedir. Özellikle SVM, doğruluk, kesinlik, geri çağırma ve F1 puanı ölçütlerinde yüksek bir performans sergilemiştir. Veri kümesi boyutları arttıkça doğruluk değerlerinin arttığı, ancak kesinlik ve geri çağırma değerlerinin değişkenlik gösterdiği gözlemlenmiştir. Bu bağlamda, Sindhi dilinde duygu analizi için denetimli makine öğrenimi yöntemlerini değerlendiren bu araştırma, özellikle SVM'nin yüksek performans gösterdiğini vurgulamıştır.

Pandya ve arkadaşları yapmış oldukları çalışmada, makine öğrenimi ve derin öğrenme modelleri kullanarak Twitter duygu analizi yapmayı hedeflemişlerdir [25]. 1.6 milyon tweet, bu çalışmanın veri kümesi olarak kullanılmış ve tweetler olumlu veya olumsuz duygulara göre sınıflandırılmıştır. Bu amaçla Naive Bayes ve Destek Vektör Makineleri gibi geleneksel makine öğrenimi modelleri kullanılmıştır. Makine öğrenimi ve derin öğrenme modellerini eğitmek için Sentiment140 veri kümesi kullanılmış olup, bu küme 800.000 negatif ve 800.000 pozitif tweet içermektedir. Veriler, ön işleme tabi tutulmuş; tweetlerden kullanıcı adı, bağlantı, noktalama işaretleri, rakamlar ve stopwords gibi ifadeler kaldırılmış ve tüm harfler küçük harfe dönüştürülmüştür. Daha sonra bu veriler köklerine ayrıştırılmıştır. Elde edilen veri kümesi

görselleştirilerek negatif ve pozitif kelime bulutu oluşturulmuştur. Veri kümesi kelimelerden oluştuğu için, vektöre dönüştürmek amacıyla Word2Vec modeli kullanılmıştır. Veri kümesi %80 eğitim ve %20 test olarak iki ayrı kümeye dönüştürülmüştür; toplam tweet sayısı üzerinden 1.28 milyon tweet eğitim için, 320.000 tweet test için kullanılmıştır. Model eğitimlerinde Naive Bayes (NB), Destek Vektör Makineleri (SVM), RNN, LSTM ve Bi-LSTM kullanılmıştır. Sonuçlar doğruluk, F1 skoru, geri çağırma skoru ve hassasiyete göre karşılaştırılmıştır. Destek Vektör Makineleri %76.53 doğrulukla daha iyi performans sergilemiştir. Ancak, derin öğrenme modelleri makine öğrenimi modellerine göre daha iyi performans göstermiştir.

Singh ve arkadaşları yapmış oldukları çalışmada, TF-IDF ve makine öğrenimi algoritmalarını kullanmışlardır [36]. Veri kümesi olarak ABD Havayolları duyarlılık tweetleri kullanılmıştır. Tweetler olumlu, olumsuz ve nötr olmak üzere üç kategoride sınıflandırılmıştır. Veriler önce ön işleme tabi tutulmuş; gereksiz kelimeler ve noktalama işaretleri gibi karakterler silinmiş ve büyük harfler küçük harflere dönüştürülmüştür. Geriye kalan metin daha küçük parçalara bölünerek tokenizasyon işlemi uygulanmıştır. Daha sonra sözcüklerin köküne indirgemek için stemming işlemi yapılmıştır. Kalan verilere TF-IDF yöntemi uygulanarak kelimelerin uygulanma sıklığı elde edilmiştir. Veri kümesi, %80 eğitim ve %20 test olarak ayrılmıştır. Bu çalışmada RF, GB, XGBoost ve SVM olmak üzere dört sınıflandırma tekniği doğruluk, hassasiyet, hatırlama ve F1 skoru ölçütlerine göre karşılaştırılmıştır. SVM'nin %83.74 oranı ile en yüksek doğruluğu elde ettiği görülmüştür. Aynı zamanda, %84 hassasiyet ve %87.85 F1 skoru ile diğer tekniklerden daha iyi sonuç verdiği tespit edilmiştir.

Kim yapmış olduğu çalışmada, müşteri yorumlarının inovasyon yöneticileri için önemli bir kaynak olduğunu vurgulayarak, duygu analizinin bu yorumlardan elde edilen istatistiksel verilerle mümkün olduğu bir perspektife odaklanmıştır [21]. Metin madenciliği yöntemlerini kullanarak müşteri duyarlılığını değerlendirmek amacıyla çeşitli ölçütler önerilmiş ve

Linguistic Inquiry and Word Count (LIWC) gibi metin madenciliği paketleri üzerindeki alternatifler incelenmiştir. Duyarlılık analizi için pozitif ve negatif eğilimli kelimelerin sayılarına dayanan bir yaklaşım benimsenmiş ve bu yöntemlerin sınırlamaları tartışılmıştır. Çalışma, duyarlılık puanları ile çevrimiçi inceleme derecelendirmeleri arasındaki dağılım farkını göstererek özellikle çevrimiçi derecelendirmelerin öznel olduğunu vurgulamıştır. Ayrıca, pazarlamacıların müşteri görüşlerini değerlendirmek ve karşılaştırmak için duyarlılık puanlarını kullanabileceği alternatif yollar sunmuştur. Sonuç olarak, metin madenciliği ve duygu analizinin pazarlamacılara müşteri geri bildirimlerini etkili bir şekilde değerlendirme ve inovasyon performanslarını ölçme konusunda yardımcı olabileceğini öne sürmüştür.

Chandra ve Jana yapmış oldukları çalışmada, duygu analizi konusundaki araştırmalara ve kullanılan yöntemlere odaklanarak insanların görüşlerini, duygularını ve tutumlarını belirli konularla ilişkilendiren bir alana dikkat çekmişlerdir [8]. Makine öğrenimi ve derin öğrenme tekniklerinin dil işleme istatistiksel yöntemleriyle birleştirilmesi, duygu analizinin gelişimine önemli katkılar sağlamıştır. Sosyal medyanın etkisi ve kullanıcıların duygularını dijital medya aracılığıyla ifade etmeleri, özellikle Twitter gibi platformlarda zengin bir veri kaynağı oluşturmuştur. Duygu analizi için kullanılan çeşitli yöntemler ve teknikler, duyarlılık sözlüklerinin algoritmik oluşturulması, duyarlılık endeksi oluşturma ve önem değerlendirmesi gibi aşamaları kapsamaktadır. Çalışma, bu yöntemlerin yanı sıra derin öğrenme modelleri ve makine öğrenimi sınıflandırıcıları kullanarak gerçekleştirilen duygu analizi için önerilen yaklaşımları detaylı bir şekilde açıklamıştır. Derin öğrenme modelleriyle yapılan deneylerde, çeşitli modellerin performansları karşılaştırılmış ve LSTM-CNN'nin en iyi sınıflandırma sonuçlarını verdiği sonucuna varılmıştır. Deneyler ve sonuçların analizi bölümünde, Twitter API ve çeşitli konulardaki veri kümeleri kullanılarak gerçekleştirilen deneylerin sonuçlarına odaklanılmıştır. Makine öğrenimi modellerinin doğruluk yüzdesinin %81 ila %90 arasında

kaldığı belirtilirken, derin öğrenme modellerinin %85 ila %97 aralığında daha iyi sonuçlar elde ettiği vurgulanmıştır. Sonuç ve gelecek kapsamı bölümünde, derin öğrenme modellerinin daha iyi performans gösterdiği ve gelecekte bu alanda yapılacak çalışmaların bu modellerin tasarımı ve uygulanması üzerine odaklanabileceği sonucuna varılmıştır. Sosyal medya verilerinin analizinin, işletmelere bilgilendirici içgörüler ve gelir sağlama potansiyeli sunduğu belirtilmiştir.

Saraswat ve arkadaşları yapmış oldukları çalışmada, ses dosyalarının makine öğrenimi kullanılarak duygu analizi yapılması ve ses verilerinin metinsel sınıflandırma ile analiz edilmesini hedeflemişlerdir [32]. Duygu analizi, genellikle metin madenciliği teknikleri ile işlenen verilere odaklanmıştır. Ancak, ses verilerinin duygu analizi alanındaki araştırmalar hala erken aşamadadır. Ses verilerinin analizinde kullanılan temel özellikler arasında ses frekansı, konuşma öznitelikleri ve zamansal öznitelikler bulunmaktadır. Projenin temelini oluşturan sistem mimarisi, ses dosyalarından metin transkriptleri oluşturmayı içermektedir. Bu transkriptler, daha sonra çeşitli metin madenciliği araçları kullanılarak işlenir. Metin verileri analize hazırlanırken önce ses dosyalarının metne dönüştürülmesi, ardından ön işleme adımları ile belirli bir özellik kümesinin çıkarılması gerçekleştirilir. Son olarak, denetimli öğrenme teknikleri kullanılarak metin transkriptleri kategorize edilir. Naïve Bayes, Lojistik Regresyon ve Sayım Vektörleştirici gibi çeşitli sınıflandırma algoritmaları kullanılarak duygu analizi gerçekleştirilir. Naïve Bayes sınıflandırıcısı kullanılarak elde edilen sonuçlar, özellikle pozitif ve nötr duyguların tanınmasında yüksek doğruluk göstermektedir. Duygu analizörü, çeşitli duyguları tanıma doğruluğuyla şu sonuçları vermektedir: Neşe %94, Üzüntü %93, Öfke %85, Korku %74, İğrenme %70, Utanç %56, Sürpriz %52. Duygu analizörünün başarı oranının, güçlü duygusal değerlere sahip kelimelerin konuşmada belirgin bir şekilde kullanılmasına bağlı olduğu gözlemlenmiştir. Bu kelimeler, tahmin skorunu ve doğruluğunu etkilemede

önemli bir rol oynamaktadır. Model, 34,000'den fazla veri noktası üzerinde eğitildiği için %86'lık bir doğruluk sağlamaktadır, bu da diğer ses duygu analizörlerine kıyasla (metinsel sınıflandırma tabanlı ve harici donanım cihazına ihtiyaç duymadan) oldukça yüksektir. Daha büyük veri setleri ve gelişmiş veri temizleme teknikleri kullanılarak gelecekteki çalışmaların, duygu analizi alanında daha kesin ve güvenilir sonuçlara ulaşabileceği belirtilmiştir.

