



**T.C**  
**OSTİM TEKNİK ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**  
**YAZILIM MÜHENDİSLİĞİ ANA BİLİM DALI**  
**YÜKSEK LİSANS PROGRAMI**

**AMAZON MÜŞTERİ YORUMLARININ DUYGU ANALİZİ**  
**YÖNTEMLERİYLE DEĞERLENDİRİLMESİ**

**YÜKSEK LİSANS TEZİ**

**HAZIRLAYAN**  
**SABUHİ YUSİFOV**

**TEZ DANIŞMANI**  
**Prof. Dr. Ali SEBETCİ**

**ANKARA-2024**

## TEZ KABUL VE ONAY

Sabuhi Yusifov tarafından hazırlanan “Amazon Müşteri Yorumlarının Duygu Analizi Yöntemleriyle Değerlendirilmesi” başlıklı bu çalışma, .../.../20... tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

**Kabul Tarihi:**...../...../.....

**Jüri Üyesi:** **Dr. Öğr. Maksat ATAGOZİEV** \_\_\_\_\_  
OSTİM Teknik Üniversitesi

**Jüri Üyesi:** **Dr. Öğr. Üyesi Venera ADANOVA** \_\_\_\_\_  
TED Üniversitesi

**Tez Danışmanı:** **Prof. Dr. Ali SEBETCİ** \_\_\_\_\_  
Ostim Teknik Üniversitesi

### ONAY

Jüri tarafından kabul edilen bu çalışmanın Yüksek Lisans/Doktora Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

...../...../20....

\_\_\_\_\_  
**Enstitü Müdürü**  
**Prof. Dr. Halil Rıdvan ÖZ**

## BİLDİRİM

Enstitü tarafından onaylanan Yüksek Lisans tezimin tamamını veya herhangi bir kısmını basılı veya dijital biçimde arşivleme ve aşağıda belirtilen koşullar dahilinde erişime açma iznini OSTİM Teknik Üniversitesine verdiğimi bildiririm. Bu izinle, Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak ve gelecekteki çalışmalar (makale, kitap, lisans, patent vb.) için tezimin tamamının veya bir bölümünün kullanım hakları yalnızca bana ait olacaktır.

Tezimin bütünüyle kendi çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Telif hakkı bulunan ve sahiplerinden yazılı izinle kullanılması zorunlu olan kaynakları, yazılı izin alarak kullandığımı ve istenildiğinde izinlerin suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayımlanan “Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge” kapsamında, tezim, aşağıda belirtilen koşullar haricince, YÖK Ulusal Tez Merkezi ve OSTİM Teknik Üniversitesi Açık Erişim Sisteminde erişime açılır.

Enstitü / Fakülte Yönetim Kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.<sup>1</sup>

Enstitü / Fakülte Yönetim Kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir.<sup>2</sup>

Tezimle ilgili gizlilik kararı verilmiştir.<sup>3,4</sup>

Tarih  
İmza

---

<sup>1</sup> MADDE 6(1) Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.

<sup>2</sup> MADDE 6(2) Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.

<sup>3</sup> MADDE 7(1) Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

<sup>4</sup> MADDE 7(2) Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

**OSTİM TEKNİK ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ YÜKSEK LİSANS TEZİ**  
**ORİJİNALLİK RAPORU**

Tezin Başlığı: Amazon Müşteri Yorumlarının Duygu Analizi Yöntemleriyle Değerlendirilmesi

Öğrencinin Adı, Soyadı: Sabuhi Yusifov

Tez Danışmanın Unvanı/Adı, Soyadı: Prof. Dr. Ali Sebetci

Anabilim Dalı: Yazılım Mühendisliği

Programı: Yüksek Lisans

Tarih: 05/07/2024

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmama ait Giriş, Ana Bölümler ve Sonuç kısmından oluşan ve toplam 38 sayfadan ibaret olan kısmı, 5 / 7 / 2024 tarihinde tez danışmanım tarafından Turnitin intihal tespit programında incelenmiştir. Orijinallik raporunda aşağıda ifade edilen filtrelemeler uygulanmıştır.

Orijinallik raporuna göre, tezin benzerlik oranı % 15'dir.

Uygulanan filtrelemeler:

1. Kaynakça (hariç)
2. Alıntılar (hariç)
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları (hariç)

Tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrencinin İmzası:

**Tarih: .../ .../ 2024**  
**Tez Danışmanı**  
**Prof. Dr. Ali Sebetci**

## ETİK BEYAN

Bu çalışmanın özgün bir çalışma olduğunu, çalışmanın hazırlık, veri toplama, analiz, bilgilerin sunumu ve diğer tüm aşamalarında bilimsel etik ve kurallara uygun davrandığımı, çalışmada bulunan tüm belge bilgileri akademik etik ve kurallar çerçevesinde elde ettiğimi, görsel, işitsel ve yazılı bütün bilgileri ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu, kullanmış olduğum verilerde herhangi bir tahrifat yapmadığımı, yararlandığım kaynaklara bilimsel normlara uygun olarak atıfta bulunduğumu, tezimin kaynak gösterdiğim durumlar dışında tarafımdan kaleme alındığını ve özgün olduğunu, Tez danışmanım Prof. Dr. Ali Sebetci danışmanlığında ve tarafımdan üretildiğini ve OSTİM Teknik Üniversitesi Tez Yazım Kılavuzuna uygun olarak yazıldığımı beyan ederim.

İmza

**Sabuhi Yusifov**

**Tarih: ... / ... / 2024**

## TEŐEKKÜR

BaŐta, bu sűreçte maddi ve manevi desteęini benden esirgemeyen canım aileme sonsuz teŐekkűr ve minnettarlıęımı bildirmekteyim. Aynı zamanda tez sűrecinde yardımlarını benden esirgemeyen deęerli hocam Prof. Dr. Ali Sebetci ve destek veren tűm hocalara sonsuz teŐekkűr ve minnettarlıęımı bildirmekteyim.

**Tarih: ... / .... / 2024**  
**Sabuhi Yusifov**



## ÖZ

|                     |                             |
|---------------------|-----------------------------|
| Yazar Adı ve Soyadı | : Sabuhi Yusifov            |
| Üniversite          | : OSTİM Teknik Üniversitesi |
| Enstitü             | : Fen Bilimleri Enstitüsü   |
| Program Adı         | : Yazılım Mühendisliği      |
| Tezin Türü          | : Yüksek Lisans Tezi        |
| Sayfa Sayısı        | : 38                        |
| Tarihi              | : 2024                      |

### AMAZON MÜŞTERİ YORUMLARININ

### DUYGU ANALİZİ YÖNTEMLERİYLE DEĞERLENDİRİLMESİ

Bu tez çalışması, Amazon müşteri yorumlarının duygu analizi üzerine odaklanmıştır. Duygu analizi, doğal dil işleme ve metin madenciliği tekniklerini kullanarak metin verilerinden duygusal eğilimleri, yani müşterilerin ürün hakkında olumlu mu olumsuz mu düşündüklerini belirleme sürecidir. Doğal dil işleme, büyük veri kümelerinden anlamlı veri çıkarmanın en önemli yöntemlerinden biridir. Tezde Amazon e-ticaret platformunun müşteri yorumlarının duygu analizinde Lojistik Regresyon, Destek Vektör Makineleri, Çok Terimli Naif Bayes ve Stokastik Gradyan İnişi gibi klasik ve Çok Katmanlı Algılayıcı gibi yarı klasik yöntemlerin yanı sıra Ekstra Rastgele Ağaçlar ve Aşırı Gradyan Artırma gibi topluluk öğrenme yöntemleri ile SpaCy kütüphanesi, kurallara dayalı VADER duygu analizi ve transformer tabanlı RoBERTa dil modellerinin performansları pozitif, nötr ve negatif şeklinde üçlü yorum sınıflaması için test edilmiş, doğruluk, kesinlik, duyarlılık ve F1 puan metrikleri tespit edilmiş ve karmaşıklık matrisleri oluşturulmuştur. VADER analizinde sınıfların bileşik ve ayrık değerleri incelenmiş, VADER ve RoBERTa modellerinin sonuçları arasındaki ilişkiler görselleştirilerek açıklanmıştır. Bu çalışma kapsamında elde edilen sonuçlara göre en iyi performansları Çok Katmanlı Algılayıcı ve Ekstra Rastgele Ağaç modelleri gösterirken Çok Terimli Naif Bayes ve VADER yöntemlerinin performans metrikleri diğer modellere kıyasla daha düşük kalmıştır. Sınıfların VADER bileşik duygu polarite değerleri, ayrık değerlerle uyum halindedir. VADER nötr ve VADER negatif değişkenleri arasında, gerçek etiketleri negatif ve nötr olan yorumlarda sıkı ve negatif doğrusal bir ilişki gözlenmiştir.

**Anahtar Sözcükler:** Yapay zekâ, derin öğrenme, doğal dil işleme, duygu analizi, veri analizi, veri odaklı kararlar, müşteri memnuniyeti, Amazon müşteri yorumları.

## ABSTRACT

Thesis : Sabuhi Yusifov  
University : OSTIM Technical University  
Institute : Graduate School of Natural and Applied Sciences  
Program's Name : Software Engineer  
Thesis Type: : Master of Science  
Pages : 38  
Year : 2024

### EVALUATING AMAZON CUSTOMER REVIEWS THROUGH SENTIMENT ANALYSIS TECHNIQUES

This thesis focuses on the sentiment analysis of Amazon customer reviews. Sentiment analysis involves using natural language processing (NLP) and text mining techniques to determine whether customers have positive or negative opinions about a product. NLP is crucial for extracting valuable insights from large data sets. In this study, we analyzed Amazon customer reviews using classical methods like Logistic Regression, Support Vector Machines, Multinomial Naive Bayes, and Stochastic Gradient Descent, as well as semi-classical methods like Multilayer Perceptron. We also used ensemble learning methods like Extra Random Trees and Extreme Gradient Boosting, along with the SpaCy library, rule-based VADER sentiment analysis, and transformer-based RoBERTa models. The comments were classified as positive, neutral, or negative. We evaluated accuracy, precision, sensitivity, and F1 score metrics and created confusion matrices. The VADER model's composite and discrete values were examined, and the relationships between the VADER and RoBERTa model outcomes were visualized and explained. According to the results, the best-performing models were Multilayer Perceptron and Extra Random Trees, while Multinomial Naive Bayes and VADER showed lower performance metrics compared to the other models. VADER's composite sentiment polarity values were consistent with the discrete values. A strong negative linear relationship was observed between VADER neutral and VADER negative for comments tagged as negative and neutral.

**Keywords:** Artificial intelligence, deep learning, natural language processing, sentiment analysis, data analysis, data-driven decisions, customer satisfaction, Amazon customer reviews.

## İÇİNDEKİLER

|   |           |
|---|-----------|
| TEZ KABUL VE ONAY .....                                       | ii        |
| BİLDİRİM .....  | iii       |
| ETİK BEYAN .....  | v         |
| TEŞEKKÜR .....  | vi        |
| ÖZ .....  | vii       |
| ABSTRACT .....  | viii      |
| TABLolar DİZİNİ .....   | xi        |
| ŞEKİLLER DİZİNİ .....   | xii       |
| SİMGELER VE KISALTMALAR .....                                 | xiii      |
| <b>1. GİRİŞ</b> .....   | <b>1</b>  |
| <b>2. LİTERATÜR TARAMASI</b> .....                            | <b>4</b>  |
| <b>3. MATERYAL VE YÖNTEM</b> .....                            | <b>7</b>  |
| <b>3.1. Problem Tanımı</b> .....                              | <b>7</b>  |
| <b>3.2. Veri Seti</b> .....                                   | <b>7</b>  |
| <b>3.3. Verinin Hazırlanması</b> .....                        | <b>8</b>  |
| <b>3.4. Yöntemler</b> .....                                   | <b>9</b>  |
| 3.4.1. TF-IDF Özellik Çıkarımı .....                          | 9         |
| 3.4.2. SVM (Destek Vektör Makinesi) Sınıflandırıcısı .....    | 10        |
| 3.4.3. MNB (Multinomial Naif Bayes) Sınıflandırıcısı .....    | 11        |
| 3.4.4. MLP (Çok Katmanlı Algılayıcı) Sınıflandırıcısı .....   | 11        |
| 3.4.5. LR (Lojistik Regresyon) Sınıflandırıcısı .....         | 12        |
| 3.4.6. SGD (Stokastik Gradyan İnişi) Sınıflandırıcısı .....   | 13        |
| 3.4.7. ERT (Ekstra Rastgele Ağaçlar) Sınıflandırıcısı .....   | 13        |
| 3.4.8. XGBoost (Aşırı Gradyan Artırma) Sınıflandırıcısı ..... | 14        |
| 3.4.9. SpaCy NLP Aracı .....                                  | 15        |
| 3.4.10. VADER NLP Aracı .....                                 | 16        |
| 3.4.11. RoBERTa NLP Aracı .....                               | 16        |
| <b>4. BULGULAR VE TARTIŞMA</b> .....                          | <b>18</b> |
| <b>4.1. Performans Metrikleri</b> .....                       | <b>18</b> |
| <b>4.2. Karmaşıklık Matrisleri</b> .....                      | <b>22</b> |
| <b>4.3. Duygu Polariteleri</b> .....                          | <b>26</b> |

|  |           |
|--|-----------|
| <b>4.4. Duygu Analizi Algoritmalarının Doğruluęu ve Güvenilirlięi.....</b> | <b>31</b> |
| <b>4.5. Tez alıřmasının Sınırlılıkları.....</b>                           | <b>31</b> |
| <b>4.6. Gelecek alıřmalar ve Geliřtirme Önerileri.....</b>                | <b>31</b> |
| <b>5. SONUÇ.....</b>   | <b>32</b> |
| <b>KAYNAKLAR.....</b>  | <b>33</b> |



## TABLÖLAR DİZİNİ

|   |    |
|---|----|
| <b>Tablo 4.1.</b> Yöntemler ve Elde Edilen Sonuçlar ..... | 20 |
|---|----|



## ŞEKİLLER DİZİNİ

|  |    |
|--|----|
| Şekil 3.1. Yorumlarda Verilen Puan Sayıları .....                                  | 8  |
| Şekil 3.2. Üç Kategorili Puan Sayıları .....                                       | 9  |
| Şekil 4.1. MNB, SVM, LR, MPL, SDG ve ERT Yöntemleri Karmaşıklık Matrisleri .....   | 25 |
| Şekil 4.2. XGBoost, SpaCy, VADER ve RoBERTa Yöntemleri Karmaşıklık Matrisleri ..   | 26 |
| Şekil 4.3. Amazon Ürün Yorumları VADER Bileşik Puanları.....                       | 28 |
| Şekil 4.4. Amazon Ürün Yorumları VADER Ayrık Puanları.....                         | 28 |
| Şekil 4.5. VADER ve RoBERTa Modellerinin Sınıf Puanları Arasındaki İlişkiler ..... | 30 |



## SİMGELER VE KISALTMALAR

|        |  |
|--------|--|
| NLP    | Doğal Dil İşleme   |
| LSTM   | Long Short-Term Memory (Uzun Kısa Süreli Bellek)                                 |
| SVM    | Support Vector Machine (Destek Vektör Makinesi)                                  |
| KNN    | K-Nearest Neighbors (K-En Yakın Komşular)  |
| DA     | Duygu Analizi  |
| MNB    | Multinomial Naive Bayes (Çok Terimli Naif Bayes)                                 |
| RNN    | Recurrent Neural Network (Tekrarlayan Sinir Ağı)                                 |
| AUC    | Area Under Curve (Eğri Altındaki Alan)   |
| FN     | False Negatives (Yanlış Negatif)   |
| FP     | False Positives (Yanlış Pozitif)   |
| NLTK   | Natural Language Toolkit (Doğal Dil İşleme Araç Takımı)                          |
| TN     | True Negatives (Doğru Negatif)   |
| TP     | True Positives (Doğru Pozitif)   |
| TF-IDF | Term Frequency-Inverse Document Frequency (Terim Frekansı-Ters Doküman Frekansı) |



## 1. GİRİŞ

İnternetin hızla gelişmesiyle birlikte artık bilgiye ulaşmadaki engeller büyük ölçüde ortadan kalmıştır. Sadece 5 yıl içinde 50 milyon kullanıcıya ulaşan internet, 7'den 70'e herkesin kullandığı bir teknolojidir ve her gün hatta her an yenilenmektedir. Dijital dönüşüm ile birlikte ortaya çıkan dijital pazarlama, işletmeler için vazgeçilmez mecralar sunmaktadır. Günümüzde insanlar akıllı telefon ve tablet gibi internet temelli cihazlar ile çevrimiçi çok fazla vakit geçirmektedir. Bu durum işletmelerin zaman ve mekândan bağımsız olarak tüketici ile etkileşim kurma kabiliyetini ortaya çıkarmıştır [1]. Dijital pazarlama geleneksel pazarlama yöntemlerinin yanı sıra daha geniş kitlelere erişim sağlamak ve hedeflenen müşterilere daha kişiselleştirilmiş içerikler sunma imkânı vermektedir. Bu da işletmelerin pazarlama stratejilerini daha etkili bir şekilde yönetmelerini sağlamaktadır. Özellikle sosyal medya platformları, arama motoru optimizasyonu, içerik pazarlaması ve e-posta pazarlaması gibi dijital kanallar, işletmelerin marka bilinirliğini artırmasına, müşteri sadakatini sağlamasına ve satışları artırmasına yardımcı olmaktadır.

