



T.C.

BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI

Yüksek Lisans Tezi

**ÖZEL GÜN İDİRİMLERİNDE TÜKETİCİ
YORUMLARININ MAKİNE ÖĞRENMESİ
ALGORİTMASIYLA KARŞILAŞTIRILMASI**

**Hatice YILMAZ İNCE
2215052008**

Tez Danışmanı:

Doç. Dr. Cemalettin HATİPOĞLU

Bandırma2025

T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI

Yüksek Lisans Tezi

**ÖZEL GÜN İDİRİMLERİNDE TÜKETİCİ
YORUMLARININ MAKİNE ÖĞRENMESİ
ALGORİTMASIYLA KARŞILAŞTIRILMASI**

Hatice YILMAZ İNCE
2215052008

Tez Danışmanı:
Doç. Dr. Cemalettin HATİPOĞLU

Bandırma 2025

ETİK BEYAN

TÜRKİYE CUMHURİYETİ BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Bu belge ile, bu tezdeki bütün bilgilerin akademik kurallara ve etik ilkelere uygun olarak toplanıp sunulduğunu beyan ederim. Bu kural ve ilkelerin gereği olarak, çalışmada bana ait olmayan tüm veri, düşünce ve sonuçları andığımı ve kaynağını gösterdiğimi ayrıca beyan ederim. (24/06/2025)

Hatice YILMAZ İNCE

ÖZET

ÖZEL GÜN İDİRLERİNDE TÜKETİCİ YORUMLARININ MAKİNE ÖĞRENMESİ ALGORİTMARIYLA KARŞILAŞTIRILMASI

Hatice YILMAZ İNCE

Dijitalleşmenin hız kazanmasıyla birlikte geleneksel ticaret anlayışı dönüşüme uğramış, çevrim içi alışveriş sistemleri tüketici davranışlarında önemli bir rol üstlenmiştir. Bu değişimle beraber, özellikle özel günlerde gerçekleştirilen kampanyalar sürecinde tüketici yorumları, satın alma kararlarını etkileyen temel unsurlardan biri hâline gelmiştir. Tüketiciler arasında artan bilgi paylaşımı sayesinde kullanıcı yorumları, bireysel kararlar üzerinde etkili olmanın ötesinde, işletmelerin stratejik planlamaları için de önemli bir veri kaynağına dönüşmüştür. Bu çalışma, özel günlerde çevrim içi alışveriş yapan üniversite öğrencilerinin ürünlere yönelik yorumlarını analiz etmeyi amaçlamaktadır. Trendyol platformundan elde edilen veriler, Python programlama dili ve doğal dil işleme teknikleri kullanılarak analiz edilmiştir. Kullanıcı yorumları olumlu ve olumsuz olmak üzere sınıflandırılmış; Naive Bayes, Lojistik Regresyon, Karar Ağacı, KNN, Gradyan Artırma, XGBoost ve Rastgele Orman olmak üzere yedi farklı denetimli makine öğrenmesi algoritmasıyla modeller oluşturulmuştur. Modeller, doğruluk oranı, ROC-AUC ve F1 skoru gibi ölçütlerle değerlendirilmiştir. En yüksek performans Gradyan Artırma ve XGBoost algoritmalarında gözlemlenmiştir. Ayrıca, farklı ürün kategorilerinde elde edilen benzer başarı oranları, modellerin genellenebilirliğini ortaya koymuştur. Elde edilen bulgular, makine öğrenmesi destekli duygu analizinin tüketici deneyimini iyileştirmede ve pazarlama stratejilerinin şekillendirilmesinde etkili bir araç olduğunu göstermektedir.

Anahtar Kelimeler: Makine Öğrenmesi, Duygu Analizi, Tüketici Yorumları, Çevrim İçi Alışveriş, Özel Gün İndirimleri

ABSTRACT

COMPARISON OF CONSUMER REVIEWS IN SPECIAL DAY DISCOUNTS USING MACHINE LEARNING ALGORITHMS

Hatice YILMAZ İNCE

With the acceleration of digitalization, traditional commerce has undergone a transformation, and online shopping systems have assumed a significant role in shaping consumer behavior. In this context, user reviews—especially during promotional periods such as special days—have become a critical factor influencing purchasing decisions. The increasing exchange of information among consumers has elevated user-generated content from a personal guide to a strategic data source for businesses.

This study aims to analyze product-related reviews of university students who shop online during special occasions. The data, collected from the Trendyol platform, were processed using the Python programming language and natural language processing (NLP) techniques. The comments were classified as positive or negative, and seven supervised machine learning algorithms—Naive Bayes, Logistic Regression, Decision Tree, K-Nearest Neighbors (KNN), Gradient Boosting, XGBoost, and Random Forest were used to develop classification models. These models were evaluated using performance metrics such as accuracy, ROC-AUC, and F1-score. Gradient Boosting and XGBoost achieved the highest performance among the tested algorithms. Furthermore, similar performance levels across different product categories supported the generalizability of the models.

The findings indicate that sentiment analysis supported by machine learning can serve as an effective tool for improving consumer experience and guiding marketing strategies.

Keywords: Machine Learning, Sentiment Analysis, Consumer Reviews, Online Shopping, Special Day Discount

ÖNSÖZ

Bu tez çalışması, dijitalleşmenin hız kazandığı günümüz dünyasında çevrim içi alışverişin tüketici davranışları üzerindeki etkilerini anlamaya yönelik bir çabanın ürünüdür. Özellikle özel günlerde yapılan indirimlerin, tüketici yorumları aracılığıyla nasıl şekillendiğini ve bu yorumların makine öğrenmesi algoritmalarıyla analiz edilebilirliğini ortaya koymayı amaçladım.

Araştırma süreci boyunca karşılaştığım teknik ve teorik zorlukların üstesinden gelmemde, yanımda olan kıymetli insanların desteği büyük önem taşımıştır. Bu süreçte, akademik birikimi ve değerli rehberliğiyle tezimin her aşamasında yol gösterici olan danışmanım Doç. Dr. Cemalettin Hatipoğlu'na en içten teşekkürlerimi sunarım.

Aynı zamanda, araştırma sürecinde akademik desteği ve moral verici katkılarıyla bana eşlik eden kıymetli arkadaşım Araştırma Görevlisi Reyhan Başbayram Öztürk'e de teşekkür ederim. Onun katkıları, bu çalışmanın ilerlemesinde değerli bir destek olmuştur.

Hayatım boyunca yanımda olan, sabır ve emekleriyle bana güç veren sevgili aileme ve en büyük dayanağım olan değerli eşim Oğuzhan İnce'ye sonsuz şükranlarımı sunarım. Onların varlığı, araştırma sürecini daha anlamlı ve kararlı bir şekilde tamamlamamı sağlamıştır.

Son olarak, akademik yolculuğumda bana ilham veren ve bilgi dünyamı şekillendiren tüm değerli hocalarıma gönülden teşekkür ederim. Bu tez çalışmasının, e-ticaret ve makine öğrenmesi alanında araştırma yapanlar için faydalı bir kaynak olmasını temenni ederim.

Hatice YILMAZ İNCE

Bandırma, 2025

İÇİNDEKİLER

ETİK BEYAN	iii
ÖZET	iv
ABSTRACT	v
TABLOLAR.....	viii
ŞEKİLLER.....	ix
KISALTMALAR.....	xi
GİRİŞ.....	4
BİRİNCİ BÖLÜM	10
1. KAVRAMSAL ÇERÇEVE	10
1.1.Litaratür Çalışmaları.....	10
1.2.Web Teknolojilerinin Gelişimi.....	12
1.2.2 Web Reklamcılığı	14
1.2.3. Google AdWords Kavramları	15
1.2.4. Google Arama Motoru Çalışma Prensibi.....	16
1.2.5. Google Trendler.....	18
1.3.Tüketici Kavramı ve Tüketici Davranışları.....	20
1.3.1. Kullanıcı Yorumu ve Yıldız Değerlendirme	21
1.4. Duygu Analizi	23
İKİNCİ BÖLÜM	25
2. VERİ, ALGORİTMA VE MAKİNE ÖĞRENMESİ	25
2.1.Veri	25
2.1.1. Nümerik Değişkenler	25
2.1.2. Veri Ön İşleme.....	26
2.1.3. Veri Temizleme.....	26
2.1.4. Veri Dönüştürme	28
2.2. Makine Öğrenmesi.....	28
ÜÇÜNCÜ BÖLÜM.....	31
MODEL, YÖNTEM, ANALİZLER	31
3.1 VERİLERİN TOPLANMASI.....	31
3.2 Çalışmanın Modeli ve Uygulanması.....	33
3.2.1 Verinin Toplanması	34
3.2.2 Veri Temizleme ve Ön İşleme	35
3.2.3 Veri Setinin Eğitilmesi.....	36
SONUÇ	71

KAYNAKÇA	75
----------------	----



TABLÖLAR

Tablo 1: Web Teknolojilerinin Kıyaslanması	13
Tablo 2: Veri Ön İşleme Sonucu	35
Tablo 3: Çam Ağacı Ürünü AUC- ROC Oranları(%)	39
Tablo 4: Tablet Ürünü AUC- ROC Oranları(%)	52
Tablo 5: Çiçek Ürünü AUC- ROC Oranları (%)	65

ŞEKİLLER

Şekil 1: Son Beş Yılın Küresel Çapta Arama Motoru Pazar Payları	17
Şekil 2: Son Beş Yıl Türkiye Geneli Arama Motoru Pazar Payları	18
Şekil 3: Son Beş Yılda Google Arama Motorunda Türkiye’de Faaliyet Gösteren E- Ticaret Sitelerine Yönelik İlgi	19
Şekil 4: Çevrimiçi Ürün Değerlendirme ve Yıldız Filtreleme	22
Şekil 5: Çevrim İçi Satın Alma Değerlendirme	22
Şekil 6: Plutchik Duygu Çarkı	24
Şekil 7: ELT Süreci.....	28
Şekil 8: Google Trendlerde Son 3 Yıl İçerisinde Aratılan E- ticaret Sitelerinin Karşılaştırılması.....	31
Şekil 9: Çalışma Modeli İçin Örnek Ürün Değerlendirmesi.....	33
Şekil 10: Çalışmanın Modeli ve Uygulanması	33
Şekil 11: Çam Ağacı Ürününün Makine Öğrenmesi Doğruluk Oranları	37
Şekil 12: Çam Ağacı Ürünün ROC-AUC Oranları.....	38
Şekil 13: Çam Ağacı Ürünü F1 Skoru	40
Şekil 14: Modellere Göre Normalize Edilmiş Performans Karşılaştırılması.....	41
Şekil 15: Çam Ağacı Ürünü Naive Bayes Tahmin Hata Matrisi	42
Şekil 16: Çam Ağacın Ürünü : Lojistik Regresyon Tahmin Hata Matrisi.....	43
Şekil 17: Çam Ağacı Ürünü Karar Ağacı Tahmin Hata Matrisi.....	44
Şekil 18: Çam Ağacı Ürünü KNN Algoritması Hata Tahmin Matrisi.....	45
Şekil 19: Çam Ağacı Ürünü Rassal Orman Tahmin Hata Matrisi.....	46
Şekil 20: Çam Ağacı Ürünü Gradyan Arttırma Tahmin Hata Matrisi	47
Şekil 21: Çam Ağacı Ürünü XBoost Hata Tahmin Matrisi	48
Şekil 22: Çam Ağacı Ürününe Ait Kelime Bulutu	49
Şekil 23: Tablet Ürününe Ait Makine Öğrenmesi Doğruluk Oranları.....	50
Şekil 24: Tablet Ürününe Ait ROC- AUC Oranları.....	51
Şekil 25: Tablet Ürünü F1 Skoru	53
Şekil 26: Tablet Ürünü Naive Bayes Hata Matrisi	54
Şekil 27: Tablet Ürünü Lojistik Regresyon Hata Matrisi	55
Şekil 28: Tablet Ürünü Karar Ağacı Hata Matrisi	56
Şekil 29: Tablet Ürünü KNN Algoritması Hata Matrisi	57
Şekil 30: Tablet Ürünü Gradyan Arttırma Hata Tahmin Matrisi.....	59
Şekil 31: Tablet Ürünü XGBoost Algoritması Hata Tahmin Matrisi	60
Şekil 32: Tablet Ürünü Rassal Orman Algoritması Hata Tahmin Matrisi.....	61
Şekil 33: Tablet Ürününe Ait Kelime Bulutu	62
Şekil 34: Çiçek Ürününe Ait Makine Öğrenmesi Doğruluk Oranları	63
Şekil 35: Çiçek Ürününe Ait ROC- AUC Oranları	64

Şekil 36: Çiçek Ürünü Naive Bayes Hata Tahmin Matrisi.....	66
Şekil 37: Çiçek Ürünü Lojistik Regresyon Hata Tahmin Matrisi.....	67
Şekil 38: Çiçek Ürünü Karar Ağacı Algoritması Hata Tahmin Matrisi	68
Şekil 39: Çiçek Ürünü KNN Algoritması Hata Tahmin Matrisi	70



KISALTMALAR

ELT	: Extraction, Transforming, Loading (Çıkart, Dönüştür, Yükle)
KNN	: K-Nearest Neighborhood (K- En Yakın Komşu Algoritması)
NLP	: Natural Language Processing (Doğal Dil İşleme)
SEM	: Search Engine Marketing (Arama Motoru Reklamcılığı)
SEO	: Search Engine Optimization (Arama Motoru Optimizasyonu)
URL	: Standart Kaynak Bulucu
SMOTE	: Synthetic Minority Over-sampling Technique (Sentetik Azınlık Aşırı Örnekleme Tekniği)

GİRİŞ

Dijital teknolojilerdeki ilerlemeler, bireyler arası etkileşimi çevrim içi platformlara taşıyarak bilgi paylaşımının yaygınlaşmasına olanak tanımış; böylece tüketiciler arasındaki mekânsal ve zamansal iletişim sınırlılıkları büyük ölçüde ortadan kalkmıştır. Tüketiciler, Web 2.0 ve teknolojileri öncesinde yalnızca kendilerine yakın konumda bulunan kişi ve kurumlarla iletişim kurabilirken Web 2.0 teknolojilerinin gelişmesi ile küresel çapta iletişim kurmaya başlamışlardır. Geçtiğimiz yıllarda yaşanan Covid-19 salgını ile fiziki konumda bulunan birçok kurum ve şirket dijital ortamda da aktif olarak faaliyet göstermeye başlamıştır. Bu da tüketicilerin satın alma şeklini ve tutumlarının değişimini de beraberinde getirmiştir. Şirketler her geçen gün sayısı artan dijital mağazalarla rekabet edebilmek için özel günleri fırsata çevirmektedir bunlara özel indirimler yapmaktadır. Tüketiciler ise yılın belli zamanlarında yapılan bu özel indirimlerle kişisel ihtiyaçlarını karşılamak ve sevdiklerine hediyeler almak için satın alma faaliyeti göstermektedir. Dijital ortamda satın alınan bu ürünler, fiziki ortamdaki satın alma davranışından farklı gerçekleştiği için diğer tüketicilerin satın alma tecrübeleri ve ürün hakkındaki yorumları diğer tüketicilerin satın alma eğilimlerini etkilemektedir.

Çalışmanın Amacı:

Özel günlerde yapılan indirim faaliyetlerinin tüketicilerin satın alama eğilimini ne yönde etkilediği ve bu satın alma faaliyeti sonrası deneyimleri ve satın aldıkları ürün hakkında yaptıkları yorumlar denetimli sınıflandırma algoritmalar tarafından öğrenilerek değer tüketiciler üzerinde nasıl bir etki bırakacağı hakkında tahminde bulunmasını sağlamak ve algoritmaların kullanıcı yorumlarının hangi nitelikler üzerinde ağırlık gösterdiğini ortaya koymaktır.

Çalışmanın Problemi ve Soruları:

Bu çalışmanın problemi Özel günlere yönelik yapılan indirimlerin, çevrim içi alışveriş yapan tüketicilerin ürün değerlendirme davranışlarını ve bu değerlendirmelerin diğer kullanıcılar üzerindeki etkisini, makine öğrenmesi algoritmalarıyla tahmin etmek mümkün müdür?” sorusuna yanıt aramaktır. Bu problemin çözümlenebilmesi için cevaplanması gereken alt sorular;

Araştırma Sorusu 1:

Özel gün indirimleri kapsamında gerçekleştirilen çevrim içi alışverişlerin, tüketici yorumlarına yansıyan duygusal ifadeler aracılığıyla değerlendirme davranışları üzerindeki etkisi ortaya konabilir mi?

Araştırma Sorusu 2:

Farklı makine öğrenmesi algoritmaları, özel günlerde indirimli olarak satın alınan ürünlere ilişkin kullanıcı yorumlarını duygu analizi açısından sınıflandırmada ne düzeyde performans göstermektedir ve hangi algoritmalar bu açıdan öne çıkmaktadır?

Araştırma Sorusu 3:

Farklı ürün kategorilerinde uygulanan makine öğrenmesi algoritmalarının duygu sınıflandırma performansları arasında; doğruluk, ROC-AUC ve F1 skoru gibi ölçütler temelinde anlamlı farklılıklar var mıdır ve bu farklılıklar hangi faktörlerden kaynaklanmaktadır?

Çalışmanın Önemi

Çevrim içi alışveriş, geleneksel mağazalara kıyasla daha fazla belirsizlik barındırmakta; özellikle ürün teslimatı ve beklenti uyumsuzlukları tüketici memnuniyetini etkileyebilmektedir. Bu nedenle, önceki kullanıcıların deneyimlerini paylaştığı yorumlar, diğer tüketicilerin karar sürecinde belirleyici bir rol üstlenmektedir.

Makine öğrenmesi algoritmalarıyla bu yorumların sistematik biçimde analiz edilmesi, yalnızca kullanıcı davranışlarını anlamayı değil, aynı zamanda marka güvenilirliğini artırma, iade oranlarını azaltma ve pazarlama stratejilerini veri temelli olarak optimize etme imkânı sunmaktadır. Bu yönüyle çalışma, e-ticaret alanında dijital tüketici içgörüsünün etkin kullanılmasına katkı sağlamaktadır.

Çalışmanın Yöntemi

Çevrim içi alışveriş platformlarında özel gün indirimleri kapsamında gerçekleştirilen alışverişlere ilişkin tüketici yorumları, kazıyıcı uygulamalar (scraper tools) aracılığıyla toplanacaktır. Toplanan veriler, doğal dil işleme teknikleri kullanılarak analiz öncesinde çeşitli veri ön işleme aşamalarından geçirilecektir. Her bir yorum, yalnızca içerdiği duygu durumu (olumlu/olumsuz) dikkate alınarak etiketlenecek ve bu doğrultuda yapılandırılan veri seti makine öğrenmesi algoritmaları ile değerlendirmeye alınacaktır. Elde edilen bu veri seti üzerinde; Naive Bayes, Lojistik Regresyon, Karar Ağacı, K-En Yakın Komşu, Rassal Orman, Gradyan Artırma ve XGBoost olmak üzere yedi denetimli makine öğrenmesi algoritması eğitilecek ve modellerin sınıflandırma performansları karşılaştırılacaktır. Analiz sonucunda, her bir algoritmanın olumlu ya da olumsuz yorumları ne ölçüde doğru sınıflandırabildiği değerlendirilecek; ayrıca sistemin işleyişini test etmek amacıyla örnek bir kullanıcı yorumu modele uygulanacaktır.

Çalışmanın Kısıtları

Araştırmanın bazı sınırlamaları bulunmaktadır. Ürün kategorilerinin belirlenmesinde, yalnızca belirli bir zaman aralığında toplanan veriler ile sınırlı bir tüketici grubunun tercihleri dikkate alınmıştır. Bu durum, farklı dönemlerde ve çeşitli demografik özelliklere sahip tüketicilerin davranış kalıplarının yeterince temsil edilememesine neden olabilir. Ayrıca, veri toplama sürecinin yalnızca tek bir e-ticaret platformuyla sınırlı tutulması, elde edilen bulguların farklı çevrim içi alışveriş ortamlarına genellenebilirliğini azaltmaktadır. Çalışmada kullanılan veri seti, kapsam itibarıyla sınırlı sayıda öznitelik içermektedir ve ilgili literatürde sıklıkla incelenen bazı niteliksel değişkenler analiz sürecine dâhil edilmemiştir. Tüm bu faktörler, araştırmanın metodolojik kapsamını daraltan başlıca sınırlayıcı unsurlar arasında yer almaktadır.

BİRİNCİ BÖLÜM

1. KAVRAMSAL ÇERÇEVE

1.1.Litaratür Çalışmaları

Fernandes ve diğerleri (2022) Tüketicilerin çevrimiçi incelemelere maruz kalması, çevrimiçi satın alma davranışları üzerinde etkilidir Satın alma kararları için ürünleri değerlendirirken yorum aradıklarını vurgulamışlardır. Çevrimiçi incelemelerin tüketici satın alma kararları üzerindeki etkisini anlamak için bir ölçek geliştirmişler ve ölçeği test etmişlerdir. Ölçeğin test edilmesi aşasına veri setini internette alışveriş yapan gençlerden oluşturmuşlardır.