Can ve Alataş yapmış oldukları çalışmada, sosyal medyanın ve genel olarak internet sosyal ağlarının günümüzde giderek önem kazanması ile beraber popüler bir konu haline gelmiş duygu analizini genel olarak anlatmış ve bu konuda yapılan çalışmalarda ortaya konulan duygu analizi algoritmalarını ve uygulamalarını göstermişlerdir [7]. Güncel bir sonuç çıkarmak amacıyla son yıllarda yapılan çalışmalar ele alınmış ve bu çalışmalar kategorize edilerek sunulmuştur. Bu çalışmalar, gerçek dünya problemlerine duygu analizi tekniklerini uygulayarak duygu analizi ile ilgili birçok alana katkı sunmuştur. Tüm bu çalışmaların incelenmesinden sonra duygu sınıflandırması ve özellik seçimi algoritmalarının geliştirilmesinin halen önemli bir çalışma alanı olduğu açıkça görülmektedir. Ayrıca duygu sınıflandırma problemlerinin çözümünde Naive Bayes ve Destek Vektör Makineleri'nin en çok kullanılan makine öğrenme algoritmaları olduğu sonucuna varılmıştır. Yapılan analizler sonucunda İngilizce dilinde birçok çalışma ve bunların doğal sonucu olarak birçok kaynak olduğu ortaya çıkmıştır. İngilizce dışındaki dillerde ise çalışma ve kaynak eksikliğinin olduğu görülmüştür. İngilizcenin dışında en çok kullanılan veri kaynağı WordNet olmuştur. Duygu analizi çalışmaları için son zamanlarda yeni bilgi kaynakları ortaya çıkmıştır. Bu kaynaklar mikro-blog, bloglar ve forumlar olmuştur. Tüm bu kaynaklar, insanların belli güncel konular hakkında ne düşündükleri ve bu konularla ilgili duygularının anlaşılması için önemli bir rol oynamaktadır. Daha derin ve detaylı analizler için sosyal ağ siteleri ve mikro-bloglar önemli kaynaklar olmaktadır. Ayrıca, algoritmaların test edilmesi için birçok veri seti de mevcuttur. Yapılan

bu çalışmada duygu analizi konusunda halen tartışmaya açık alanlar olduğu görülmüştür. Genel çalışma alanlarına ek olarak, Türkçe konusunda yapılan çalışmalar halen yetersiz kalmaktadır.

Satrya ve arkadaşları yapmış oldukları çalışmada, Twitter üzerindeki kripto para birimleri ile ilgili olumlu ve olumsuz eğilimleri inceleyerek, toplumun bu dijital varlıklara yönelik tutumlarını anlamayı amaçlamıştır [34]. Duygu analizi veya fikir madenciliği yöntemleri kullanılarak Twitter'daki veriler sınıflandırılmış ve olumlu ve olumsuz eğilimler ortaya çıkarılmaya çalışılmıştır. Bu analiz, özellikle Bitcoin'e odaklanarak kripto para birimlerinin genel algısını belirleyebilir ve yatırımcılar ile işletmeler için önemli bir referans olabilir. Araştırmacılar, daha önceki çalışmaların incelenmesine dayanarak bu araştırmada Destek Vektör Makinesi (SVM) algoritmasını kullanmışlardır. Literatürde SVM'nin duygu analizi konusunda yüksek doğruluk sağladığı belirtilmiştir. Elde edilen Twitter verileri SVM kullanarak sınıflandırılarak olumlu ve olumsuz eğilimlere dair sonuçlar elde edilmeye çalışılmıştır. Veri toplama işlemi, Twint kütüphanesi kullanılarak gerçekleştirilmiş ve Twitter API'sine gerek duyulmamıştır. Elde edilen veriler öncelikle temizlenmiş ve duygu analizi için uygun hale getirilmiştir. Veri kümesinin dengesizliği, alt örnekleme yöntemi kullanılarak ele alınmış ve analiz için hazırlanan veriler SVM algoritmasıyla işlenmiştir. Sonuçlar, kullanılan yöntemin yüksek doğruluk değerleri elde ettiğini göstermiştir. 80:20 oranında eğitim-test bölme durumunda en yüksek doğruluk %94,64 olarak bulunmuştur. Hassasiyet, özgüllük ve F1 skoru gibi değerler de incelendiğinde, SVM algoritmasının kripto para duygu analizi konusunda etkili bir araç olduğu görülmüştür. Bu sayısal değerler, analizin güvenilirliğini ve gerçek dünya uygulamalarındaki potansiyel etkisini daha fazla vurgulamaktadır.

Sarkar ve Bhowmick tarafından yapılan çalışmada Bengalce tweetler üzerinde duygu analizi gerçekleştirilmiştir [33]. Çalışmada, Multinomial Naive Bayes ve SVM gibi sınıflandı-

ıcılar kullanılarak duygu kutupluluğu tespiti yapılmıştır. N-gram Tokenizer ve SentiWordnet gibi yöntemler kullanılarak veriler işlenmiştir. Veriler, olumluluk, olumsuzluk ve tarafsızlık olmak üzere üç duyarlılık puanına göre sınıflandırılmıştır. Öncelikle veriler ön işleme tabi tutulmuş, ardından etiketli vektörlerle bir sınıflandırıcı eğitilmiş ve bu model tweetler üzerinde test edilmiştir. Kullanılan Bengalce veri kümesinde eğitim veri kümesi olarak 356 negatif, 368 pozitif ve 276 nötr tweet, test veri kümesi olarak ise 500 etiketlenmemiş tweet olmak üzere toplamda 1500 tweet bulunmaktadır. Deney sonuçlarına göre, Naive Bayes multinomial unigram ve bigram kullanılarak SentiWordnet ile %44.20 doğrulukla en iyi sonuçları vermiştir. SVM ise unigram ve SentiWordnet özelliklerini kullanarak %45 doğrulukla en iyi performansı göstermiştir. Ayrıca, önerilen SVM tabanlı sistem, SAIL 2015 yarışmasında kullanılan en iyi tweet duyarlılık tespit sistemleriyle karşılaştırıldığında daha iyi performans sergilemiştir.

Jayakody ve Kumara yapmış oldukları çalışmada, Twitter'daki ürün incelemelerini kullanarak duygu analizi üzerinde çalışmışlardır [17]. Veri olarak 3548 açıklamalı tweet kullanılmıştır. İki farklı vektörleştirme yöntemi kullanılarak üç denetimli makine öğrenimi modeli karşılaştırılmıştır. Veriler; SVM, Logistic Regression, K-Nearest Neighbor algoritmaları kullanılarak sınıflandırılmıştır. Vektörleştirici olarak TF-IDF ve Sayım Vektörleştirici kullanılmıştır. Veriler olumlu-olumsuz duygulara göre ayrıştırılmıştır. Tweetler pozitif olanlar 1, negatif olanlar ise 0 şeklinde etiketlenmiştir. Daha sonra boş tweetler kaldırılarak kalan tweetler üstünden istenmeyen içerikler silinmiştir. 2 kelimedenden daha az tweetlerde silinerek kalan metinler küçük harfe dönüştürülüp veri hazırlama aşaması tamamlanmıştır. Eldeki veri kümesi eğitim ve test olmak üzere iki veri kümesine ayrılmıştır. Eğitim veri kümesinin oranı %70, test veri kümesinin oranı ise %30 olarak bildirilmiştir. Makine Öğrenimi algoritmaları kullanılarak sonuçlar elde edilmiştir. Elde edilen sonuçlara bakıldığında Sayım Vektörleştiri-



rici kullanılarak yapılan çalışmada Lojistik Regresyonun diğer algoritmalarla göre daha iyi sonuç verdiği gözlemlenmiştir. TF-IDF kullanılarak yapılan çalışmada ise SVM algoritmasının %87.41 ile diğer algoritmalarla kıyasla daha iyi sonuç verdiği gözlemlenmiştir. Diğer bir çalışma olarak veri kümesinin bölünmesinin sonuca etkisini belirlemek için eğitim ve test olmak üzere her ikisi de %50 oranla bölünmüş ve çalışma uygulanmıştır. Sayım vektörleştirici kullanıldığında Lojistik Regresyon, TF-IDF kullanıldığında SVM algoritması daha iyi sonuç vermiştir. KNN algoritması her iki çalışma içinde benzer sonuçlar vermiştir. En doğru sonuçları ise %87.2 doğruluk oranı ile Sayım Vektörleştirici kullanılan Lojistik Regresyon algoritması vermiştir. Genel sonucun veri bölme işleminden etkilenmediği görülmektedir.

Bhasin ve Das yapmış oldukları çalışmada, Apache Hadoop ile makine öğrenimi teknikleri incelenmiştir [6]. Çalışmada ilk olarak hashtag temelinde tweetler toplanmış daha sonra bir ön işlemden geçirilmiştir. Özellik çıkarımı yapılmıştır. Makine öğrenimi modellerinden LR, NB ve SVM kullanılmıştır. Hadoop büyük verileri işlemek için kullanılan açık kaynaklı bir framework olarak tanımlanmıştır. Map-Reduce algoritması bir Hadoop modülü olarak üç aşamadan oluşmaktadır; Eşleyici aşaması, Karıştırma aşaması ve İndirgeyici aşaması. Eşleyici aşaması tweetlerin girdi olarak alınıp işlenir. Karıştırma aşaması karıştırma, kopyalama ve sıralama olarak üç aşamaya ayrılmaktadır. İndirgeyici aşama ise tweetlerin pozitif ve negatif olarak belirlendiği aşamadır. Hadoop Dağıtılmış Dosya Sistemi, bir dosya sisteminin bilgisayarlar üzerinde çalışmasını sağlayan sistemdir. Apache Spark Map Reduce'un kısıtlamalarını kaldırmak için tasarlanan toplu reduce seçimi olarak tanımlanmaktadır. Hızlı sorgu yapmak için kullanılmaktadır. Çalışmanın sonucunda veri kümesinin boyutu çok büyük olduğu zaman Naive Bayes makine öğrenimi algoritması diğer algoritmalarla göre daha sağlamdır.

Jagadeesan ve arkadaşları tarafından yapılan çalışmada, Twitter duygu analizinin gerçekleştirilmesi için makine öğrenimi yöntemi kullanılmıştır [16]. Çalışmanın amacı, Twitter'daki

ırkçı, kadın düşmanı, dine ve mültecilere karşı yapılan nefret söylemlerinin makine öğrenimi kullanılarak tespit edilip tanımlanmasıdır. Metodoloji dört adımdan oluşmaktadır. İlk adım olarak, veri koleksiyonları toplanmıştır. Daha sonra, bu koleksiyonlar ön işleme tabi tutulmuş ve istenmeyen karakterler metinden kaldırılmıştır. Tokenizasyon uygulanmış, lemmatizasyon ve stemming işlemleri uygulanarak bir kelimenin bütün alternatifleri bir kategoride sınıflandırılıp kelimenin köküne yoğunlaştırılmıştır. Bu iki işlem Doğal Dil İşleme yaklaşımıdır. Üçüncü adım olarak, özellik çıkarma uygulanmıştır. Son adımda ise duygu sınıflandırması yapılmıştır. Bu adımda, Rastgele Orman Sınıflandırıcı, Lojistik Regresyon, Karar Ağacı Sınıflandırıcı, Destek Vektör Makinesi ve XGBoost Sınıflandırıcı olmak üzere beş algoritma ile deneyler yapılmıştır. Hata matrisi oluşturularak doğruluk oranları incelenmiştir. Sonuçlara bakıldığında, doğruluk oranı %95 ile Rastgele Orman ve Destek Vektör Makinesi algoritmaları en iyi sonucu verdiği gözlemlenmiştir.