E-ticaret siteleri günümüzde tüketicilerin alışveriş yaparken kullandıkları en önemli mecralardan biridir. İnsanlar evlerinde oturarak dünyanın herhangi bir yerinden istedikleri ürünü sipariş edebiliyorlar. Dijital çağın getirdiği en önemli değişikliklerden biri de tüketicilerin ürün ve hizmet hakkında bilgi alma sürecini önemli ölçüde değiştirmiş olmasıdır. Geleneksel mağazaların yerini e-ticaret platformları almış ve bu platformalar tüketicilere geniş ürün yelpazesi sunmuştur. Bununla beraber tüketicilerin ürün hakkındaki fikir ve değerlendirmelerini diğer tüketicilerle paylaşmasına olanak tanımıştır. Bu geri dönüşler tüketicilerin ürün satın alma kararlarını önemli ölçüde etkilemektedir. E-ticaret devi Amazon ve diğer e-ticaret sitelerinde milyonlarca ürün değerlendirmesi ve yorumu bulundurmaktadır.

Amazon.com, Amerika merkezli bir e-ticaret ve bulut bilişim şirkettir. Jeff Bezos tarafından 1994 yılında kurulan şirket, dünya genelindeki en büyük perakendecidir ve piyasa değeri olarak da önemli bir konuma sahiptir. Başlangıçta online bir kitapçı olarak faaliyet gösteren Amazon, zamanla ürün yelpazesini genişleterek elektronikten giyime, gıdadan oyuncaklara kadar birçok kategoride satış yapmaya başlamıştır. Ayrıca tüketici elektroniği üretimi ve bulut altyapı hizmetleriyle de öne çıkmaktadır.

Küreselleşme ve teknolojik gelişmelerin etkisi ile rekabet baskısı işletmeler tarafından daha fazla hissedilir hale gelmiştir. Neredeyse her faaliyetin başına getirilebilen “e” ön ekini alarak dijital platforma taşınan ticari faaliyetler şirketleri hem tehdit etmekte hem de büyük fırsatlar sunmaktadır. İşletmeler bir yandan yeni pazarlara ulaşma imkânı yakalarken, diğer yandan stok ve haberleşme gibi giderlerini azaltma şansı bulmaktadır. İnternete erişebilen her işletmenin eşit imkanlara sahip olması rakip sayısını artırmaktadır. Müşteri perspektifinden bakılacak olursa bol çeşit ve ucuz fiyatlı ürünlere zaman ve mekândan bağımsız olarak ulaşabilme imkânı sunmaktadır [2].

Günümüzün e-ticaret kullanıcıları sadece var olan bilgilere yorum yapmak, sayfalara yer işareti koymak ve puan vermekle kalmamakta, aynı zamanda fikirlerini, haberlerini ve bilgilerini de toplumla paylaşmaktadır. İnternet üzerinde kişilerin kendilerini ifade edebileceği birçok sosyal ağ bulunmaktadır [3]. Bu kullanıcı yorumları, işletmeler için değerli bilgiler içermekte, doğru bir şekilde analiz edildiğinde pazarlama stratejilerinin belirlenmesinde ve müşteri memnuniyetinin artırılmasında kritik bir rol oynamaktadır.

Amazon müşteri yorumları, ürünler hakkında geniş ve çeşitli perspektifler sunar. Bu yorumlar, kullanıcıların ürünlerle ilgili memnuniyetlerini, hayal kırıklıklarını, önerilerini ve duygusal tepkilerini içerir. İnsanlar en basit kararları alırken dahi benzer deneyimlerden geçmiş diğer insanların o konu hakkındaki duygu ve düşüncelerini öğrenmek isteyebilir. Yapılan bir araştırmaya göre internet kullanıcılarının %84’ü çevrimiçi yorumları kişisel tavsiye olarak almakta ve %68’i de 1 ila 6 arasında çevrimiçi yorum okuduktan sonra bir yargıya vardığını belirtmektedir [3]. Bu nedenle, bu yorumların doğru bir şekilde analiz edilmesi, işletmelerin müşteri memnuniyetini artırmak için izlemesi gereken yolları belirlemelerine yardımcı olur.

Bu tez çalışması, Amazon müşteri yorumlarının duygu analizi üzerine odaklanmıştır. Duygu analizi, doğal dil işleme (NLP) ve metin madenciliği tekniklerini kullanarak metin verilerinden duygusal eğilimleri, yani müşterilerin ürün hakkında olumlu mu olumsuz mu düşündüklerini belirleme sürecidir. Doğal dil işleme (NLP) büyük veri kümelerinden anlamlı veri çıkarmanın en önemli yöntemlerinden biridir.

İnsanların kullandığı dilin anlaşılması, analiz edilmesi ve işlenmesi için yapay zekâ ve makine öğrenmesi algoritmalarının bir arada kullanıldığı, metin ve seslerin dillerin kurallı yapısına göre işlenmesini sağlayan alana Doğal Dil İşleme denir. Doğal dil işlemeyi,

dođal dil anlama ve dođal dil üretme olarak iki sınıfa bölmek mümkündür. Dođal dil anlama metinde geçen bağlamın anlaşılması, dođal dil üretme ise yeni bir metin oluşturmak anlamına gelir. Birçok dođal dil işleme uygulamasında bu iki alanın beraber kullanıldığı görülür. Dođal dil işleme, arama motorları, sanal asistanlar, metin sınıflandırma, metin özetleme, metin çevirisi, soru cevaplama, bilgi getirme, kelime işleme, duygu analizi, varlıkların ismini tanıma gibi birçok popüler alanı içermektedir [4, 5].

Sosyal medya platformları ve diđer internet ortamları özellikle hedef kitleden geri dönüşüm alabilmek için önemli ve yeterince geniş kaynaklardır. Ancak bunları insan eliyle analiz etmek neredeyse imkansızdır. Bu noktada Duygu (Sentiment) Analizi (DA) araçları devreye girerler ki bunların, sosyal platformları gözlemlemek için en işlevsel araçlar olduğu söylenebilir. Bir metinde ilgili konu hakkındaki tutum, ancak DA yapılarak anlaşılabilir. DA bir metnin duygu barındırıp barındırmadığını ve bu duygunun olumlu mu olumsuz mu olduğunu saptama sürecidir. Duygu barındıran metinler genellikle görüş ya da değerlendirme içerirler. Bu görüş ve değerlendirmeler herhangi bir konu, şahıs, marka ya da siyasi görüş hakkında olabilir [6].

Tezde öncelikle, duygu analizinin temel prensipleri ve teknikleri hakkında genel bir bilgi verilecek, ardından çeşitli duygu analizi yöntemleri ve bu yöntemlerin Amazon müşteri yorumları üzerinde uygulanması ele alınacaktır. Söz konusu analizlerin sonuçları tartışılarak, elde edilen bulguların işletmeler ve tüketiciler için taşıdığı potansiyel faydalar değerlendirilecektir. Tezin amacı, duygu analizinin müşteri yorumlarının değerlendirilmesindeki önemini vurgulamak ve bu alanda yapılacak çalışmalara ışık tutmaktır.

## 2. LİTERATÜR TARAMASI

Doğal Dil İşlemenin bir alt alanı olan ve aynı zamanda Görüş Analizi veya Görüş Madenciliği olarak da bilinen Duygu Analizi, insanların çeşitli konu, ürün ve hizmetlere ilişkin görüş, düşünce ve izlenimlerini metin madenciliği kullanarak toplama ve analiz etme sürecidir. İnsanların görüşleri şirketlere, hükümetlere ve bireylere bilgi toplama ve görüşe dayalı kararlar alma konusunda fayda sağlar. Pratik hayatın birçok konusu ayrıntılı duygu analizi gerektirir. Örneğin ürün analizi, bir ürünün hangi bileşenlerinin veya niteliklerinin müşterilere hitap ettiğini keşfeder. Duygu analizi, metindeki öznel bilgileri tanımlar ve çıkarır. Ancak bu değerlendirme prosedürü duyguların doğru yorumlanmasına engel teşkil eden birçok zorlukla karşı karşıyadır.

Yakın zamanda yayımlanan birkaç tarama makalesinde yazarlar duygu analizinin zorluklarını tanımlamış ve uygulanabilir çözümler önermiştir [7, 8]. Subhashini ve ark. [9], gürültülü veya belirsiz görüşlerden metin özelliklerinin nasıl çıkarılacağı, görüşlerdeki bilginin nasıl temsil edileceği ve bunların nasıl kategorize edileceğini anlattıkları makalelerinde çağdaş fikir madenciliği literatürünün kapsamlı bir incelemesini sundu. Yadav ve Vishwakarma [10] ile Süleymani ve ark. [11] duygu sınıflandırması üzerine makaleler yayınladılar. Yue ve ark. [12] ve Liu ve Zhang [13] internetin etkililiği üzerine çalışmalar yaptı. Piryani ve ark. [14] temel amacı, yapılandırılmamış veriden görüş çıkarmak ve ruh halini belirlemek olan hesaplamalı bir metodoloji önerdi. Jain ve ark. [15] duygu sınıflaması, tahmine dayalı karar verme, spam ve sahte yorumların tespitinde çevrimiçi incelemeleri birleştiren makine öğrenimi uygulamalarını tartıştı. Balaji ve ark. [16], gelişmiş makine öğrenimi algoritmalarını kullanarak çeşitli sosyal medya analizi uygulamalarının kapsamlı bir raporunu sundu. Mowlaei ve ark. [17], duygu sınıflandırması için uyarlanabilir bakış açısına dayalı sözlükler içeren bir teknik önerdi. Duyguların bakış açısına göre sınıflandırılmasına yardımcı olacak istatistiklere ve genetik algoritmalara dayalı iki dinamik sözlük oluşturma stratejisi tanımladılar.

Duygu analizi daha önce otellerden havayollarına, sağlık hizmetlerinden borsaya kadar çeşitli alanlarda uygulanmaktaydı [18]. Sonra (Valencia ve ark. 2019), piyasa duyarlılığına dayalı olarak borsa ve kripto para birimlerinin eğilimlerini belirlemek [19], çeşitli alanlarla ilgili tweet'leri analiz etmek [20], sağlık hizmeti alanların görüş analizini

yapmak [21] için de kullanılmaya başladı. İş sektörü, itibar yönetimi, pazar araştırması, rakip ve ürün analizi gibi konularda duygu analizinden yararlanmaya devam etmektedir.

Türkçe bağlamında son yıllarda yapılan bazı çalışmalar şunlardır: Ciplak ve Yıldız [22], Twitter'da Türkçe paylaşım yapan kullanıcıların meslek gruplarını makine öğrenmesi yöntemleriyle tahmin etmeyi amaçladı ve optimum özellik sayısı yöntemiyle %97,3 başarı elde etti. Alawi ve Bozkurt [23], Türkçe Twitt'leri geleneksel makine öğrenme modelleri, derin öğrenme teknikleri ve BERT tabanlı dönüştürücüler kullanarak Türk Üniversiteleri hakkındaki memnuniyeti tahmin etmek için analiz etti ve 0,9632'lik bir İşlem Karakteristik Eğrisine (ROC) ulaştı. Bilgin ve Kara [24], sosyal medyadaki yorumların analizini yaparak Türk toplumunun yeşil enerjiye geçişe ilişkin zaman içinde gelişen algısını tespit etti ve risk, çevre ve maliyet etiketleri için sırasıyla 0,896, 0,902 ve 0,923 RO puanlarına ulaştı.

Başarslan ve Kayaalp [25], Karar Ağacı, K-En Yakın Komşu, Lojistik Regresyon, Naif Bayes ve Rastgele Orman geleneksel sınıflandırıcıları ile Çoğunluk Oylaması, Olasılığa Dayalı Oylama ve Yığınlama topluluk öğrenimi yöntemlerini kullanarak Korona virüs salgınıyla ilgili tweet'lere duygu analizi yaptı. Küçüklerli ve Ulusoy [26], Uzun Kısa Süreli Bellek (LSTM), Aşırı Gradyan Artırma (XGBoost) ve Yinelemeli Sinir Ağı (RNN) tekniklerini kullanarak TL/USD döviz kuru verilerinin yanı sıra Türkiye ekonomisiyle ilgili anahtar kelimeler içeren Twitter verilerini analiz etti. Gökalp ve Subay [27], Twitter'daki sosyal medya kullanıcılarının Güçlendirilmiş Parlamenter Sistem hakkındaki tartışmalarını analiz ederek Türkiye'de demokrasinin zorluklarını inceledi.

Göçgün ve Onan [28], Amazon ürün yorumları veri seti üzerinde Karar Ağacı, Naif Bayes, Destek Vektör Makineleri ve Lojistik Regresyon makine öğrenmesi yöntemlerinin duygu analizi sınıflandırma performanslarını analiz etti. En yüksek doğruluk oranını (%94) karar ağacı yöntemi elde etti. Aynı yöntem %75 F1-puanı aldı. Maurya ve Pratap [29], bir Topluluk Sınıflandırıcısını Amazon ürün yorumları veri seti üzerinde, Lojistik Regresyon, SVM, Çok Terimli Naif Bayes ve Karar Ağacı gibi algoritmalarla karşılaştırdı ve Topluluk Sınıflandırıcısının bahsedilen makine öğrenme algoritmalarından daha iyi performans gösterdiğini buldu. Meghana ve ark. [30], belli bir ürün hakkındaki genel değerlendirmeyi tespit etmek için LSTM ve NB yöntemlerini kullandı. Tufchi ve ark. [31], Amazon ürün yorumlarını analiz etmek için Çok Terimli Naif Bayes, Lojistik

Regresyon (LR), Lineer Destek Vektör (SVC) ve Çok Terimli Rastgele Orman gibi makine öğrenimi modellerini kullandı ve LR ve Lineer SVC'nin sırasıyla %87,3 ve %87,4 doğrulukla iyi performans gösterdiğini buldu.