Chatrath, Batra ve Chaba (2022) tarafından yapılan çalışmada tüketicilerin ikinci el ürün satın alırken karşılaştıkları sorunları değerlendirerek güvenlerinin artırılmasına vurgu yapılmıştır.Güvenlik açıklarını bulmak için görüntü işleme ve makine öğrenimi yaklaşımları kullanılmıştır. Çalışma sonucunda E-ticaret sitelerinin kullanıcılarının ihtiyaçlarına yönelik hizmette bulunmayı tahattüt etmelerine rağmen, çalışmada kullanılan bazı mağza ve satıcıların ürünlerinin incelemelerini ve orijinal fotoğraflarını tarif ederek hileli davranışlarda bulundukları gözlemlenmiştir. Bu durum müşterilerin söz konusu e-ticaret sitesine olan güven ve itimadını kaybetmesine yol açabileceği ve söz konusu e-ticaret sitesine yönelik olumsuz bir algının oluşmasına neden olabileceği öngörülmüştür.

Liu ve diğerleri (2023) Yaptıkları Çalışmada diyalektik düşünmenin, ürün inceleme dağılımının satın alma niyetleri üzerindeki etkilerini ve ilgili temel mekanizma ve sınır koşulları üzerindeki etkilerini incelemişlerdir.

Daha fazla inceleme dağılımının, artan güvenilirlik algıları yoluyla yüksek ya da düşük diyalektik düşünceye sahip tüketiciler arasında satın alma niyetlerini arttırması gerektiğini vurgulamışlardır ve ek olarak ürün türü, inceleme dağılımının etkilerini hafifletmelidir ;yüksek diyalektik düşünceye sahip tüketiciler, yalnızca hedonik ürünler için daha fazla dağınık incelemeyi tercih etmektedir faydacı ürünler için, son derece diyalektik tüketicilerin yüksek yorum dağılımına yönelik tercihi azalmalıdır. Hipotezleri ile diyalektik düşüncenin farklı operasyon eleştirmelerini kullanan çok sayıda deneyi ampirik olarak test etmişlerdir

Li,Yuan ve Tian (2023) Tüketici memnuniyeti aracı değişkenini etkileşim perspektifinden tanımlamakta, etkileşim ve güven arasında bir bağ oluşturmakta ve çevrimiçi etkileşimin çevrimiçi alışverişte tüketici memnuniyeti üzerindeki iç etki mekanizmasını incelemektedirler JD etkileşimli alışveriş platformunda kadın giyim etkileşimli alışveriş platformunun hedef tüketici memnuniyetini incelemişler. Etkileşimin satıcı nitelikleri ve hizmet memnuniyeti değerlendirmeleri, mağaza büyüklüğü ve satın alınan malların lojistiği üzerindeki etkisi analiz edilmiştir.

Ağmeka, Wathoni ve Santoso (2019) Endonezya da faaliyet gösteren bir e- ticaret şirketinde alışveriş yapmış tüketicilerin incelendiği çalışmada Marka imajının yanı sıra marka itibarını etkileyen birçok indirimin uygulaması bulunduğunu ve bu iki niteliğin tüketicilerin satın alma niyetini ve gerçek davranışı etkilediğini vurgulamışlardır.

Lucas, Lunardi ve Dolci (2023) Ücretsiz bir simülasyon deneyi aracılığıyla kullanıcının erişilen platformdaki deneyimini öngören farklı faktörleri geleneksel e-ticaret sitesi ve m-ticaret platformları arasındaki kullanıcı deneyimlerini analiz etmişler. İncelenen platformlar arasında cinsiyet, yaş ve kullanıcı deneyimi arasında farklılıklar bulunduğunu ve bununda satın alma niyetleri üzerinde önemli etkili olduğunu gözlemlemişlerdir.

Hu ve Liu (2004) Tüketicilerin ürünler üzerindeki yorum ve değerlendirmelerini olumlu olumsuz olarak etiketlendirerek incelemişler. Ürünler üzerinde duygu içeren özellikleri tespit ederek bu özelliklere göre cümleler çıkararak ürün incelemeleri özeti oluşturmuşlar.

Olumlu veya olumsuz etiki verilirken yorumlardaki kutupsal kavramları tespit ederek her bir ürün için sıfatlardan yararlanılmıştır. Ürün yorum ve değerdendirmelerinin özetlenmesinin yalnızca tüketiciler için değil aynı zamanda üreticiler açısından oldukça önemli olduğunun vurgusunu yapmışlar.

1.2.Web Teknolojilerinin Gelişimi

Word Wide Web 1990' lerde Avrupa Nükleer Araştırmalar Merkezi (CERN) 'de kullanıcılara internette bilgi aramaları amacı ile Tim Berners-Lee tarafından geliştirilmiştir (Berners-Lee, 1998). Web 1.0 internet teknolojilerinin başlangıç dönemidir, bu dönemdeki web siteleri günümüz web sitelerine kıyasla daha basit bir görünüme sahiptir. Web 1.0 teknolojileri sayfa yöneticileri tarafından kontrol edilebilen ve kullanıcılara yalnızca okunabilir (read- only) yapıda bilgi aktarımı sağlamaktadır.(Choudhury, 2014: 8096)

- Web 1.0 teknolojilerinin özellikleri;
- Sadece okunabilir ve doküman odaklı yapı
- Web siteleri HTML programlama dili ile kodlanmıştır
- Web sitesi kullanıcıların sayfa yöneticisi ve diğer kullanıcılar ile iletişim kurmasına imkan sağlanmamaktadır. (Özden, 2022: 30)

Web 2.0 teknolojileri, 2004 yılında Tim O'Reilly'nin öncülüğünde Web 1.0'ın sınırlı etkileşim sunan yapısına karşılık, daha dinamik ve kullanıcı merkezli bir yaklaşımı benimseyen yeni bir internet anlayışı olarak ortaya konmuştur.

Web 2.0, kullanıcılara yalnızca içerik tüketme değil, aynı zamanda içerik üretme ve paylaşma imkânı da tanımaktadır. Bu yaklaşımla birlikte, artık yalnızca site yöneticileri değil, kullanıcılar da aktif şekilde içerik üretimine katkı sağlamaya başlamıştır.

Bu da sosyal medyanın gelişimine ortam hazırlamıştır, sosyal medyada kullanıcılar kişileştirilmiş hesap oluşturarak diğer kullanıcılarla arkadaşlık kurabilir onlarla fotoğraflarını, mevcut konumlarını, fikirlerini, video ve müzik ve birçok eklentiği paylaşabilme imkanı bulabilmektedirler. (Barsky ve Purdon, 2006).

Tablo 1: Web Teknolojilerinin Kıyaslanması

Web 1.0(1995-2000)	Web 2.0(1995-2010)	Web 3.0(2010-2020)	Web 4.0 (2020-)
Salt okunabilir	Okunup yazılabilir	Taşınabilir ve kişisel	Kendi kendine öğrenebilen, kendi kendini organize edebilen
Doküman esaslı	İnsan esaslı	Bilgi esaslı	Sanal gerçeklik ve büyük veri odaklı
Bireysel bilgiye erişim	İnternet deneyimi başkalarıyla paylaşım	Sanal bilgisayar asistanları yardımcı olur	Bilgisayarlar isteklerinizi anlar.
Statik web	Sosyal web	Semantik web	Faaliyetlerinizi yönetir.
Web formları	Web uygulamaları	Akıllı uygulamalar	Akıllı web
Şirket odaklı	Toplum odaklı	Bireysel odaklı	Yapay zeka, makine öğrenmesi, Sanal gerçeklik, IOT, Otonom hizmetler, Üretici yapay zeka araçları
Şirket siteleri	Blog siteler, Sosyal medya	Yarı sosyal ağ yarı blog siteler	Yapay zeka odaklı yapay zeka ile otonom çalışan siteler
Anonim kullanıcılar	Kayıtlı kullanıcılar	Canlı yayınlar	Kişiselleştirilmiş verilere sahip kullanıcılar
Haber bağlantı ağı	İnsan bağlantı ağı	Yetkili kullanıcılar	
		Bilgi bağlantı ağı ve Açık bilgi	

Kaynak: Yazar Tarafından Oluşturulmuştur

Web 3.0 ya da semantik web kullanıcıların ve makinelerin birbiriyle etkileşimde bulunduğu bir yapıdadır. Web 3.0'ın temel mantığı internette bulunan tüm verileri tek bir ortamda toparlayarak birbirleri ile bağdaştırılarak yönetilmesidir. (Demirli ve Kütük, 2010: 100). Web 3.0 teknolojileri kullanıcılarının log kayıtlarını tutarak onlara bireysel sonuçlar üretmektedir. Kullanıcıların e- ticaret yapmak istediklerinde e- ticaret sitesine cinsiyet, renk, fiyat aralığı, ürün tipi gibi nitelikler girerek kişiselleştirilmiş alışveriş yapma imkanı sunulmaktadır.

Web 4.0 teknolojileri günümüzde kullanılan temelde Yapay zeka tabanlı makinelerin kendi aralarında ve kullanıcılar ile iletişiminden oluşan bir yapıya sahiptir. Web 4.0, makinelerin sadece öğrendikleri değil aynı zamanda sorunlara uygun çözümler üretebildikleri bir gelişmedir. (Ersöz, 2020: 62) Web 4.0 teknolojisinde her geçen gün daha çok insanlar gibi düşünen ve davranan makineler geliştirilmektedir. Günümüzde Web 4.0' a örnek olarak OpenAI tarafından geliştirilen ChatGPT, DALL-E, Robot Süpürgeler, Facebook's agent M, WebOS, Google Docs ve Spreadsheets verilebilir (Aktaş, Kurgun H.ve Kurgun O.,2018: 56).

1.2.2 Web Reklamcılığı

İnternet teknolojilerindeki gelişmeler, reklamların dijital mecralar aracılığıyla hedef kitleye daha kolay ulaşmasını sağlamış; bu durum, kullanıcı odaklı ürün ve hizmet sunumlarının daha etkin ve kişiselleştirilebilir biçimde yürütülmesine olanak tanımıştır (Gümüş Sungur ve Gümüş Bilim, 2014:41). Web 2.0 teknolojileri, reklamcılık faaliyetlerinin çevrim içi ortama taşınmasına öncülük etmiş; bu sayede geniş kitlelere ulaşmak kolaylaşmış ve tüketiciyle etkileşim süreci hız kazanmıştır (Ayaz, 2023:101). Reklamların web ortamında kullanılması, işletmelere hem maliyetleri azaltma hem de daha geniş hedef gruplarına ulaşma fırsatı sunmuştur (Tosun, 2004:159–166).

Google AdWords, Reklam veren işletmelerin yeni müşterilere ulaşmasını ve mevcut işlerini geliştirmelerini sağlayan dijital reklamcılık platformudur. Reklam veren işletmeler ürün ya da hizmeti içeren anahtar sözcükler belirlerler ve kullanıcılar bu anahtar kelimeleri Google arama motoru üzerinden artıklarında, Google bu reklamları kullanıcılara sunar. Kullanıcılar sunulan sayfayı ziyaret ettiklerinde reklam veren işletme her bir tıklanma için ücret almaktadır (Eskicumalı ve Sarsılmaz, 2019 :196).

Google Ads SEM (Arama Motoru Reklamcılığı) sistemi ile çalışmaktadır.

Şirketler SEM sisteminde reklam vererek, SEO (Arama Motoru Optimizasyonu) sisteminde ise ilgili web sitesini arama motoru kriterlerine uygun şekilde düzenleyerek reklamlarda üst listelerde yer alabilirler (Cömert, 2016: 69)

1.2.3. Google AdWords Kavramları

Tıklama Sayısı: Hedeflenen kullanıcı sayısının reklamları kaç kez tıklayarak web sitesine yönlendirilme sayısıdır.

Tıklanma Oranı : Reklamı Gören kullanıcıların ilgili bağlantıyı kaç kez tıkladığını gösteren orandır. Tıklanma oranı, anahtar kelimelerin ve reklamların gösterdiği performansı ölçmede kullanılan orandır. (support.google.com, 2023).

Tıklanma oranı, bir web sayfasındaki bir reklamın tıklanma sayısının reklamın gösterim sayısına bölünmesi ile hesaplanır. Örneğin, 10 tıklanma ve 100 gösterim varsa tıklanma oranı %10' dur (Gupta ve Tuladhar, 2014:9).

Ortalama Tıklanma Başı Maliyet: Reklamlara yönelik gerçekleştirilen her bir tıklanma için ödenen ortalama ödeme tutarıdır.

Ortalama Tıklanma Başı Maliyet, toplam tıklanma maliyetinin, toplam tıklanma sayısına bölünmesiyle elde edilir (support.google.com, 2023).

Kalite Puanı: Reklam kalitesi rakip reklam verenlerle karşılaştırmasını yapan ve ne düzeyde olduğunu gösteren araçtır. Kalite Puanı 1-10 aralığında değer alır. Kalite puanının yüksek olması belirlenen anahtar kelimeyi arayan kullanıcı açısından reklamınız üst sıralarda yer almasını sağlayabilir (support.google.com, 2023).

Bin gösterim Başı Maliyet: Google reklamları kümesinde bin gösterim başı maliyet için ne kadar ödeme yapmak istenildiğini belirtilen tutardır.

Gösterim Sayısı: Reklamın Google arama motorunda gösterilen sonuçlarda ve bağlantı sağlanan web sitelerinde kaç kez görüntülediğidir.

Yatırım Getirisi: Reklam için yapılan yatırımın elde edilen kazanca oranıdır.

Anahtar Kelime: Sunulan ürün ya da hizmetin kullanıcıların yaptığı aramalar ile eşleştirilmesi için kullanılan kelime ya da kelime gruplarıdır. Uygun anahtar kelime kullanımı daha büyük kitlelere erişmek için oldukça önemlidir (Özbek, 2021:15).

Dönüşüm: Reklamlarda belirlenen hedefin dönüşünü ifade eder. Örneğin bir işletmenin kaç kez ziyaret edildiği, işletmeden ürün ya da hizmet alma gibi hedeflerin gerçekleşmesidir.

Dönüşüm Oranı: Reklam için belirlenen hedefleri gerçekleştiren ziyaretçilerin oranıdır. Web sitesini ziyaret eden kullanıcıların tüm kullanıcılara bölünüp 100 ile çarpımı sonucunda elde edilir (support.google.com, 2023).

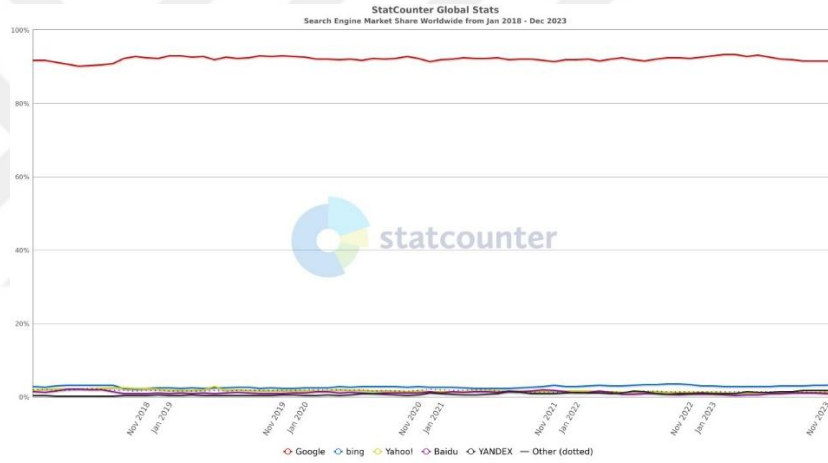
1.2.4. Google Arama Motoru Çalışma Prensibi

Arama motoru, internette bulunan milyonlarca web sitesi içerisinde kullanıcı tarafından aranan içerikle örtüşen sonuçları sıralayan bir tür yazılımdır (Aydemir, 2011). Arama Motorlarında arama işlemini üç adımda gerçekleştirir; tarama, dizin ekleme ve sunuştur (Ünal, 2022: 244) Arama motorlarında

Web sayfası URL' si web sitesini tanımlayan bazı anahtar sözcükler, web sayfasının oluşturulduğu yazılım kodları ve web sayfası içerisinde verilen linkler yer alır bu eklentiler indexlenerek veri tabanlarında saklanır. Bu yapının oluşturulması için örümcekler ya da bot adı verilen yazılımlar kullanır. (Eskicumalı ve Sarsılmaz, 2019: 197). Bu yazılımlar Web'in bağlantı(link) yapısını kullanırlar. Belirli aralıklarla web sitelerini ziyaret eder ve en son ziyaret ettikleri siteye kadar yapılan eylemleri büyük kapasiteli veri tabanlarına kaydederler, bu veri tabanı arama motorlarının indexidir. (Eskicumalı ve Sarsılmaz, 2019:197).

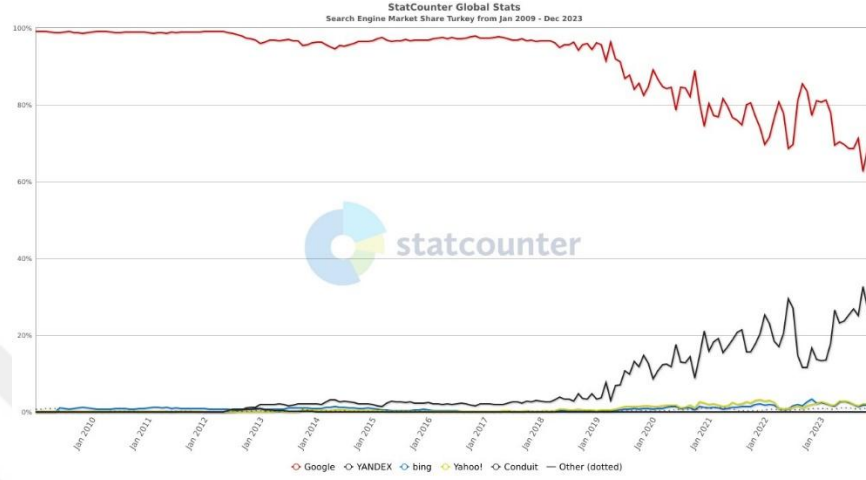
Kullanıcı internet üzerinden araştırma yaptığında, araştırma eylemi arama motoru dizinine aktarılır ve sonuç ekranı alaka düzeyine göre sıralanmış şekilde araştırmacıya sunulur. (Rouhola, 2021:6)

StatCounter, 2023 Kasım ayı güncel verilerinde Google' nin küresel pazarda payı %91,54, Türkiye pazarında %68,6 pazar payı ile sektördeki en yüksek paya sahiptir(<https://gs.statcounter.com/>).



Şekil 1:Son Beş Yılın Küresel Çapta Arama Motoru Pazar Payları

Kaynak: <https://gs.statcounter.com/> Erişim Tarihi: 21/12/2023



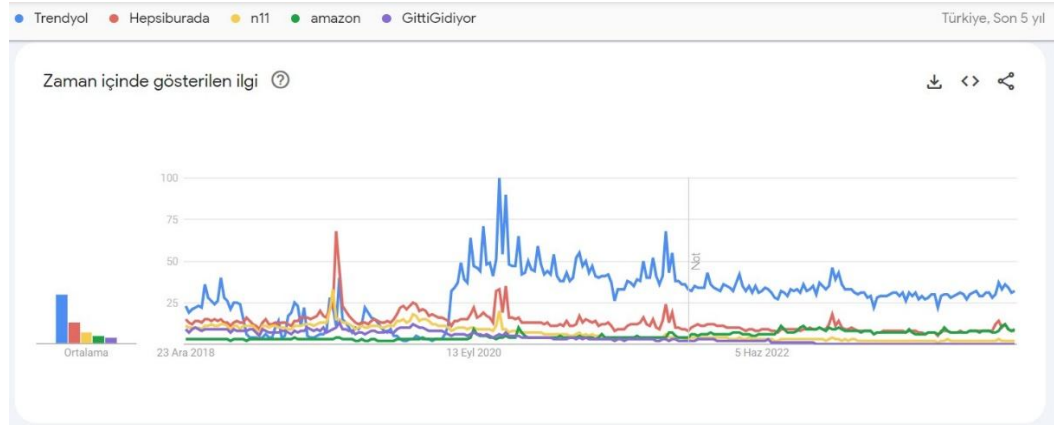
Şekil 2: Son Beş Yıl Türkiye Geneli Arama Motoru Pazar Payları
Kaynak: <https://gs.statcounter.com/> Erişim Tarihi: 21/12/2023

1.2.5. Google Trendler

Google Trends, küresel ölçekte gerçekleştirilen arama sorgularını coğrafi bölgeler ve diller temelinde analiz ederek, kullanıcı ilgilerinin zamansal ve bölgesel dağılımını istatistiksel olarak sunan web tabanlı bir analiz aracıdır (Prabhu, 2020).

Google Trends, internet kullanıcılarının Google arama motoru üzerinden gerçekleştirdiği sorgulara ilişkin eğilimleri; zamana, bölgeye ve konu başlıklarına göre analiz ederek sunan bir veri gözlemlleme aracıdır (Page, 2016). Platform, yıl içerisinde milyonlarca kez yapılan arama sorgularından elde edilen veriler doğrultusunda oluşturulmaktadır. Google Trends, kullanıcıların arama davranışlarına dair kişisel bilgileri sistemde bulundurmakta; bu veriler, kullanıcıların bulundukları bölgelere ve gerçekleştirdikleri sorguların temalarına göre sınıflandırılarak saklanmaktadır (Işık ve Koz, 2022: 27).