Aytuğ tarafından yapılan çalışmada, Twitter üzerinde duygu analizi için psikodilbilimsel bir yaklaşım benimsenmiş ve çeşitli metodolojik adımlar takip edilmiştir [4]. Veri seti, İngilizce tweet'lerden oluşmakta olup olumlu, olumsuz ve nötr duygular içermektedir. Toplamda 13,361 tweet içeren veri seti, dengeli bir şekilde etiketlenmiş 4200 negatif, 4200 pozitif ve 4200 nötr tweet içermektedir. Veri Seti Koleksiyonu aşamasında, Twitter Streaming API ve Twitter4J kütüphanesi kullanılarak tweet'ler toplanmıştır. Veri Ön İşleme aşamasında, tweet'lerin düzensiz ve gayri resmi doğası dikkate alınarak, baş harfler, gereksiz tekrarlar ve yanlış kullanılan harfler gibi belirli sorunlar ortadan kaldırılmıştır. Özellik Mühendisliği aşamasında, duygu analizi için psikodilbilimsel özellikler çıkarmak amacıyla LIWC (Dil bilimsel Sorgulama ve Kelime Sayısı) kullanılmıştır. Bu çerçevede, dilin duygusal, bilişsel ve yapısal yönlerini tanımlamak için kullanılan bir metin analizi uygulamasıdır. Sınıflandırma Algoritmaları aşamasında, çeşitli özellik mühendisliği şemalarının tahmin per-

formansını değerlendirmek için Naïve Bayes, Destek Vektör Makineleri, k-En Yakın Komşu algoritması ve Lojistik Regresyon kullanılmıştır. Naïve Bayes algoritması, en yüksek tahmin performansını sağlamıştır. Ensemble Öğrenme Yöntemleri aşamasında, topluluk öğrenme algoritmaları olan Bagging, AdaBoost ve Rastgele Altuzay yöntemleri kullanılmıştır. Bagging, farklı eğitim setleri üzerinde eğitilmiş zayıf öğrenme algoritmalarını birleştirerek daha yüksek tahmin performansı elde etmeyi amaçlayan bir yöntemdir. AdaBoost, sınıflandırılması zor olan veri noktalarına odaklanarak sağlam bir sınıflandırma şeması elde etmeyi amaçlamaktadır. Rastgele Altuzay yöntemi, rastgele seçilen özellik altuzayları üzerinde eğitilen birden fazla sınıflandırıcıyı birleştiren bir topluluk öğrenme algoritmasıdır. Deney analizi ve sonuçlar aşamasında, 10 kat çapraz doğrulama yöntemi kullanılarak çeşitli özellik kümeleri ve sınıflandırıcılarla yapılan deneyler değerlendirilmiştir. Psikodilbilimsel özelliklerin ve topluluk öğrenme yöntemlerinin duygu analizi performansını arttırdığı gözlemlenmiştir. Özellikle, dilbilimsel süreçler, psikolojik süreçler ve kişisel kaygıların birleştirildiği topluluk özellik kümeleri, Naïve Bayes sınıflandırıcısı ile birlikte kullanıldığında %89,10'luk bir sınıflandırma doğruluğu elde edilmiştir. F-skoru sonuçları da sınıflandırma doğrulukları ile paralellik göstermiştir. Bu çalışmanın sonuçları, Twitter verilerinde duygu analizi için psikodilbilimsel özelliklerin etkili bir şekilde kullanılabileceğini ve topluluk öğrenme yöntemlerinin sınıflandırma algoritmalarının performansını arttırabileceğini göstermektedir.

Kemaloğlu ve arkadaşları yapmış oldukları çalışmada, Facebook, Twitter, Instagram, YouTube ve Rss verilerini kullanarak duygu analizi üzerinde çalışmıştır [20]. LSTM modeli, özellikle Twitter verileriyle eğitilen modeller arasında en yüksek doğruluk oranını elde etmiştir. Multinomial Naive Bayes ise %86.06 başarı oranıyla dikkat çekmektedir. Rastgele Orman algoritması, kelime tabanlı yöntemlerde %89.5 başarı sağlamıştır. Ancak özellikle TF-IDF kullanılarak geliştirilen LSTM modeli %83.3 başarı oranı sağlamıştır. Bu çalışmada,

5 farklı sosyal medya platformundan API'ler kullanılarak veri toplanmış ve veri seti %30 test, %70 eğitim verisi olarak ayrılmıştır. Python tabanlı bir yazılım kullanılarak geliştirilen sistemde, numpy, pandas, keras ve scikit-learn gibi kütüphaneler kullanılmıştır. Analiz sonuçlarına göre, en yüksek başarı oranı %84.46 ile LSTM modelinden elde edilmiş ve bu nedenle geliştirilen sistemle birlikte kullanılmak üzere LSTM modeli tercih edilmiştir.

Hassan ve arkadaşları yapmış oldukları çalışmada, tweetlerin duyarlılık analizini kullanarak COVID-19 salgınına tespit etmeye yönelik bir araştırma gerçekleştirilmiştir [14]. Bu bağlamda Twitter veri kaynağı olarak seçilmiştir ve COVID-19 ile ilgili anahtar kelimeler içeren birçok tweet toplanmıştır. Toplanan bu veriyi daha sonra filtrelemek ve iki etikete ayırmak amacıyla farklı ön işleme teknikleri uygulanmıştır. Bu tekniklerden biri geçerli veriyi temsil ederken diğeri geçersiz veriyi ifade etmektedir. Ön işlemeden sonra ait olduğu verinin bir dizi kategorisini tanımlamak için sınıflandırma algoritmaları kullanılmıştır. Sınıflandırma için Karar Ağacı Algoritmasını, Naïve Bayes Algoritmasını ve Destek Vektör Makinesi Algoritmasını kullanılmıştır. Deneylerimiz için en uygun yolu belirlemek amacıyla dört farklı k-kat değeri kullanılmıştır. Bu adımı, araştırmanın doğruluk testlerini uyguladığı performans değerlendirme adımı izlemiştir. Deneylerden, SVM'nin verilerin doğruluk açısından sınıflandırılmasına en uygun olduğu gözlemlenmiştir. Bu araştırma, COVID-19 hastası olan ve herhangi bir nedenle test yapılmayan hastaları tespit etmek için bu tür bir yaklaşımın kullanabileceğini göstermektedir.

Li ve arkadaşlarının üzerinde çalıştığı çalışma, Çince üzerinden yürütülmüş ve duygular 7 ana kategoride sınıflandırılmıştır [23]. Bu ana kategoriler, iki ana duygu kategorisinde birleştirilmiştir. “Mutlu” ve “iyi” ifadeleri olumlu duygu; “üzgün”, “kızgın”, “korkmuş” ve “şaşkın” olumsuz duygu olarak sınıflandırılmıştır. Olumlu duyguların sıfır puandan büyük, olumsuz duyguların sıfır puandan küçük ve nötr duyguların ise sıfır puan olarak değerlendirildi-

rilmesi sağlanmıştır. Bu çalışmada bazı kurallar uygulanmıştır. İlk kural olarak, olumsuz bir kelimenin bir duygudan sonra gelmesi durumunda cümlelerin duygusunun tam tersine çevrilmesi ve ona göre puanlandırma yapılması sağlanmıştır. İkinci kural olarak, karşıt bağlaç kullanıldığında cümlelerin olumsuza çevrilmesi ve puan ağırlığının eksi olması sağlanmıştır. Üçüncü kural ise derece zarflarına farklı ağırlıklar verilmesi üzerinedir. Bu kuralların birçok olumlu ve olumsuz yanlarının olduğu belirtilmiştir. Bu çalışmada, duyarlılık sözlüğü ve TF-IDF yöntemini birleştiren yeni bir yöntem tasarlanmıştır. Tüm duygu kelimelerinin ağırlıkları E olarak değerlendirilmiş ve bir metnin toplam ağırlığı 1 olarak sınırlandırılmıştır. Daha sonra, ağırlığın ikiye bölünmesi ve tüm duygu kelimelerine a, nötr kelimelere ise 1-a ifadesinin TF-IDF değerleriyle orantılı olarak atanması sağlanmıştır. Bu verilerle bir denklem elde edilmiştir. Duygu analizi için Yu Hua'nın "Çiseleyen Yağmurda Çılgınlıklar" adlı romanı 16 bölüm şeklinde ayrıştırılarak veri olarak kullanılmıştır. Sonuçlar incelendiğinde, toplamda 2079 cümlelerin analiz edildiği gözlemlenmiştir. Metindeki verilerin duygu dağılımı incelendiğinde, romanın daha çok olumsuz duygular içerdiği tespit edilmiştir. Değerlendirme olarak kesinlik, geri çağırma ve f1 puanı hesaplanmıştır. Duygu sözlüğüne dayalı kurallar, TF-IDF ve bu çalışmada önerilen yöntemlerin doğruluk oranı karşılaştırıldığında en yüksek oranı önerilen yöntemin sağladığı belirtilmiştir. Geri çağırma için yöntemler karşılaştırıldığında, pozitif duygularda en yüksek oranı duygu sözlüğüne dayalı kurallar, nötr ve negatif duygularda ise önerilen yöntemin sağladığı gözlemlenmiştir. F1 skoru karşılaştırıldığında, üç durumda da en yüksek oranı önerilen yöntemin verdiği tespit edilmiştir. Sonuç olarak, önerilen yöntemin en iyi performansı gösterdiği belirtilmiştir.

Başarlan ve Kayaalp yapmış oldukları çalışmada, TripAdvisor ve Rotten Tomatoes gibi web sitelerindeki açık kullanıcı geri bildirimlerine duygu analizi uygulanmıştır [5]. Word2Vec ve TF-IDF yöntemleri kullanılarak dört makine öğrenme ve iki topluluk öğrenme yöntemiyle

sonuçlar elde edilmiştir. Topluluk öğrenme, özellikle çoğunluk oylaması, tekli öğrenme yöntemlerinden daha etkili sonuçlar gösterdiği belirtilmiştir. Eğitim-test ayrımları %80-%20 ve %70-%30 olarak iki şekilde yapılmıştır. Metin ön işleme adımları arasında yorum temizleme, özel karakterleri çıkarma, küçük harfe dönüştürme ve stemming yer almaktadır. LR, SVM, NB ve KNN sınıflandırıcıları kullanılarak gerçekleştirilen deneylerde, scikit-learn kütüphanesi Python ile uygulanmıştır. Bu çalışma, topluluk öğrenme yöntemlerinin duyarlılık sınıflandırma sürecine önemli katkı sağladığını göstermektedir.