Görmez ve ark. [32] duygu analizi için, evrişimsel sinir ağı, transformer ve hibrit olmak üzere üç farklı derin öğrenme yaklaşımı önermiş ve bu modelleri iki farklı veri kümesi üzerinde test etmiştir. Modellerin doğruluğunu artırmak için hiper-parametreler ve model derinlikleri Bayes optimizasyon yöntemi ile optimize edilmiştir. Ön işlem süreçlerinin model performansına etkisine bakılmış, ön işlem uygulanmasının doğruluk oranını artırdığı, ancak örnek sayısı, dolayısıyla doğruluğu daha yüksek olan veri kümesindeki artışın diğerine göre daha az olduğu tespit edilmiştir. Güner ve ark. [33], Kaggle'ın sunduğu 4 milyon Amazon.com ürün incelemesinden rastgele seçilen 60 bin satırlık veri kümesi üzerinde Multinomial Naif Bayes (MNB), Doğrusal Destek Vektör Makinesi (LSVM) ve Uzun Kısa Süreli Bellek Ağı (LSTM) algoritmalarını karşılaştırdı ve LSTM'in en yüksek performansı sağladığını (Doğruluk = 0,90, AUC = 0,96) gördü.



### 3. MATERYAL VE YÖNTEM

#### 3.1. Problemin Tanımı

Tezde araştırılan problem, Amazon örneğindeki gibi bir e-ticaret platformunun müşteri yorumlarının duygu analizinde Lojistik Regresyon, Destek Vektör Makineleri, Gauss Naif Bayes ve Stokastik Gradyan İnişi gibi klasik ve Çok Katmanlı Algılayıcı gibi yarı klasik yöntemlerin yanı sıra Ekstra Rastgele Ağaçlar ve Aşırı Gradyan Artırma gibi topluluk öğrenme yöntemleri ile SpaCy NLP kütüphanesi, kurallara dayalı VADER duygu analizi modeli ve transformer tabanlı RoBERTa dil modeli gibi NLP araçlarının performanslarını test etmek ve söz konusu yöntemlerin güçlü ve zayıf yönlerini tespit etmektir.

#### 3.2. Veri Seti

Çalışmada Kaggle'dan [34] alınan Amazon tüketici ürün yorumları veri kümesi (Products\_May2019) kullanılmıştır. Amazon tüketici incelemeleri geniş bir ürün yelpazesini kapsadığından, bu veri kümesi birden fazla sektörden ve ürün türünden yorumlar içerir. Bu çeşitlilik, görüşleri analiz ederken farklı tüketici gruplarının tepkilerini anlamamıza olanak tanıyor. Veri kümesi 28332 satır ve 24 sütundan oluşmaktadır.

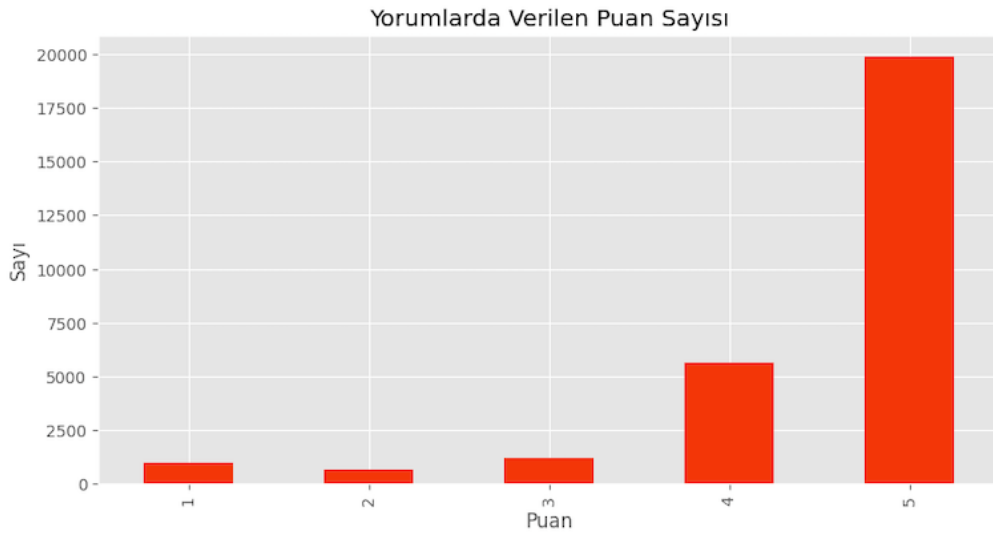
Sütunların adları ve içerdikleri bilgi aşağıdaki gibidir:

- 'id': Ürünlerin benzersiz kimlik numaralarını içerir. Bu sütunun veri türü metindir.
- 'dataAdded': Ürünün eklendiği tarihi içerir. Tarih veri türündedir.
- 'dataUpdated': Ürünün güncellendiği tarihi verir. Tarih veri türündedir.
- 'name': Ürünün ismini içerir. Metin veri türündedir.
- 'asins': Ürünün ASIN numarasını içerir. Metin veri türündedir.
- 'brand': Ürünün markasını içerir. Metin veri türündedir.
- 'categories': Ürünün kategori bilgisini içerir. Metin veri türündedir.
- 'primaryCategories': Ürünün ana kategori bilgisini içerir. Metin veri türündedir.
- 'imageURLs': Ürünle ilgili görsellerin URL adresini içerir. Nesne veri türündedir.
- 'keys': Ürünün anahtar kelime bilgisini içerir. Metin veri türündedir.
- 'manufacturer': Ürünün üretici bilgisini içerir. Metin veri türündedir.
- 'manufacturerNumber': Ürünün üretici numarasını içerir. Metin veri türündedir.
- 'reviews.date': Yorumun yapıldığı tarihi içerir. Tarih veri türündedir.
- 'reviews.dateSeen': Yorumun görüntülediği tarihi içerir. Tarih veri türündedir.

- ‘reviews.didPurchase’: Yorumu yapan kişinin ürünü satın alıp almadığını içerir. Metin veri türündedir.
- ‘reviews.doRecommend’: Yorumu yapan kişinin ürünü tavsiye edip etmediğini içerir. Metin veri türündedir.
- ‘reviews.id’: Yorumun kimlik numarasını içerir. Sayısal bir veri türüdür.
- ‘reviews.numHelpful’: Yorumun kaç kişiye yardımcı olduğu bilgisini içerir. Sayısal bir veri türüdür.
- ‘reviews.rating’: Yorumda ürüne verilen puan bilgisini içerir. Sayısal bir veri türüdür.
- ‘reviews.sourceURLs’: Yorumun kaynak URL’lerini içerir. Metin veri türündedir.
- ‘reviews.text’: Yorumun metnini içerir. Metin veri türündedir.
- ‘reviews.title’: Yorumun başlığını içerir. Metin veri türündedir.
- ‘reviews.username’: Yorumcunun kullanıcı adını içerir. Metin veri türündedir.
- ‘sourceURLs’: Kaynak URL’lerini içerir. Metin veri türündedir.

### 3.3. Verinin Hazırlanması

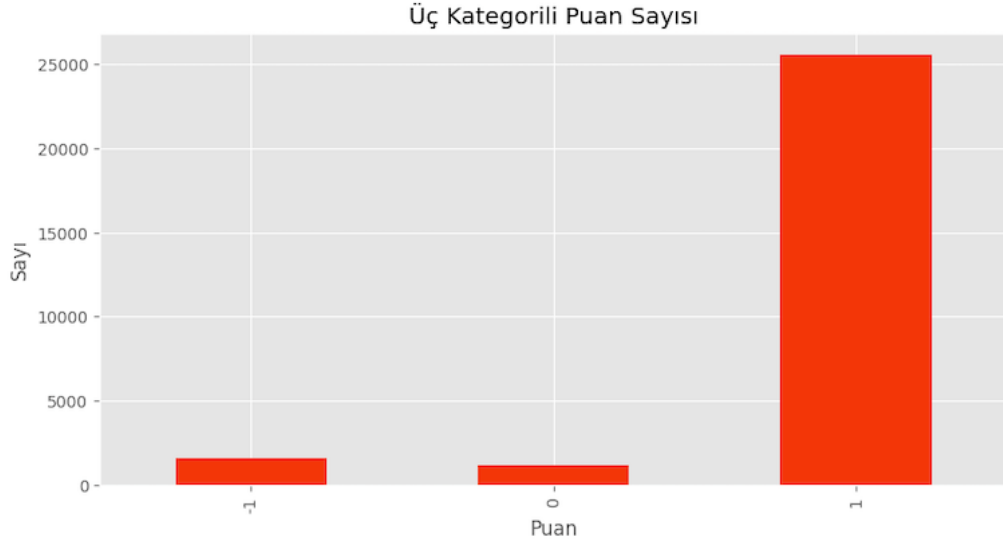
Veri kümesinin “reviews.rating” ve “reviews.text” sütunları seçilerek yeni bir DataFrame oluşturuldu ve indis numaraları “reviews.id” sütununa atandı. Yorum puanları 1’den 5’e kadar tam sayılardır. Yüksek puan yüksek memnuniyet anlamına gelir. Şekil 3.1’de her puan için kaç ürün seçildiği görülmektedir.



Şekil 3.1. Yorumlarda Verilen Puan Sayıları.

İkinci aşamada 1 ve 2 puanlarının olumsuz, 3 puanının nötr ve 4 ve 5 puanlarının olumlu yorumlar olduğu düşüncesiyle ürünlerin puanları sırasıyla -1, 0 ve 1 olarak değiştirildi.

Dolayısıyla puan kategori sayısı 5’den 3’e (negatif, nötr ve pozitif) inmiş oldu. Üç kategorili veri kümesinde her kategoriye düşen ürün sayısı Şekil 3.2’de verilmiştir.



Şekil 3.2. Üç Kategorili Puan Sayıları.

Üçüncü aşamada verinin “reviews.text” sütununa NLTK (Natural Language Toolkit) [35] kütüphanesi ile şu ön işlemler uygulanmıştır:

- Noktalama işaretleri ve alfa-sayısal olmayan karakterlerin kaldırılması
- Metnin kelime olarak token'lere ayrılması
- Durak kelimelerinin kaldırılması
- Kelime köklerinin bulunması (stemming)
- Kök haline getirilen kelimelerin tekrar birleştirilerek normalleştirilmesi

### 3.4. Yöntemler

#### 3.4.1. TF-IDF Özellik Çıkarımı

TF-IDF (Term Frequency-Inverse Document Frequency) doğal dil işleme ve bilgi çıkarımında kullanılan bir özellik çıkarım yöntemidir. Bu yöntem, belgelerdeki kelimelerin önem derecelerini belirlemek için kullanılır [36]. Bu algoritma, bir belgedeki bir kelimenin önemini hesaplamak için iki bileşen kullanır: terim sıklığı (TF) ve ters belge frekansı (IDF).

Terim Sıklığı (TF), bir terimin bir metinde kaç kez geçtiğiyle ilgilidir. Bir kelimenin daha sık kullanılması, o kelimenin belgedeki öneminin daha yüksek olduğunu gösterir.

$$TF(t, d) = \frac{\text{Terim Frekansı}(t, d)}{\text{Toplam Kelime Sayısı}(d)}$$

(3.1)

Burada  $t$ : terimleri (kelimeleri),  $d$ : dokümanları temsil eder.

Ters Belge Sıklığı (IDF) ise, hangi terimin belgeler arası ayırt edici özelliğinin daha önemli olduğunu belirlemeye yönelik bir tekniktir. TF hesaplanırken analiz edilen tüm terimlerin anlamlılıkları eşit kabul edilir. Söz konusu kelimelerin durdurma kelimeleri mi yoksa önemsiz bağlaçlar mı olduğu önemli değildir. Oysa bir kelime derlemde ne kadar çok belgede geçiyorsa o kelimenin IDF değeri o kadar küçük, ne kadar az belgede geçiyorsa IDF değeri o kadar büyük olur.

$$IDF(t, D) = \log \left( \frac{Doküman Sayısı(D)}{1 + Terimin Geçtiği Doküman Sayısı(t, D)} \right) \quad (3.2)$$

Burada  $D$ : derlemdeki tüm dokümanları temsil eder.

TF ile IDF'nin birleşiminden oluşan TF-IDF bir kelimenin belgeler arası ayırt edici özelliği ile her bir belge için önemini birlikte değerlendirmiş olur [37, 38]. TF-IDF, TF ve IDF değerlerinin çarpımı ile hesaplanır.

#### 3.4.2. SVM (Destek Vektör Makinesi) Sınıflandırıcısı

Destek Vektör Makineleri hem sınıflandırma hem de regresyon analizinde kullanılan güçlü ve popüler bir makine öğrenimi algoritmasıdır. Veri noktalarını farklı sınıflara ayıran en iyi hiper düzlemi bulmayı amaçlar. İki boyutlu bir uzayda bu bir doğru olurken, daha yüksek boyutlarda hiper düzlem olur. En iyi hiper düzlem, sınıflar arasındaki marjin adı verilen boşluğu maksimize eder. Söz konusu marjin, farklı sınıflara ait veri noktalarına en yakın mesafedir. Destek vektörleri, hiper-düzleme en yakın veri noktalarıdır ve hiper-düzlemin yönünü belirlemede önemli rol oynarlar. Algoritma bu noktaları kullanılır ve karar sınırlarını belirler.

Tüm veri kümeleri lineer olarak ayrılabilir olmayabilir. SVM, çekirdek (kernel) fonksiyonları kullanarak verileri daha yüksek boyutlu bir uzaya dönüştürür ve bu uzayda lineer olarak ayrılabilir hale getirir. Yaygın çekirdek fonksiyonları arasında lineer çekirdek, polinom çekirdek, radyal bazlı fonksiyon (RBF) çekirdek ve sigmoid çekirdek bulunur.

SVM, yüksek boyutlu uzaylarda iyi performans gösterir ve özellikle sınıflar arasındaki ayırım belirgin olduğunda güçlüdür. Aşırı öğrenme (overfitting) sorununu minimize eder ve genellikle iyi genelleme sağlar. Çekirdek fonksiyonları kullanarak lineer olmayan problemleri çözme yeteneği sağlar. Öte yandan büyük veri setleri üzerinde çalışırken eğitim süreci yavaş olabilir. Uygun çekirdek fonksiyonunu ve hiperparametreleri seçmek zor olabilir. SVM'nin kararları, başka bazı algoritmalara göre daha zor yorumlanabilir.

### 3.4.3. MNB (Multinomial Naif Bayes) Sınıflandırıcısı

Multinomial Naif Bayes (Çok Terimli Naif Bayes), özellikle metin sınıflandırma ve belge kategorizasyonu gibi doğal dil işleme görevlerinde yaygın olarak kullanılan olasılık temelli bir sınıflandırma algoritmasıdır. Naif Bayes sınıflandırıcıları, Bayes Teoremi'ne dayanır ve özelliklerin birbirinden bağımsız olduğunu varsayar (bu nedenle “naif” olarak adlandırılır). Multinomial Naive Bayes, belirli bir sınıfa ait bir belgenin olasılığını hesaplamak için özelliklerin (kelimelerin) çok terimli (multinomial) dağılımını kullanır. Bayes Teoremi şu şekilde ifade edilir:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (3.3)$$

Burada  $P(C|X)$ , X özelliği verildiğinde C sınıfının olasılığı;  $P(X|C)$ , C sınıfı verildiğinde X özelliğinin olasılığı;  $P(C)$ , C sınıfının öncül olasılığı;  $P(X)$ , X özelliğinin öncül olasılığıdır.

Multinomial Naive Bayes, metin sınıflandırma ve belge kategorizasyonu için basit ama etkili bir algoritmadır. Metindeki kelimelerin çok terimli dağılımlarını kullanarak sınıflandırma yapar ve özellikle kelime frekanslarının önemli olduğu uygulamalarda başarılı sonuçlar verir.