Google arama trendlerine erişim üç temel araç üzerinden sağlanmaktadır: Google Trends'in genel kullanıma açık web arayüzü, Google Trends API'si ve yalnızca belirli uzmanlık alanları için geliştirilen özel API'ler. Ana web arayüzü, tüm kullanıcıların erişimine açık genel verilere ulaşılmasını sağlarken, Google Trends Pazarlama API'si ise pazarlama ve özellikle e-ticaret alanlarında faaliyet gösteren profesyonellerin kullanımına sunulmaktadır. Bu özel API aracılığıyla, e-ticaret sektöründeki araştırmacılar belirli ürün veya kategoriye ilişkin kullanıcı arama davranışlarını daha ayrıntılı olarak inceleyebilmektedir (Işık ve Koz, 2022: 27). Ayrıca, sistem kullanıcı sorgularını 25 farklı konu başlığı altında filtreleme yeteneğine sahiptir. Bu başlıklar arasında alışveriş, sağlık, bilim, çevrim içi topluluklar ile bilgisayar ve elektronik ürünler gibi kategoriler yer almaktadır (Google Trends, 2023).(<https://trends.google.com>, 2023).



Şekil 3: Son Beş Yılda Google Arama Motorunda Türkiye’de Faaliyet Gösteren E-Ticaret Sitelerine Yönelik İlgi

Kaynak: <https://trends.google.com> Erişim Tarihi: 21/12/2023

1.3.Tüketici Kavramı ve Tüketici Davranışları

Tüketici, bireysel istek ve ihtiyaçlarını karşılamak amacı ile bir bedel karşılığında satın alıp, tüketen tüzel kişidir (Altunışık, Özdemir ve Torlak, 2006: 50). Günümüz pazarında tüketiciler farklı türlerdedir. Tüketici türlerinden bazıları fiyat ve performansa odaklanırken, bazıları tüketim esnası ve sonrasına odaklanır (Yaman, 2013: 80). E-ticarete ise tüketiciler satın alma süreci, kargolama süreci ve satın alma sonrası deneyim ve satın alan diğer tüketicilerin yorumlarına odaklanmaktadır.

Tüketici davranışları, tüketicilerin kişisel istek veya arzularını karşılama beklentisi ile yaptıkları satın alma, ürünü değerlendirme, tüketme ve tüketim sonrası elden çıkartma ile ilgili süreçlere etki eden faktörlerin incelendiği davranış bilimidir. (Solomon, 2009:33.). Tüketici ürün ya da hizmeti satın alma eylemini gerçekleştirmeden önce düşünür, alternatifleri belirler, belirlediği alternatifleri değerlendirir, seçenekler arası benzerlik ve farklılıkları karşılaştırır, geçmiş tecrübelerinden yararlanır ve satın alma eylemini gerçekleştirir. (Barış ve Odabaşı, 2003: 30)

Tüketici davranışlarının özellikleri;

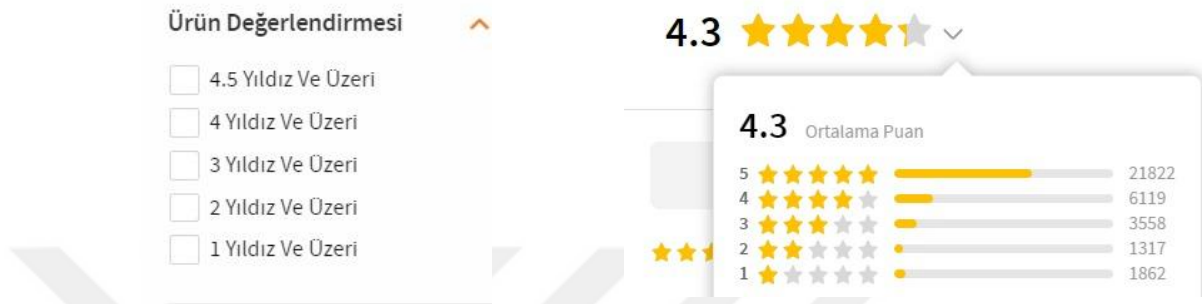
- Tüketici davranışı belirli bir amaca yönelik olarak şekillenir.
- Tüketici davranışı sürekli değişen ve gelişen bir süreçtir
- Tüketici davranışı çok yönlü bir yapıya sahiptir ve zaman içerisinde değişkenlik gösterir
- Tüketici davranışı, farklı eylemler ve karar mekanizmaları içerir
- Tüketici bulunduğu sosya-kültürel ve ekonomik çevreden etkilenir.
- Tüketici davranışı bireysel farklılıklara bağlı olarak değişkenlik gösterir

1.3.1. Kullanıcı Yorumu ve Yıldız Değerlendirme

Kullanıcı yorumları, bireylerin internet ortamında yaptığı alışverişler sonrasında satın aldıkları ürün ya da hizmet hakkındaki memnuniyet ya da memnuniyetsizliklerini ifade ettiği yorumlardır(Hwang, M. Kim ve T. Kim 2022: 676).Ürün değerlendirmeleri kullanıcı ürün/ hizmeti beğendi ise olumlu beğenmedi ise olumsuz ya da tarafsız tutumda olabilir. Ürün /hizmeti beğenen kullanıcı ürünün taahhüt edilen nitelikleri karşıladığını, satıcının sorunlara olumlu ve makul çözümler yarattığını ve kargolama sürecinin hızlı ve sağlam geldiği gibi memnuniyetlerini, ürün/ hizmeti beğenmeyen kullanıcılar ise beklentilerini karşılamadığını, satıcıdan ve kargolamadan memnun olmadıklarını çevrim içi platformda paylaşmamaktadır.

Mudambi ve Schuff, (2010) Müşteri analizlerini işletme/ diğer üçüncü taraf e-ticaret sitelerinde tüketicilerin ürün değerlendirmeleri olarak ifade etmektedirler.Her geçen gün artan internet kullanımı ve tüketicilerin fikirlerini beyan ettiği ortamların artması ile tüketicilerin doğru ve güvenilir bilgiye ulaşması güçleşmektedir. E- ticaret siteleri bu duruma alternatif çözüm olarak kullanıcı yorum ve değerlendirmelerini filtrelemede bazı yöntemleri kullanmaktadır.

Buna örnek olarak ise değerlendirdikleri ürün/ hizmete 1 ile 5 arasında derece vermelerini talep etmektedir. Burada 1 yıldız olarak değerlendirme ürün/ hizmeti hiç beğenmediklerini, 5 yıldız olarak değerlendirmede ise çok beğendiklerini ifade etmektedir. (Dokoohaki ve Krestel, 2011)



Şekil 4: Çevrimiçi Ürün Değerlendirme ve Yıldız Filtreleme
Kaynak: Trendyol.com Ürün Değerlendirmeleri Filtreleme Ekranı

Tüketicilerin ürün/ hizmet satın alma davranışı sürecinde kargo tüketiciye ulaştığında alışveriş yaptığı e- ticaret sitesi tarafından ürün hakkında, kargolama hizmeti ve satıcı hakkında yorum yapmaları için değerlendirme ekranı sunmaktadır.

Satıcıyı Değerlendir

2.98 Takip Et

Teslimat Tarihi:

Teslimat No:

Satıcıyı değerlendiriniz:

Kargoyu değerlendiriniz:

Yorum Yap

Yayınlanma Kriterleri

Fiyatı uygundu. Kurulumunda zorlanmadım ama ürün elime çok geç ulaştı.

0/2000

Ürün Fotoğrafı Ekle

En fazla 3 görsel ekleyebilirsiniz.

Gönder

Ürün Değerlendirmeleri

Ürünü aşağıdan puanlayabilir ve yorum yazabilirsin

Yorum Yap

Yayınlanma Kriterleri

Ürüne ilgili yorumlarını fotoğraf ekleyerek daha iyi açıklayabilirsin.

0/2000

Ürün Fotoğrafı Ekle

En fazla 3 görsel ekleyebilirsiniz.

Fotoğraf Ekle

Fotoğraf Ekle

Fotoğraf Ekle

Yorumlarda isminin gözükmesine ve yorum detaylarının site genelinde kullanılmasına izin veriyorum. Aydınlatma Metni'ne ulaşmak için tıklayınız.

Gönder

Şekil 5: Çevrim İçi Satın Alma Değerlendirme
Kaynak: Trendyol.com Satın Alınan Ürünü değerlendirme Ekranı

1.4. Duygu Analizi

İnternet, toplumda yaş ve meslek gözetmeksizin her birey tarafından görüş ve fikirlerini açıkça paylaştığı bir iletişim aracıdır. İnternet kullanımının her geçen gün daha da yaygınlaşmasıyla birlikte bireyler sosyal medya araçları ve e- ticaret platformlarında aldıkları ürünler hakkındaki fikirlerini, seyahat ettikleri yerleri ve güncel bir konu hakkındaki fikirlerini dijital platformlarda paylaşmaktadır.

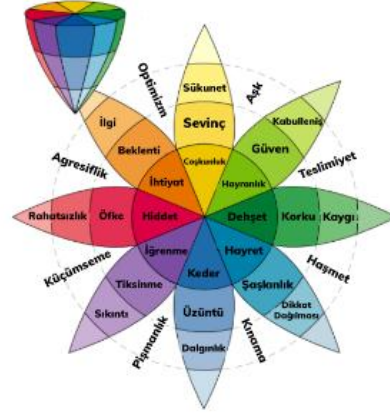
Günümüzde bireyler bir ürünü/ hizmeti satın almadan önce, bir yere seyahat etmeden önce, bir etkinliğe katılmadan önce diğer bireylerin ya da influencerların görüş ve önerilerine bakarak tercih yapmaktadır. Örneğin bir birey bir ürün/ hizmeti satın alırken ürün ya da hizmet hakkında daha önce yorum yapıldı mı, yapılan yorumlar olumlu yönde mi ya da olumsuz yönde mi diye araştırma yapmaktadır. Yapılan yorumların büyük bir kısmı olumlu yönde ise bireyin satın alma davranışını da o yönde etkilemektedir.

Bireylerin yaptığı yorumlar işletmeler açısından ürün/ hizmete yönelik tutum ve öneri niteliği taşımaktadır, ürün/ hizmeti o yönde yeniden düzenlen işletmeler satışlarını artırma yönünde pazarlama stratejileri geliştirerek iade maliyetlerini de en aza indirebilirler. (Chan vd, 2023)

NLP' de (Doğal Dil İşleme) duygu ve duyarlılık birbiriyle ilişkilidir fakat tam olarak aynı şeyi ifade etmezler, Duyarlılık, bir metinde aktarılan duygusal tonu veya tutumu ifade ederken, duygu belirli duygusal olguları ifade eder. Duygu, bir metinde yer alan tutumu olumlu, olumsuz ve nötr olarak ifade ederken, duygu analizi bu bilgiyi metinsel verilerde etiketleme ve çıkarma işlemidir. (Lee ve Pang, 2008)

Plutchik R. (2001) yılında yaptığı bir çalışmada duyguları birinci, ikincil ve üçüncül gruptandırmış ve sekiz duygunun birincil duyguyu ifade ettiğini belirtmiştir. Çalışmasında yer verdiği birincil duygular, güven, korku, şaşkınlık, beklenti, sevinç, üzüntü, tiksinti ve öfkedir.

Duygu analizinde, metin içerisinde duygu ifade eden sözcükler doğru ve düzgün olarak saptanmalıdır. Bu saptama yapılmadan önce belirli sözcük kütüphanelerinden ya da öznitelik seçimi yöntemi ile metin içerisindeki her bir sözcüğün hangi duyguyu ifade ettiği saptama işleminden yararlanılabilir.



Şekil 6: Plutchik Duygu Çarkı
Kaynak: (Plutchik R.,2001)

İKİNCİ BÖLÜM

2. VERİ, ALGORİTMA VE MAKİNE ÖĞRENMESİ

2.1. Veri

Veri (Data), bilgi parçacığının yapılandırılarak depolanabilen yapıya dönüştürülmesidir (Canber ve Sağıroğlu, 2006: 166). Gündelik hayatta veriler sınıflandırılarak analiz edilirler, sınıflandırma işlemi verinin türlerine göre yapılır. Bu durum makine öğrenmesi algoritmaları için de oldukça önem arz etmektedir.

2.1.1. Nümerik Değişkenler

Nümerik değişkenler, sayısal değer alabilen ve matematiksel işlemlerin yapılabildiği değişken tipleridir. Nümerik değişkenler, Aralık (Interval) ve Oran (Ratio) Ölçeği olarak iki grupta değerlendirilmektedir. Aralık ölçeği değişkenlerinde başlangıç değeri bulunmaz. Bu değişkende bulunan sıfır değeri yokluğu ifade etmemektedir. Sıcaklık değeri, IQ puanı, üniversite giriş sınav puanı aralıklı ölçek değişkenlerindendir (Can, 2019). Oran ölçeği, satın alınan ürün sayısı, aylık gelir miktarı kargo teslimat süresi (gün- cinsinden), boy uzunluğu, vücut ağırlığı ve yaş gibi nicel özelliklerin ölçülmesinde kullanılan bir veri düzeyidir. Bu tür değişkenlerde mutlak sıfır noktası anlamlıdır ve sıfır değeri, ölçülen özelliğin yokluğunu ifade eder (Gürbüz ve Şahin, 2015).

Sıfır değeri yalnızca başlangıç noktası olarak kabul edilir ve sıfır değeri alan değişken o olgunun ya da durumun mevcut olmadığını ifade eder (Harker ve Vargas, 1987).

2.1.2. Veri Ön İşleme

Veri elde edildiği kaynaktan alınırken ya da veri mevcut sisteme aktarılırken bazı sorunlar olabilir. Bu tarz sorunlar olmasa bile veri eksik ya da hatalı girilmiş olabilir, mevcut değer aralığından çok uzak değer ya da ifadeler alabilir.

Veri ön işleme işlemi, eksik, hatalı ya da yanlış verilerin düzenlenmesi aşamasında kullanılır ayrıca verilerin kullanılan algoritmayla uyumlu şekilde getirilmesi için de kullanılabilir (Arı, 2022: 25)

2.1.3. Veri Temizleme

Veri temizleme, verinin niteliğini arttırmak, veri içerisinde uygun olmayan verileri düzenlemek ve veri tutarlılığını sağlamak için yapılan birtakım işlemlerden oluşur (Crivat ,MacLennan ve Tang, 2005). Veri tek bir noktadaki veri tabanından sağlanıyorsa verinin yanlış değerlerde kaydedilmiş olması, eksik ya da hatalı kaydedilmesi, veri tipi uyumsuzluğu gibi sorunlarla karşılaşırken, Veri farklı veri tabanlarından sağlanıyorsa veri tabanlarının yanlış birleştirilmesi, veri standardının sağlanamamış olması gibi sorunları gidermek için veri temizleme işlemi yapılmaktadır (Do ve Rahm, 2000).

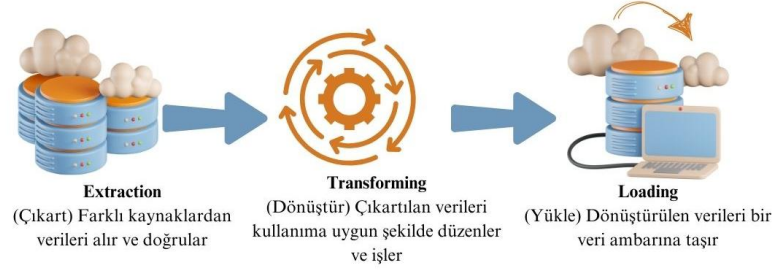
Veri yanlışlıkları, tek veri tabanından elde edilen ile bir çok veri tabanının birleşmesi ile oluşan veri setleri üzerinde farklılaşmaktadır. Tek bir veri tabanından elde edilen verilerde verilerin eşsiz olmaması, veri girişindeki yazım yanlışları, veri tutarlılığının sağlanmaması, aynı veri parçacığının birden fazla yazılması ve uç değerlerin bulunması gibi problemler yaşanabilmektedir. Birden çok veri tabanının birleşiminde ise tek veri tabanında olan sorunların yaşanmasının yanı sıra veri isimlendirme uyumsuzluğu, zaman senkronizasyonu sorunu, yanlış biçimde birleştirme ve yapısal uyumsuzluk yaşanabilmektedir.

İsimlendirme sorununa örnek olarak bir veri tabanında doğum yılı olarak girilmiş iken ötekinde doğum tarihi olarak girilebilmektedir.

Yapısal karmaşıklar, bir veri kümesinde doğum yeri ve doğum tarihi ayrı sütunlarda verilmişken, bir diğer veri kümesinde doğum tarihi ve yeri aynı sütünde verilmiş olabilir, bu veri kümeleri arasında yapısal uyumsuzluksorunları yaratır, bir veri kümesinde iki ayrı sütünü birleştirmesi ya da diğer veri kümesinde doğum yeri ve doğum tarihi sütunlarının ayrıştırılması gerekmektedir.

Yanlış biçimde birleştirme, bir veri tabanında veriler sırası ile adı, soyadı, doğum tarihi, doğum yeri, mesleği olarak etiketlenmiş iken ötekinde adı, soyadı, mesleği, doğum yeri, doğum tarihi şeklinde etiketlenmiş olabilir.

Bu veri tabanları birleştirilirken, etiketleme sırasına dikkat edilmeden yapılan entegrasyonlar, veri eşleşmesinde hatalara neden olabilir. Zamanlama uyumsuzluğu sorunu, veri kaynaklarının farklı zaman dilimlerinde güncellenmesinden kaynaklanabilir. Bu tür durumlarda, veri bütünlüğünü koruyabilmek amacıyla zaman temelli tutarlılığın sağlanması büyük önem taşımaktadır. (Do ve Rahm, 2000). Veri temizleme işlemi genellikle birkaç adımda gerçekleştirilir. Veri analizi, veri dönüşümünü sağlamak için iş akışı/ kurallarını saptama, doğrulama, dönüştürme ve temizlenmiş verinin geri gönderimi (feedback) işlemleri başlıca genel aşamalardır. Veri analizi süreci veri hatalarını ya da tutarsızlıkları saptama açısından önem arz eder. Verinin bütünü ya da bir parçası alınarak yapılan analiz ile üst veri (meta data)' ya ulaşılır ve ayrıntılı bir şekilde sunulur. Veri dönüşümü verinin elde edilmesi için bir çok veri tabanı kullanılmasıyla orantılı olarak veriler heterojen yapıda ya da dağınık yapıda olduğunda veri dönüştürme işlemi gerçekleştirilmeden önce uyumluluk analizi yapılmaktadır. Verileri homojen yapıya dönüştürmek için ELT (Extraction, Transforming, Loading) çıkar, dönüştür ve yükle sürecinden geçerler. Analiz aşamasında eksik veriler ya veri kümesi içerisinde çıkarılır ya da genel veri dağılımının ortalaması alınarak doldurulur. Dönüşüm aşamasından sonra veriler yüklenir (Do ve Rahm, 2000).



Şekil 7: ELT Süreci

2.1.4. Veri Dönüştürme

Veri Dönüştürme işlemi, veri kümesindeki veriyi farklı yapılara ya da farklı değerlere dönüştürme işlemidir. (Crivat ,MacLennan ve Tang, 2005).

Örneğin veri setinde müşterinin duygusunu ifade eden metin string yapıda iken algoritmanın onu anlayabilmesi açısından -1(olumsuz), 0(nötr) ve 1(olumlu) olarak integer yapıya dönüştürülmesi.

2.2. Makine Öğrenmesi

Günümüzde verinin hacmi, geleneksel yöntemlerle işlenmesi mümkün olmayacak düzeyde hızla artış göstermektedir. Bu denli büyük veri kümelerinin makineler aracılığıyla işlenmesi, işlemleri hem daha hızlı ve hatasız hâle getirmekte hem de insan gücüne duyulan ihtiyacı en aza indirmektedir

(Schmidhuber, 2015).

Samuel (1959) Makine öğrenmesini “bilgisayarlara açıkça programlamadan öğrenme yeteneği veren bir çalışma alanı” olarak ifade etmiştir.

Makine öğrenmesi; sağlık alanında hastalıkların erken saptanması ve tanısının konulması ile tedavi yaklaşımlarının başarı düzeylerinin hesaplanmasında (Sevli, 2019), tarım sektöründe üretim verimliliğinin artırılmasında (Arıkan, Bilban ve Uzun, 2018), hukuk alanında dava süreçlerinin analiz edilmesi, yorumlanması ve olası sonuçların öngörülmesinde (Kemaloğlu, Küçüksille ve Turan, 2020), müşteri kaybı olasılığının tahmin edilmesinde (Kaynar ve diğerleri, 2017), müşteri şikayetlerinin sınıflandırılmasında ve tüketici satın alma davranışlarının öngörülmesinde (Yaman, 2021) ile eğitimde öğrenci performansının tahminine yönelik uygulamalarda (Selvi, 2020) olmak üzere çok çeşitli alanlarda etkili bir şekilde kullanılmaktadır.

Sahip olduğu esneklik ve hesaplama gücü sayesinde, makine öğrenmesi yöntemlerinin uygulama alanları sürekli olarak genişlemekte ve disiplinler arası araştırmalarda önemli bir yer edinmektedir.