Karayığit ve arkadaşlarının yapmış oldukları çalışmanın amacı, sistem hakkındaki düşünceleri öğrenmek amacıyla, öğrenciler tarafından sistem hakkında toplanan Türkçe tweetlerin duygu analizinin yapılması hedeflenmiştir [19]. Türkçe dilinde duygu analizi için makine öğrenmesi yaklaşımlarının daha iyi performans gerçekleştirdiği yapılan araştırmalarla ortaya konulmuştur. Bu nedenle bu çalışmada makine öğrenimi tabanlı yöntemler tercih edilmiştir. Twitter aracılığıyla konu ile ilgili hashtag kullanılarak 63699 adet tweet toplanmıştır. Veri ön işleme dört adıma ayrılmıştır: temizleme, normalizasyon, tokenizasyon ve durak kelimeleri kaldırma. Temizleme kısmında gereksiz kelimeler, bağlantılar veri kümesinden çıkarılmıştır. Yabancı dil kullanılan tweetler “Language Detection API” ile tespit edilip aynı şekilde veri kümesinden çıkarılmıştır. Geriye 4652 adet tweet kalmıştır. Normalleştirme kısmında ise noktalama işaretleri kaldırılmış ve tüm harfler küçük harfe dönüştürülmüştür. Tokenizasyon kısmında kalan veriyi anlamlı parçalara bölme işlemi yapılmıştır. Son adımda durak kelimeler çıkarılarak kalan metin daha anlamlı hale gelmiştir. Duygu analizinde üç ayrı duygu kullanılmıştır; olumlu, olumsuz ve nötr. Vektör uzayı modeli olarak ise Bag of Words ve TF-IDF kullanılmıştır. Duygu sınıflandırması için bu çalışmada Destek Vektör Makinesi, Lojistik Regresyon, K-En Yakın Komşu ve Yapay Sinir Ağları kullanılmıştır. Veri seti 1075 negatif, 3345 nötr ve 232 pozitif veriden oluşmaktadır. Veri setinde açık bir şekilde görülen dengesizlik

vardır. Bu sebeple karşılaştırma için F1 skoru kullanılmıştır. Sonuçlar incelendiğinde Lojistik Regresyon sınıflandırıcının TF-IDF ve BoW vektör uzaylarını kullanarak diğer makine öğrenimi yöntemlerine göre daha iyi sonuç verdiği görülmüştür. Detaylı incelemede SVM, ANN ve kNN yöntemlerinin BoW vektör uzayı modeli kullanılan çalışmada, LR yönteminin ise TF-IDF modeli kullanan çalışmada çok daha iyi sonuç verdiği gözlemlenmiştir.

Aydın ve arkadaşları yapmış oldukları çalışmada, duygu analizi ve fikir madenciliği arasındaki ilişkiyi açıklayarak bunlar üzerinde bir inceleme yapmak amaçlanmıştır [3]. Ayrıca bu çalışmada Türkçe'nin diğer dillerden daha farklı olmasına ve bir kelime ile bir cümle kurulabilmesine dikkat çekilmiş, duygu analizi yaparken başka bir dilin modellerinin kullanılamayacağı fakat duygu analizinde kullanılan yöntemlerin kullanılabileceği vurgulanmıştır. Duygu analizi ve fikir madenciliği incelendiğinde fikir madenciliğinin akademik alanda, duygu analizinin ise endüstride kullanıldığı görülmektedir. Farklılıklar olsa dahi iki alanında birlikte çalıştığı gözlemlenmiştir. Sonuçlar kısmında Türkçe duygu analizi yapmak için veri kümesi bulunmaması gibi zorluklar yaşandığına, bu zorlukların sosyal medyadan veri elde edilmesi ile aşılabileceğine değinilmiştir. Aynı zamanda bu konu ile ilgili yapılan projelerin yetersiz olması gereken bütçenin sağlanması ile aşılabileceği belirtilmiştir. Eğitim setinin kategori olarak düzgün ayarlanmasının analiz üzerinde doğrudan etkili olduğu söylenmiştir.

Salur ve Aydın yapmış oldukları çalışmada, Türkçe tweetler için yeni bir duygu sınıflandırma modeli tasarlanmıştır [31]. Veri ön işlemede silinen karakterlerden, emojiler gibi sözcüklerden özellik çıkarma ve diğer özelliklerle birleştirilmesine dayalı bir çalışma yapılmıştır. Veri kümesinde 6887 adet pozitif ve 7890 adet negatif tweet bulunmaktadır. Tweetler anlamlı veya anlamsız birçok karakter, kelime içermektedir. Bu kelimelerin daha anlamlı olabilmesi için veri ön işleme yapılmıştır. Veri ön işlemede URL bilgisi, fazla boşluklar, noktalama işaretleri, kullanıcı adı, anlamsız kelimeler, rakamlar ve tek karakterler silinmiş

karakter tekrarı içerdiği için anlamsız olan kelimeler düzeltilmiştir. Ayrıca daha önce yapılan çalışmalarda Türkçe tweetlerin analizinde kelime gövdeleme işlemi, duygu analizin performansına katkı sağlamadığı için bu çalışmada kullanılmadığına değinilmiştir. Kalan veriler FastText kullanılarak metin temsili oluşturulmuş ve bu veri kümesinden LSTM, BiLSTM ve GRU kullanılarak derin özellik çıkarımı yapılmıştır. Veri ön işleme kısmında silinen içeriklerin aslında duygu belirttiği ve Twitter'a özgü ifadeler olduğu görülmüştür. Bu içeriklerden 12 adet elle özellik çıkarımı yapılmıştır. Bu özellikler, sözdizimsel olarak; toplam karakter sayısı, büyük harf sayısı, küçük harf sayısı, rakam sayısı, noktalama işareti sayısı ve URL sayısı olarak toplam 6 tane, Twitter'a özgü olarak; Bahsetme (mention) sayısı ve Hashtag sayısı olarak toplam 2 tane ve anlamsal olarak; pozitif emoji sayısı, negatif emoji sayısı, pozitif ünlem sayısı, negatif ünlem sayısı olarak toplam 4 tane özellikten oluşmaktadır. Elle çıkarılan özellikler LSTM, BiLSTM ve GRU kullanılarak derin özellikler çıkarılmıştır. Daha sonra FastText kullanılarak yapılan özellik çıkarımı ile elle yapılan özellik çıkarımı birleştirilerek duygu sınıflandırma katmanına iletilmiştir. Bu çalışmada veriler eğitim için %90 ve test işlemi için %10 olarak ayrılmıştır. Modelin başarısını değerlendirmek için karmaşıklık matrisi kullanılmıştır. Karmaşıklık matrisinden elde edilen değerlerle doğruluk, kesinlik, duyarlılık ve F1 skoru hesaplanmıştır. Derin özellik çıkarma modelinde doğruluk oranları incelendiğinde BiLSTM'in daha yüksek doğruluk verdiği gözlemlenmiştir. Genel sonuçlarda en yüksek doğruluk oranı %75.71 oran ile elle özellik çıkarma ve BiLSTM'in birleşiminden oluşan modelde gözlemlenmiştir. Çıkan sonuçlar incelendiğinde en yüksek sonuçların özellik birleştirme modellerinde olduğu görülmektedir. Bu çalışma daha önce yapılan çalışmalar ile karşılaştırıldığında önerilen modelin daha iyi performans gösterdiği gözlemlenmiştir.

Tan ve arkadaşları yapmış oldukları çalışmada, duygu analizi sırasında alaycılık tespitini artırmaya yönelik bir çerçeve sunulmuştur [37]. Alaycı ifadelerin duygu sınıflandırmasını



olumsuz etkilediği belirtilmiş ve bu sorunu çözmek adına çoklu görev öğrenimini içeren bir derin sinir ağı kullanılmıştır. Çerçeve, duygu sınıflandırması ve alaycılık sınıflandırması görevlerini aynı anda gerçekleştirmek üzere Çift Yönlü LSTM ağını eğitmeyi içermektedir. Çalışma, bu yöntemin diğer modellere göre daha iyi performans sergilediğini göstermektedir. Veri kümesi olarak Kaggle'dan elde edilen duyarlılık ve alaycılık veri kümeleri kullanılmıştır. Her iki veri kümesi için aynı doğal dil işleme teknikleri ve ön işleme yöntemleri uygulanmıştır. Çerçeve, özellikle paylaşılan bir katman kullanarak aşırı uyum riskini azaltmayı amaçlayarak duygu analizi performansını artırmayı hedeflemiştir. Önerilen çerçeve, duyarlılık sınıflandırması için %91, alaycılık sınıflandırması için %92 F1 puanları elde edilmiştir. Bağımsız Bi-LSTM modeline kıyasla, çerçeve duyarlılık ve alaycılık sınıflandırmalarında sırasıyla %3 ve %1 iyileşme sağlanmıştır. Çerçeve, çoklu görev öğrenimi kullanarak eğitildiğinden hem eğitim süresini hem de hesaplama gücünü daha verimli bir şekilde kullanıldığı sonucuna varılmıştır.

Tsai ve Huang yapmış oldukları çalışmada, evcil hayvanların duygu analizini yapay zeka sistemlerinde derin öğrenme teknolojileriyle gerçekleştirmeyi amaçlamıştır [38]. İnsan duygu analizi çerçevesini evcil hayvanlar için uyarlayarak görüntü ve ses özelliklerini birleştirerek duygu analizi yapılmıştır. Evcil hayvanların yüz kaslarının sınırlamaları nedeniyle duruşları, ses özellikleri ise spektrogram kullanılarak değerlendirilmiştir. Nesne etiketlerini tespit etmek ve tanımlamak için Mask R-CNN kullanılarak derin öğrenme teknolojileriyle evcil hayvanların duygu analizi gerçekleştirilmiştir. Kullanılan teknolojiler arasında Logitech Webcam C925 kamera, gömülü sistemlerde çekirdek bilgi işlem platformu, Python programlama dili ve Tensorflow derin öğrenme kütüphanesi bulunmaktadır. Ortalama yürütme süresi 1,32 dakika olan önerilen yöntemin doğruluğu %85,71'dir. Bu yöntemin, temassız bir şekilde ev-

cil hayvan duygu tanımlamasında başarılı bir doğruluk oranına sahip olduğu ve sahibine bilgi sağladığı vurgulanmıştır.

Chen ve arkadaşları yapmış oldukları çalışmada, askeri alandaki duygu analizi için özel olarak geliştirilen MILSentic adlı duygu sözlüğünün performansını değerlendirmişlerdir [9]. Yapılan deneylerde, mevcut duygu sözlükleri ile MILSentic'in kombinasyonları kullanılarak gerçekleştirilen analizlerde, MILSentic'in eklenmesinin kutupluluk tahmininde %1,5 ve %1,4 daha yüksek doğruluk ve F1-ölçümü sağladığı belirlenmiştir. Ayrıca, çalışma, LSTM ve Bi-LSTM modelleri kullanılarak gerçekleştirilen parametre kalibrasyonunu içermektedir. Bu bağlamda, ağ katmanları ve aktivasyon fonksiyonları üzerinde yapılan ayarlamalarda, doğruluk ve F1-ölçümü değerlerindeki değişiklikler gözlemlenmiştir. Bi-LSTM modeli, özellikle Tanh aktivasyon fonksiyonu ve belirli bir ağ katmanı sayısı ile birlikte kullanıldığında en iyi performansı sergilemiştir. Bu çalışma, askeri duygu analizi için özel veri kaynaklarının etkili bir şekilde kullanılmasının, duygu sözlüklerinin geliştirilmesinin ve derin öğrenme modellerinin parametre ayarlarının performans artışına nasıl katkı sağlayabileceğini detaylı bir şekilde ele almaktadır.

Proje tasarımı dersi kapsamında incelenen çalışmaların genel özetlerinin yanında karşılaştırma tablosu da hazırlanmıştır. Karşılaştırma tabloları Çizelge 1'de verilmektedir.