### 3.4.4. MLP (Çok Katmanlı Algılayıcı) Sınıflandırıcısı

Çok Katmanlı Algılayıcı, yapay sinir ağlarının temel yapı taşlarından biri olan ve özellikle sınıflandırma ve regresyon problemlerinde kullanılan bir tür derin öğrenme modelidir. Giriş verilerinin alındığı katmanda her nöron, giriş veri kümesindeki bir özelliği temsil eder. Bir veya daha fazla gizli katman bulunabilir. Her gizli katmandaki nöronlar, bir önceki katmandan gelen verileri alır, işleyip bir sonraki katmana iletir. Bu katmanlar, ağın öğrenme kapasitesini ve doğruluğunu artırır. Çıkış katmanı sınıflandırma problemlerinde her bir sınıfı temsil eden nöronlardan oluşur. Çıkış katmanı, problem türüne bağlı olarak genellikle softmax veya sigmoid fonksiyonu kullanır.

Tüm nöronlar, bir ağırlık ve bir bias terimi ile ilişkili girişleri alır, bunları birleştirir ve bir aktivasyon fonksiyonu kullanarak çıktı üretir. Ağırlıklar ve bias terimleri, eğitim sürecinde verilerden öğrenilir ve modelin performansını optimize eder. Aktivasyon fonksiyonları, nöronların çıktısını belirler ve modeli doğrusal olmayan problemlerde etkin kılar. Yaygın olarak kullanılan aktivasyon fonksiyonları arasında sigmoid, tanh (hiperbolik tanjant) ve ReLU (Rectified Linear Unit) bulunur.

MLP, ileri besleme (feedforward) ve geri yayılım (backpropagation) algoritmaları kullanılarak eğitilir. İleri besleme esnasında giriş verileri, ağırlık katmanlarından geçirilerek tahmin edilen çıktı üretilir. Geri yayılımda ise tahmin edilen çıktı ile gerçek çıktı arasındaki hata hesaplanır ve bu hata, ağırlıkları ve bias terimlerini güncellemek için kullanılır. Bu süreç, hata fonksiyonunu minimize etmek için gradyan inişi gibi optimizasyon algoritmalarıyla yürütülür.

MLP, geniş bir veri türü ve problem tipiyle çalışabilir. Karmaşık ilişkileri ve desenleri öğrenme yeteneğine sahiptir. Doğru yapılandırıldığında ve eğitildiğinde iyi genelleme performansı gösterir. Öte yandan büyük veri setlerinde ve derin ağ yapılarında eğitim süreci uzun olabilir. Eğitim verisine aşırı uyum sağlayabilir, bu nedenle düzenleme (regularization) teknikleri kullanılması gerekebilir. Katman sayısı, nöron sayısı, öğrenme hızı gibi hiper-parametrelerin ayarlanması zor olabilir.

#### 3.4.5. LR (Lojistik Regresyon) Sınıflandırıcısı

LR, istatistik ve makine öğrenmesinde yaygın olarak kullanılan bir sınıflandırma algoritmasıdır. Özellikle ikili sınıflandırma (binary classification) problemlerinde kullanılır, ancak çok sınıflı (multiclass) problemler için de genelleştirilebilir. Bağımsız değişkenler ile bağımlı değişken (sınıf etiketi) arasındaki ilişkiyi modeller. Belirli bir giriş verisinin belirli bir sınıfa ait olma olasılığını tahmin eder. Bu olasılık, 0 ile 1 arasında bir değer alır. Olasılık tahmini sırasında sigmoid (lojistik) fonksiyonunu kullanır. Sigmoid fonksiyonu, her gerçek değeri (örneğin, lineer regresyonun çıktısı) 0 ile 1 arasında bir değere dönüştürür. Sigmoid fonksiyonu bu şekilde tanımlanır:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.4)$$

Burada  $z$ , bağımsız değişkenlerin lineer kombinasyonudur:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (3.5)$$

Tahmin edilen olasılık, belirli bir eşik değeri (genellikle 0.5) ile karşılaştırılarak sınıf etiketi belirlenir. Örneğin, olasılık 0.5'ten büyükse, veri noktası pozitif sınıfa, aksi takdirde negatif sınıfa atanır.

LR, basit ve hızlı olma, yüksek bir yorumlanabilirliğe sahip olma, az sayıda parametreyle çalışma ve aşırı öğrenmeye yatkın olmama gibi avantajlara sahiptir. Ancak doğrusal ayrılabilirlik varsayımına dayandığı için doğrusal olmayan ilişkileri iyi yakalayamayabilir. İki'den fazla sınıflarda multinomial lojistik regresyon kullanılır. Bu

yöntem tıbbi teşhis, finansal risk analizi, pazarlama gibi birçok alanda sınıflandırma problemlerinin çözümünde yaygın olarak kullanılır.

#### 3.4.6. SGD (Stokastik Gradyan İnişi) Sınıflandırıcısı

Stokastik Gradyan İnişi sınıflandırıcısı, büyük veri kümeleriyle çalışan makine öğrenimi algoritmalarını optimize etmek için kullanılan bir öğrenme modelidir. Özellikle doğrusal modellerde ve derin öğrenme modellerinin eğitiminde yaygın olarak kullanılır. Temel prensibi olan gradyan inişi, bir kayıp (hata) fonksiyonunu minimize etmek için kullanılan bir optimizasyon algoritmasıdır. Bu algoritma, kayıp fonksiyonunun gradyanını hesaplayarak modelin parametrelerini (ağırlıklar ve bias) günceller.

Stokastik Gradyan İnişi, klasik gradyan inişinin bir varyantıdır. Klasik gradyan inişinde, model parametreleri tüm veri kümesine göre güncellenir, bu da büyük veri setlerinde hesaplama açısından maliyetli olur. SGD, her adımda yalnızca bir veya birkaç veri örneği kullanarak parametreleri günceller. Böylece hesaplama maliyetini düşürür ve daha hızlı güncellemeler sağlar.

Öğrenme hızı, model parametrelerinin ne kadar güncelleneceğini belirleyen bir hiper-parametredir. Öğrenme hızı dikkatle ayarlanmalıdır; çok büyük bir öğrenme hızı, modelin doğru şekilde öğrenmesini engellerken, çok küçük bir öğrenme hızı ise öğrenme sürecini yavaşlatır. SGD özellikle doğrusal regresyon, lojistik regresyon ve destek vektör makineleri gibi doğrusal modellerde etkilidir. L1 (Lasso) ve L2 (Ridge) düzenleme gibi düzenleme teknikleriyle birleştirilebilir.

Büyük veri kümeleriyle çalışırken hızlı ve verimli güncellemeler sağlanması, yeni veriler geldikçe modelin gerçek zamanlı olarak güncellenmesine olanak tanınması, uygulama ve anlama açısından kolay bir algoritma olması avantajlarıdır. Ancak her adımda yalnızca bir veri örneği kullanılması nedeniyle optimizasyon süreci gürültülü ve dalgalı olabilir. Öğrenme hızı gibi hiper-parametrelerin doğru ayarlanması zor olabilir. Dalgalanmalar nedeniyle modelin global minimum yerine yerel minimumda takılıp kalma olasılığı vardır.

#### 3.4.7. ERT (Ekstra Rastgele Ağaçlar) Sınıflandırıcısı

ERT, topluluk öğrenme yöntemlerinden biri olan ve Rastgele Ormanlara (Random Forests) benzer şekilde çalışan bir makine öğrenimi algoritmasıdır. Daha fazla rastgelelik ekleyerek modelin çeşitliliğini artırmayı amaçlar. Birden fazla karar ağacından oluşan bir topluluk modelidir. Her bir ağaç, bağımsız olarak eğitilir ve sınıflandırma veya regresyon

problemleri için tahminlerde bulunur. Topluluk modeli, bireysel ağaçların tahminlerini birleştirerek daha doğru ve genelleştirilebilir sonuçlar elde eder.

ERT’de her ağaç için eğitim veri kümesinden bir rastgele alt küme seçilir ve ağaç bu alt küme üzerinde eğitilir. Ek olarak, her bir düğümde bölünme noktaları da rastgele seçilir ve en iyi bölünme noktası, rastgele seçilen adaylar arasından belirlenir. Bu, her ağacın daha fazla çeşitliliğe sahip olmasını, modelin aşırı uyumluluğa karşı dirençli, daha hızlı ve hesaplama açısından daha verimli olmasını sağlar.

Sınıflandırma problemlerinde, her bir ağaç bir sınıf tahmini yapar ve en çok oy alan sınıf (majority voting) nihai tahmin olarak seçilir. Regresyon problemlerinde, her bir ağacın tahminlerinin ortalaması alınarak nihai tahmin yapılır.

Avantajları yanları, daha fazla rastgelelik ekleyerek modelin genelleme yeteneğini artırmak ve aşırı uyumu önlemek, rastgele bölünme noktaları seçimi sayesinde daha hızlı ve hesaplama açısından daha verimli olması, bireysel ağaçlar arasındaki çeşitliliği artırması sayesinde daha güçlü olmasıdır. Diğer yandan, ekstra rastgelelik, bazı durumlarda modelin performansını düşürebilir ve belirli veri kümelerinde daha az doğru sonuçlar verebilir. Doğru hiper-parametrelerin belirlenmesi zor olabilir ve modelin performansını optimize etmek için dikkatli ayar gerektirir.

#### 3.4.8. XGBoost (Aşırı Gradyan Artırma) Sınıflandırıcısı

XGBoost, makine öğrenimi modellerinde yüksek performans ve hız sağlayan güçlü bir gradyan artırma algoritmasıdır. Bu algoritma, özellikle yapılandırılmış (tabular) veri kümelerinde ve endüstri uygulamalarında yaygın olarak kullanılır. Gradyan artırma, zayıf öğrencilerin (genellikle karar ağaçları) ardışık olarak eklenmesiyle çalışan bir topluluk (ensemble) öğrenme yöntemidir. Her bir yeni ağaç, önceki ağaçların hatalarını düzelterek modeli iyileştirir. Model, her adımda hata fonksiyonunun negatif gradyanını minimize ederek öğrenir. Böylece modelin hataları azalır, doğruluğu artar.

XGBoost, genellikle karar ağaçları kullanır. Her bir ağaç, veri setindeki özelliklere göre veri noktalarını bölerek kararlar verir. Ağaç derinliği, yaprak sayısı gibi hiper-parametreler modelin kapasitesini ve karmaşıklığını kontrol eder. Aşırı uyumu önlemek için L1 (Lasso) ve L2 (Ridge) düzenleme tekniklerini kullanır. Bu, modelin genelleme yeteneğini artırır ve daha kararlı tahminler yapmasını sağlar. Bu düzenleme terimleri, modelin karmaşıklığını kontrol ederek ağaçların aşırı derinleşmesini ve öğrenilmemiş verilerde düşük performans göstermesini önler.

Paralel hesaplama ve veri parçalama teknikleri ile hesaplama hızını artırır. Bu, büyük veri kümeleriyle çalışırken önemli bir avantaj sağlar. Dahili olarak kullanılan optimizasyon

teknikleri, bellek kullanımını azaltır ve hesaplama verimliliğini artırır. Eksik verileri işleyebilme yeteneğine sahiptir. Böylece eksik değerlerle başa çıkabilir ve doğru tahminler yapabilir. Veri örneklerine ağırlıklar atanarak, belirli örneklerin modelde daha önemli hale gelmesi sağlanabilir. Bu, dengesiz veri setlerinde performansı artırır. Sahip olduğu geniş hiper-parametre seti modelin esnekliğini artırır. Ancak modelin performansını optimize etmek için çok sayıda hiper-parametrenin dikkatlice ayarlanması gerekir. Büyük veri setlerinde hesaplama ve bellek açısından yoğun kaynaklar gerektirebilir.

#### 3.4.9. SpaCy NLP Aracı

SpaCy, Python dilinde yazılmış doğal dil işleme (NLP) görevleri için açık kaynaklı bir kütüphanedir. Metni anlamayı ve işlemeyi kolaylaştıran şu araç ve işlevleri sunar:

- Tokenizasyon: Metni kelimeler, fiiller ve diğer anlamlı birimlere (tokenler) ayırır.
- Morfolojik analiz: Kelimelerin temel formlarını ve dilbilgisi öğelerini (sıfat, fiil, zamir vb.) belirler (Part-of-Speech Etiketleme).
- Söz dizimsel analiz: Kelimelerin cümledeki sırasını ve işlevlerini analiz eder.
- Semantik analiz: Kelimelerin anlamlarını ve birbirleriyle olan ilişkilerini (chunking) belirler.
- Adlandırılmış varlık tanımı: Metindeki kişi isimleri, yer adları, kuruluş isimleri gibi adlandırılmış varlıkları tanıır (Named Entity Recognition - NER).
- Metin benzerliği: İki metnin ne kadar benzer olduğunu hesaplar.
- Makine çevirisi: Metni bir dilden diğerine çevirir.
- Kullanıcıların özel durumları işlemek için terim listeleri ve özel kural tabanlı belirteçler eklemesine olanak tanıır.
- İşlevselliği modüler bileşenler halinde sunar. Bu, kullanıcıların ihtiyaçlarına uyacak şekilde farklı bileşenleri bir araya getirerek özelleştirilmiş NLP işlemleri oluşturmasını sağlar.
- Açık kaynak kodlu bir projedir ve güçlü bir topluluk desteğine sahiptir. Bu, sürekli olarak geliştirilmesini ve iyileştirilmesini sağlar.

SpaCy, NLP'nin çeşitli alanlarında kullanılacak güçlü ve çok yönlü bir araçtır. Hız ve verimlilik odaklı tasarlanmıştır. Performansı yüksek olup büyük veri kümelerinde bile hızlı çalışabilir. Önceden eğitilmiş dil modelleri sunar. Bu modeller, geniş bir dil yelpazesini kapsar ve farklı görevler için kullanılabilir.

Metnin duygusal tonunu belirlemek (duygu analizi), metnin hangi konuya ait olduğunu belirlemek (konu sınıflandırması), metnin ana fikrini ve içerdiği bilgileri özetlemek (özetleme), metne dayalı soruları cevaplamak (soru cevaplama), Makaleler, raporlar, diyaloglar vb. gibi metinler oluşturmak (metin üretme) gibi görevleri yerine getirebilir.

#### 3.4.10. VADER NLP Aracı

VADER (Valence Aware Dictionary and sEntiment Reasoner), doğal dil işleme alanında duygu analizi yapmak için kullanılan bir kütüphanedir. Büyük bir duygu sözlüğüne dayanır. Bu sözlük, binlerce kelime ve kelime öbeği içerir ve hangisinin olumlu, olumsuz veya nötr bir duyguya sahip olduğunu belirtir.

Metinlerde duygusal ifadeleri tespit etmek için kurallı bir yaklaşım kullanır. Kelimelerin ve ifadelerin pozitif, negatif veya nötr duygusal eğilimlerini değerlendirir. Duygusal içeriği analiz ederek her bir metnin pozitif, negatif ve nötr duygusal bileşenlerini sınıflandırır. Bu analiz sonucunda metnin toplam duygu skoru elde edilir. Hızlı ve etkin bir şekilde çalışır. Büyük metin verileri üzerinde duygu analizi yapabilir ve sonuçları hızla döndürebilir.

Dilbilimsel inceliklere ve sözcüklerin kontekstine dikkat eder. Örneğin, “not bad” ifadesini olumlu olarak yorumlar çünkü “not” kelimesi genellikle olumsuz bir ifade ile birlikte kullanılırken “bad” kelimesi de olumsuz bir anlam taşıdığından “not bad” olumlu bir duygu taşır. Açık kaynak kodlu bir projedir ve kullanımı kolaydır. Python üzerinde bir kütüphane olarak sunulur ve diğer Python projelerine kolayca entegre edilebilir.