Makine öğrenmesi algoritmaları denetimli ve denetimsiz olmak üzere iki temel gruba ayrılmaktadır. Denetimli makine öğrenmesi algoritmalarında her bir parametrenin etiketlenilmiş başka bir parametre üzerinden eğitilerek yeni gelen değişken üzerinde tahminde bulunmaktadır. Algoritmanın tahminleme yaparken ne derece doğru tahmin yaptığını, yaptığı tahminin hangi doğrultuda olduğunu ölçümlemede birden fazla teknik kullanılmaktadır (Arı, 2022: 27). Denetimli öğrenme algoritmalarından en yaygın kullanılanları, Lojistik Regresyon (Logistics Regression), Naive Bayes, Karar Ağaçları (Decision Tree), K-En Yakın Komşu Algoritması (KNN-K-Nearest Neighborhood), Rastgele Ağaçlar (Random Forest), Gradient Boosting Classifier ve XGB Classifier'dır. Denetimli makine öğrenmesi algoritmalarına örnek olarak, Müşterinin verilen banka kredisini ödemeyeceği, Islak imza tanıma, görüntü tanıma, hastalık tespiti verilebilir (Arslan,2020).

Denetimsiz öğrenme, etiketsiz veriler arasındaki belirsiz bağıntıyı keşfederek gruplandırmaktadır. Denetimsiz öğrenme algoritmaları tahminlemede bulunurken algoritmaların belirli bir sonuç üretmemesinden dolayı başarı sınaması yapılması oldukça güçtür. (Lecun vd., 2015; Pedregosa vd., 2011).

Takviyeli Öğrenme (RL), denetimli ve denetimsiz öğrenme algoritmalarından farklılaşan yönü tahminlerde bulunurken referans bir nokta belirlenir ve algoritmanın bu hedef için tahminlerde bulunması beklenir. Doğada yaygın olarak kullanılan bu öğrenme yöntemi ceza ve ödüle dayanır. Tahminler hedefe yakın sonuçlar veriyorsa ödüllendirilir, tahminler hedefe uzak sonuçlar verir ise cezalandırılır. (Shakya, Pillai ve Chakrabarty, 2023)

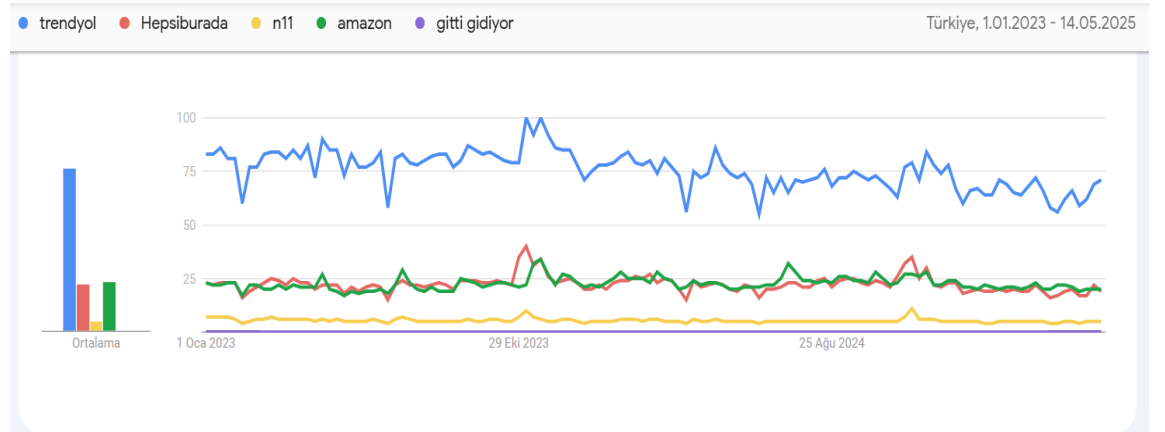
Aşırı Öğrenme (Overfitting), Algoritmanın belirli bir veri setinin sonuçlarını öğrenerek yalnızca o veri setinde başarılı sonuçlar vermesi sorunsalıdır. Bu problem uygulanan modelin üzerinde öğrendiği veri setinde olumlu sonuçlar verirken bilmediği veri setinde olumsuz sonuçlar üretir. Aşırı öğrenme değişken sayısının fazla ve farklı yapıda olmasından, veri sayısının yeterli sayıda olmamasından dolayı olabilir. (Şenol Çelik, 2022: 11) Aşırı öğrenmeyi engellemek için değişken sayısını minimal düzeye getirme, veri setindeki veri sayısını arttırma ve veri üzerinde bazı düzenlemeler yapılabilir.

ÜÇÜNCÜ BÖLÜM

MODEL, YÖNTEM, ANALİZLER

3.1 VERİLERİN TOPLANMASI

Bu araştırmada incelenen veriler, Türkiye'deki en popüler e-ticaret platformlarından biri olan Trendyol'un son üç yılda Google Trends verilerine göre sürekli yüksek düzeyde aranma hacmine sahip olması, Türkiye e-ticaret pazarındaki güçlü pazar payı, geniş kullanıcı kitlesi ve sosyal medya entegrasyonundaki etkinliği doğrultusunda seçilmiştir. Platform, dijital reklamcılık alanında yenilikçi stratejiler benimseyerek yalnızca geleneksel televizyon reklamcılığıyla sınırlı kalmamakta, aynı zamanda sosyal medya kanalları üzerinden yüksek etkileşimli ve hedef odaklı reklam kampanyaları yürütmektedir. Bu durum, Trendyol'u yalnızca bir çevrim içi alışveriş platformu olmanın ötesine taşıyarak dijital pazarlama alanında öne çıkan bir örnek haline gelmektedir.



Şekil 8:Google Trendlerde Son 3 Yıl İçerisinde Aratılan E- ticaret Sitelerinin Karşılaştırılması

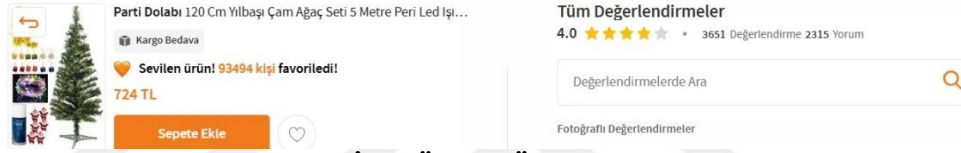
Kaynak: <https://trends.google.com> Erişim tarihi: 14/05/2025

Şekil 8’ de görüldüğü üzere, “Trendyol” anahtar kelimesi 1 Ocak 2023 – 14 Mayıs 2025 tarihleri arasında Türkiye genelinde diğer e-ticaret platformlarıyla kıyaslandığında istikrarlı biçimde en yüksek aranma hacmine sahip olmuştur. Özellikle kampanya dönemlerinde örneğin “Efsane Cuma”, “11.11” ve “Anneler Günü” gibi özel günlerde bu aranma yoğunluğunun kayda değer biçimde arttığı gözlemlenmektedir. Buna karşın, Hepsiburada, n11, Amazon ve GittiGidiyor gibi rakip platformların aranma hacimlerinin daha sınırlı ve dalgalanmasız bir seyir izlediği dikkat çekmektedir. Bu eğilim, Trendyol’un dijital farkındalık düzeyinin yüksekliğini ve tüketici zihninde güçlü bir marka algısı oluşturduğunu ortaya koymaktadır. Bu bağlamda, Trendyol, çalışmada kullanılan verilerin elde edildiği başlıca dijital ticaret platformu olarak değerlendirilmiş; dijital pazarlama stratejilerinin etkinliğini ve çevrim içi tüketici davranışlarını analiz etmek için anlamlı bir veri kaynağı olarak ele alınmıştır.

Bu veriler göz önünde bulundurularak özel günler olarak anneler günü, sevgililer günü, Yeni yıl seçilerek kategorileri ise ev ve yaşam, hobi ve elektronikler.

Veri seçimi aşamasında, elektronik, ev ve yaşam gibi kategorilerin tercih edilmesinin temel gerekçesi, bu alanların özel günlerde tüketici harcamalarının yoğunlaştığı, yüksek hacimli ve çeşitlilik arz eden ürün gruplarını içermesidir. Ayrıca, hobi kategorisinin dahil edilmesi, özellikle özel günlerde hediye ve kişisel ilgi alanlarına yönelik tüketim davranışlarının analizine olanak sağlamaktadır. Seçilen özel günler; Anneler Günü, Sevgililer Günü ve Yeni Yıl ise Türkiye’de tüketici harcamalarının önemli ölçüde arttığı dönemler olması nedeniyle, dijital pazarlama kampanyalarının ve tüketici geri bildirimlerinin yoğunlaştığı zaman dilimlerini temsil etmektedir. Bu kapsamda, belirlenen kategori ve zaman dilimleri, tüketici davranışlarını kapsamlı biçimde değerlendirmek ve dijital ticaretin dinamiklerini derinlemesine incelemek için uygun ve stratejik bir temel oluşturmuştur.

Yeni yıl için en çok tercih edilen ürün olan çam ağacı tercih edilmiştir. Ürünün 3651 değerlendirme ,2315 yorumu ve 4.0 yıldızı bulunmaktadır. Sevgililer günü için en çok tercih edilen ürün olan çiçek buketi tercih edilmiştir. Ürünün 944 değerlendirmesi, 550 yorumu ve 4.3 yıldız oranı bulunmaktadır. Elektronik kategorisinde en çok satılan ürün olarak tablet tercih edilmiştir. Ürünün 1721 değerlendirmesi, 1031 yorumu ve 4.7 yıldızı bulunmaktadır.



Şekil 9: Çalışma Modeli İçin Örnek Ürün Değerlendirmesi

3.2 Çalışmanın Modeli ve Uygulanması



Şekil 10: Çalışmanın Modeli ve Uygulanması

Çalışmanın modeli verilerin çevrim içi e- ticaret sitesinden kullanıcı yorumları ve yıldız değerlendirmeleri alınarak CSV uzantılı dosya formatına dönüştürülerek veri toplama süreci tamamlanacaktır.

Verilerin temizlenme ve ön işleme süreci verinin normalizasyon işlemi uygulanarak makine öğrenmesi için uygun formata dönüştürülmesi hedeflenmektedir. Normalize edilmiş verilerin duygu durumu (pozitif ise 1- negatif ise 0) olarak etiketlendirilecektir. Veri ön işleme sürecinden geçen veri kümesi eğitim ve test verisi olarak ikiye ayrılarak.

Eğitim verisi ile algoritmanın öğrenmesi ve test verisi ile de algoritma başarı durumu değerlendirilecektir.

3.2.1 Verinin Toplanması

Makine öğrenmesi algoritmasının eğitilmesi ve analizi için e- ticaret sitelerini en aktif kullanan yaş skalası olan üniversite öğrencilerinin özel gün ve kategori tercihleri referans alınarak, Şekil 11' de belirtildiği gibi en çok ziyaret ettikleri e- ticaret sitesi olan Trendyol tüketici yorumları veri setini oluşturacaktır. Belirlenen kategorideki veri Python programlama dilinde yazılan kod ile e- ticaret sitesi otomatik olarak açılıp, sayfayı aşağı doğru kaydırıp tüm yorumları toplayarak CSV uzantılı bir dosyaya kaydederek veri setini oluşturacaktır. Web sayfasının içindeki yazıları etiketlendirmek için BeautifulSoup kütüphanesi, verileri tablo halinde saklamak için ve dosyaya yazmak için Pandas kütüphanesi, sayfayı otomatik olarak kontrol edebilmek içinse Selenium kütüphanelerinden yararlanılacaktır. Kod Fotoğraflı yorumlarda yalnızca metne odaklanan görsel değerlendirmeyi metinsel olarak kaydedecektir.

3.2.2 Veri Temizleme ve Ön İşleme

Veri Ön işleme süreci için Microsoft Visual Studio Code 1.99.3 ortamında Pandas kütüphanesinden yararlanılarak veriler normalize edilecektir. Olumlu veriler için 1 olumsuz veriler için 0 nötr veri için değer girmeden etiketlendirme yapacaktır.

Bu işlemi yaparken Türkçe karakterlere dikkat ederek onları dilimizin kurallarına uygun hale getirmesi için UTF-8 kullanılacaktır.

Duygu etiketlendirme yapılırken BERT kütüphanesinden yararlanılacaktır. Veri Ön işleme adımlarından sonra verinin son hali Tablo 2’de verilmiştir.

	comment	duygu_metin	duygu_numeric
41	görseldeki her ş...	olumlu	1
42	çok güzel kulla...	olumlu	1
43	ürün sorunsuz ...	olumlu	1
44	süper çocuklar ...	olumlu	1
45	küçük salonlar i...	olumlu	1
46	bayıldım çocukl...	olumlu	1
47	eksiksiz elime u...	olumsuz	0
48	boyut ölçülerle ...	olumlu	1
49	çocuklar heves ...	olumlu	1
50	güzel beğendim	olumlu	1

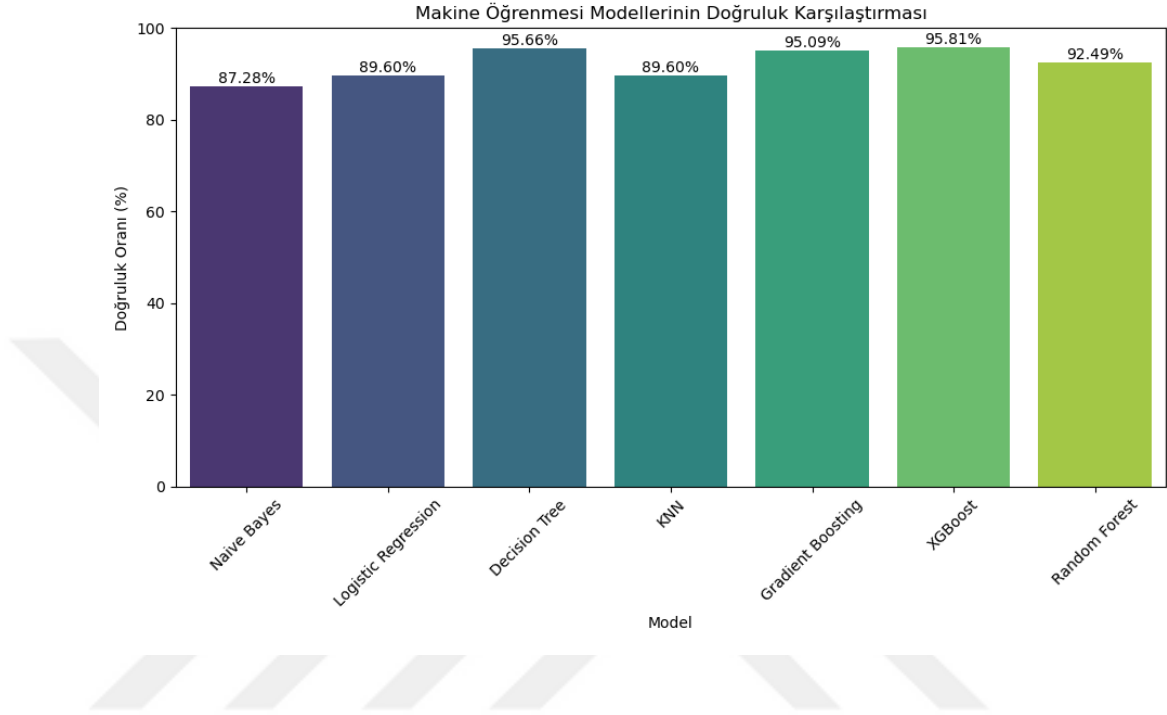
Tablo 2:Veri Ön İşleme Sonucu

3.2.3 Veri Setinin Eğitilmesi

Naive Bayes, Lojistik Regresyon, Karar Ağacı, K-En Yakın Komşu (KNN), Gradyan Artırma, XGBoost ve Rastgele Orman algoritmalarının performans analizinde kullanılacak olan model, ön işlem aşamalarından geçirilmiş veri kümesi temel alınarak oluşturulacaktır. Hazırlanan eğitim verisi, Python programlama dili içerisinde yer alan scikit-learn (sklearn) kütüphanesinin `train_test_split` işlevi kullanılarak; %70'i eğitim, %30'u test olmak üzere iki alt kümeye ayrılacaktır. X ile gösterilen veri kümesinde "comment (kullanıcı yorumu)", Y ile temsil edilen veri kümesinde ise "duygu_comment (0-1 arasında sayısal biçimde etiketlenmiş)" sütunu bulunmaktadır.

Algoritmanın, `X_train` niteliği ile `y_train` hedef değişkeni arasındaki örüntüyü öğrenmesi ve daha önce görmediği test verileri üzerinde doğru sınıflandırmalar yapması amaçlanmaktadır. Belirtilen algoritmalar, oluşturulan eğitim kümesiyle eğitilecek ve elde edilen doğruluk değerlerine göre aralarındaki başarı düzeyleri karşılaştırılacaktır.

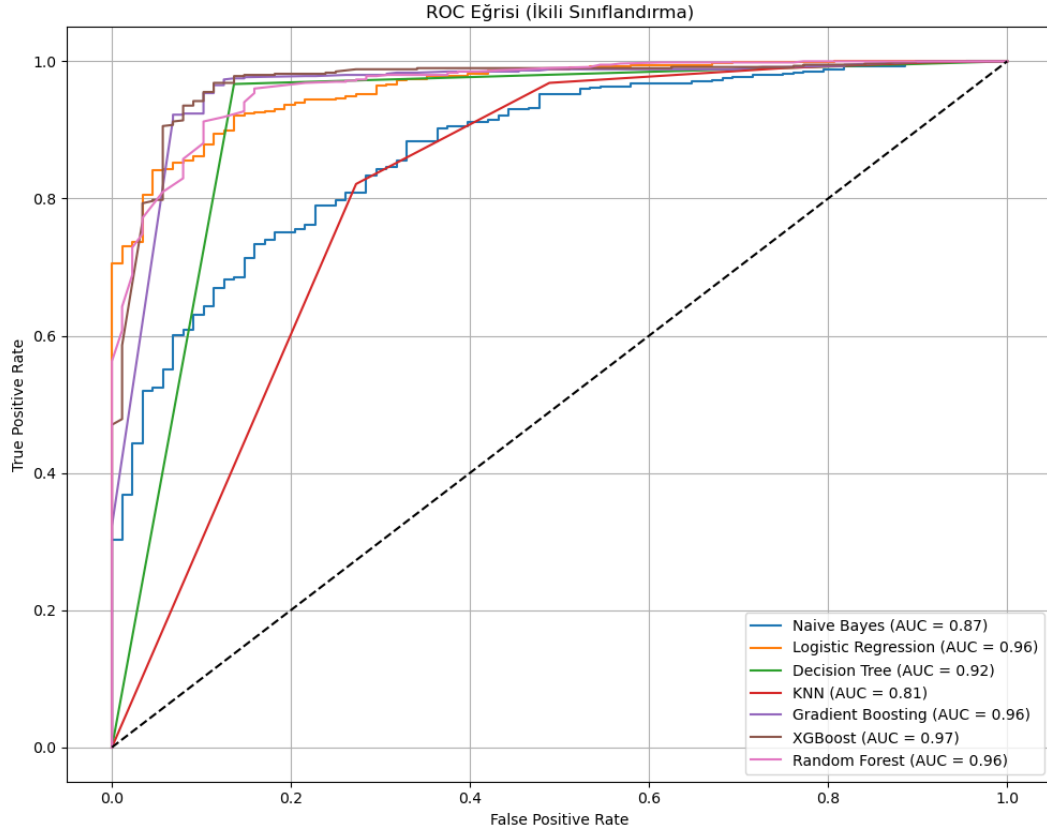
Öğrenme setinde yer alan veriler rastgele dağılmaktadır ve çam ağacı ürünü için toplamda 2305, sevgililer günü için yapay led ışıklı çiçek buketi isimli ürün için 551 değerlendirme ve tablet ürünü için 1032 değerlendirme bulunmaktadır. Çam ağacı ürünü için algoritmaların test verisi üzerindeki başarısı Şekil 11' de yer almaktadır.



Şekil 11: Çam Ağacı Ürününün Makine Öğrenmesi Doğruluk Oranları

Algoritmaların doğruluk oranına göre XGBoost en yüksek doğruluk oranına sahip model %95,81 iken Naive Bayes %87,28 ile en düşük doğruluğa sahip algoritmadır.

Karar Ağacı ve Gradyan Arttırma' da %95'in üzerinde doğruluk oranlarıyla çok başarılı performans göstermiştir Buna karşın Naive Bayes %87,28 oranı ile en düşük doğruluğa sahiptir, benzer şekilde Lojistik Regresyon ve KNN modelleri %89,60 ile düşük doğruluk oranlarına sahiptir. Doğruluk oranı, algoritmanın yapmış olduğu tahminlerin ne derece doğru olduğunu ifade etmektedir. ROC- AUC oranları Şekil 12' de yer almaktadır.



Şekil 12: Çam Ağacı Ürünün ROC-AUC Oranları

ROC-AUC oranları algoritmanın, test setinde bulunan “olumlu ve olumsuz”duygu etiketlerini ne derece doğru tahmin edebildiğini göstermektedir. En yüksek AUC değerine sahip algoritma %97 (0.97) ile XGBoost algoritmasıdır.

KNN algoritması %81 (0,81) ve Naive Bayes %87 (0.87) diğer algoritmalar içerisinde en düşük AUC- ROC oranına sahip algoritmalarıdır.