Çizelge 1: Çalışmaların Karşılaştırılması

| Çalışma        | Yöntem | Sonuç | Çalışılan Konu |
|----------------|--------|-------|----------------|
| Rawat vd. [28] | BNB    | %81   | Aşı ile atılan |
|                | SVC    | %97   | tweet verileri |
|                | LR     | %95   | incelenmiştir. |

Çizelge 1

| Çalışma         | Yöntem     | Sonuç | Çalışılan Konu  |
|-----------------|------------|-------|---|
| Demir vd. [10]  | DistilBERT | %80   | Türkçe haber<br>metinleri<br>işlenmiştir.   |
|                 | RoBERT     | %80   |   |
|                 | ALBERT     | %77   |   |
|                 | NB         | %71   |   |
|                 | SVC        | %68   |   |
|                 | RF         | %68   |   |
| Öge vd. [24]    | LR         | %87   | IMDb film<br>yorumları<br>üzerinde duygu<br>analizi<br>yapılmıştır.                             |
| Fu vd. [12]     | LSTM       | %78.5 | Belirli duygu<br>analizi veri<br>kümeleri üzerinde<br>model başarımları<br>karşılaştırılmıştır. |
|                 | LE-LSTM    | %80.8 |   |
| Rumeli vd. [30] | NB         | %73.2 | Türkçe metinler<br>için bir duygu<br>analizi modeli<br>geliştirilmiştir.                        |
|                 | SVM        | %73.4 |   |
|                 | RF         | %73.3 |   |
|                 | k-NN       | %73.8 |   |

Çizelge 1

| Çalışma                  | Yöntem  | Sonuç  | Çalışılan Konu   |
|--------------------------|---------|--------|--|
| Wang vd. [39]            | CNN     | %89.9  | Geleneksel kelime gösterimine duygu bilgisi eklemek amacıyla çalışılmıştır.                      |
|                          | Bİ-LSTM | %90.6  |  |
|                          | Bİ-GRU  | %91.1  |  |
| Hassan vd. [14]          | kNN     | %59.7  | Tweetler kullanılarak COVID-19 salgınının tespiti araştırılmıştır.                               |
|                          | DT      | %61.5  |  |
|                          | SVM     | %80.9  |  |
|                          | NB      | %51.6  |  |
| Başarslan ve Kayaalp [5] | NM      | %82.5  | Açık kullanıcı geri bildirimlerine duygu analizi uygulanmıştır.                                  |
|                          | kNN     | %81.2  |  |
|                          | LR      | %83.2  |  |
|                          | SVM     | %84.6  |  |
| Poornima vd. [26]        | MNB     | %83.54 | Twitter verilerini kullanarak cümle sınıflandırılmasında performans karşılaştırması yapılmıştır. |
|                          | SVM     | %85.69 |  |
|                          | LR      | %86.23 |  |

Çizelge 1

| Çalışma             | Yöntem  | Sonuç   | Çalışılan Konu   |
|---------------------|---|---|--|
| Zhang ve Zheng [41] | SVM<br>ELM with kernels                         | %88.54<br>%88.74                              | Çince metinlerde duygu analizi ele alınmıştır.   |
| Sharma vd. [35]     | CNN<br>LSTM                                     | %93.5<br>%77.5                                | Çok terimli makine öğrenimi yaklaşımlarının duygu analizindeki performansları karşılaştırılmıştır. |
| Xu ve Song [40]     | CNN<br>LSTM<br>GRU<br>Bert+BiLSTM<br>Bert + CNN | %97.9<br>%97.68<br>%97.89<br>%97.37<br>%98.34 | Derin öğrenme ve geleneksel yöntemler ile duygu analizi yapılmıştır.                               |
| Aydın vd. [3]       | SVM<br>LR<br>kNN<br>ANN                         | %73.5<br>%75.32<br>%62.03<br>%74.11           | Duygu analizi ve fikir madenciliği arasındaki ilişkiyi açıklamışlardır.                            |

Çizelge 1

| Çalışma                | Yöntem                              | Sonuç  | Çalışılan Konu   |
|------------------------|-------------------------------------|--|--|
| Salur ve Aydın<br>[31] | SVM<br>MNB<br>Bİ-BİLSTM             | %65.2<br>%66.6<br>%75.71                       | Türkçe tweetler için yeni bir duygu sınıflandırma modeli tasarlanmıştır. |
| Singh vd. [36]         | RF<br>SVM                           | %77.90<br>%83.76                               | TF-IDF ve makine öğrenimi algoritmaları ile duygu analizi yapılmıştır.   |
| Saraswat vd. [32]      | ANN                                 | %84  | Ses ile duygu analizi yapılmıştır.                                       |
| Pandya vd. [25]        | SVM<br>NB<br>Bi-LSTM<br>RNN<br>LSTM | %76.67<br>%75.67<br>%82.13<br>%78.36<br>%80.83 | Tweetler üzerinde duygu analizi yapmak hedeflenmiştir.                   |

Çizelge 1

| Çalışma                 | Yöntem                 | Sonuç                            | Çalışılan Konu  |
|-------------------------|------------------------|----------------------------------|---|
| Kemaloğlu vd. [40]      | MLR<br>RF<br>LSTM      | %83.07<br>%84.24<br>%84.46       | Facebook, Twitter, Instagram, YouTube ve Rss verilerini kullanarak duygu analizi yapılmıştır. |
| Satrya vd. [34]         | SVM                    | %96.64                           | Toplumun dijital varlıklara yönelik tutumlarını anlamak amaçlanmıştır.                        |
| Sarkar ve Bhowmick [33] | SVM<br>NB              | %45<br>%44.2                     | Bengalce tweetler üzerine araştırma yapmıştır.  |
| Hammad ve Anwar [13]    | SVM<br>NB<br>DT<br>kNN | %71.3<br>%61.8<br>%65.1<br>%66.1 | Sindhi dilinde makine öğrenimiyle duygu analizi gerçekleştirilmiştir.                         |

Çizelge 1

| Çalışma           | Yöntem | Sonuç  | Çalışılan Konu  |
|-------------------|--------|--------|---|
| Jayakody vd. [17] | CNN    | %90.90 | Ürün inceleme yorumları üzerinde duygu analizi yapılmıştır. |
|                   | NB     | %57.90 |   |
|                   | SVM    | %67.7  |   |



BÖLÜM: İİ

MATERYAL

---

## MOTİVASYON VE KATKI

---

Sosyal medya platformlarının yaygınlaşması ve kullanıcıların bu platformlarda yoğun şekilde veri üretmesi, duygu analizine olan ilgiyi artırmıştır. Duygu analizi, sosyal medya verilerinden anlamlı bilgi çıkarımı sağlayarak çeşitli alanlarda önemli uygulamalara olanak tanımaktadır. Bu çalışmanın motivasyonu, makine öğrenimi ve derin öğrenme tekniklerini kullanarak duygu analizinde yüksek doğruluk oranları elde etmek ve farklı algoritmalar ile veri temsili yöntemlerini karşılaştırarak en etkili yaklaşımları belirlemektir.

Önceki çalışmalar genellikle sınırlı sayıda veri kümesi ve algoritma kullanarak gerçekleştirilmiştir. Bu çalışmada ise daha kapsamlı bir yaklaşım benimsenmiş, farklı veri kümeleri ve çeşitli algoritmalar kullanılarak duygu analizinin performansı değerlendirilmiştir. Bu sayede, duygu analizinde hangi yöntemlerin daha etkili olduğunu belirlemek ve bu alandaki literatüre katkıda bulunmak amaçlanmıştır.

Bu çalışma, duygu analizi alanına önemli katkılar sağlamaktadır:

**Kapsamlı Literatür İncelemesi:** Çalışma kapsamında 39 farklı makale incelenmiş ve mevcut yöntemler ile yaklaşımlar hakkında derinlemesine bilgi sunulmuştur. Bu literatür taraması, çalışmanın temellerini güçlendirerek literatürdeki mevcut boşlukları ve geliştirme alanlarını belirlemeye yardımcı olmuştur.

**Çeşitli Veri Kümelerinin Kullanımı:** İki farklı veri kümesi kullanılarak, farklı veri kaynaklarından elde edilen sonuçlar karşılaştırılmıştır. Bu yaklaşım, duygu analizi modellerinin genellenebilirliğini ve farklı veri kümelerindeki performansını değerlendirme imkanı sunmaktadır.

**Çok Sınıflı Karşılaştırmalar:** Hem iki sınıflı hem de üç sınıflı sınıflandırma yapılmış, böylece farklı sınıflandırma düzeylerinde elde edilen sonuçlar incelenmiştir. Bu, duygu analizinin farklı sınıflandırma senaryolarındaki performansını karşılaştırma imkanı sunmaktadır.

**Çeşitli Metin Temsil Yöntemleri:** Çalışmada, beş farklı metin temsil yöntemi kullanılmış ve metinlerin farklı şekillerde temsil edilmesinin duygu analizine etkisi araştırılmıştır. Bu yöntemlerin karşılaştırılması, hangi temsil yöntemlerinin daha etkili olduğunu belirlemeye yardımcı olmaktadır.

**Geniş Algoritma Yelpazesi:** Dokuz makine öğrenimi algoritması ve üç derin öğrenme algoritması kullanılarak sınıflandırma yapılmıştır. Bu geniş algoritma yelpazesi, farklı algoritmaların duygu analizi performansını karşılaştırma ve en etkili olanları belirleme imkanı sunmaktadır.

Bu katkılar sayesinde, duygu analizi alanında daha geniş ve derinlemesine bir anlayış sağlanmakta ve bu alanda yapılacak gelecekteki çalışmalara önemli bir referans noktası oluşturulmaktadır.

BÖLÜM: İİİ

YÖNTEM

---

## DENEYSEL AYARLAMALAR

---

Bu bölümde, duygu analizi çalışmasının altyapısını oluşturan tekniklere değinilecektir.

### 4.1 Programlama Dili Ve Kütüphaneler

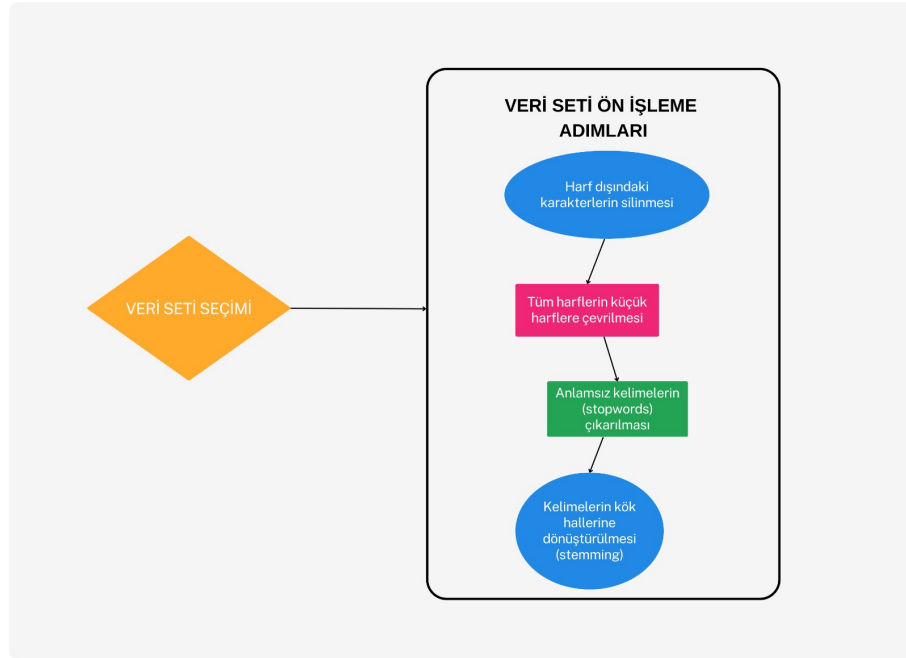
Bu çalışmada Python programlama dili ve çeşitli Python kütüphaneleri kullanılmıştır. Veri analizi ve işleme için Pandas, bilimsel hesaplamalar için NumPy, makine öğrenimi algoritmaları için Scikit-Learn, derin öğrenme modelleri için Tensorflow, metin madenciliği için Gensim bu kütüphanelerden bazılarıdır.