Sosyal medya analizi, ürün incelemeleri, duygu analizi, haber analizi ve daha pek çok alanda kullanılabilir. Özellikle sosyal medya verileri üzerinde hızlı ve etkili bir şekilde duygu analizi yapmak için popülerdir.

#### 3.4.11. RoBERTa NPL Aracı

RoBERTa (Robustly optimized BERT approach), doğal dil işleme alanında kullanılan başka bir güçlü dil modelidir. RoBERTa, BERT (Bidirectional Encoder Representations from Transformers) modelinin geliştirilmiş bir versiyonudur ve özellikle büyük ölçekli dil modeli eğitimi için optimize edilmiştir.

Transformer mimarisine dayanır. Bu model, özellikle uzun metinlerdeki dil yapısını daha iyi anlamak ve temsil etmek için etkili bir şekilde çalışır. BERT modelinin temel fikirlerini devralır ancak eğitim süreci ve modelin mimarisi üzerinde bir dizi iyileştirme yapar. Bu, daha iyi performans ve daha tutarlı sonuçlar sağlar.

Büyük ölçekli metin verileri üzerinde eğitilmiştir. Bu, genel dil yapısını ve kelime ilişkilerini daha iyi kavramasına ve genelleme yeteneğinin artmasına olanak tanır. Önceden eğitilmiş bir dil modelidir ve geniş bir dil anlayışına sahiptir. Bu model, özellikle metin sınıflandırması, duygu analizi, metin üretimi ve daha pek çok NLP görevi için kullanılabilir. Farklı dillerdeki metinleri işlemek için çoklu dil desteği sunar. Bu, çok dilli NLP uygulamaları için idealdir. Pre-trained modelleri, transfer öğrenme prensiplerine dayanır. Bu, önceden eğitilmiş bir modeli, belirli bir görev için hızlı ve etkili bir şekilde ince ayar (fine-tuning) yapmak için kullanmanıza olanak tanır. Açık kaynak kodlu bir projedir ve güçlü bir araştırma topluluğuna sahiptir.



## 4. BULGULAR ve TARTIŞMA

### 4.1. Performans Metrikleri

Makine öğrenmesi ve veri madenciliği alanında, model performansını değerlendirmek için çeşitli metrikler kullanılır. Bu metrikler, modelin doğruluğunu, kesinliğini, duyarlılığını, F1 skorunu ve daha fazlasını ölçer.

Doğruluk (accuracy), modelin doğru sınıflandırdığı örneklerin toplam örnek sayısına oranıdır:

$$\text{Doğruluk} = \frac{\text{Doğru Tahminler}}{\text{Tüm Tahminler}} \quad (4.1)$$

Kesinlik (precision), pozitif tahminlerin ne kadar doğru olduğunu ölçer.

$$\text{Kesinlik} = \frac{\text{Doğru Pozitifler}}{\text{Doğru Pozitifler} + \text{Yanlış Pozitifler}} \quad (4.2)$$

Duyarlılık (recall), gerçek pozitif örneklerin ne kadarını doğru tahmin ettiğinizi ölçer.

$$\text{Duyarlılık} = \frac{\text{Doğru Pozitifler}}{\text{Doğru Pozitifler} + \text{Yanlış Negatifler}} \quad (4.3)$$

F1 skoru, hassasiyet ve duyarlılığın harmonik ortalamasıdır. Bu metrik, özellikle dengesiz sınıflar için faydalıdır.

$$\text{F1 Skoru} = 2 \times \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4.4)$$

Metriklerin ağırlıklı ortalaması, her sınıfın metrik skoruna, o sınıfın örnek sayısına göre bir ağırlık verilmesiyle hesaplanır. Bu yöntem, sınıf dengesizliği olan veri setlerinde daha doğru bir performans değerlendirmesi sağlar, zira sık görülen sınıfların ağırlığı daha fazladır. Ağırlıklı ortalama kesinlik, duyarlılık ve F1 skoru şu şekilde hesaplanır:

$$\text{Ağırlıklı Ortalama Kesinlik} = \frac{\sum_{i=1}^N (\text{Kesinlik}(i) \times \text{Örnek Sayısı}(i))}{\sum_{i=1}^N \text{Örnek Sayısı}(i)} \quad (4.5)$$

$$\text{Ağırlıklı Ortalama Duyarlılık} = \frac{\sum_{i=1}^N (\text{Duyarlılık}(i) \times \text{Örnek Sayısı}(i))}{\sum_{i=1}^N \text{Örnek Sayısı}(i)} \quad (4.6)$$

$$\text{Ağırlıklı Ortalama F1 Skoru} = \frac{\sum_{i=1}^N (\text{F1 Skoru}(i) \times \text{Örnek Sayısı}(i))}{\sum_{i=1}^N \text{Örnek Sayısı}(i)} \quad (4.7)$$

Burada  $i$ : veri kümesindeki her bir sınıfı temsil eder.

Bu metriklerin değerlendirilmesi modelin genel performansını anlamak için önemlidir. Yüksek doğruluk, kesinlik, duyarlılık ve F1 puanları, modelin etkili bir şekilde çalıştığını ve Amazon müşteri yorumlarındaki duygu durumlarını doğru bir şekilde sınıflandırdığını gösterir.

Tezde çalışılan sınıflandırıcı ve NLP araçlarının doğruluk, kesinlik, duyarlılık ve F1 puan metrikleri Tablo 1’de verilmiş. Bu metrikler, modelin performansını değerlendirmek için kullanılır. Elde edilen sonuçlara göre SVM modeli, yüksek bir doğruluk ve duyarlılığa (0.93) sahiptir. Kesinlik (0.92) ve F1 puanı da (0.91) oldukça yüksektir. Bu durum modelin hem doğru pozitifleri hem de doğru negatifleri iyi ayırt edebildiğini gösterir.

MNB modeli, doğruluk (0.90) açısından biraz daha düşük performans göstermiştir. Kesinlik (0.86) ve F1 puanı (0.85), diğer modellerle kıyaslandığında daha düşüktür ve bu sonuç modelin özellikle yanlış pozitif ve yanlış negatiflerden kaçınmakta zorlandığı anlamına gelir. MLP modeli, tüm metriklerde yüksek performans (Doğruluk, Duyarlılık, F1 puanı 0.95, Kesinlik 0.94) göstermiştir. Bu model hem doğru pozitifleri hem de doğru negatifleri çok iyi tahmin etmektedir.

LR modeli, genel olarak yüksek bir doğruluk ve duyarlılığa (0.92) sahiptir. Kesinlik ve F1 puanı da (0.90) oldukça iyi olduğundan bu modelin dengeli bir performans sergilediği söylenebilir. SGD modeli, doğruluk, kesinlik ve duyarlılık (0.92) açısından güçlü bir performans göstermektedir. Ancak F1 puanı (0.89), diğer metriklere göre biraz daha düşük olduğundan bazı yanlış pozitif veya negatif tahminlerin olduğu öngörülebilir.

ERT modeli, yüksek doğruluk, kesinlik, duyarlılık (0.95) ve F1 puanı (0.94) ile bu çalışmada en iyi performans gösteren modellerden biri olmuştur. Bu, modelin büyük oranda doğru tahminler yaptığı anlamına gelir. XGBoost modeli, yüksek performans metrikleri ile (doğruluk, kesinlik, duyarlılık 0.93, F1 puanı 0.92) güçlü bir modeldir. Tüm metriklerde dengeli ve yüksek değerler sunmaktadır.

SpaCy modeli, yüksek doğruluk ve duyarlılığa (0.93) sahiptir. Kesinlik ve F1 puanı da oldukça yüksektir (0.91), bu da modelin yukarıdaki yöntemlere eşdeğer performans gösterdiğini işaret eder. VADER modeli, diğer modellerle kıyaslandığında daha düşük doğruluk ve duyarlılık (0.83) sergilemektedir. Ancak kesinlik değeri (0.88) görece yüksektir, bu da modelin doğru pozitifleri iyi tahmin edebildiğini ancak bazı yanlış negatifler yaptığı anlamına gelir. RoBERTa'nın sergilediği yüksek kesinlik (0.91), doğruluk, duyarlılık ve F1 puanı (0.90), modelin genel olarak iyi performans gösterdiğini ve hataları minimize ettiğini göstermektedir.

**Tablo 4.1.** Yöntemler ve Elde Edilen Sonuçlar

| Sınıflandırıcı       | Doğruluk | Kesinlik <sup>1</sup> | Duyarlılık <sup>1</sup> | F1 Puanı <sup>1</sup> |
|----------------------|----------|-----------------------|-------------------------|-----------------------|
| SVM                  | 0.93     | 0.92                  | 0.93                    | 0.91                  |
| MNB                  | 0.90     | 0.86                  | 0.90                    | 0.85                  |
| MLP                  | 0.95     | 0.94                  | 0.95                    | 0.95                  |
| LR                   | 0.92     | 0.90                  | 0.92                    | 0.90                  |
| SGD                  | 0.92     | 0.92                  | 0.92                    | 0.89                  |
| ERT                  | 0.95     | 0.95                  | 0.95                    | 0.94                  |
| XGBoost              | 0.93     | 0.93                  | 0.93                    | 0.92                  |
| SpaCy                | 0.93     | 0.91                  | 0.93                    | 0.91                  |
| VANDER <sup>2</sup>  | 0.83     | 0.88                  | 0.83                    | 0.85                  |
| RoBERTa <sup>2</sup> | 0.90     | 0.91                  | 0.90                    | 0.90                  |

<sup>1</sup>Ağırlıklı ortalamalar alınmıştır; <sup>2</sup>Veri kümesine ön işlem uygulanmamıştır

Sonuç olarak en iyi performans gösteren modeller MLP ve ERT, tüm metriklerde en yüksek değerlere ulaşmıştır. Bu değerler aynı veri kümesi üzerinde yakın zaman önce yapılmış benzer bir çalışmada [28] elde edilen en iyi değerlerden daha yüksektir. Bu iki modelin karmaşıklığı ve esnekliği, veri setindeki karmaşık ilişkileri iyi öğrenmelerini sağlamıştır. Orta performans gösteren SVM, XGBoost, LR, SpaCy ve RoBERTa, yüksek doğruluk ve dengeli sonuçlar üretmiştir. Bu modeller, genellikle yüksek doğruluk ve duyarlılık sağlar. Düşük performans gösteren MNB ve VANDER, diğer modellere kıyasla daha düşük değerler sergilemiştir. Bu durum, söz konusu iki modelin veri setindeki karmaşık ilişkileri yeterince iyi öğrenememesinden kaynaklanmaktadır.

Her modelin performansı, veri setinin özelliklerine ve modelin karmaşıklığına bağlı olarak değişir. Bu yüzden, hangi modelin en iyi performansı gösterdiğini belirlemek için farklı modelleri denemek ve değerlendirmek önemlidir.

Modellerin performans farklarının altında yatan muhtemel nedenler şu şekilde sıralanabilir:

**Modelin Doğası ve Karmaşıklığı:** MLP (Çok Katmanlı Algılayıcı) ve ERT (Ekstrem Rastgele Ağaçlar) gibi modeller, daha karmaşık veri yapıları ve ilişkileri öğrenme yeteneklerine sahiptir. Bu, onların daha yüksek doğruluk ve duyarlılık elde etmelerini sağlar. MNB (Multinomial Naif Bayes) gibi daha basit modeller, veri setindeki karmaşık ilişkileri yeterince yakalayamayabilir ve bu nedenle performansları düşük olabilir.

**Veri Setinin Özellikleri:** Bazı modeller, belirli veri setlerinde daha iyi performans gösterebilir. Örneğin, Naif Bayes modelleri bağımsızlık varsayımı yapar, bu nedenle özellikler arasında bağımlılık varsa performans düşebilir. XGBoost ve ERT gibi modeller, veri setindeki özellikler arasında karmaşık etkileşimleri daha iyi öğrenebilir.

**Hiperparametre Ayarları:** Modellerin hiperparametreleri doğru şekilde ayarlandığında performansları önemli ölçüde artabilir. Örneğin, SVM ve MLP gibi modellerin doğru hiperparametre ayarları, daha yüksek performans sağlar. Hiperparametre optimizasyonu yapılmamış veya yetersiz yapılmış modeller, potansiyel performanslarını gösteremeyebilir.

**Veri Ön İşleme ve Özellik Mühendisliği:** Verilerin düzgün bir şekilde ön işlenmesi ve uygun özelliklerin seçilmesi, model performansını etkileyebilir. Örneğin, SpaCy gibi doğal dil işleme araçları, metin verilerinin doğru şekilde temsil edilmesini sağlar. MNB gibi modeller, özellikle metin sınıflandırma görevlerinde belirli ön işleme adımlarına daha duyarlı olabilir.

**Veri Setindeki Dengesizlikler:** Eğer veri setinde sınıf dengesizlikleri varsa, bu bazı modellerin performansını olumsuz etkileyebilir. ERT ve XGBoost gibi modeller, sınıf dengesizliklerine karşı daha dayanıklı olabilir. MNB ve LR gibi modeller, sınıf dengesizliklerine karşı daha hassas olabilir ve bu durum performanslarını düşürebilir.

**Modelin Öğrenme Yeteneği ve Hızlı Öğrenme Kapasitesi:** MLP ve ERT gibi modeller, büyük veri setlerinde hızlı ve etkili bir şekilde öğrenebilirler. Bu modeller, veri setindeki karmaşıklığı ve varyasyonu daha iyi öğrenebilir. VANDER gibi daha az bilinen veya optimize edilmemiş modeller, öğrenme kapasitesi açısından sınırlı olabilir ve bu da düşük performansa yol açabilir.

**Aşırı Uyum ve Genelleme:** Bazı modeller, eğitim verilerine aşırı uyum sağlayabilir ve bu durum test verilerinde düşük performansa neden olabilir. ERT ve XGBoost gibi güçlü modeller, aşırı uyumdan kaçınmak için daha iyi yöntemlere sahiptir. SVM ve MLP gibi modeller, genelleme kapasitesi açısından daha iyidir ve bu da daha yüksek doğruluk ve F1 puanına katkıda bulunur.

Yukarıda kısaca değinilen nedenler, modellerin performans farklarını açıklamaya yardımcı olur. Her modelin kendi avantajları ve dezavantajları vardır ve performans, kullanılan veri setine ve problemin doğasına göre değişiklik gösterir.

#### 4.2. Karmaşıklık Matrisleri

Karmaşıklık matrisi (confusion matrix) analizi, modelin gerçek ve tahmin edilen duygu durumları arasındaki uygunluğu gösteren önemli bir araçtır. Modelin hangi sınıfları iyi tahmin ettiğini ve hangi sınıflarda zorluk yaşadığını anlamaya yarar. Bu analiz, dört ana kategoride değerlendirilir: Doğru Pozitif (True Positives), Doğru Negatif (True Negatives), Yanlış Pozitif (False Positives) ve Yanlış Negatif (False Negatives). Bu kategorilerin değerleri, modelin performansını daha ayrıntılı bir şekilde açıklar. Karmaşıklık matrisinde, satırlar gerçek sınıfları, sütunlar ise modelin tahmin ettiği sınıfları temsil eder.

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$$

(4.8)

Doğru Pozitif değeri, modelin gerçek etiketi pozitif olan yorumlardan doğru sınıflandırdığı örneklerin sayısını temsil eder. Yani, gerçekte pozitif olan yorumların doğru bir şekilde tahmin edildiği durumları ifade eder. Yüksek bir Doğru Pozitif değeri, modelin olumlu yorumları başarıyla tanıdığını gösterir.