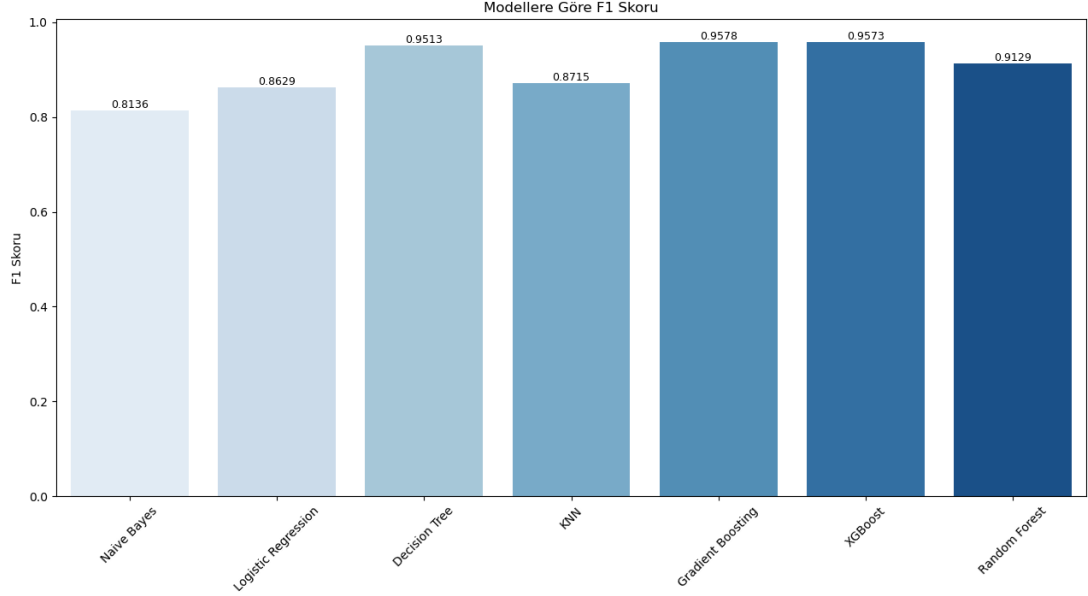
Algoritmaların doğruluk oranlarının yüksek başarı sağlaması ile doğru orantılı olarak algoritmaya bağlı olan eğrilerinde grafikte kapladıkları alanda artış görülmektedir. Şekil 12’ de yer aldığı gibi en yüksek orana sahip olan XGBoost en büyük alana sahip iken, en düşük orana sahip olan KNN algoritması en düşük alana sahiptir. ROC- AUC oranlarının yüzdelik gösterimi Tablo 3’ de yer almaktadır.

Tablo 3:Çam Ağacı Ürünü AUC- ROC Oranları(%)

Model	Doğruluk (%)	AUC-ROC Skoru
Naive Bayes	87.283237	0.871538
Lojistik Regresyon	89.595376	0.960923
Decision Tree (Karar Ağacı)	95.086705	0.913606
KNN	89.595376	0.810158
Gradient Boosting (Gradyan Arttırma)	95.809249	0.960651
XGBoost	95.809249	0.969634
Random Forest (Rassal Orman)	92.485549	0.964084

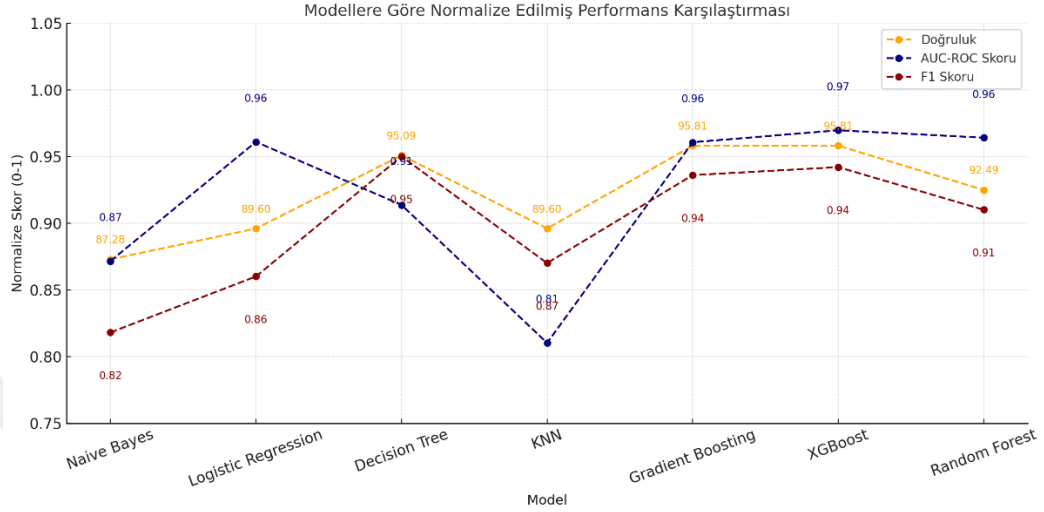
AUC (Area Under Curve) değeri 1'e ne kadar yakın değere sahip ise, model o kadar iyi sınıflandırma yapmaktadır.

Algoritmalara ait F1 skoru Şekil 15' de yer almaktadır.



Şekil 13: Çam Ağacı Ürünü F1 Skoru

Gradyan Arttırma (%95,78) ve XGBoost(%95.73), F1 skorları açısından en başarılı modeller olarak öne çıkmasına karşın Naive Bayes, en düşük F1 skoruna sahip modeldir (0.8136). Algoritmaların doğruluk oranı, duyarlılık ve F1 skorları önemli metrikler olsa da bu metrikler tek başına sınıflar arası hata dağılımı hakkında bilgi vermez. Bu nedenle, hata matrisleri, her bir algoritmanın sınıflandırma hatalarını detaylı görmek için kritik bir araçtır.



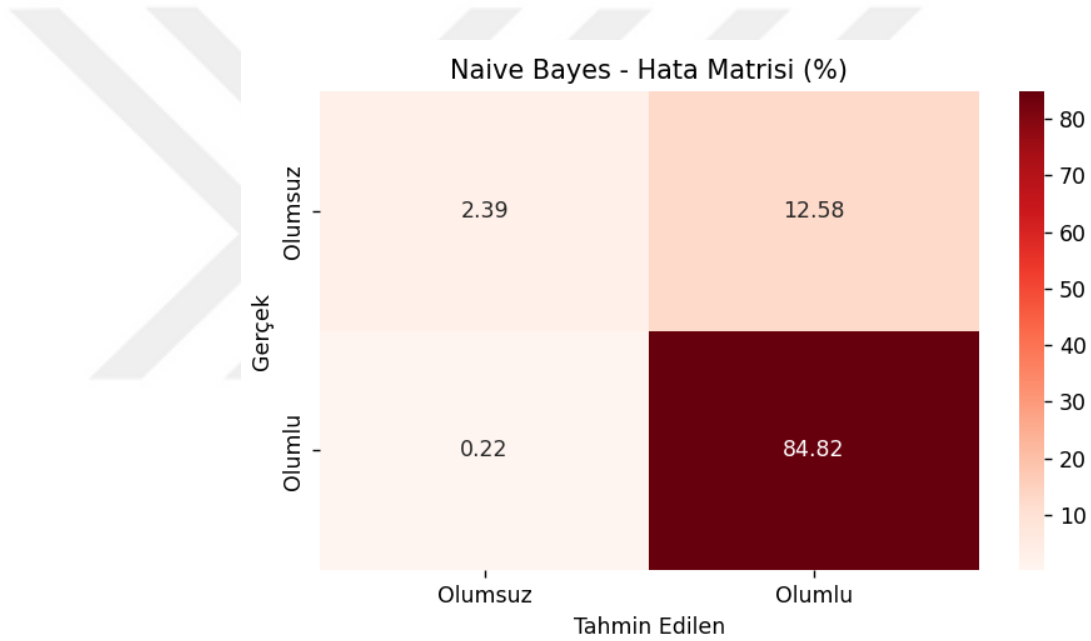
Şekil 14: Modellere Göre Normalize Edilmiş Performans Karşılaştırılması

Algoritmalar arasında elde edilen en düşük doğruluk değeri %87,28 olup, ROC eğrisi altında kalan alan oranı %81, F1 skoru ise %82 olarak hesaplanmıştır. Modellerin karar mekanizmaları ve mimari çeşitlilikleri dikkate alındığında, ortalama olarak %92,39 doğruluk, %93,57 AUC (ROC eğrisi altında kalan alan) ve %88,29 F1 skoru elde edilmiştir. Genel başarı ortalamasına bakıldığında en iyi performansı XGBoost algoritması göstermiştir; ancak, F1 skorunun görece düşük kalmasından dolayı en zayıf sonuçlardan biri Naive Bayes algoritması tarafından üretilmiştir. F1 skoru ile ROC eğrisi altındaki alan değeri, algoritmanın öngörülerinin doğruluğunu ve sınıflandırma kabiliyetini yansıtmaktadır. Algoritmalarından herhangi birinin veri kümesindeki tüm girdileri “olumlu” olarak etiketlemesi durumunda, sistem %92 doğruluk oranına ulaşabilirdi. Çünkü incelenen veri kümesinde %92 oranında olumlu, %8 oranında ise olumsuz geri bildirim yer almaktadır.

Buradan yola çıkılacak olunur ise F1 skoru ve ROC eğrisi altında kalan oranı %0 olacaktır. Bu yüzden algoritmanın değerlendirilmesinde yalnızca doğruluk oranı yeterli değildir.

3.2.4.1 Naive Bayes Tahmin Hata Matrisi Değerlendirmesi

Naive Bayes algoritmasının doğruluk oranı %87,28, ROC eğrisi altında kalan oranı %0.8715 ve F1 skoru %81,36'dır. Algoritmanın Hata matrisi Şekil 17'de yer almaktadır.

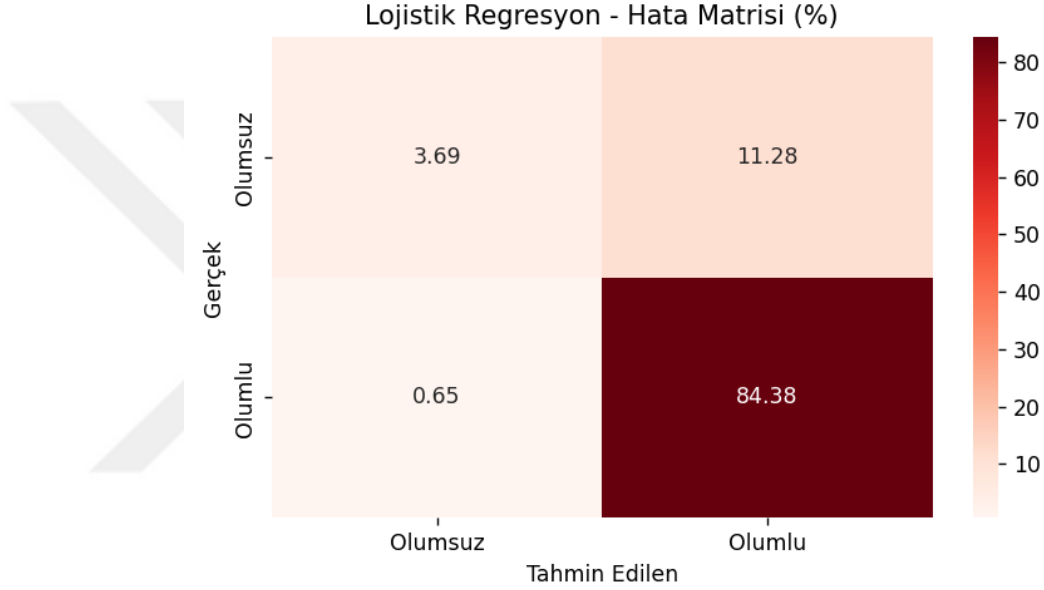


Şekil 15: Çam Ağacı Ürünü Naive Bayes Tahmin Hata Matrisi

Model, müşterilerin olumlu yönde yaptığı yorumların %84,82'sini doğru şekilde olumlu olarak, gerçekte olumsuz olan yorumların ise %2,39'unu doğru şekilde olumsuz olarak sınıflandırmıştır. Buna karşılık, müşterilerin olumsuz yöndeki yorumlarının %12,58'i yanlışlıkla olumlu olarak sınıflandırılmıştır. Öte yandan, gerçekte olumlu olan yorumların yalnızca %0,22'si hatalı şekilde olumsuz olarak tahmin edilmiştir.

3.2.4.2 Lojistik Regresyon Tahmin Hata Matrisi Değerlendirmesi

Lojistik Regresyon algoritmasının doğruluk oranı %89,60, ROC eğrisi altında kalan oranı %96 .09 ve F1 skoru %86,29’dur. Hata matrisi Şekil 18’ de yer almaktadır.

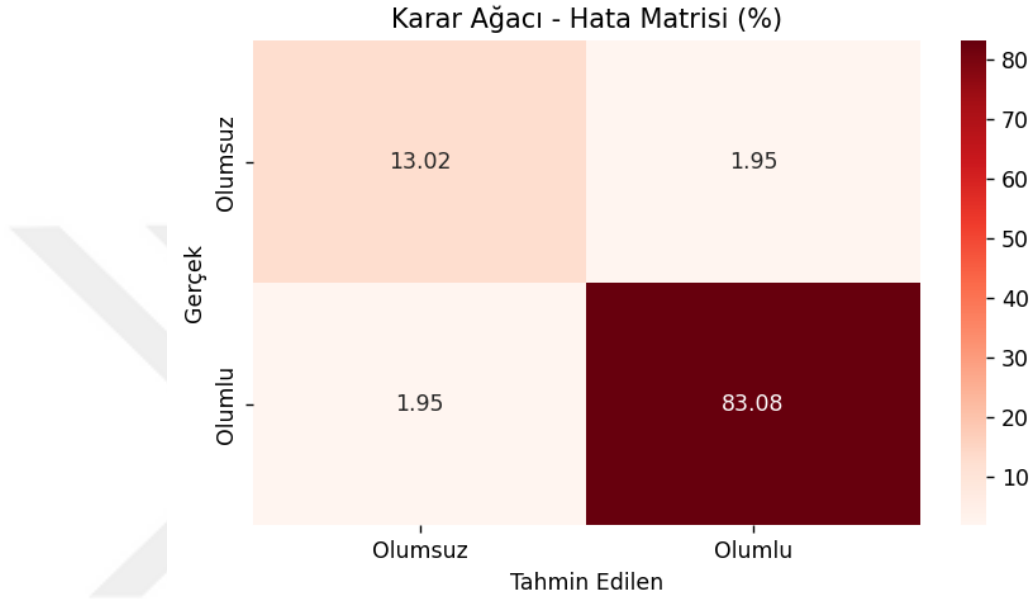


Şekil 16: Çam Ağacın Ürünü : Lojistik Regresyon Tahmin Hata Matrisi

Lojistik Regresyon modeli, müşterilerin olumlu yönde yaptığı yorumlarının %84,38’ini doğru biçimde olumlu olarak sınıflandırabilmiştir. Buna karşılık, modelin olumsuz sınıfa yönelik ayırt edici gücü oldukça sınırlı kalmış ve yalnızca %3,69 oranında doğru sınıflandırma gerçekleştirebilmiştir. Gerçekte olumsuz olan yorumların %11,28’i model tarafından hatalı biçimde olumlu sınıflandırılırken, olumlu sınıfa ait yorumların yalnızca %0,65’i yanlışlıkla olumsuz olarak tahmin edilmiştir.

3.2.4.3 Karar Ağacı Tahmin Hata Matrisi Değerlendirmesi

Karar Ağacı algoritmasının doğruluk oranı %95.08, ROC eğrisi altında kalan oranı %91,36 ve F1 skoru %95,13’tür. Hata matrisi Şekil 19’ da yer almaktadır.

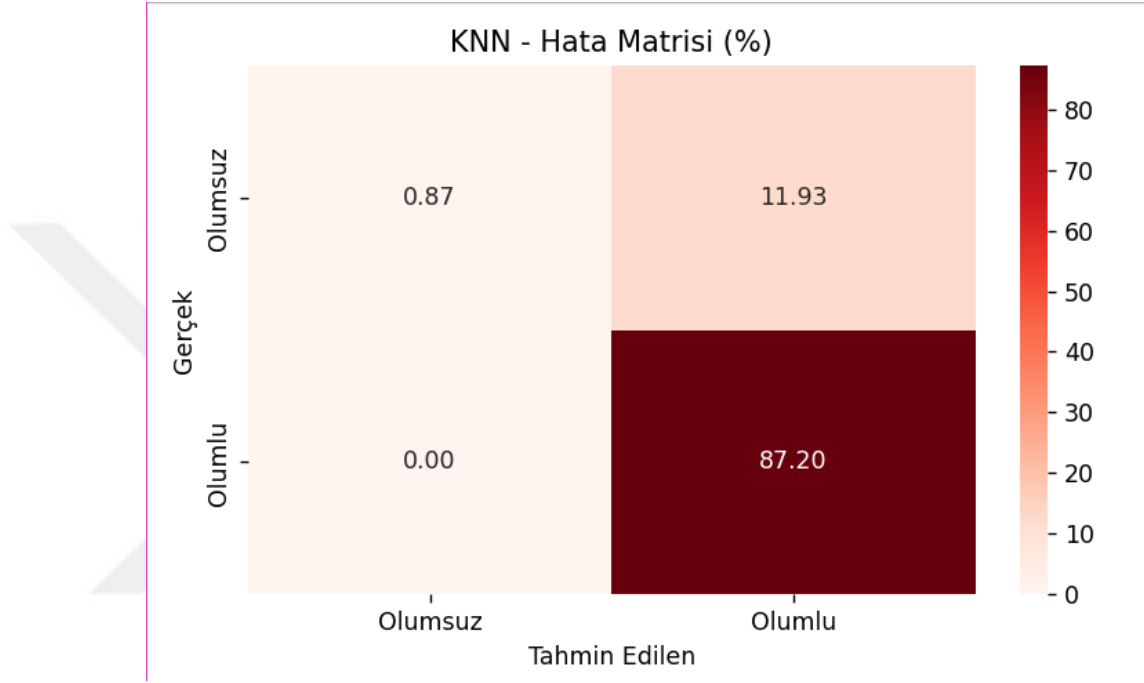


Şekil 17: Çam Ağacı Ürünü Karar Ağacı Tahmin Hata Matrisi

Karar Ağacı modeli, müşterilerin olumlu yönde yaptığı yorumların %83,08'ini doğru biçimde olumlu olarak sınıflandırabilmiştir. Benzer şekilde, modelin olumsuz sınıfı tanıma performansı da dikkate değer düzeyde olup, gerçekte olumsuz olan yorumların %13,02'si doğru şekilde olumsuz olarak tahmin edilmiştir. Olumsuz yorumların yalnızca %1,95'i hatalı biçimde olumlu olarak değerlendirilmiş; olumlu yorumların da aynı oranda (%1,95) yanlışlıkla olumsuz olarak tahmin edilmiştir.

3.2.4.4 KNN Algoritması Tahmin Hata Matrisi Değerlendirmesi

KNN algoritmasının doğruluk oranı %89,59, ROC eğrisi altında kalan oranı %81,01 ve F1 skoru %87,15'tir. Hata matrisi Şekil 20' de yer almaktadır.

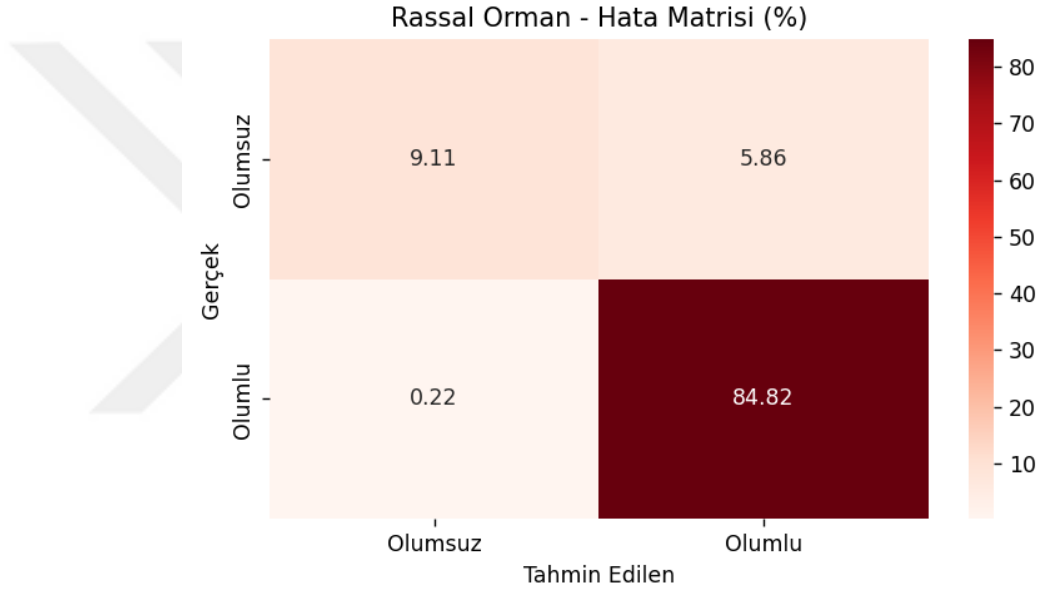


Şekil 18: Çam Ağacı Ürünü KNN Algoritması Hata Tahmin Matrisi

KNN algoritması, müşterilerin olumlu yönde yaptığı yorumların %87,20'sini doğru şekilde olumlu sınıfa atayarak, pozitif sınıfı tanıma konusunda güçlü bir performans ortaya koymuştur. Öte yandan, modelin olumsuz sınıfa yönelik duyarlılığı daha sınırlı kalmış ve gerçekte olumsuz olan yorumların yalnızca %0,87'si doğru biçimde olumsuz olarak tahmin edilebilmiştir. Bu, modelin negatif sınıfı tanımakta zorlandığını ve sınıf dengesizliği karşısında pozitif sınıfa eğilim gösterdiğini ortaya koymaktadır. Modelin yaptığı sınıflandırma hatalarına bakıldığında, olumsuz yorumların %11,93'ünün hatalı biçimde olumlu olarak değerlendirilmiş olması dikkat çekicidir. Buna karşın, olumlu yorumların hiçbirinin (%0,00) yanlışlıkla olumsuz sınıfa atanmamış olması, modelin pozitif sınıfı ayırt etme konusundaki yüksek doğruluğunu pekiştirmektedir.