### 4.2 Veri Kümeleri

Çalışmada iki farklı veri kümesi kullanılmıştır. Bu veri kümeleri Kaggle sitesinden alınmıştır. İlk veri kümesi Amazon sitesindeki müzik ürünlerine aittir ve olumlu, olumsuz ve nötr duyguları içeren toplam 10.000 veriden oluşmaktadır [1]. Diğer veri kümesi ise IMDb film yorumlarından oluşmakta olup, olumlu ve olumsuz duyguları içeren ilk 20.000 veriden meydana gelmektedir [2].

### 4.3 Veri Ön İşleme

Veri ön işleme aşaması, metin verilerinin analiz için uygun hale getirilmesini sağlar. Bu çalışmada, Şekil 1’de gösterilen adımlar izlenerek boş satırların temizlenmesi, her bir metin ögesinin küçük harfe dönüştürülmesi, sayıların ve noktalama işaretlerinin kaldırılması, anlamsız kelimelerin (stopwords) çıkarılması ve kelimelerin kök hallerinin bulunması işlemleri gerçekleştirilmiştir. Bu adımlar, metin verilerinin analiz edilebilir ve tutarlı bir formata getirilmesine katkıda bulunarak daha doğru sonuçlar elde edilmesini sağlamaktadır.



Şekil 1: Veri Ön İşleme Adımları

### 4.4 Özellik Çıkarım Teknikleri

Metin temsil yöntemleri, ham metni makine öğrenimi ve derin öğrenme algoritmalarının işleyebileceği sayısal vektörlere dönüştürmek amacıyla uygulanmaktadır. Bu çalışmada,

TF-IDF, Bag of Words, GloVe, FastText ve Word2Vec gibi yöntemler kullanılmıştır. Her bir yöntem, farklı prensiplere dayanarak kendine özgü avantajlar sunmaktadır.

#### 4.5 Sınıflandırma Algoritmaları

Bu çalışmada, sosyal medya verilerinin duygu analizi için çeşitli makine öğrenimi ve derin öğrenme algoritmaları kullanılmıştır. Makine öğrenimi algoritmaları kapsamında, Scikit-learn kütüphanesinden KNeighbors Classifier, Random Forest Classifier, Logistic Regression, Decision Tree Classifier, Bernoulli Naive Bayes, Support Vector Machine ve Multi-Layer Perceptron Classifier algoritmaları uygulanmıştır. Ayrıca, LightGBM ve XGBoost gibi gelişmiş gradient boosting algoritmaları kullanılarak metin verilerinin sınıflandırılması gerçekleştirilmiştir. Derin öğrenme algoritmaları ise TensorFlow Keras kütüphanesi kullanılarak gerçekleştirilmiş olup, LSTM, SimpleRNN ve Conv1D katmanları ile modeller oluşturulmuştur. Bu geniş algoritma yelpazesi, duygu analizi performansını optimize etmek ve farklı yöntemlerin etkinliğini karşılaştırmak amacıyla tercih edilmiştir.

#### 4.6 Performans Metrikleri

Bu çalışmada, duygu analizi modellerinin performansını değerlendirmek amacıyla Doğruluk, Hassasiyet, Duyarlılık ve F1-Skor metrikleri kullanılmıştır.

Doğruluk, modelin doğru tahmin ettiği örneklerin toplam örnek sayısına oranıdır. Formülü 1'de verilmiştir.

$$Dogruluk = \frac{DogruTahminSayisi}{TumTahminlerinSayisi} \quad (1)$$

Hassasiyet, pozitif sınıf olarak tahmin edilen örnekler arasından gerçekten pozitif olanların oranını gösterir. Formülü 2’de verilmiştir.

$$Hassasiyet = \frac{DogruPozitif}{DogruPozitif + YanlisPozitif} \quad (2)$$

Duyarlılık, gerçek pozitif örneklerin model tarafından doğru tahmin edilme oranını gösterir. Formülü 3’te verilmiştir.

$$Duyarlilik = \frac{DogruPozitif}{DogruPozitif + YanlisNegatif} \quad (3)$$

F1-Skor, hassasiyet ve duyarlılığın harmonik ortalamasıdır ve sınıflandırma modelinin genel performansını değerlendirir. Formülü 4’te verilmiştir. Özellikle dengesiz veri setlerinde faydalıdır.

$$F1Skor = 2 \times \frac{Hassasiyet * Duyarlilik}{Hassasiyet + Duyarlilik} \quad (4)$$

Bu metrikler, modelin doğru ve yanlış tahminlerini detaylı bir şekilde analiz ederek, duygu analizi performansının kapsamlı bir değerlendirmesini sağlar. Projede kullanılan algoritmaların her biri için bu metrikler hesaplanmış ve sonuçlar karşılaştırılarak en etkili yöntemlerin belirlenmesi amaçlanmıştır.



---

## ELDE EDİLEN SONUÇLAR

---

Amazon veri seti üzerinde derin öğrenme teknikleri kullanılarak elde edilen sonuçlar Tablo 3’de, makine öğrenimi algoritmaları kullanılarak elde edilen sonuçlar ise Tablo 2’de verilmiştir. IMDb veri seti üzerinde derin öğrenme teknikleri kullanılarak elde edilen sonuçlar Tablo 4’de, makine öğrenimi algoritmaları kullanılarak elde edilen sonuçlar ise Tablo 5’de verilmiştir.

Bu tablolar, her iki veri seti üzerinde farklı öğrenme tekniklerinin ve algoritmaların performansını karşılaştırmak için kullanılmıştır. Her bir yöntemin doğruluğu, hassasiyeti, geri çağırımı ve F1 skoru gibi metriklerle değerlendirilmiştir.

Çizelge 2: Amazon Veri Seti Makine Öğrenimi Sonuçları

| Metin Temsil<br>Yöntemi | Sınıflandırma<br>Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|-------------------------|------------------------------|--------------|-----------|--------|----------|
| TF-IDF                  | BNB                          | 84.05        | 0.79      | 0.84   | 0.81     |
|                         | LR                           | 88.34        | 0.82      | 0.88   | 0.83     |
|                         | RF                           | 88.24        | 0.78      | 0.88   | 0.83     |
|                         | DT                           | 83.32        | 0.82      | 0.83   | 0.83     |
|                         | SVM                          | 88.29        | 0.84      | 0.88   | 0.83     |

Çizelge 2 Amazon Veri Seti Makine Öğrenimi Sonuçları

| Metin Temsil<br>Yöntemi | Sınıflandırma<br>Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|-------------------------|------------------------------|--------------|-----------|--------|----------|
|                         | ANN                          | 87.03        | 0.84      | 0.87   | 0.85     |
|                         | KNN                          | 87.51        | 0.79      | 0.88   | 0.83     |
|                         | XGBOOST                      | 88.20        | 0.83      | 0.88   | 0.84     |
|                         | LightGBM                     | 88.29        | 0.84      | 0.88   | 0.84     |
| WORD2VEC                | BNB                          | 88.10        | 0.81      | 0.88   | 0.83     |
|                         | LR                           | 88.05        | 0.81      | 0.88   | 0.83     |
|                         | RF                           | 88.34        | 0.87      | 0.88   | 0.83     |
|                         | DT                           | 80.30        | 0.82      | 0.80   | 0.81     |
|                         | SVM                          | 88.29        | 0.84      | 0.88   | 0.83     |
|                         | ANN                          | 83.71        | 0.83      | 0.84   | 0.83     |
|                         | KNN                          | 86.44        | 0.81      | 0.86   | 0.83     |
|                         | XGBOOST                      | 88.59        | 0.84      | 0.89   | 0.84     |
|                         | LightGBM                     | 88.39        | 0.83      | 0.88   | 0.83     |
| GLOVE                   | BNB                          | 84.73        | 0.79      | 0.85   | 0.82     |
|                         | LR                           | 87.95        | 0.80      | 0.88   | 0.83     |
|                         | RF                           | 88.20        | 0.83      | 0.88   | 0.83     |
|                         | DT                           | 77.96        | 0.80      | 0.78   | 0.79     |
|                         | SVM                          | 88.29        | 0.90      | 0.88   | 0.83     |
|                         | ANN                          | 87.56        | 0.81      | 0.88   | 0.83     |
|                         | KNN                          | 87.03        | 0.80      | 0.87   | 0.83     |
|                         | XGBOOST                      | 88.10        | 0.83      | 0.88   | 0.83     |

Çizelge 2 Amazon Veri Seti Makine Öğrenimi Sonuçları

| Metin Temsil<br>Yöntemi | Sınıflandırma<br>Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|-------------------------|------------------------------|--------------|-----------|--------|----------|
|                         | LightGBM                     | 88.20        | 0.83      | 0.88   | 0.83     |
| FASTTEXT                | BNB                          | 82.35        | 0.80      | 0.82   | 0.81     |
|                         | LR                           | 88.05        | 0.82      | 0.88   | 0.83     |
|                         | RF                           | 88.29        | 0.85      | 0.88   | 0.83     |
|                         | DT                           | 77.86        | 0.80      | 0.78   | 0.79     |
|                         | SVM                          | 88.29        | 0.90      | 0.88   | 0.83     |
|                         | ANN                          | 87.46        | 0.81      | 0.87   | 0.83     |
|                         | KNN                          | 86.15        | 0.80      | 0.86   | 0.83     |
|                         | XGBOOST                      | 88.20        | 0.80      | 0.88   | 0.83     |
|                         | LightGBM                     | 88.10        | 0.78      | 0.88   | 0.83     |
| BINARY                  | BNB                          | 83.95        | 0.79      | 0.84   | 0.81     |
|                         | LR                           | 86.78        | 0.84      | 0.87   | 0.85     |
|                         | RF                           | 88.29        | 0.78      | 0.88   | 0.83     |
|                         | DT                           | 83.08        | 0.82      | 0.84   | 0.83     |
|                         | SVM                          | 88.29        | 0.78      | 0.88   | 0.83     |
|                         | ANN                          | 86.25        | 0.83      | 0.86   | 0.84     |
|                         | KNN                          | 88.29        | 0.85      | 0.88   | 0.83     |
|                         | XGBOOST                      | 88.34        | 0.83      | 0.88   | 0.84     |
|                         | LightGBM                     | 88.34        | 0.84      | 0.88   | 0.85     |

Tablo 2, Amazon veri setinde yapılan makine öğrenimi çalışmalarının sonuçlarını kapsamlı bir şekilde sunmaktadır. Makine öğrenimi modellerinin doğruluk oranları %77.86 ile %88.59 arasında değişmektedir. En yüksek doğruluk oranı %88.59 ile Word2Vec temsil yöntemi ve XGBOOST algoritması kombinasyonu tarafından elde edilmiştir. Genel olarak, XGBOOST, LightGBM, ve SVM algoritmaları, farklı metin temsil yöntemleriyle kullanıldığında yüksek doğruluk oranları sağlamıştır.