Doğru Negatif değeri, modelin gerçek etiketi negatif olan yorumlardan doğru sınıflandırdığı örneklerin sayısını temsil eder. Yani, gerçekte negatif olan yorumların doğru bir şekilde tahmin edildiği durumları ifade eder. Yüksek bir Doğru Negatif değeri, modelin olumsuz yorumları başarıyla tanıdığını gösterir.

Yanlış Pozitif değeri, modelin gerçek etiketi pozitif olan yorumlardan yanlış sınıflandırdığı örneklerin sayısını temsil eder. Yani, gerçekte pozitif olan yorumların yanlış bir şekilde olumsuz olarak tahmin edildiği durumları ifade eder. Düşük bir Yanlış Pozitif değeri, modelin olumlu yorumları yanlış sınıflandırma eğiliminde olmadığını gösterir.

Yanlış Negatif değeri, modelin gerçek etiketi negatif olan yorumlardan yanlış sınıflandırdığı örneklerin sayısını temsil eder. Yani, gerçekte negatif olan yorumların yanlış bir şekilde olumlu olarak tahmin edildiği durumları ifade eder. Düşük bir Yanlış Negatif değeri, modelin olumsuz yorumları yanlış sınıflandırma eğiliminde olmadığını gösterir.

Tezde test edilen yöntemlerden elde edilen karmaşıklık matrisleri Şekil 3 ve Şekil 4'te verilmiştir. MNB yöntemi karmaşıklık matrisi bu yöntemin negatif ve nötr yorumlar için performansının oldukça düşük olduğunu, bu sınıfları pozitif olarak yanlış tahmin ettiğini, dolayısıyla modelin bu sınıfları ayırt etmekte zorlandığını gösteriyor. Model pozitif sınıfı çok yüksek bir doğrulukla tahmin etmektedir. Bu, modelin veri setinde en fazla örneği bulunan sınıfın bu olduğunu göstermektedir.

SVM yöntemi, negatif sınıfı çoğunlukla doğru tahmin etmiş (187 doğru tahmin), ancak pozitif sınıf olarak önemli bir miktarda yanlış tahminde bulunmuştur (138 yanlış tahmin). Model nötr sınıfta oldukça zorlanmış, doğru tahmin sayısı düşük kalmış (12 doğru tahmin) ve çoğunlukla pozitif sınıf olarak yanlış tahminler yapmıştır (207 yanlış tahmin). Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı oldukça yüksektir (5055 doğru tahmin) ve nadiren yanlış tahminler yapmıştır (30 ve 3 yanlış tahmin).

LR yöntemi, negatif sınıfı doğru tahmin etmede orta derecede başarılıdır (150 doğru tahmin), ancak birçok örneği pozitif sınıf olarak yanlış tahmin etmiştir (167 yanlış tahmin). Model nötr sınıfta oldukça zorlanmıştır. Doğru tahmin sayısı düşüktür (20 doğru tahmin) ve çoğunlukla pozitif sınıf olarak yanlış tahminler yapmıştır (213 yanlış tahmin). Ayrıca bazı örnekleri negatif sınıf olarak tahmin etmiştir (21 yanlış tahmin). Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (5061 doğru tahmin) ve nadiren yanlış tahminler yapmıştır (21 ve 6 yanlış tahmin).

MLP yöntemi, negatif sınıfı doğru tahmin etmede iyi bir başarı göstermiştir (237 doğru tahmin), ancak bazı örnekleri nötr (19 yanlış tahmin) ve pozitif sınıf (69 yanlış tahmin) olarak tahmin etmiştir. Model, nötr sınıfı doğru tahmin etmede oldukça iyidir (131 doğru tahmin), ancak bazı örnekleri negatif (17 yanlış tahmin) ve pozitif sınıf (106 yanlış tahmin) olarak tahmin etmiştir. Model, pozitif sınıfını çok yüksek bir doğrulukla tahmin etmiştir (5010 doğru tahmin), ancak nadiren diğer sınıflar olarak yanlış tahminde bulunmuştur (39 ve 39 yanlış tahmin).

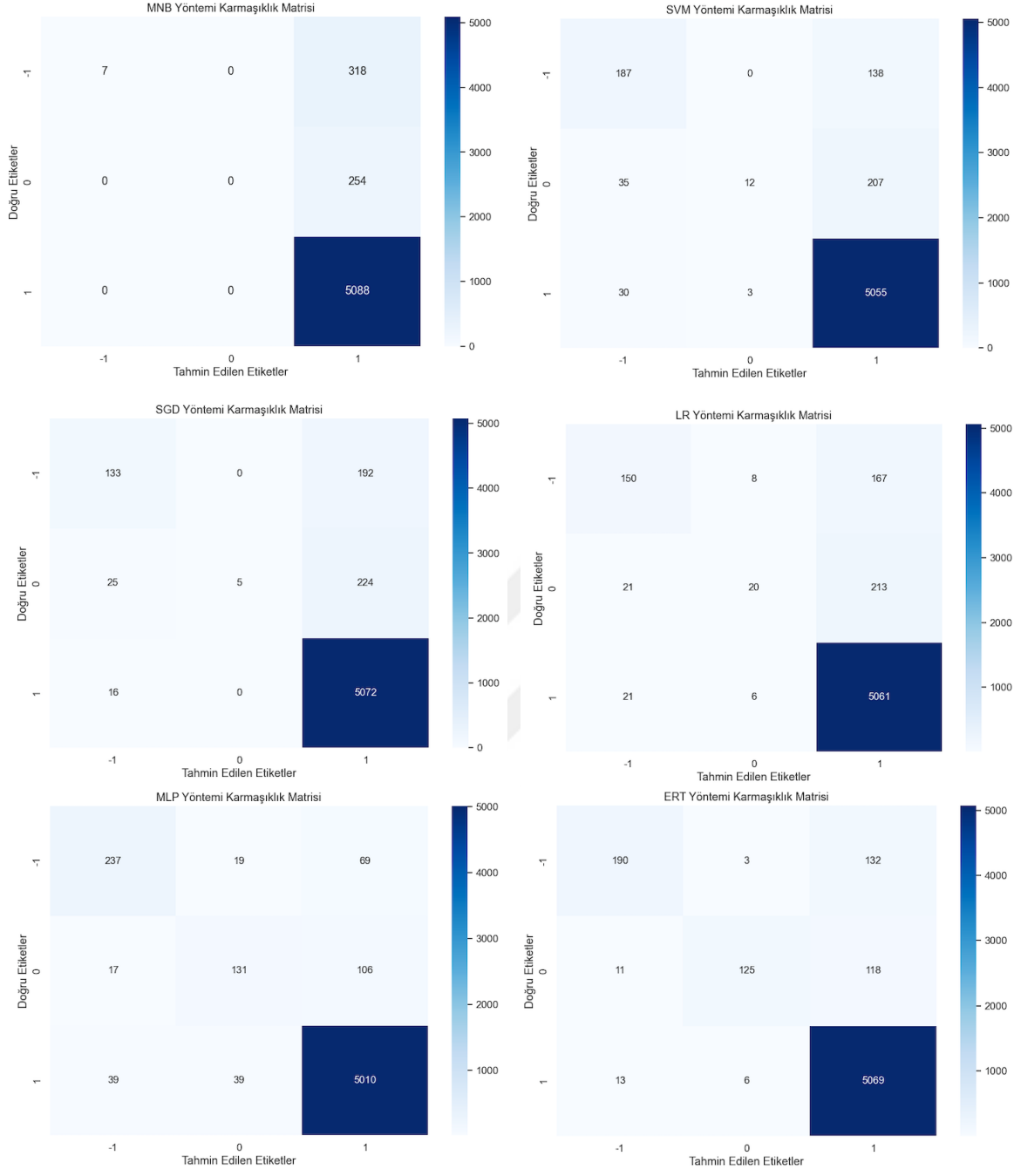
SGD yöntemi, negatif sınıfı doğru tahmin etmede orta derecede başarılıdır (133 doğru tahmin), ancak birçok örneği (192) pozitif olarak yanlış tahmin etmiştir. Model nötr sınıfta oldukça zorlanmıştır. Doğru tahmin sayısı çok düşüktür (5 doğru tahmin) ve çoğunlukla pozitif sınıf olarak yanlış tahminler yapmıştır (224 yanlış tahmin). Ayrıca bazı örnekleri negatif olarak tahmin etmiştir (25 yanlış tahmin). Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (5072 doğru tahmin), nadiren negatif olarak yanlış tahmin yapmış (16 yanlış tahmin) ve hiç nötr sınıf olarak tahminde bulunmamıştır.

ERT yöntemi, negatif sınıfını doğru tahmin etmede orta derecede başarılıdır (190 doğru tahmin), ancak birçok örneği (132) pozitif sınıf olarak yanlış tahmin etmiş. Model, nötr sınıfta orta derecede başarılıdır. Doğru tahmin sayısı makuldür (125 doğru tahmin), ancak birçok örneği (118) pozitif olarak ve bazı örnekleri (11) negatif olarak yanlış tahmin etmiştir. Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (5069 doğru tahmin), nadiren negatif sınıfı olarak (13 yanlış tahmin) ve bazen nötr sınıfı olarak (6 yanlış tahmin) yanlış tahminlerde bulunmuştur.

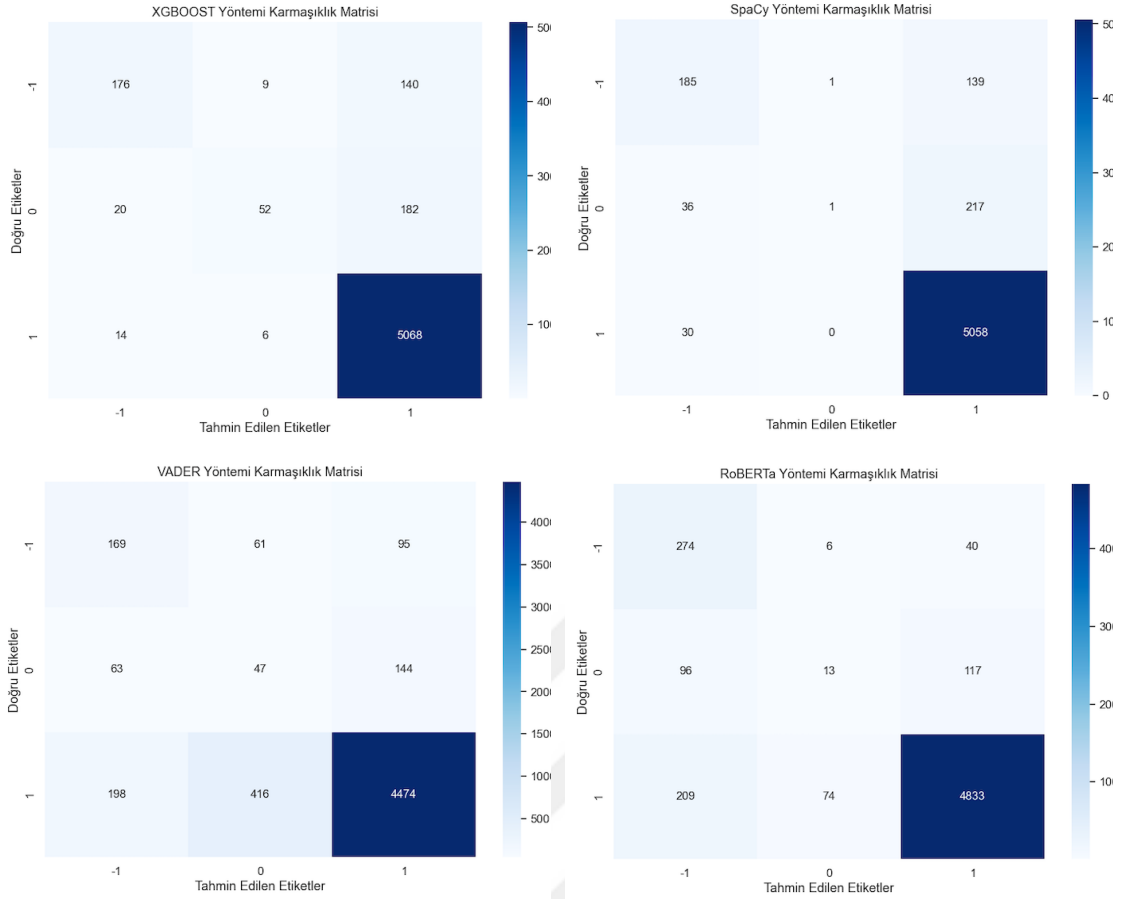
XGBoost yöntemi, negatif sınıfını doğru tahmin etmede orta derecede başarılıdır (176 doğru tahmin), ancak birçok örneği (140) pozitif ve bazı örnekleri (9) nötr olarak yanlış tahmin etmiştir. Model nötr sınıfta orta derecede başarılıdır. Doğru tahmin sayısı makuldür (52 doğru tahmin), ancak birçok örneği (182) pozitif, bazı örnekleri (20) ise negatif olarak yanlış tahmin etmiştir. Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (5068 doğru tahmin), nadiren negatif (14 yanlış tahmin) ve nötr (6 yanlış tahmin) olarak yanlış tahminlerde bulunmuştur.

SpaCy yöntemi, negatif sınıfın doğru tahmin etmede orta derecede başarılıdır (185 doğru tahmin), ancak birçok örneği (139) pozitif ve bazı örnekleri (1) nötr olarak yanlış tahmin etmiştir. Model nötr sınıfta zorlanmıştır. Doğru tahmin sayısı düşüktür (1 doğru tahmin), birçok örneği (217) pozitif, bazı örnekleri ise (36) negatif olarak yanlış tahmin etmiştir. Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (5058 doğru tahmin), nadiren negatif (30 yanlış tahmin) olarak yanlış tahminler yapmıştır.

VADER yöntemi, negatif sınıfı doğru tahmin etmede orta derecede başarılıdır (169 doğru tahmin), ancak birçok örneği (95) pozitif sınıf olarak ve bazı örnekleri (61) nötr sınıf olarak yanlış tahmin etmiştir. Model nötr sınıfta zorlanmıştır. Doğru tahmin sayısı düşüktür (47 doğru tahmin), birçok örneği (144) pozitif sınıf olarak ve bazı örnekleri (63) negatif sınıf olarak yanlış tahmin etmiştir. Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (4474 doğru tahmin), ancak dikkate değer miktarda örneği (198) negatif sınıf ve (416) nötr sınıf olarak yanlış tahmin etmiştir.



**Şekil 4.1.** MNB, SVM, LR, MPL, SDG ve ERT Yöntemleri Karmaşıklık Matrisleri.



**Şekil 2.2.** XGBoost, SpaCy, VADER ve RoBERTa Yöntemleri Karmaşıklık Matrisleri

RoBERTa yöntemi, negatif sınıfı doğru tahmin etmede oldukça başarılıdır (274 doğru tahmin). Ancak bazı örnekleri (40) pozitif sınıf olarak ve nadiren (6) nötr sınıf olarak yanlış tahmin etmiştir. Model nötr sınıfta oldukça zorlanmıştır. Doğru tahmin sayısı düşüktür (13 doğru tahmin). Çoğu örneği (117) pozitif sınıf olarak ve bazı örnekleri (96) negatif sınıfı olarak yanlış tahmin etmiş. Model pozitif sınıfta çok başarılıdır. Doğru tahmin oranı çok yüksektir (4833 doğru tahmin), ancak dikkate değer miktarda örneği (209) negatif sınıf ve (74) nötr sınıf olarak yanlış tahmin etmiştir.