3.2.4.5 Rassal Orman Tahmin Hata Matrisi Değerlendirmesi

Rassal Orman algoritmasının doğruluk oranı %92,48, ROC eğrisi altında kalan oranı %96,40 ve F1 skoru %91,29'dur. Hata matrisi Şekil 21' de yer almaktadır.

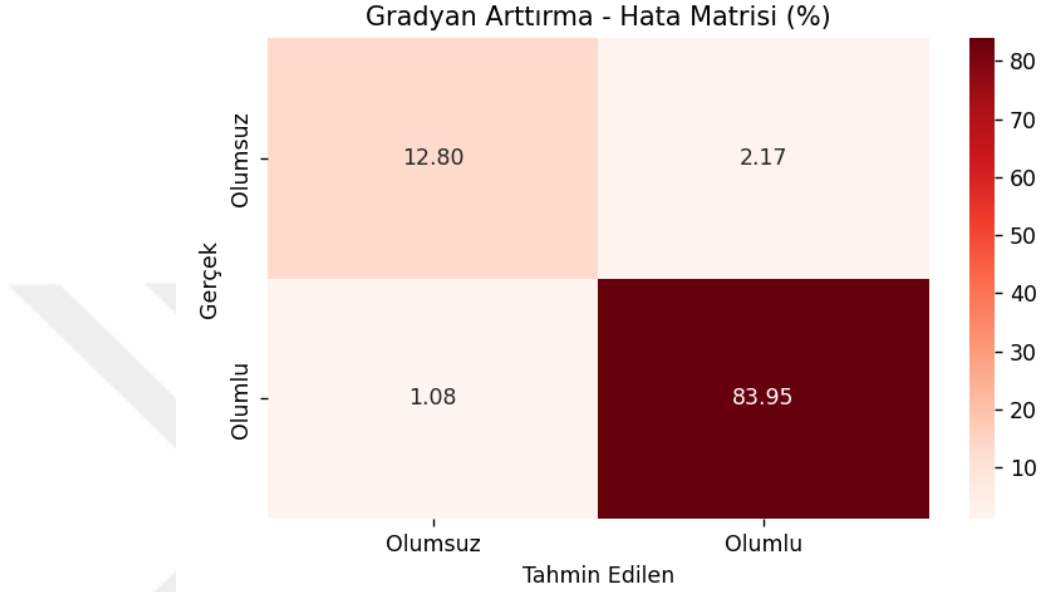


Şekil 19: Çam Ağacı Ürünü Rassal Orman Tahmin Hata Matrisi

Rassal Orman modeli, müşterilerin olumlu yönde yaptığı yorumların %84,82'sini doğru bir şekilde olumlu olarak tahmin etmiştir. Aynı şekilde, model gerçekte olumsuz olan yorumların %9,11'ini doğru biçimde olumsuz tahmin etmiştir. Olumsuz yorumların %5,86'sı hatalı biçimde olumlu olarak değerlendirilmiş; olumlu yorumların %0,22'si yanlışlıkla olumsuz olarak tahmin edilmiştir.

3.2.4.6. Gradyan Arttırma Tahmin Hata Matrisi Değerlendirmesi

Gradyan Arttırma algoritmasının doğruluk oranı %95,80, ROC eğrisi altında kalan oranı %96.06 ve F1 skoru %95,78'dir. Hata matrisi Şekil 22' de yer almaktadır.

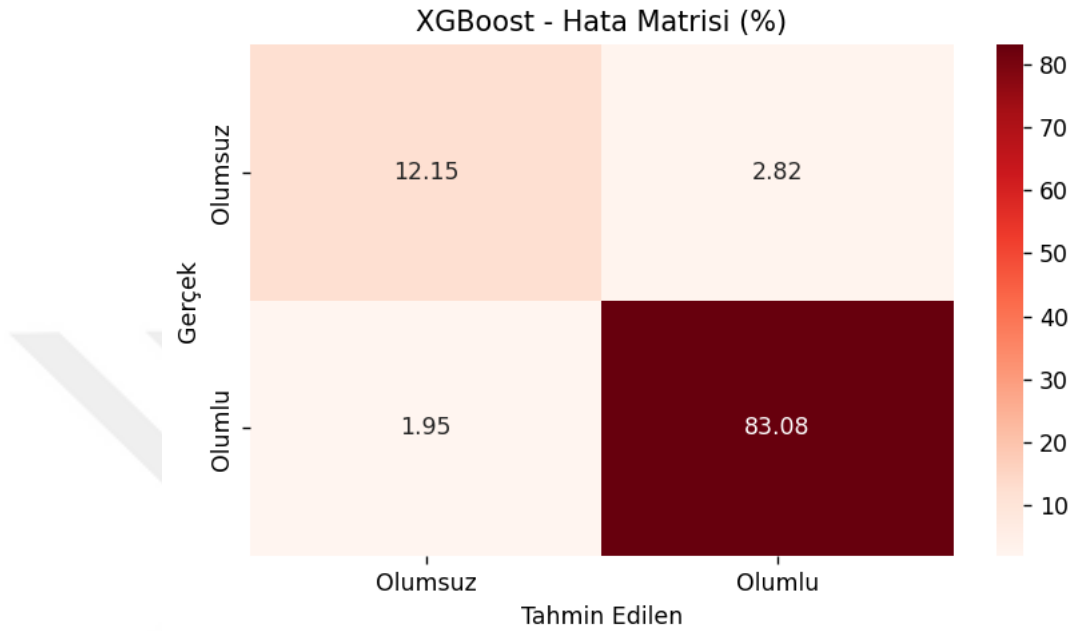


Şekil 20:Çam Ağacı Ürünü Gradyan Arttırma Tahmin Hata Matrisi

Gradyan Arttırma modeli, müşterilerin olumlu yönde yaptığı yorumların %83,95’ini doğru bir şekilde olumlu olarak tahmin etmiştir. Model, olumsuz içerikli yorumların ise %12,80’ini doğru biçimde olumsuz tahmin etmiştir. Olumsuz yorumların %2,17’si hatalı biçimde olumlu değerlendirilmiştir; olumlu yorumların %1,08’i ise yanlışlıkla olumsuz olarak tahmin edilmiştir.

3.2.4.7. XBoost Tahmin Hata Matrisi Değerlendirmesi

XBoost algoritmasının doğruluk oranı %95,80, ROC eğrisi altında kalan oranı %96,96 ve F1 skoru %95,73’ dür. Hata matrisi Şekil 23’ de yer almaktadır.



Şekil 21: Çam Ağacı Ürünü XBoost Hata Tahmin Matrisi

XGBoost modeli, müşterilerin olumlu yönde yaptığı yorumların %83,08'ini doğru şekilde olumlu olarak tahmin etmiştir. Model, olumsuz içerikli yorumların ise %12,15'ini doğru bir biçimde olumsuz tahmin etmiştir. Olumsuz yorumların %2,82'si hatalı biçimde olumlu değerlendirilmiştir; olumlu yorumların %1,95'i ise yanlışlıkla olumsuz tahmin edilmiştir.



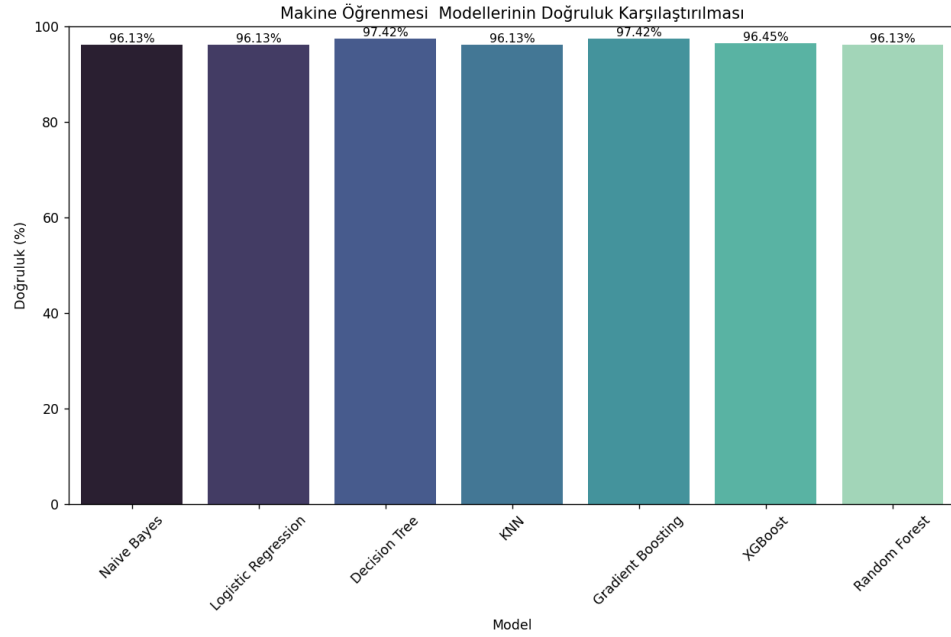
Şekil 22: Çam Ağacı Ürününe Ait Kelime Bulutu

Çam ağacı ürününe ait kelime bulutu Şekil 24 ‘de yer almaktadır. Kelime bulutu, kullanıcıların Trendyol platformu üzerinden yaptıkları değerlendirmelerinin metin madenciliği (text mining) yöntemleriyle analizi ile elde edilmiştir. Kelime bulutunda büyük puntolarla yer alan kelimeler,” güzel”,” ağaç”,” süsler”, “çok güzel” “ürün”, “geldi”, “teşekkürler”, “fiyatına göre”,” fiyat performans” gibi olumlu değerlendirmeler ön plana çıkmaktadır. Bu ifadeler müşterilerin büyük çoğunluğunun ürünü beğendiklerinin ve memnun kaldıklarını göstermektedir Buna karşın “eksik geldi”, “süsler gelmedi”, “seyrek”, “iade”, “kötü”, “almayın”,” kırık geldi” gibi olumsuz ifadelerde ön plana çıkmaktadır. Bu değerlendirmeler ise teslimat süreci ve kalite problemlerinin yaşandığını ifade etmektedir.

Ayrıca kelime bulutunda sıkça yer alan “ve”, “ama”, “fakat”, “gibi”, “değil” gibi bağlaç ve kıyaslama ifadeleri, kullanıcı yorumlarının yalnızca olumlu ya da olumsuz nitelendirmelerden ibaret olmadığını, çoğu zaman ürünle ilgili deneyimlerin olumlu ve olumsuz yönlerinin birlikte değerlendirildiğini göstermektedir.

Bu tür dengeli ifadeler, özellikle “fiyat-performans” vurgusuyla birlikte değerlendirildiğinde, markanın fiyat esnekliğine ve müşteri beklentisi yönetimine dayalı stratejik kararlar almasına imkân tanımaktadır. Tüketicilerin ürünü “güzel ama küçük”, “fiyatına göre iyi fakat süsleri zayıf” gibi yapılarla değerlendirmesi, ürün iyileştirme sürecinin veri temelli yürütülmesini mümkün kılmaktadır.

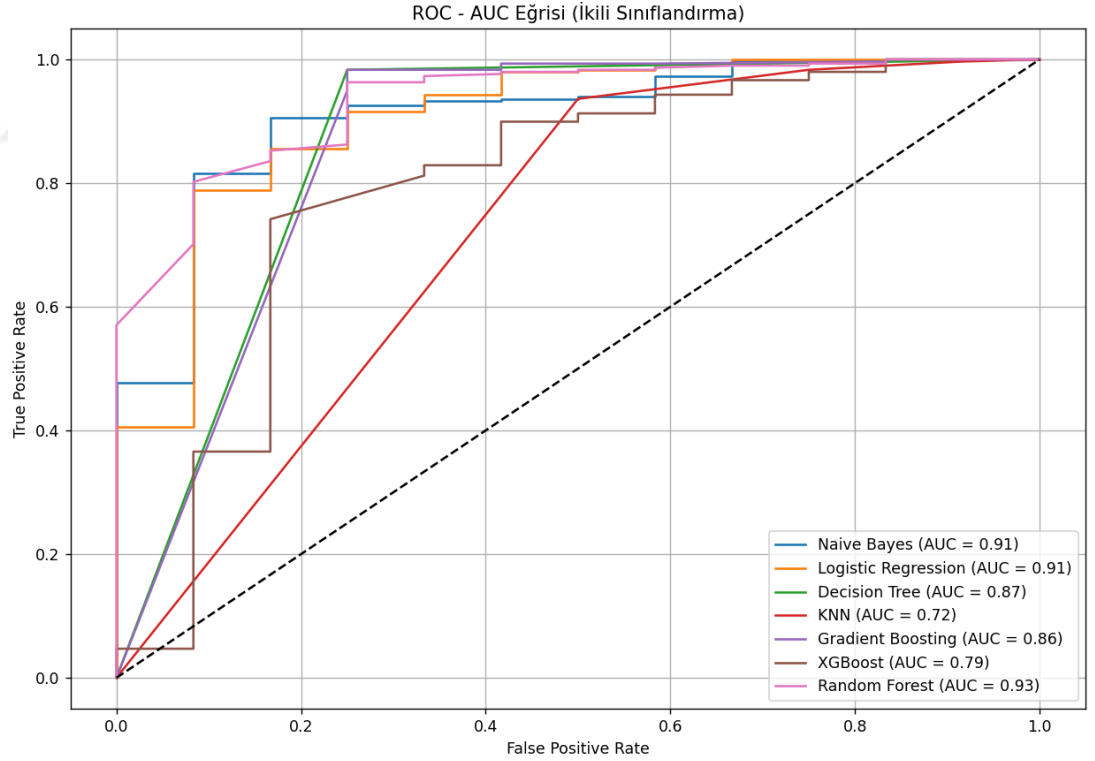
Tablet ürününün verisi ürünü için algoritmaların test verisi üzerindeki başarısı Şekil 25’te yer almaktadır.



Şekil 23: Tablet Ürününe Ait Makine Öğrenmesi Doğruluk Oranları

Algoritmaların doğruluk oranlarına göre Karar Ağacı ve Gradyan Artırma modelleri %97,42 ile en yüksek doğruluğa sahip algoritmalarlardır.

XGBoost modeli ise %96,45 doğruluk oranı ile bu modelleri takip etmektedir. Buna karşın Naive Bayes, Lojistik Regresyon, En Yakın Komşu (KNN) ve Rastgele Orman algoritmaları %96,13 doğruluk oranı ile daha düşük performans göstermiştir. Doğruluk oranı, modellerin tahminlerinin ne ölçüde doğru olduğunu göstermektedir. ROC-AUC değerleri



Şekil 24: Tablet Ürününe Ait ROC- AUC Oranları

En yüksek AUC değerine sahip algoritma %93 (0.93) ile Rastgele Orman algoritmasıdır. KNN algoritması %72 (0.72) ve XGBoost %79 (0.79) oranları ile diğer algoritmalar içerisinde en düşük AUC-ROC oranına sahip algoritmalarıdır.

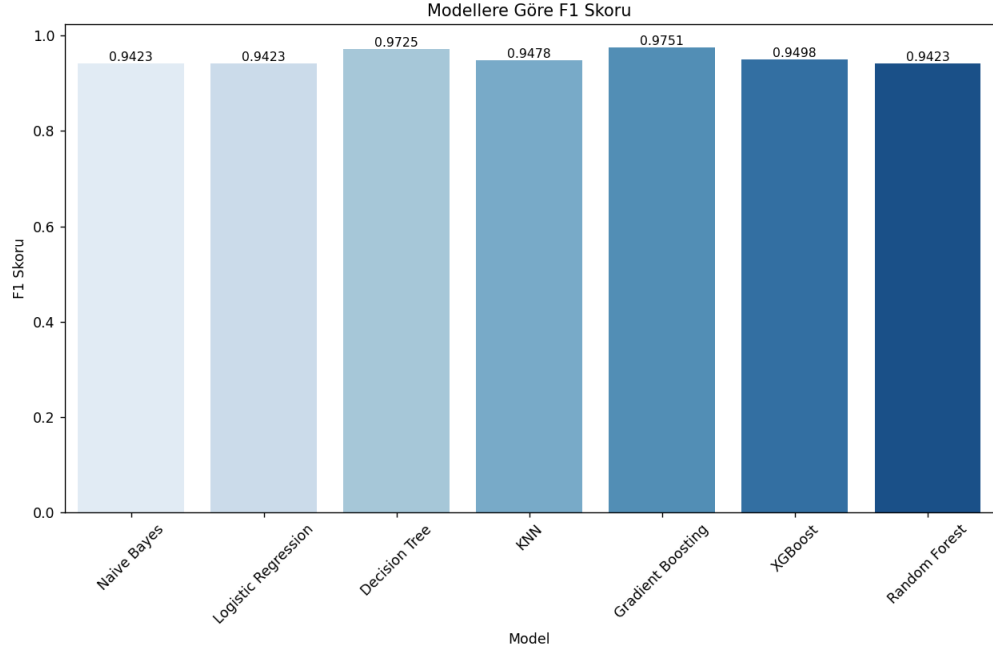
Algoritmaların doğruluk oranlarının yüksek başarı sağlaması ile doğru orantılı olarak algoritmaya bağlı olan eğrilerin grafikte kapladıkları alanda artış görülmektedir.

Şekil 25'te yer aldığı gibi, en yüksek orana sahip olan Rastgele Orman en büyük alana sahipken, en düşük orana sahip olan KNN algoritması en düşük alana sahiptir. ROC-AUC oranlarının yüzdelik gösterimi Tablo 4'te yer almaktadır.

Tablo 4: Tablet Ürünü AUC- ROC Oranları(%)

Model	Doğruluk (%)	AUC-ROC Skoru
Naive Bayes	96.129032	90.0827
Lojistik Regresyon	96.129032	90.5480
Decision Tree (Karar Ağacı)	97.096774	86.4932
KNN	96.129032	72.2175
Gradient Boosting (Gradyan Arttırma)	97.419354	86.2975
XGBoost	96.451612	79.1387
Random Forest (Rassal Orman)	96.129032	93.2046

Algoritmalara ait F1 skoru Şekil 27' de yer almaktadır.

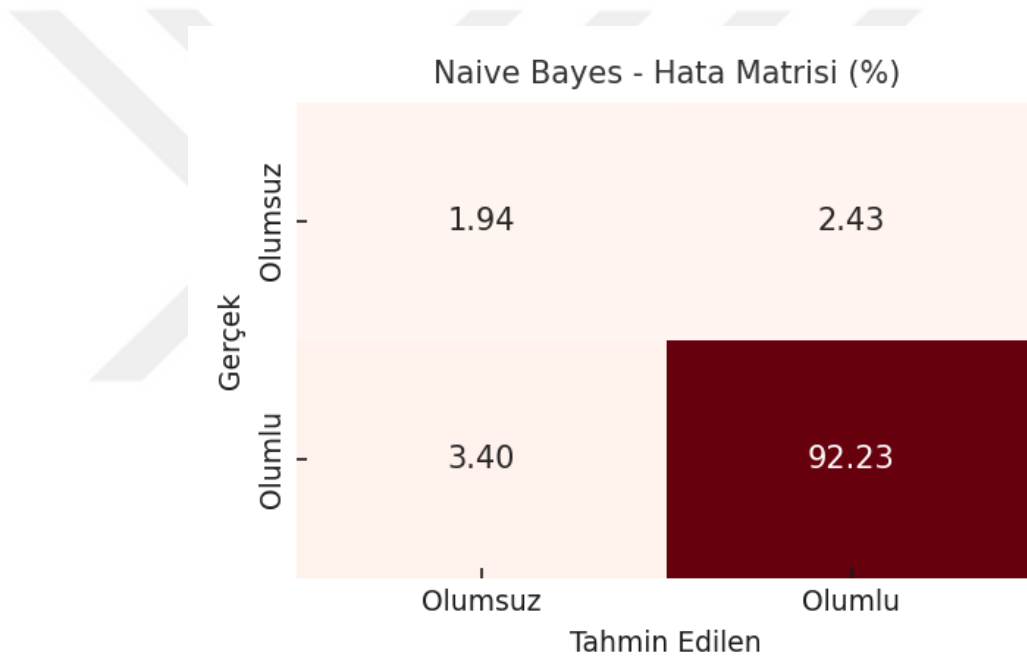


Şekil 25: Tablet Ürünü F1 Skoru

Gradyan Artırma (%97,51) ve Decision Tree (%97,25), F1 skorları açısından en başarılı modeller olarak öne çıkmasına karşın, Naive Bayes en düşük F1 skoruna sahip modeldir (%94,23). Algoritmaların doğruluk oranı, duyarlılık ve F1 skorları önemli performans metrikleri olsa da bu göstergeler tek başına sınıflar arasındaki hata dağılımını bütüncül biçimde yansıtmakta yetersiz kalabilmektedir. Bu nedenle, sınıflandırma sürecinde oluşan hataların daha ayrıntılı biçimde değerlendirilmesi amacıyla kullanılan hata matrisleri, her bir algoritmanın güçlü ve zayıf yönlerini ortaya koymada kritik bir rol üstlenmektedir.