Çizelge 3: Amazon Veri Seti Derin Öğrenme Sonuçları

| Metin Temsil Yöntemi | Sınıflandırma Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|----------------------|---------------------------|--------------|-----------|--------|----------|
| TF-IDF               | CNN                       | 88.29        | 0.90      | 0.88   | 0.83     |
|                      | RNN                       | 88.00        | 0.83      | 0.88   | 0.83     |
|                      | LSTM                      | 88.29        | 0.90      | 0.88   | 0.83     |
| WORD2VEC             | CNN                       | 88.10        | 0.79      | 0.88   | 0.83     |
|                      | RNN                       | 88.15        | 0.78      | 0.88   | 0.83     |
|                      | LSTM                      | 88.00        | 0.81      | 0.88   | 0.84     |
| GLOVE                | CNN                       | 88.29        | 0.90      | 0.88   | 0.83     |
|                      | RNN                       | 88.29        | 0.90      | 0.88   | 0.83     |
|                      | LSTM                      | 88.24        | 0.80      | 0.78   | 0.83     |
| FASTTEXT             | CNN                       | 88.29        | 0.90      | 0.88   | 0.83     |
|                      | RNN                       | 88.29        | 0.87      | 0.88   | 0.83     |
|                      | LSTM                      | 88.30        | 0.90      | 0.88   | 0.83     |
| BINARY               | CNN                       | 88.05        | 0.90      | 0.88   | 0.83     |
|                      | RNN                       | 88.29        | 0.90      | 0.88   | 0.83     |
|                      | LSTM                      | 88.29        | 0.78      | 0.88   | 0.83     |

Tablo 3, Amazon veri setinde gerçekleştirilen derin öğrenme çalışmalarının sonuçlarını detaylı bir şekilde sunmaktadır. Bu çalışmanın sonuçlarına göre, derin öğrenme modellerinin doğruluk oranları %88.00 ile %88.30 arasında değişmektedir. En yüksek doğruluk oranı

%88.30 ile FastText temsil yöntemi ve LSTM algoritması kombinasyonu tarafından elde edilmiştir. Genel olarak, TF-IDF, GloVe, ve FastText temsil yöntemleri, CNN, RNN, ve LSTM algoritmaları ile birlikte kullanıldığında yüksek doğruluk oranları sergilemiştir. Özellikle CNN ve LSTM algoritmaları, çoğu temsil yöntemiyle yüksek performans göstermiştir. Bununla birlikte, LSTM algoritması, özellikle FastText ve TF-IDF ile birlikte kullanıldığında, genellikle yüksek doğruluk oranları sağlamıştır.

Çizelge 4: IMDb Veri Seti Derin Öğrenme Sonuçları

| Metin Temsil Yöntemi | Sınıflandırma Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|----------------------|---------------------------|--------------|-----------|--------|----------|
| TF-IDF               | CNN                       | 87.24        | 0.88      | 0.87   | 0.87     |
|                      | RNN                       | 51.80        | 0.75      | 0.49   | 0.32     |
|                      | LSTM                      | 51.20        | 0.75      | 0.51   | 0.35     |
| WORD2VEC             | CNN                       | 84.77        | 0.85      | 0.85   | 0.85     |
|                      | RNN                       | 76.67        | 0.80      | 0.77   | 0.76     |
|                      | LSTM                      | 85.37        | 0.85      | 0.85   | 0.85     |
| GLOVE                | CNN                       | 75.02        | 0.77      | 0.75   | 0.75     |
|                      | RNN                       | 69.57        | 0.71      | 0.70   | 0.83,,69 |
|                      | LSTM                      | 80.54        | 0.82      | 0.81   | 0.80     |
| FASTTEXT             | CNN                       | 85.07        | 0.86      | 0.85   | 0.85     |
|                      | RNN                       | 73.65        | 0.79      | 0.74   | 0.73     |
|                      | LSTM                      | 87.84        | 0.88      | 0.88   | 0.88     |
| BINARY               | CNN                       | 78.50        | 0.79      | 0.79   | 0.78     |
|                      | RNN                       | 48.00        | 0.49      | 0.48   | 0.33     |
|                      | LSTM                      | 52.00        | 0.75      | 0.52   | 0.36     |

Tablo 4, IMDb film incelemeleri veri seti üzerinde gerçekleştirilen derin öğrenme çalışmalarının sonuçlarını detaylandırmaktadır. Çalışma %48.00 ile %87.84 arasında değişen

doğruluk oranları göstermiştir. En yüksek doğruluk oranı %87.84 ile FastText temsil yöntemi ve LSTM algoritması kombinasyonu tarafından elde edilmiştir. Genel olarak, TF-IDF ve FastText temsil yöntemleri CNN ve LSTM algoritmaları ile yüksek doğruluk oranları sağlamıştır. GloVe ve Binary temsil yöntemleri ise genellikle daha düşük doğruluk oranları sergilemiş, özellikle RNN algoritmaları ile düşük performans göstermiştir.

Çizelge 5: IMDb Veri Seti Makine Öğrenimi Sonuçları

| Metin Temsil Yöntemi | Sınıflandırma Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|----------------------|---------------------------|--------------|-----------|--------|----------|
| TF-IDF               | BNB                       | 85.65        | 0.86      | 0.86   | 0.86     |
|                      | LR                        | 88.37        | 0.88      | 0.88   | 0.88     |
|                      | RF                        | 84.05        | 0.84      | 0.84   | 0.84     |
|                      | DT                        | 70.17        | 0.70      | 0.70   | 0.70     |
|                      | SVM                       | 88.85        | 0.89      | 0.89   | 0.89     |
|                      | ANN                       | 87.45        | 0.87      | 0.87   | 0.87     |
|                      | KNN                       | 74.52        | 0.75      | 0.75   | 0.74     |
|                      | XGBOOST                   | 83.52        | 0.84      | 0.84   | 0.84     |
|                      | LightGBM                  | 84.70        | 0.85      | 0.85   | 0.85     |
| WORD2VEC             | BNB                       | 51.25        | 0.59      | 0.51   | 0.35     |
|                      | LR                        | 86.65        | 0.87      | 0.87   | 0.87     |
|                      | RF                        | 83.02        | 0.83      | 0.83   | 0.83     |
|                      | DT                        | 71.97        | 0.72      | 0.72   | 0.72     |
|                      | SVM                       | 86.87        | 0.87      | 0.87   | 0.87     |
|                      | ANN                       | 84.47        | 0.85      | 0.85   | 0.85     |

Çizelge 5 IMDb Veri Seti Makine Öğrenimi Sonuçları

| Metin Temsil<br>Yöntemi | Sınıflandırma<br>Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|-------------------------|------------------------------|--------------|-----------|--------|----------|
|                         | KNN                          | 80.55        | 0.81      | 0.81   | 0.81     |
|                         | XGBOOST                      | 85.20        | 0.85      | 0.85   | 0.85     |
|                         | LightGBM                     | 84.77        | 0.85      | 0.85   | 0.85     |
| GLOVE                   | BNB                          | 67.00        | 0.67      | 0.67   | 0.67     |
|                         | LR                           | 75.75        | 0.76      | 0.76   | 0.76     |
|                         | RF                           | 73.15        | 0.73      | 0.73   | 0.73     |
|                         | DT                           | 61.65        | 0.62      | 0.62   | 0.62     |
|                         | SVM                          | 75.05        | 0.75      | 0.75   | 0.75     |
|                         | ANN                          | 72.00        | 0.73      | 0.73   | 0.73     |
|                         | KNN                          | 66.12        | 0.66      | 0.66   | 0.66     |
|                         | XGBOOST                      | 74.02        | 0.74      | 0.74   | 0.74     |
|                         | LightGBM                     | 73.77        | 0.74      | 0.74   | 0.74     |
| FASTTEXT                | BNB                          | 71.75        | 0.72      | 0.72   | 0.72     |
|                         | LR                           | 88.07        | 0.85      | 0.85   | 0.85     |
|                         | RF                           | 79.85        | 0.80      | 0.80   | 0.80     |
|                         | DT                           | 69.35        | 0.69      | 0.69   | 0.69     |
|                         | SVM                          | 85.22        | 0.85      | 0.85   | 0.85     |
|                         | ANN                          | 81.70        | 0.82      | 0.82   | 0.82     |
|                         | KNN                          | 77.05        | 0.77      | 0.77   | 0.77     |
|                         | XGBOOST                      | 82.82        | 0.83      | 0.83   | 0.83     |
|                         | LightGBM                     | 82.67        | 0.83      | 0.83   | 0.83     |

Çizelge 5 IMDb Veri Seti Makine Öğrenimi Sonuçları

| Metin Temsil Yöntemi | Sınıflandırma Algoritması | Accuracy (%) | Precision | Recall | F1-Score |
|----------------------|---------------------------|--------------|-----------|--------|----------|
| BINARY               | BNB                       | 85.72        | 0.86      | 0.86   | 0.86     |
|                      | LR                        | 87.12        | 0.87      | 0.87   | 0.87     |
|                      | RF                        | 84.22        | 0.84      | 0.84   | 0.84     |
|                      | DT                        | 70.70        | 0.71      | 0.71   | 0.71     |
|                      | SVM                       | 88.07        | 0.88      | 0.88   | 0.88     |
|                      | ANN                       | 86.15        | 0.86      | 0.86   | 0.86     |
|                      | KNN                       | 59.87        | 0.62      | 0.60   | 0.57     |
|                      | XGBOOST                   | 84.10        | 0.84      | 0.84   | 0.84     |
|                      | LightGBM                  | 85.07        | 0.85      | 0.85   | 0.85     |

Tablo 5, IMDb veri setinde gerçekleştirilen makine öğrenimi çalışmalarının derinlemesine incelenmesini sunmaktadır. Çalışma %51.25 ile %88.85 arasında doğruluk oranları göstermiştir. En yüksek doğruluk oranı %88.85 ile TF-IDF ve SVM kombinasyonu tarafından elde edilmiştir. Genel olarak, TF-IDF ve Binary metin temsil yöntemleri, SVM ve LR algoritmaları ile yüksek doğruluk oranları sağlamıştır. TF-IDF ve SVM en iyi performansı verirken, GloVe temsil yöntemi genellikle daha düşük doğruluk oranları göstermiştir.