Tezde incelenen tüm yöntemlerin karmaşıklık matrisinde görülen ortak özellik, söz konusu veri kümesinin pozitif sınıf için hepsinin iyi çalışıyor olmasıdır. Bunun nedeni, pozitif sınıf örneklerinin diğer sınıflardan çok daha fazla olması ve böylece modellerin bu sınıfa daha fazla odaklanmasıdır. Bu durum, bazı modellerde diğer sınıfların örneklerinin yanlış tahmin edilmesine yol açabilir. Modellerin düşük performans gösterdiği sınıflar için daha iyi özellikler çıkartmak, hiper-parametre optimizasyonu yapmak, veriyi yeniden gözden geçirmek, temizleme işlemlerini kontrol ederek hatalı

veya eksik verilerin model performansını olumsuz etkilemesini engellemek söz konusu modellerin performansını artırmak adına faydalı olabilir.

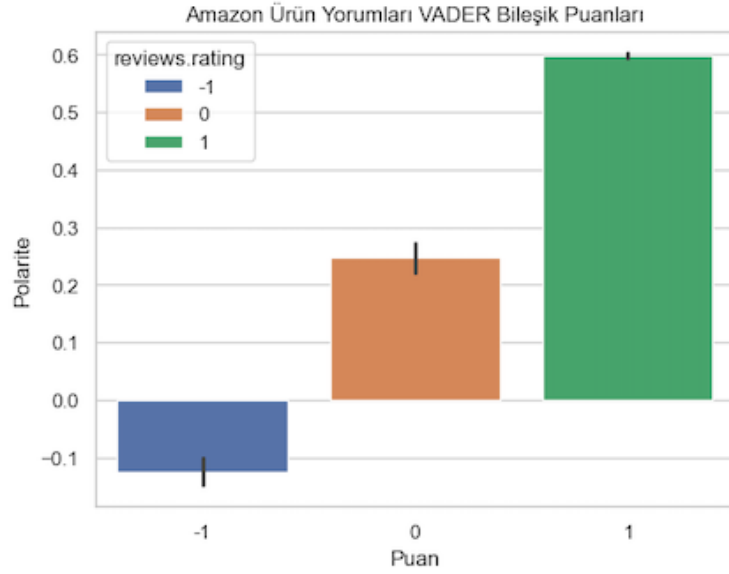
### 4.3. Duygu Polariteleri

VADER duygu analizinde dört ana skor vardır: pozitif, nötr, negatif ve bileşik duygu puanları. Bu puanlar, bir metnin duygusal içeriğini farklı yönlerden değerlendirmek için kullanılır. İlk üç puan, metnin sırasıyla ne kadar olumlu, nötr ve olumsuz duygular içerdiğini gösterir. Model, nötr kelimeleri bir referans noktası olarak kullanır ve bu kelimelerin duygusal yoğunluğunu 0 olarak kabul eder. Bileşik duygu puanı ise metnin genel duygusal içeriğini belirler. VADER'in en önemli çıktılarından biridir. Bileşik duygu skoru, pozitif, negatif ve nötr skorların ağırlıklı ortalaması olarak hesaplanır ve -1 ile +1 arasında bir değer alır. Bileşik duygu skoru, metnin duygusal içeriğini genel olarak özetler. Örneğin, +1 çok pozitif bir duygusal içeriği gösterirken, -1 çok negatif bir içeriği gösterir. 0, nötr bir duygusal içeriği gösterir. Bu skorlar, VADER'in metinlerdeki duygusal tonu ve içeriği belirlemek için nasıl çalıştığını anlamamıza yardımcı olur. Özellikle sosyal medya metinleri gibi kısa ve hızlı değişen metinler için VADER, hızlı ve etkili bir çözümlenme aracı olarak yaygın olarak kullanılır.

Amazon ürün yorumları VADER bileşik puanları Şekil 5'de verilmiştir. Negatif sınıf bileşik puanı (-0.12), tüm veri kümesinin hafifçe negatif bir duygusal içerik taşıdığını gösterir. Güçlü negatif içerik puanı -1'e daha yakın değerler üretir. Nötr sınıf bileşik puanı, veri kümesinin genel olarak hangi ölçüde nötr bir duygusal içeriğe sahip olduğunu gösterir. Bu çalışmada elde edilen 0.25 değeri veri kümesinin yaklaşık  $\frac{1}{4}$ 'ünün nötr ifadelerden oluştuğu anlamına gelir. Pozitif sınıf bileşik puanı (0.6), veri kümesinin belirgin bir şekilde pozitif duygusal içeriği olduğunu gösterir. +1'e yaklaşan değerler, metnin daha güçlü pozitif duygular içerdiğini ifade ederken, 0.6 pozitif duyguların ön planda olduğunu göstermektedir.

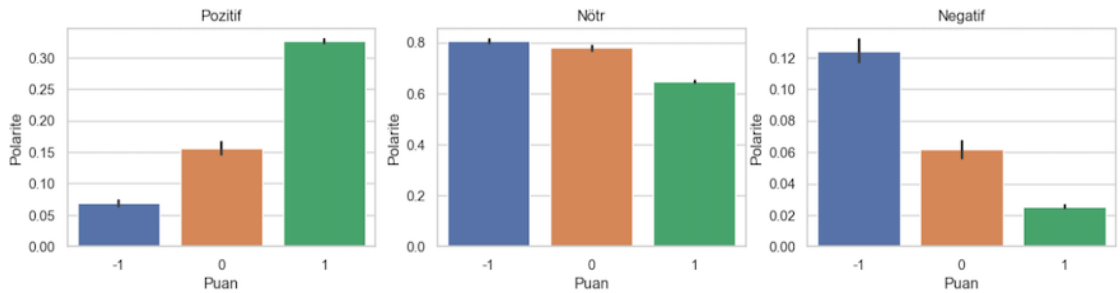
Amazon ürün yorumları VADER ayrık puanları Şekil 6'da verilmiştir. Bu sonuçlar, negatif sınıfın kendi içinde genel olarak hafif bir negatif eğilime sahip olduğunu gösterir. Ayrık negatif sınıf içindeki negatif sınıf puanı (0.12) diğer sınıf puanlarından (0.06 ve 0.025) yüksektir, bu da negatif sınıf içinde negatif duygusal içeriklerin varlığına işaret eder. Ayrık nötr sınıfın iç puanları negatiften pozitive doğru 0.8, 0.75 ve 0.65 olarak sıralanmıştır. Normalde ayrık nötr sınıf içindeki nötr sınıf puanının en yüksek olması beklenir. Elde edilen sonuçlar, 0.8 ve 0.75 puanları arasındaki küçük farkın bileşik puanlara tam olarak aktarılamadığı şeklinde yorumlanabilir. Ayrık pozitif sınıf içindeki puanlar (negatif = 0.06, nötr = 0.15, pozitif = 0.325), pozitif sınıfın genel olarak, hafif de

olsa beklendiği gibi pozitif eğilime sahip olduğunu gösterir, zira pozitif puan diğerlerinden biraz daha yüksektir. Bu değerler, her bir sınıfın duygusal içeriğindeki negatif, nötr veya pozitif unsurların yoğunluğunu belirtir.



**Şekil 4.3.** Amazon Ürün Yorumları VADER Bileşik Puanları.

Dil yapısını daha karmaşık bir şekilde anlayan ve kapsayan RoBERTa, BERT'in geliştirilmiş bir versiyonudur ve doğal dil işleme görevlerinde (duygu analizi dahil) kullanılmak üzere önceden eğitilmiştir. Büyük bir metin verisi üzerinde eğitilmiş olduğundan genel dil modeli olarak geniş bir dil bilgisine sahiptir. Kelime ve cümleleri derin öğrenme tabanlı yöntemlerle temsil eder. Ön eğitiminden sonra, duygu analizi gibi özel görevler için ince ayar yapılabildiğinden daha esnektir ve daha iyi performans gösterir. BERT ve türevlerinin performansı, duygu analizi de dahil olmak üzere, birçok doğal dil işleme görevinde endüstri standardını oluşturur.



**Şekil 4.4.** Amazon Ürün Yorumları VADER Ayrık Puanları.

Şekil 7'de VADER ve RoBERTa modellerinin sınıf puanları arasındaki ilişkiler ve dağılımlar, negatif, nötr ve pozitif sınıflar farklı renklerle temsil edilerek görselleştirilmiştir. Şekil 7'deki diyagonal (köşegen) hücreler her bir değişkenin dağılımını göstermektedir. VADER negatif değişkeni ile RoBERTa negatif ve nötr

değişkenleri küçük ve tek tepe değerleri (unimodal) etrafında yoğunlaşmışken RoBERTa pozitif değişkeni yüksek ve yine tek değer etrafında yoğunlaşmıştır. VADER nötr değişkeni birden fazla tepe değeri (multimodal) etrafında yoğunlaşmış ve sağa çarpıktır. Yine multimodal dağılım gösteren VADER pozitif değişkeni ise sola çarpıktır.

Diyagonal dışındaki hücrelerdeki dağılım grafikleri (scatter plot), iki değişken arasındaki ilişkiyi gösterir. Noktalar düz bir çizgi boyunca yer alıyorsa, iki değişken arasında doğrusal bir ilişki vardır. Pozitif doğrusal ilişki, bir değişken arttıkça diğerinin de arttığı anlamına gelir. Negatif doğrusal ilişki, bir değişken arttıkça diğerinin azaldığı anlamına gelir. Noktalar ne kadar sıkı bir şekilde bir çizgi etrafında toplanmışsa, ilişki o kadar kuvvetlidir. Dağınık noktalar zayıf bir ilişkiyi gösterir. Eğer noktalar bir eğri boyunca yer alıyorsa, iki değişken arasında doğrusal olmayan bir ilişki vardır. Farklı renklerin belirgin bir şekilde ayrıldığı yerler, kategoriler arasında belirgin bir fark olduğunu gösterir. Farklı renklerin birbirine karıştığı yerler, kategoriler arasında fark olmadığını veya çok az fark olduğunu gösterir.

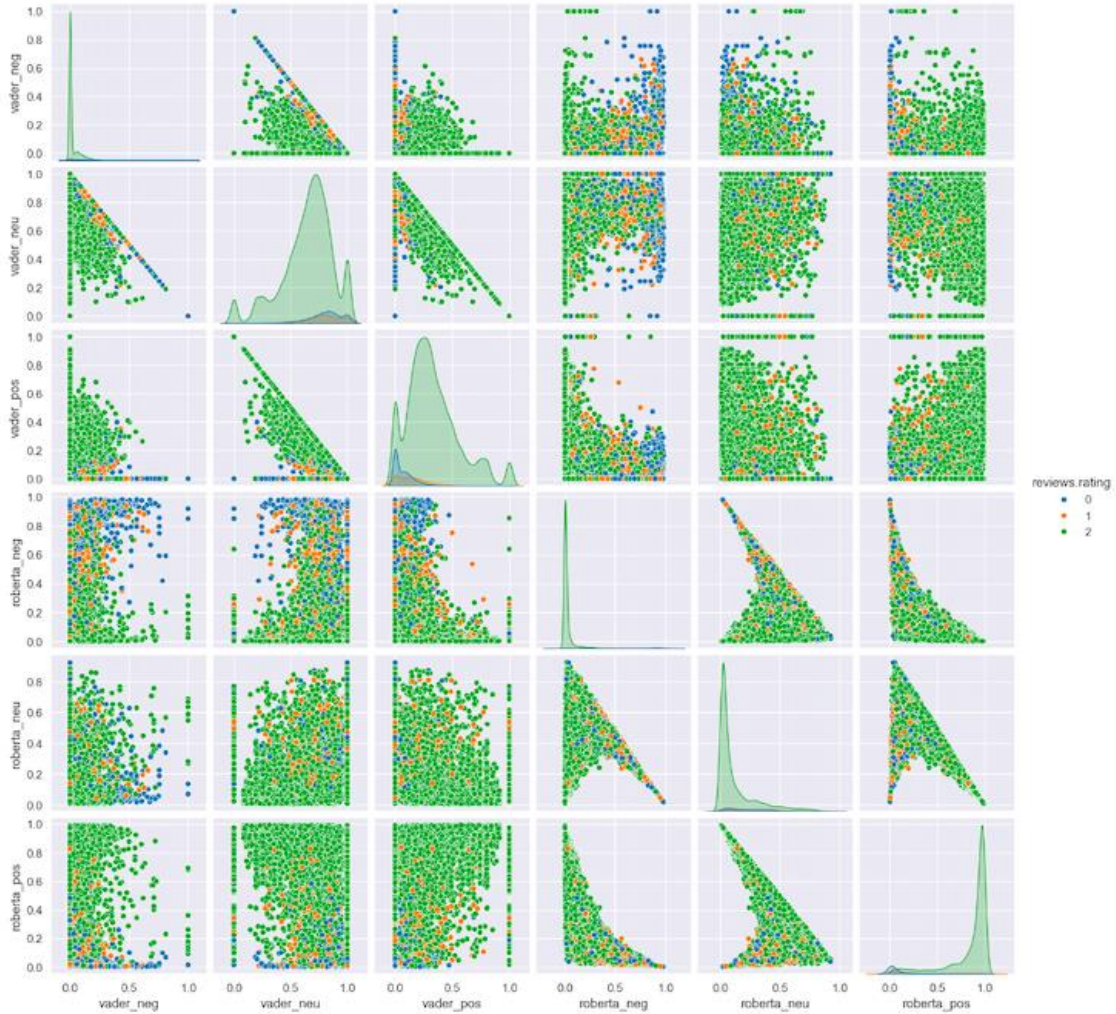
Veri setinde gerçek etiketleri pozitif olan (Şekil 7'deki yeşil noktalar) yorumların VADER nötr ve negatif değişkenleri arasında sıkı bir ilişki yoktur (Şekil 7'nin 2. satır 1. sütunundaki grafiğe bakınız). Gerçek etiketleri negatif (mavi noktalar) ve nötr (turuncu noktalar) olan yorumlarda ise aynı değişkenler arasında sıkı ve negatif doğrusal bir ilişki gözlenmektedir.

Veri setinde gerçek etiketleri pozitif olan yorumların VADER pozitif ve negatif değişkenleri (Şekil 7'nin 3. satır 1. sütununa bakınız) ile VADER pozitif ve nötr değişkenleri (Şekil 7'nin 3. satır 2. sütununa bakınız) arasında sıkı bir ilişki yoktur. Gerçek etiketleri nötr olan yorumlarda ise her iki grafikte de doğrusal olmayan kısmi bir ilişki gözlenmektedir.

Her üç sınıf yorum için RoBERTa negatif değişkeni ile VADER negatif, nötr ve pozitif değişkenleri arasında (sırasıyla 4. satır, 1.-3. sütunlarındaki grafikler) herhangi bir ilişki gözlenmemiştir. Ancak bu üç grafikte RoBERTa negatif değerleri, genel olarak gerçek etiketleri pozitif olan yorumlarda küçük, nötr yorumlarda orta seviyede, negatif yorumlarda ise büyüktür.

Her üç sınıf yorum için RoBERTa nötr değişkeni ile VADER negatif, nötr ve pozitif değişkenleri arasında (sırasıyla 5. satır, 1.-3. sütunlarındaki grafikler) herhangi bir ilişki gözlenmemiştir. RoBERTa nötr değerleri, RoBERTa negatif değerlerinde gözlenen sınıfsal sıralama eğilimine sahip değildir. RoBERTa nötr ile RoBERTa negatif değişkenleri (5. satır, 4. sütundaki grafik) arasında, VADER nötr – VADER negatif

değişkenleri arasındakine benzer bir ilişki vardır. Ancak bu kez gerçek etiketleri negatif ve nötr olan yorumlardaki negatif doğrusal ilişki sıkı değil gevşektir.