3.2.5.1 Naive Bayes Tahmin Hata Matrisi Değerlendirmesi

Naive Bayes algoritmasının doğruluk oranı %96,13, ROC eğrisi altında kalan oranı %90.08 ve F1 skoru %94,23'tür. Algoritmanın Hata matrisi Şekil 28'de yer almaktadır.

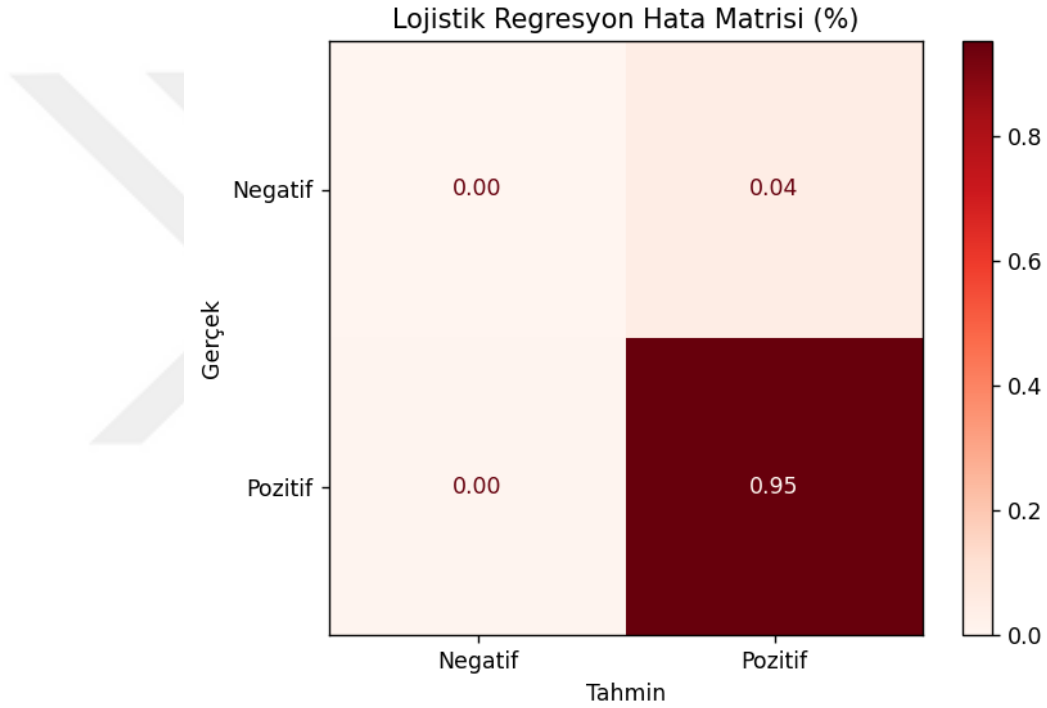


Şekil 26: Tablet Ürünü Naive Bayes Hata Matrisi

Model, müşterilerin olumlu yönde yaptığı yorumların %92,23'ünü doğru şekilde olumlu olarak, gerçekte olumsuz olan yorumların ise %1,94'ünü doğru şekilde olumsuz olarak sınıflandırmıştır. Buna karşılık, müşterilerin olumsuz yöndeki yorumlarının %2,43'ü yanlışlıkla olumlu olarak sınıflandırılmıştır. Öte yandan, gerçekte olumlu olan yorumların %3,40'ı hatalı şekilde olumsuz olarak tahmin edilmiştir.

3.2.5.2 Lojistik Regresyon Tahmin Hata Matrisi Değerlendirmesi

Lojistik Regresyon algoritmasının doğruluk oranı %96,12, ROC eğrisi altında kalan oranı %96,54 ve F1 skoru %94,23. Hata matrisi Şekil 29’ da yer almaktadır.



Şekil 27: Tablet Ürünü Lojistik Regresyon Hata Matrisi

Model, gerçek pozitif örneklerin %95’ini doğru şekilde “pozitif” sınıfına atayarak, olumlu yorumları tanıma konusunda yüksek bir başarı sergilemiştir. Bu durum, modelin pozitif sınıf üzerindeki ayırıştırma yetisinin oldukça güçlü olduğunu göstermektedir.

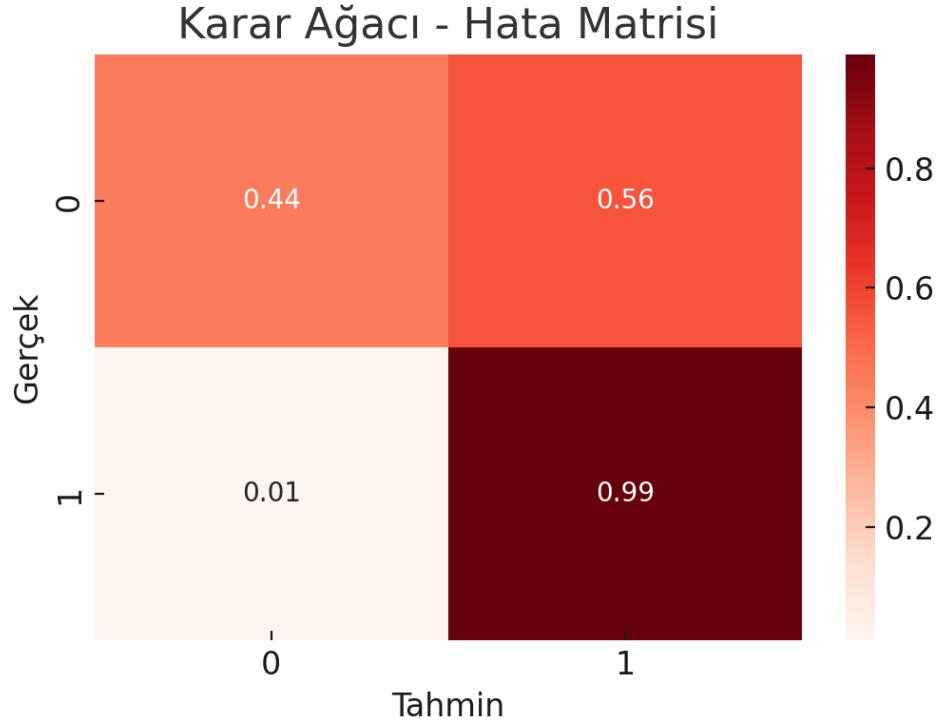
Bununla birlikte, gerçek negatif örneklerin tamamını hatalı biçimde pozitif sınıfa dahil edilmiştir. Başka bir ifadeyle, model olumsuz sınıfa ait hiçbir örneği doğru biçimde tanımlayamamış, bu da negatif sınıf için duyarlılığın (recall) sıfır olduğunu ortaya koymaktadır.

Bu durum, modelin olumsuz örnekleri ayırt etme kabiliyetinin son derece zayıf olduğu ve ciddi bir sınıf dengesizliği sorunu bulunduğu yönünde güçlü bir gösterge niteliğindedir.

Pozitif sınıfa ilişkin sınıflandırma hatalarının yok denecek düzeyde (%0) kalması, modelin yalnızca baskın sınıfa (pozitif) odaklandığını ve bu nedenle dengesiz veri dağılımının modelin öğrenme sürecini olumsuz etkilediğini düşündürmektedir.

3.2.5.3 Karar Ağacı Tahmin Hata Matrisi Değerlendirmesi

Karar ağacı algoritmasının doğruluk oranı %97,06,12, ROC eğrisi altında kalan oranı %86 ,29 ve F1 skoru %97,25. Hata matrisi Şekil 30’ da yer almaktadır.

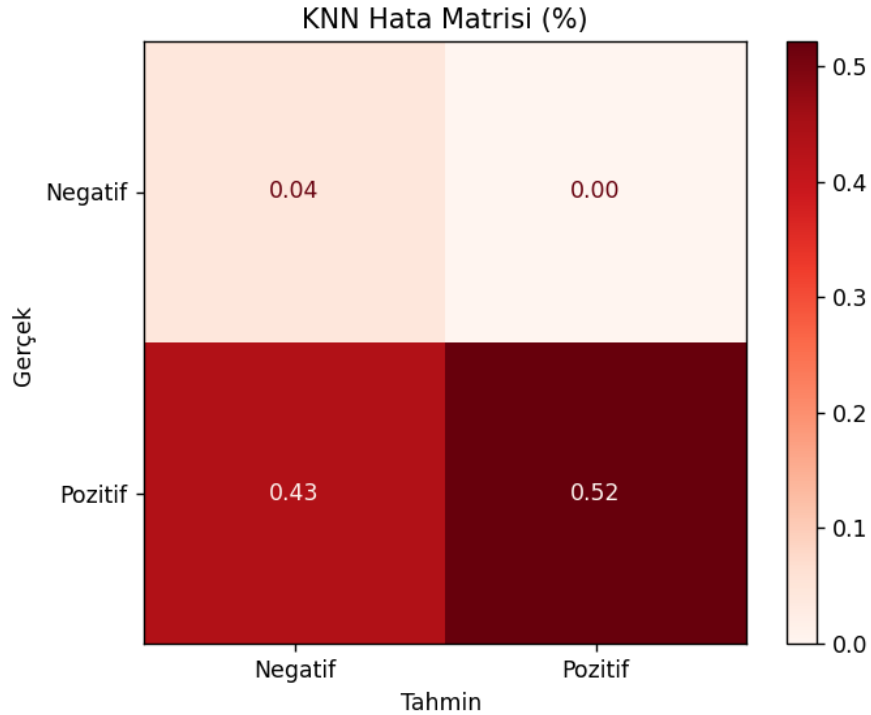


Şekil 28: Tablet Ürünü Karar Ağacı Hata Matrisi

Model, Müşterilerin olumlu yönde yaptığı yorumların %98,41'ini doğru biçimde olumlu olarak sınıflandırabilmiştir. Ancak modelin olumsuz yorumları ayırt etme performansı düşüktür; bu sınıfa ait yorumların yalnızca %5,53'ü doğru olarak tahmin edilmiştir. Gerçekte olumsuz olan yorumların %94,47'si model tarafından hatalı biçimde olumlu olarak değerlendirilirken, olumlu içerikli yorumların %1,59'u yanlışlıkla olumsuz olarak sınıflandırılmıştır.

3.2.5.4 KNN Algoritması Tahmin Hata Matrisi Değerlendirmesi

KNN algoritmasının doğruluk oranı %96,12 ROC eğrisi altında kalan oranı %72,21 ve F1 skoru %94,78 Hata matrisi Şekil 31' de yer almaktadır.



Şekil 29: Tablet Ürünü KNN Algoritması Hata Matrisi

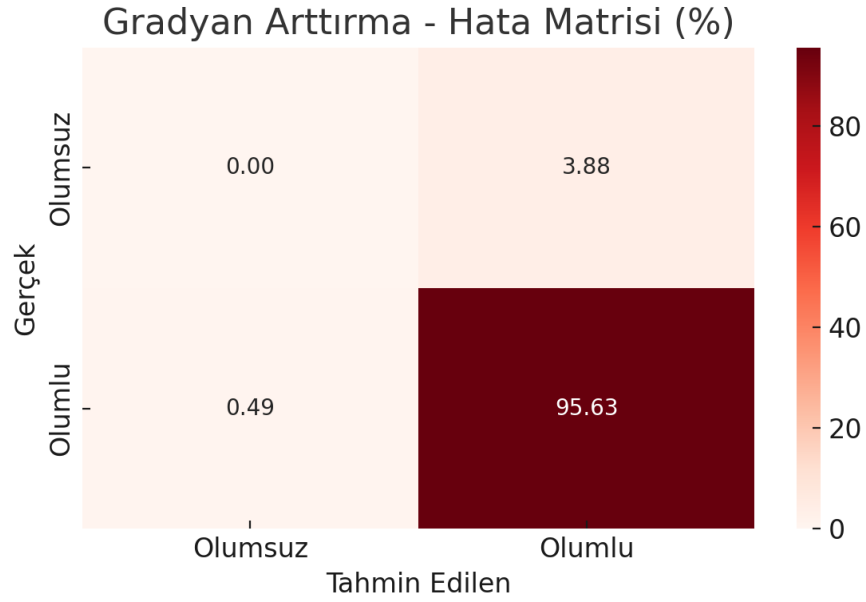
K-En Yakın Komşu (KNN) algoritması, sınıflar arası dengesizliği azaltmak amacıyla SMOTE (Synthetic Minority Over-sampling Technique) yöntemi uygulanarak eğitilmiştir. Bu yaklaşım sayesinde, eğitim verisindeki azınlık sınıf olan olumsuz (negatif) yorumların temsil gücü artırılmış ve modelin bu sınıfı daha sağlıklı öğrenmesi hedeflenmiştir.

SMOTE ile dengeleme sonrası eğitilen KNN modeli, müşterilerin olumlu nitelikteki yorumlarının %52'sini doğru biçimde “pozitif” sınıfında tahmin edebilmiştir. Model, olumsuz yorumların ise yalnızca %0,04'lük kısmını doğru şekilde “negatif” olarak sınıflandırmıştır.

Öte yandan, olumlu yorumların %43'ü yanlışlıkla “negatif” olarak; olumsuz yorumların %0'ı ise “pozitif” sınıfa hatalı biçimde atanmıştır. Bu sonuçlar, SMOTE uygulamasına rağmen KNN algoritmasının özellikle azınlık sınıfın (negatif yorumlar) tanımlanmasında yetersiz kaldığını ortaya koymaktadır. Doğruluk oranının büyük kısmının pozitif sınıfa ait olması, KNN'nin karmaşık sınırları tanımlama konusunda daha sınırlı bir başarı gösterdiğine işaret etmektedir. Bu durum, özellikle metin sınıflandırma gibi yüksek boyutlu problemlerde KNN gibi örnek tabanlı yöntemlerin, daha gelişmiş sınıflandırıcılarla karşılaştırıldığında performans açısından dezavantajlı olabileceğini göstermektedir.

3.2.5.5 Gradyan Arttırma Tahmin Hata Matrisi Değerlendirmesi

Gradyan Arttırma algoritmasının doğruluk oranı %97,41, ROC eğrisi altında kalan oranı %86,29 ve F1 skoru %97,51'dir. Hata matrisi Şekil 32' de yer almaktadır.

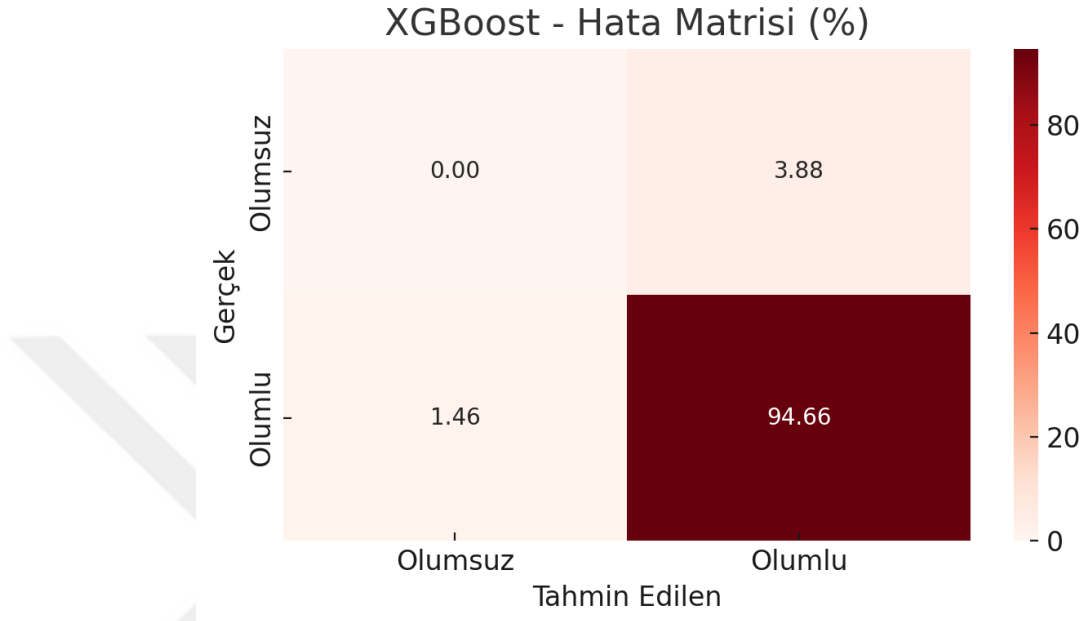


Şekil 30: Tablet Ürünü Gradyan Arttırma Hata Tahmin Matrisi

Model, müşterilerin olumlu yönde yaptığı yorumların %95,63'ünü doğru biçimde olumlu olarak sınıflandırabilmiştir. Ancak modelin olumsuz yorumları ayırt etme performansı oldukça sınırlıdır; bu sınıfa ait yorumların yalnızca %0,00'ı doğru şekilde tahmin edilmiştir. Gerçekte olumsuz olan yorumların %100'ü model tarafından hatalı biçimde olumlu olarak değerlendirilirken, olumlu içerikli yorumların %0,49'u yanlışlıkla olumsuz sınıfa atanmıştır. Bu sonuçlar, modelin olumlu sınıfı tanımada yüksek başarı gösterdiğini, ancak olumsuz geri bildirimleri tespit etmede yetersiz kaldığını göstermektedir.

3.2.5.6. XBoost Tahmin Hata Matrisi Değerlendirmesi

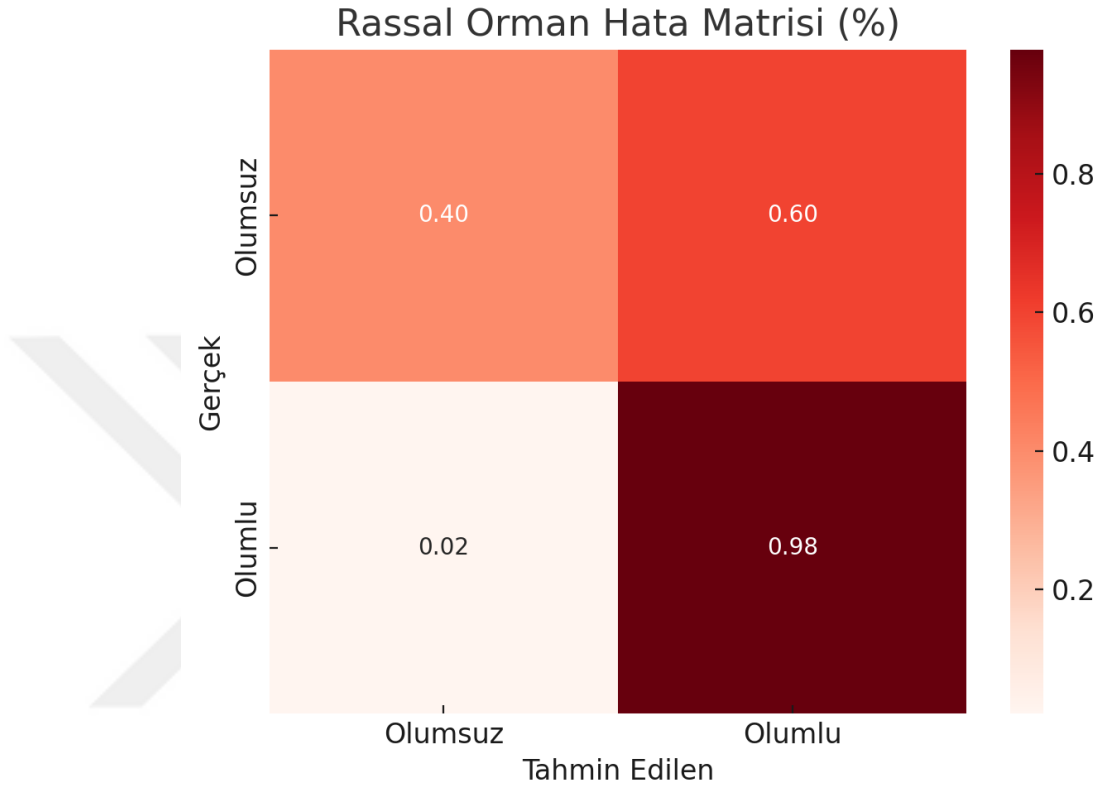
XBoost algoritmasının doğruluk oranı %96,45, ROC eğrisi altında kalan oranı %79,13 ve F1 skoru %94,98' dir. Hata matrisi Şekil 33' de yer almaktadır.



Şekil 31: Tablet Ürünü XGBoost Algoritması Hata Tahmin Matrisi

3.2.5.7. Rassal Orman Tahmin Hata Matrisi Değerlendirmesi

Rassal Orman algoritmasının doğruluk oranı %96,12, ROC eğrisi altında kalan oranı %93,20 ve F1 skoru %94,23'tür. Hata matrisi Şekil 34' de yer almaktadır.



Şekil 32: Tablet Ürünü Rassal Orman Algoritması Hata Tahmin Matrisi

Model, gerçekte olumlu olan müşteri yorumlarının %98,00'ini doğru biçimde olumlu olarak sınıflandırabilmiştir. Bu durum, modelin olumlu sınıfa ilişkin duyarlılığının oldukça yüksek olduğunu ve olumlu yorumları doğru tanıma konusunda güçlü bir performansa sahip olduğunu göstermektedir.