## BÖLÜM: İV

### SONUÇ

---

## GENEL DEĞERLENDİRME

---

Bu çalışmada, Amazon ve IMDb veri setleri kullanılarak duygu analizi gerçekleştirilmiş ve çeşitli metin temsilleri ile sınıflandırma algoritmalarının performansları karşılaştırılmıştır. Amazon veri setinde üç sınıflı (olumlu, olumsuz, tarafsız), IMDb veri setinde ise iki sınıflı (olumlu, olumsuz) sınıflandırma yapılmıştır. Elde edilen sonuçlar, derin öğrenme ve makine öğrenmesi yöntemlerinin farklı sınıflandırma problemlerindeki performansını ortaya koymaktadır.

Amazon veri seti üzerinde yapılan analizlerde, derin öğrenme yöntemleri arasında en yüksek doğruluk oranı %88.30 ile TF-IDF ve FastText kullanan LSTM algoritması tarafından elde edilmiştir. Bu sonuç, LSTM'in metinlerdeki uzun süreli bağıntıları yakalamadaki üstün başarısını ortaya koymaktadır. CNN ve RNN modelleri de benzer doğruluk oranlarına ulaşmış, ancak LSTM biraz daha yüksek performans sergilemiştir. Öte yandan, makine öğrenmesi yöntemleri arasında LightGBM algoritması %88.39 doğruluk oranı ile en yüksek performansı gösterirken, Logistic Regression (LR) algoritması %88.34 doğruluk oranı ile dikkat çekici bir performans sergilemiştir. Bu bulgular, makine öğrenmesi algoritmalarının da yüksek doğruluk oranlarına ulaşabileceğini ve bazı durumlarda derin öğrenme yöntemlerine kıyasla daha hızlı ve verimli olabileceğini göstermektedir.

IMDb veri seti üzerinde gerçekleştirilen analizlerde, derin öğrenme yöntemleri arasında en yüksek doğruluk oranı %87.84 ile FastText kullanan LSTM algoritması tarafından elde edilmiştir. Bu LSTM'in metinlerdeki uzun süreli bağıntıları başarılı bir şekilde yakaladığını göstermektedir. Ayrıca, Word2Vec ile LSTM %85.37 doğruluk oranı ile yüksek bir performans sergilemiştir. Makine öğrenmesi yöntemleri arasında ise Logistic Regression %88.85 doğruluk oranı ile en yüksek performansı gösterirken Support Vector Machine algoritması %88.07 doğruluk oranı ile dikkat çekici bir başarı elde etmiştir. Bu sonuçlar, iki sınıflı duygu analizinde makine öğrenmesi algoritmalarının oldukça etkili olduğunu ortaya koymaktadır.

Sonuç olarak, bu çalışmada Amazon ve IMDb veri setlerinde duygu analizi için derin öğrenme ve makine öğrenmesi yöntemleri karşılaştırılmıştır. Amazon veri setinde en yüksek doğruluk oranı LightGBM algoritması ile elde edilirken (%88.39) IMDb veri setinde Logistic Regression algoritması en yüksek doğruluk oranına (%88.85) ulaşmıştır. Derin öğrenme modelleri, özellikle büyük ve karmaşık veri setlerinde daha avantajlı olmasına rağmen, bazı durumlarda makine öğrenmesi modelleri daha hızlı ve verimli sonuçlar sunabilmektedir. İki sınıflı ve üç sınıflı sınıflandırma problemleri, model performansı ve sınıflandırma zorluğu açısından farklılıklar göstermektedir. Bu bulgular, duygu analizi projelerinde kullanılacak yöntemi belirlerken dikkate alınması gereken önemli faktörleri ortaya koymaktadır.

---

## KAYNAKÇA

---

- [1] Amazon Veri seti: <https://www.kaggle.com/datasets/eswarchandt/amazon-music-reviews>.
- [2] IMDB Veri seti: <https://www.kaggle.com/datasets/hrishikeshpable/imdb-movies-sentiment-dataset>.
- [3] AYDIN, Z. E., OZTURK, Z. K., AND ÇIÇEK, Z. İ. E. Turkish sentiment analysis for open and distance education systems. *Turkish Online Journal of Distance Education* 22, 3 (2021), 124–138.
- [4] AYTUĞ, O. Sentiment analysis on twitter based on ensemble of psychological and linguistic feature sets. *Balkan Journal of Electrical and Computer Engineering* 6, 2 (2018), 69–77.
- [5] BAŞARSLAN, M. S., AND KAYAALP, F. Sentiment analysis with ensemble and machine learning methods in multi-domain datasets. *Turkish Journal of Engineering* 7, 2 (2023), 141–148.
- [6] BHASIN, A., AND DAS, S. Twitter sentiment analysis using machine learning and hadoop: A comparative study. 267–272.
- [7] CAN, U., AND ALATAS, B. Duygu analizi ve fikir madenciliği algoritmalarının incelenmesi. *International Journal of Pure and Applied Sciences* 3, 1 (2017), 75–111.

- [8] CHANDRA, Y., AND JANA, A. Sentiment analysis using machine learning and deep learning. 1–4.
- [9] CHEN, L.-C., LEE, C.-M., AND CHEN, M.-Y. Exploration of social media for sentiment analysis using deep learning. *Soft Computing* 24 (2020), 8187–8197.
- [10] DEMIR, E., AND BILGIN, M. Sentiment analysis from turkish news texts with bert-based language models and machine learning algorithms. 01–04.
- [11] DHARIWAL, N., AKUNURI, S. C., BANU, K. S., ET AL. Audio and text sentiment analysis of radio broadcasts. *IEEE Access* 11 (2023), 126900–126916.
- [12] FU, X., YANG, J., LI, J., FANG, M., AND WANG, H. Lexicon-enhanced lstm with attention for general sentiment analysis. *IEEE Access* 6 (2018), 71884–71891.
- [13] HAMMAD, M., AND ANWAR, H. Sentiment analysis of sindhi tweets dataset using supervised machine learning techniques. 1–6.
- [14] HASSAN, S. M., JAWAD, K., KHAN, M. A., KHAN, M. S., AHMAD, I., AND MOHSIN, K. Detecting covid-19 pandemic using sentiment analysis of tweets. *Artificial Intelligence Theory and Applications* 1, 2 (2021), 39–47.
- [15] HE, H., ZHOU, G., AND ZHAO, S. Exploring e-commerce product experience based on fusion sentiment analysis method. *IEEE Access* 10 (2022), 110248–110260.
- [16] JAGADEESAN, M., SARAVANAN, T., SELVARAJ, P., ALI, U. A., ARUNSIVARAJ, J., AND BALASUBRAMANIAN, S. Twitter sentiment analysis with machine learning. 681–686.
- [17] JAYAKODY, J., AND KUMARA, B. Sentiment analysis on product reviews on twitter using machine learning approaches. 1056–1061.

- [18] JOSHI, V., PATEL, S., AGARWAL, R., AND ARORA, H. Sentiments analysis using machine learning algorithms. 1425–1429.
- [19] KARAYİĞİT, H., ÇİĞDEM, A., AND AKDAĞLI, A. A review of turkish sentiment analysis and opinion mining. *Balkan Journal of Electrical and Computer Engineering* 6, 2 (2018), 94–98.
- [20] KEMALOĞLU, N., KÜÇÜKSİLLE, E., AND ÖZGÜNSÜR, M. Turkish sentiment analysis on social media. *Sakarya University Journal of Science* 25, 3 (2021), 629–638.
- [21] KIM, R. Y. Using online reviews for customer sentiment analysis. *IEEE Engineering Management Review* 49, 4 (2021), 162–168.
- [22] LI, Z., LI, R., AND JIN, G. Sentiment analysis of danmaku videos based on naïve bayes and sentiment dictionary. *Ieee Access* 8 (2020), 75073–75084.
- [23] LIU, H., CHEN, X., AND LIU, X. A study of the application of weight distributing method combining sentiment dictionary and tf-idf for text sentiment analysis. *IEEE Access* 10 (2022), 32280–32289.
- [24] ÖĞE, B. C., AND KAYAALP, F. Farklı sınıflandırma algoritmaları ve metin temsil yöntemlerinin duygu analizinde performans karşılaştırılması. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi* 9, 6 (2021), 406–416.
- [25] PANDYA, V., SOMTHANKAR, A., SHRIVASTAVA, S. S., AND PATIL, M. Twitter sentiment analysis using machine learning and deep learning techniques. 1–5.
- [26] POORNIMA, A., AND PRIYA, K. S. A comparative sentiment analysis of sentence embedding using machine learning techniques. 493–496.

- [27] PORIA, S., HAZARIKA, D., MAJUMDER, N., AND MIHALCEA, R. Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research. *IEEE Transactions on Affective Computing* (2020).
- [28] RAWAT, A., MAHESHWARI, H., KHANDUJA, M., KUMAR, R., MEMORIA, M., AND KUMAR, S. Sentiment analysis of covid19 vaccines tweets using nlp and machine learning classifiers. 225–230.
- [29] RIZK, Y. E., AND ASAL, W. M. Sentiment analysis using machine learning and deep learning models on movies reviews. 129–132.
- [30] RUMELLI, M., AKKUŞ, D., KART, Ö., AND ISIK, Z. Sentiment analysis in turkish text with machine learning algorithms. 1–5.
- [31] SALUR, M. U., AND AYDIN, İ. Türkçe tweetler için derin özellik çıkarımı tabanlı yeni bir duygu sınıflandırma modeli. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi* 34, 1 (2022), 1–13.
- [32] SARASWAT, S., BHARDWAJ, S., VASHISTHA, S., AND KUMAR, R. Sentiment analysis of audio files using machine learning and textual classification of audio data. 1–5.
- [33] SARKAR, K., AND BHOWMICK, M. Sentiment polarity detection in bengali tweets using multinomial naïve bayes and support vector machines. 31–36.
- [34] SATRYA, R. N., PRATIWI, O. N., FA'RIFAH, R. Y., AND ABAWAJY, J. Cryptocurrency sentiment analysis on the twitter platform using support vector machine (svm) algorithm. 01–05.

- [35] SHARMA, S., PANDEY, A., KUMAR, V., OHDAR, D., PILLAI, A. R., AND MAHAJAN, M. Recent trends in sentiment analysis using different machine learning based models: A short review. 474–481.
- [36] SINGH, S., KUMAR, K., AND KUMAR, B. Sentiment analysis of twitter data using tf-idf and machine learning techniques. 252–255.
- [37] TAN, Y. Y., CHOW, C.-O., KANESAN, J., CHUAH, J. H., AND LIM, Y. Sentiment analysis and sarcasm detection using deep multi-task learning. *Wireless personal communications* 129, 3 (2023), 2213–2237.
- [38] TSAI, M.-F., AND HUANG, J.-Y. Sentiment analysis of pets using deep learning technologies in artificial intelligence of things system. *Soft Computing* 25, 21 (2021), 13741–13752.
- [39] WANG, Y., HUANG, G., LI, J., LI, H., ZHOU, Y., AND JIANG, H. Refined global word embeddings based on sentiment concept for sentiment analysis. *IEEE Access* 9 (2021), 37075–37085.
- [40] XU, L., AND SONG, Y. Comparison of text sentiment analysis based on traditional machine learning and deep learning methods. 692–695.
- [41] ZHANG, X., AND ZHENG, X. Comparison of text sentiment analysis based on machine learning. 230–233.