**Şekil 4.5.** VADER ve RoBERTa Modellerinin Sınıf Puanları Arasındaki İlişkiler.

Son olarak, her üç sınıf yorum için RoBERTa pozitif değişkeni ile VADER negatif, nötr ve pozitif değişkenleri arasında (sırasıyla 6. satır, 1.-3. sütunlarındaki grafikler) herhangi bir ilişki gözlenmemiştir. RoBERTa pozitif değerleri, RoBERTa negatif değerlerinde gözlenen sınıfsal sıralama eğiliminin tam tersine sahiptir, yani RoBERTa pozitif değerleri, genel olarak gerçek etiketleri pozitif olan yorumlarda büyük, nötr yorumlarda orta seviyede, negatif yorumlarda ise küçüktür. RoBERTa pozitif ile RoBERTa negatif değişkenleri (6. satır, 4. sütundaki grafik) arasında, VADER pozitif – VADER negatif değişkenleri arasındakine benzer ancak daha gevşek bir ilişki vardır. RoBERTa pozitif ile RoBERTa nötr değişkenleri (6. satır, 5. sütundaki grafik) arasındaki ilişki de VADER pozitif – VADER nötr değişkenleri arasındakine benzer ancak daha gevşektir.

#### **4.4. Duygu Analizi Algoritmalarının Doğruluğu ve Güvenilirliği**

Tezde kullanılan duygu analizi algoritmalarının genel olarak yüksek doğruluk oranlarına sahip olduğu bilinmektedir. Ancak bazı inceleme ve yorumların karmaşıklığı yanlış sınıflandırmalara neden olabilmektedir. Özellikle, ironik veya karmaşık dil yapıları içeren yorumların doğru analiz edilmesi zordur. Örneğin, “Bu ürün o kadar harika ki kullanamadan bozuldu!” gibi ironik ifadeler, algoritmalar tarafından yanlış sınıflandırılabilir. Bu durum, duygu analizi algoritmalarının sürekli geliştirilmesi gerektiğini göstermektedir. Algoritmaların, dilin inceliklerini ve bağlamını daha iyi anlayabilmesi için ileri makine öğrenimi teknikleri ve daha geniş veri kümeleriyle eğitilmesi gerekir.

#### **4.5. Tez Çalışmasının Sınırlılıkları**

Tezde kullanılan veri kümesinin belirli bir döneme ve belirli ürün kategorilerine odaklanması, sonuçların genelliğini sınırlayabilir. Örneğin, sadece elektronik ürünler üzerinden yapılan bir analiz, giyim veya gıda ürünleri için aynı sonuçları vermeyebilir. Ayrıca, belirli bir dönemde yapılan yorumlar, o dönemin özel koşullarını yansıtabilir ve uzun vadeli eğilimleri tam olarak göstermeyebilir. Gelecekteki çalışmalar, daha geniş ve çeşitli veri setleri kullanarak bu sınırlılıkları aşabilir. Farklı dönemlerde ve çeşitli ürün kategorilerinde yapılacak analizler, daha kapsamlı ve genellenebilir sonuçlar sağlayacaktır. Ayrıca, farklı coğrafi bölgelerdeki müşteri yorumlarının analiz edilmesi de kültürel farklılıkların müşteri memnuniyeti üzerindeki etkilerini anlamak için önemlidir.

#### **4.6. Gelecek Çalışmalar ve Geliştirme Önerileri**

Gelecek çalışmalar, duygu analizi algoritmalarının doğruluğunu artırmak için ileri tekniklerin kullanımını ve daha geniş veri setleri ile eğitilmesini hedefleyebilir. Özellikle, derin öğrenme ve doğal dil işleme tekniklerinin entegrasyonu, algoritmaların performansını önemli ölçüde artırabilir. Ayrıca, müşteri memnuniyetinin daha derinlemesine incelenmesi için, yorumların yanı sıra kullanıcı profilleri, satın alma geçmişi ve diğer müşteri verileri de analiz edilebilir. Bu tür entegre yaklaşımlar, müşteri memnuniyetinin daha kapsamlı bir şekilde anlaşılmasını sağlayacaktır.

## 5. SONUÇ

Bu tez çalışmasında Amazon müşteri yorumlarının duygu analizi için on bir farklı model test edilmiştir. Bu modellerin dördü (Lojistik Regresyon, Destek Vektör Makineleri, Çok Terimli Naif Bayes ve Stokastik Gradyan İnişi) klasik makine öğrenimi yöntemleri arasından, biri (Çok Katmanlı Algılayıcı) yarı klasik yöntemler arasından, ikisi (Ekstra Rastgele Ağaçlar ve Aşırı Gradyan Artırma) topluluk öğrenme yöntemleri arasından ve üçü (SpaCy kütüphanesi, kurallara dayalı VADER duygu analizi ve transformer tabanlı RoBERTa dil modeli) modern doğal dil işleme araçları arasından seçilmiştir. Tüm modellerin doğruluk, kesinlik, duyarlılık ve F1 puan metrikleri tespit edilmiş, pozitif, nötr ve negatif şeklinde üçlü yorum sınıflaması için karmaşıklık matrisleri oluşturulmuştur. VADER analizinde sınıfların bileşik ve ayrık değerleri incelenmiş, VADER ve RoBERTa modellerinin sonuçları arasındaki ilişkiler görselleştirilerek açıklanmıştır.

Bu çalışma kapsamında elde edilen sonuçlara göre en iyi performansı Çok Katmanlı Algılayıcı ve Ekstra Rastgele Ağaç modelleri göstermiş, bu lider grubu Lojistik Regresyon, Destek Vektör Makineleri, Stokastik Gradyan İnişi, Aşırı Gradyan Artırma, SpaCy, VADER ve RoBERTa modelleri takip etmiş, Çok Terimli Naif Bayes ve VADER yöntemlerinin performans metrikleri ise diğer modellere kıyasla düşük kalmıştır. Sınıfların VADER bileşik duygu polarite değerlerinin, ayrık değerlerle uyum halinde olduğu gözlenmiştir. Ayrıca VADER nötr ve VADER negatif değişkenleri arasında, gerçek etiketleri negatif ve nötr olan yorumlarda sıkı ve negatif bir doğrusal ilişki olduğu tespit edilmiştir.

Bu çalışma, e-ticaret platformları ve satıcılar için önemli ve değerli öngörüler sunabilecek duygu analizi algoritmalarının doğruluğu, güvenilirliği ve etkinliği konusunda somut ve uygulanabilir bilgiler sunmaktadır. Duygu analizi, işletmelerin müşteri memnuniyetini anlamalarına, ürün ve hizmetlerini geliştirmelerine, müşteri sadakatini artırmalarına ve genel olarak daha rekabetçi bir pozisyon elde etmelerine yardımcı olmaktadır. Duygu analizi algoritmalarının doğru ve etkili bir şekilde çalışması, elde edilen sonuçların güvenilirliğini ve geçerliliğini artırır. Çalışmada incelenen algoritmaların güçlü ve zayıf yanları anlatılmış, duygu analizlerinde karşılaşılan dengesiz veri kümeleri, veri kümesinin belirli bir döneme ve belirli ürün kategorilerine odaklanması, veri temizliği, müşteri yorumlarının karmaşıklığı, seçilen modellerin hiper-parametre ayarları, aşırı uyum gibi genel sorunlara değinilmiş ve olası çözümler önerilmiştir.

## KAYNAKLAR

- [1] R. R. Durucasu, "Covid-19 pandemi sürecinde e-ticaret ve dijital pazarlamanın türkiye'deki önemi," *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, cilt 42, pp. 137-154, 2024.
- [2] A. Akyazı, "Dijitalleşen Ticaret: Yaşlı Dostu E-Ticaret Siteleri Üzerine Bir Araştırma," *Türk Çevrimiçi Tasarım Sanatı ve İletişim Dergisi*, cilt 8, no. 4, pp. 602-614, 2018.
- [3] S. Tuzcu, "Çevrimiçi kullanıcı yorumlarının duygu analizi ile sınıflandırılması," *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi*, cilt 1, no. 2, pp. 1-5, 2020.
- [4] H. Yılmaz ve S. Yumuşak, "Açık kaynak doğal dil işleme kütüphaneleri," *İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, cilt 3, no. 1, pp. 81-85, 2021.
- [5] A. E. Özmutlu, "Bilgisayar Bilimlerinde Teorik Ve Uygulamalı Araştırmalar," %1 içinde *Doğal dil işleme*, İstanbul, Efe Akademik Yayıncılık, 2021, pp. 129-154.
- [6] C. Türkmenoğlu, *Türkçe metinlerde duygu analizi*, Yüksek Lisans Tezi: İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2015.
- [7] M. Birjali, M. Kasri ve A. Beni-Hssane, "A comprehensive survey on sentiment analysis: approaches, challenges and trends," *Knowl-Based Syst*, cilt 226, p. 107134, 2021.
- [8] A. Yousif, Z. Niu, Z. K. Tarus ve A. Ahmad, "A survey on sentiment analysis of scientific citations," *Artif Intell Rev*, cilt 52, no. 3, pp. 1805-1838, 2019.
- [9] L. Subhashini, Y. Li, J. Zhang, A. S. Atukorale ve Y. Wu, "Mining and classifying customer reviews: a survey," *Artif Intell Rev*, cilt 54, pp. 6343-6389, 2021.
- [10] A. Yadav ve D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artif Intell Rev*, cilt 53, no. 6, pp. 4335-4385, 2020.
- [11] M. Soleymani, D. Garcia, B. Jou, B. Schuller, S. F. Chang ve M. Pantic, "A survey of multimodal sentiment analysis," *Image Vis Comput*, cilt 65, pp. 3-14, 2017.
- [12] L. Yue, W. Chen, X. Li, W. Zuo ve M. Yin, "A survey of sentiment analysis in social media," *Knowl Inf Syst*, cilt 60, no. 2, pp. 617-663, 2019.
- [13] B. Liu ve L. Zhang, "A survey of opinion mining and sentiment analysis," *Aggarwal C, Zhai C (eds) Mining text data*, pp. 415-463, 2012.
- [14] R. Piryani, D. Madhavi ve V. K. Singh, "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015," *Inf Process Manag*, cilt 53, no. 1, pp. 122-150, 2017.
- [15] P. K. Jain, R. Pamula ve S. Ansari, "A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," *Comput Sci Rev*, cilt 41, p. 100413, 2021.
- [16] T. Balaji, C. S. R. Annavarapu ve A. Bablani, "Machine learning algorithms for social media analysis: a survey," *Comput Sci Rev*, cilt 40, p. 100395, 2021.
- [17] M. E. Mowlaei, M. S. Abadeh ve H. Keshavarz, "Aspect-based sentiment analysis using adaptive aspect-based lexicons," *Expert Syst Appl*, cilt 148, p. 113234, 2020.
- [18] K. Zvarevashe ve O. O. Olugbara, "A framework for sentiment analysis with opinion mining of hotel reviews," *Conference on information communications technology and society (ICTAS). IEEE*, pp. 1-4, 2018.

- [19] F. Valencia, A. Gómez-Espinosa ve B. Valdés-Aguirre, “Price movement prediction of cryptocurrencies using sentiment analysis and machine learning,” *Entropy*, cilt 21, no. 6, p. 589, 2019.
- [20] S. Ahmad, M. Z. Asghar, F. M. Alotaibi ve I. Awan, “Detection and classification of social media-based extremist affiliations using sentiment analysis techniques,” *Hum Centric Comput Inf Sci*, cilt 9, no. 1, pp. 1-23, 2019.
- [21] A. Arora, P. Chakraborty, M. Bhatia ve P. Mittal, “Role of emotion in excessive use of twitter during COVID-19 imposed lockdown in India,” *J Technol Behav Sci*, cilt 6, no. 2, pp. 370-377, 2021.
- [22] Z. Ciplak ve K. Yildiz, “Occupational groups prediction in Turkish Twitter data by using machine learning algorithms with multinomial approach,” *Expert Systems With Applications*, cilt 252, p. 124175, 2024.
- [23] A. B. Alawi ve F. Bozkurt, “A hybrid machine learning model for sentiment analysis and satisfaction assessment with Turkish universities using Twitter data,” *Decision Analytics Journal*, cilt 11, p. 100473, 2024.
- [24] U. Bilgin ve S. S. Kara, “Identification of Perceived Challenges in the Green Energy Transition by Turkish Society through Sentiment Analysis,” *Sustainability* 2, cilt 16, p. 3367, 2024.
- [25] M. S. Başarslan ve F. Kayaalp, “Sentiment analysis of coronavirus data with ensemble and machine learning methods,” *Turkish Journal of Engineering*, cilt 8, no. 2, pp. 175-185, 2024.
- [26] K. B. Küçüklerli ve V. Ulusoy, “Sentiment-Driven Exchange Rate Forecasting: Integrating Twitter Analysis with Economic Indicators,” *Journal of Applied Finance & Banking*, cilt 14, no. 3, pp. 75-79, 2024.
- [27] B. Gökalp ve Ö. Ö. Subay, “The reflections on social media of the statement of the nation alliance’s strengthened parliamentary system in Turkey,” *Zeitschrift für Vergleichende Politikwissenschaft*, <https://doi.org/10.1007/s12286-024-00593-z>, 2024.
- [28] Ö. F. Göçgün ve A. Onan, “Amazon Ürün Değerlendirmeleri Üzerinde Derin Öğrenme/Makine Öğrenmesi Tabanlı Duygu Analizi Yapılması,” *Avrupa Bilim ve Teknoloji Dergisi*, cilt 24, pp. 445-448, 2021.
- [29] S. Maurya ve V. Pratap, “Sentiment Analysis on Amazon Product Reviews,” %1 içinde *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, 26-27 May 2022, Faridabad, India, 2022.
- [30] M. S. Meghana, P. K. Kollu, D. Abhijith ve S. Aysha, “Sentiment Analysis on Amazon Product Reviews using LSTM and Naive Bayes,” %1 içinde *Proceedings of the 7th International Conference on Computing Methodologies and Communication (ICCMC-2023)*, Erode, India, 2023.
- [31] S. Tufchi, A. Yadav, V. K. Rai ve A. Banerjee, “Sentiment Analysis on Amazon Product Review: A Comparative Study,” %1 içinde *Khanna, A., Polkowski, Z., Castillo, O. (eds) Proceedings of Data Analytics and Management . Lecture Notes in Networks and Systems, vol 572. Springer*, Singapore, 2023.
- [32] Y. Görmez, H. Arslan ve B. Atak, “Türkçe Metinlerde Duygu Analizi: Derin Öğrenme Yaklaşımlarının ve Ön İşlem Süreçlerinin Model Performansına Etkisi,” *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, cilt 36, no. 1, pp. 509-520, 2024.

- [33] L. Güner, E. Coyne ve J. Smith, “Sentiment analysis for Amazon.com reviews,” *Big Data in Media Technology (DM2583)*, KTH Royal Institute of Technology, Stockholm, 2019.
- [34] “Kaggle,” [Çevrimiçi]. Available: <https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products>.
- [35] “NLTK,” [Çevrimiçi]. Available: <https://www.nltk.org/>. [Erişildi: 29 May 2024].
- [36] A. I. Kadhim, “Twitter'da özellik çıkarımı için terim ağırlıklandırma: BM25 ve TF-IDF arasında bir karşılaştırma,” %1 içinde *2019 Uluslararası İleri Bilim ve Mühendislik Konferansı (ICOASE), IEEE.*, 2019.
- [37] K. Shah, H. Patel, D. Sanghvi ve M. A. Shah, “A comparative analysis of logistic regression, random forest and KNN models for the text classification,” *Augment. Hum. Res.*, cilt 5, p. 12, 2020.
- [38] T. Hossain, H. Z. Mauni ve R. Rab, “Reducing the effect of imbalance in text classification using SVD and GloVe with ensemble and deep learning,” *Comput. Inform.*, cilt 41, pp. 98-115, 2022.