Buna karşın, gerçekte olumsuz olan yorumların yalnızca %40,00'ı doğru şekilde olumsuz olarak tahmin edilebilmiştir. Kalan %60,00'lık kısım ise yanlışlıkla olumlu sınıfa atanmıştır. Bu sonuç, modelin olumsuz sınıfı ayırt etme konusundaki başarısının sınırlı olduğunu ortaya koymaktadır.



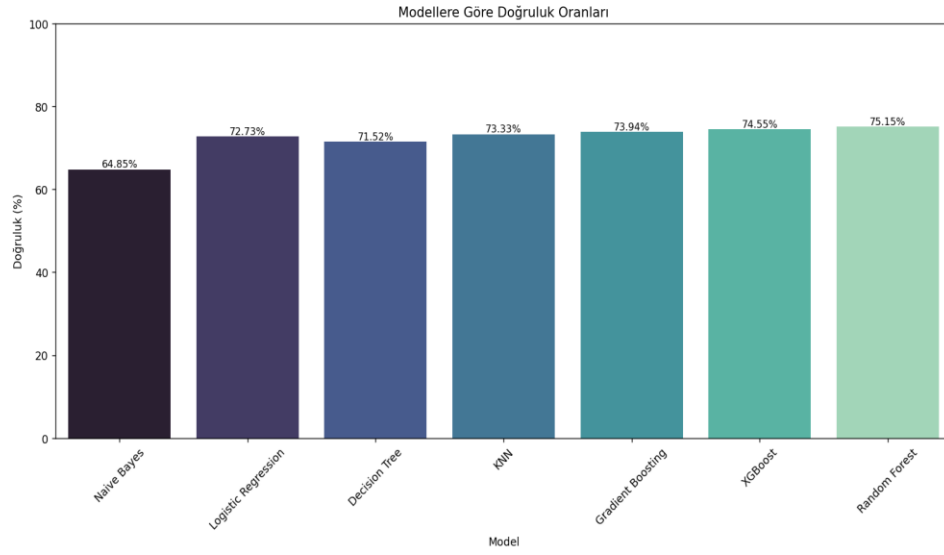
Şekil 33: Tablet Ürününe Ait Kelime Bulutu

Şekil 35’te yer alan tablet ürününe ait kelime bulutu incelendiğinde, kullanıcı yorumlarında sıkça tekrar eden ifadelerin görsel olarak öne çıktığı görülmektedir. “Tablet”, “güzel”, “performans”, “hızlı”, “ekran”, “kaliteli”, “aldım”, “kullanışlı”, “memnun”, “çok” ve “teşekkür” gibi kelimelerin büyük puntolarla belirtilmiş olması, kullanıcıların genel olarak ürünün teknik özellikleri ve kullanım deneyiminden memnuniyet duyduklarını göstermektedir.

Kelime bulutundaki bu terimler, kullanıcıların tabletin performansına, ekran kalitesine ve işlevselliğine yönelik olumlu algılarının yanı sıra, satın alma süreçlerinde yaşadıkları memnuniyeti de yansıtmaktadır. “Paketleme”, “dayanıklı” ve “fiyat-performans” gibi ifadelerin yer alması, kullanıcıların yalnızca yüzeysel deneyimler değil, aynı zamanda ürünün sağlamlığı ve ekonomik değeri üzerinde de durduklarını ortaya koymaktadır.

Bununla birlikte, kelime bulutunda “ama”, “bazı”, “fakat” gibi bağlaçların bulunması, yorumların tamamen olumlu olmadığını, kullanıcıların bazı özelliklere yönelik eleştirel veya dengeli görüşler sunduğunu göstermektedir.

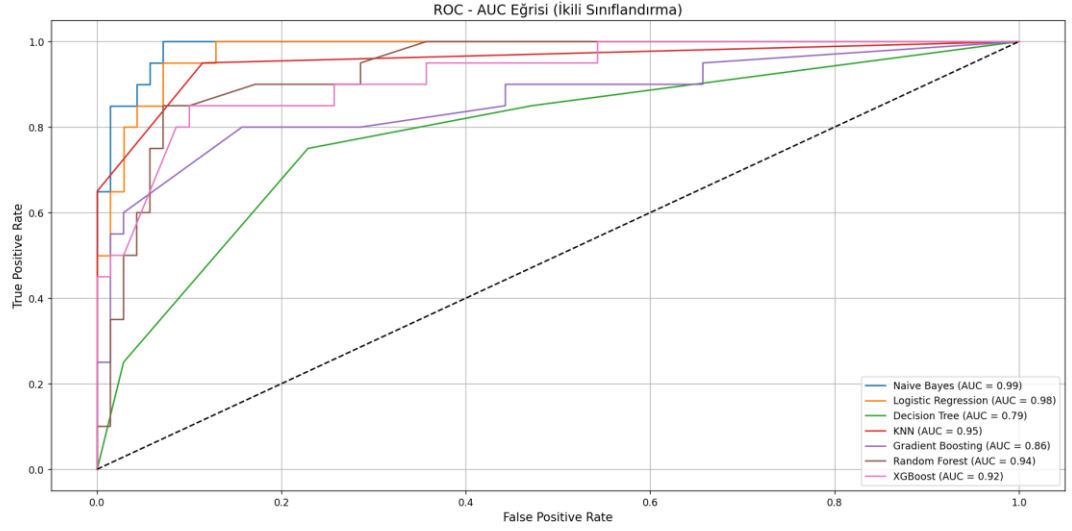
Bu bağlamda, tablet ürününe ait kelime bulutu, kullanıcıların büyük oranda olumlu değerlendirmelerde bulunduğunu; ancak değerlendirmelerin çok boyutlu ve kapsamlı olduğunu ortaya koymaktadır. Bu durum, ürünün hem avantajlarının hem de geliştirilebilir yönlerinin tüketiciler tarafından fark edildiğine işaret etmektedir.



Şekil 34: Çiçek Ürününe Ait Makine Öğrenmesi Doğruluk Oranları

Şekil 36’da yer alan doğruluk oranlarına göre, makine öğrenmesi algoritmaları arasında en yüksek başarıyı %75,15 ile Random Forest modeli göstermektedir; onu sırasıyla %74,55 ile XGBoost, %73,94 ile Gradient Boosting ve %73,33 doğruluk oranı ile K-En Yakın Komşu (KNN) takip etmektedir. Logistic Regression ve Decision Tree algoritmaları sırasıyla %72,73 ve %71,52 oranlarında doğruluk sunarken, en düşük performans %64,85 ile Naive Bayes modeline aittir.

Naive Bayes algoritmasının bu düşük doğruluk oranı, büyük olasılıkla veri setindeki örnek sayısının yetersiz olmasından kaynaklanmaktadır; çünkü bu tür istatistiksel modeller, sınıflar arası dağılımları doğru tahmin edebilmek için daha fazla örneğe ihtiyaç duyar.



Şekil 35: Çiçek Ürününe Ait ROC- AUC Oranları

En yüksek AUC değerine sahip algoritma %99 (0.99) oranıyla Naive Bayes algoritmasıdır. Bunu sırasıyla %98 (0.98) ile Logistic Regression, %95 (0.95) ile K-Nearest Neighbors (KNN), %94 (0.94) ile Random Forest, %92 (0.92) ile XGBoost, %86 (0.86) ile Gradient Boosting, ve son olarak %79 (0.79) ile Decision Tree algoritmaları takip etmektedir.

Algoritmaların AUC değerlerinin yüksekliği, sınıflandırma başarısıyla doğrudan ilişkilidir. Bu doğrultuda, algoritmaların ROC eğrileri altında kalan alan büyüklüğü de AUC skorlarına paralel olarak değişmektedir. Grafik üzerinde en geniş alanı kapsayan eğri Naive Bayes algoritmasına aittir. Buna karşılık, Decision Tree algoritması daha düşük bir AUC skoruna sahip olduğundan, eğrisinin altında kalan alan daha dar olarak gözlemlenmektedir.

Şekil 37'de sunulan ROC eğrileri, her bir algoritmanın pozitif sınıfı ayırt etme performansını görsel olarak karşılaştırmalı biçimde ortaya koymaktadır. ROC-AUC oranlarının sayısal karşılıkları Tablo 'te detaylandırılmıştır.

Tablo 5:Çiçek Ürünü AUC- ROC Oranları (%)

Model	Doğruluk (%)	AUC-ROC Skoru
Naive Bayes	64.0855	99,4936
Lojistik Regresyon	72,0739	98,1469
Decision Tree (Karar Ağacı)	70,0304	79,3955
KNN	73,0336	95,1256
Gradient Boosting (Gradyan Arttırma)	74,0459	86.2975
XGBoost	74,0554	94,2564
Random Forest (Rassal Orman)	75,0152	92.2046

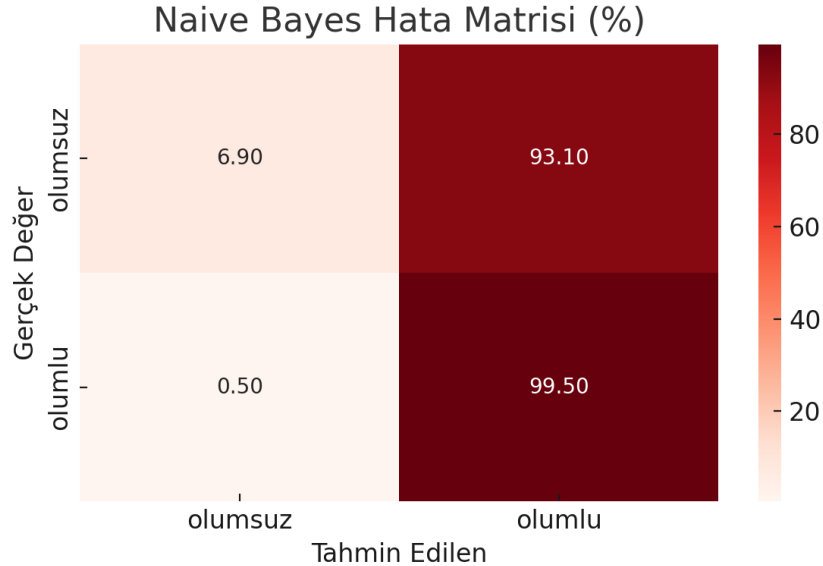
Algoritmalarla ait F1 skoru Şekil 27' de yer almaktadır. F1 skoru karşılaştırmasına göre, en yüksek başarıyı gösteren algoritma %72,20'lik oranla XGBoost olmuştur. Bu modeli sırasıyla Rassal Orman(%71,26), , Gradyan Arttırma(%70,47) ve Karar Ağacı (%70,27) algoritmaları takip etmektedir. Görece daha düşük başarı sergileyen modeller ise KNN (%68,51) ve Lojistik Regresyon (%66,89) olarak belirlenmiştir. En düşük F1 skoruna sahip algoritma ise %52,13 ile Naive Bayes olmuştur.

F1 skoru, sınıflandırma modellerinin hem duyarlılık (recall) hem de özgüllük (precision) performanslarını dengeli bir şekilde yansıtmaları açısından önemli bir metriktir.

Ancak, bu gösterge tek başına sınıflar arası hata dağılımını ve modelin genel hatalarını detaylı olarak ortaya koymakta yetersiz kalabilmektedir. Bu nedenle, her bir modelin sınıflandırma sürecindeki başarısının daha bütüncül ve ayrıntılı biçimde değerlendirilebilmesi adına hata matrisi gibi tamamlayıcı performans analizlerine de ihtiyaç duyulmaktadır. Bu tür analizler, algoritmaların güçlü ve zayıf yönlerinin anlaşılmasında kritik rol oynamaktadır.

3.2.6.1 Naive Bayes Tahmin Hata Matrisi Değerlendirmesi

Naive Bayes algoritmasının doğruluk oranı %64.09, ROC eğrisi altında kalan oranı %99,49 ve F1 skoru %52,13'tür. Algoritmanın Hata matrisi Şekil 38'de yer almaktadır.

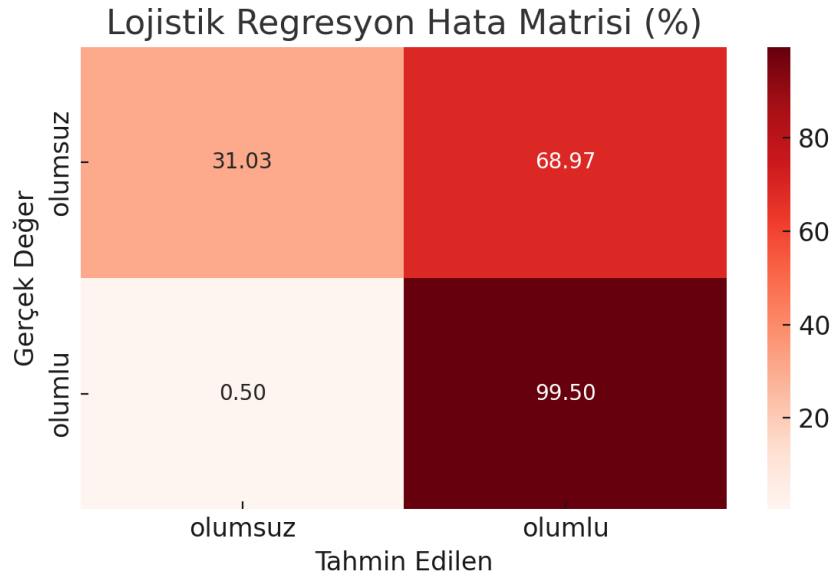


Şekil 36: Çiçek Ürünü Naive Bayes Hata Tahmin Matrisi

Naive Bayes modeli, kullanıcıların olumlu yorumlarını büyük oranda doğru tahmin etmiş; bu yorumların yaklaşık %91,62'si doğru şekilde olumlu olarak sınıflandırılmıştır. Ancak olumsuz yorumları tanıma konusunda aynı başarıyı gösterememiş, bu yorumların yalnızca %2,58'ini doğru biçimde olumsuz olarak etiketlemiştir. Buna karşılık, model bazı olumsuz yorumları yanlışlıkla %5,80 oranında olumlu olarak değerlendirmiştir. Benzer şekilde, olumlu yorumların da %8,38'lik bir kısmı hatalı biçimde olumsuz olarak sınıflandırılmıştır. Bu tablo, modelin özellikle olumlu ifadeleri algılamada başarılı olduğunu, fakat olumsuz duyguları ayırt etmede zorlandığını ortaya koymaktadır.

3.2.6.2 Lojistik Regresyon Tahmin Hata Matrisi Değerlendirmesi

Lojistik Regresyon algoritmasının doğruluk oranı %72,07, ROC eğrisi altında kalan oranı %98,14 ve F1 skoru %66,89. Hata matrisi Şekil 29' da yer almaktadır.



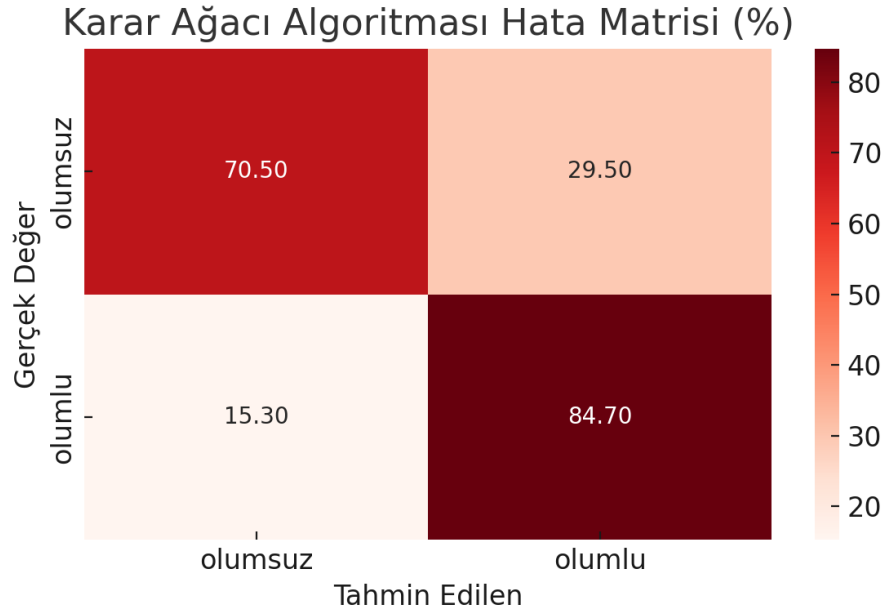
Şekil 37: Çiçek Ürünü Lojistik Regresyon Hata Tahmin Matrisi

Lojistik Regresyon modeli, olumlu yorumları sınıflandırmada yüksek bir doğruluk oranına ulaşmış; bu tür yorumların %92,49'unu doğru bir şekilde olumlu olarak tanımıştır. Öte yandan, olumsuz yorumların yalnızca %1,93'ü doğru şekilde olumsuz olarak belirlenebilmiştir.

Model, olumsuz yorumların %5,58'ini yanlışlıkla olumlu olarak sınıflandırırken, olumlu yorumların da %7,51'lik kısmını hatalı biçimde olumsuz olarak değerlendirmiştir. Genel olarak model, olumlu duyguları başarıyla yakalarken, olumsuz içerikleri doğru şekilde analiz etme konusunda daha sınırlı bir performans sergilemiştir.

3.2.6.3 Karar Ağacı Tahmin Hata Matrisi Değerlendirmesi

Karar ağacı algoritmasının doğruluk oranı 70,03, ROC eğrisi altında kalan oranı %79,39 ve F1 skoru %%70,27 Hata matrisi Şekil 40' da yer almaktadır.



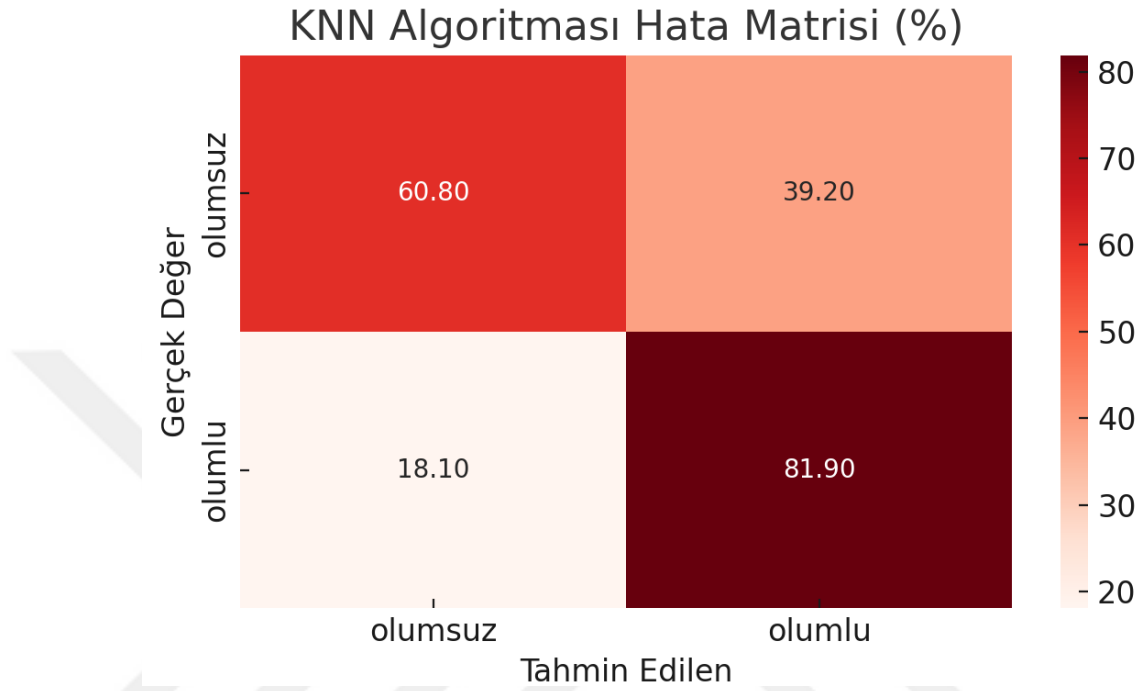
Şekil 38: Çiçek Ürünü Karar Ağacı Algoritması Hata Tahmin Matrisi

Karar Ağacı algoritması, olumlu sınıfı tanıma konusunda benzer biçimde güçlü bir performans sergilemiştir. Gerçekte olumlu olan yorumların %97,62'si, doğru bir şekilde olumlu olarak tahmin edilmiştir. Bu oran, modelin olumlu duyguları içeren metinleri başarılı bir şekilde öğrenebildiğini göstermektedir. Ancak olumsuz sınıf açısından performans düşüktür.

Gerçekte olumsuz olan yorumların yalnızca %9,52'si, doğru biçimde olumsuz olarak sınıflandırılabilmiştir. Geriye kalan %90,48'lik bölüm hatalı şekilde olumlu olarak değerlendirilmiştir. Ayrıca modelin gerçekte olumlu olan yorumların %2,38'ini yanlış biçimde olumsuz olarak sınıflandırdığı gözlenmiştir. Bu da modelin sınıflar arası ayrımı yaparken, özellikle olumsuz sınıfa ait örneklerde yeterince ayrıştırıcı özelliklere ulaşamadığını göstermektedir.

3.2.6.4 KNN Algoritması Tahmin Hata Matrisi Değerlendirmesi

KNN algoritmasının doğruluk oranı %73,03 ROC eğrisi altında kalan oranı %95,12 ve F1 skoru % (%68,51 Hata matrisi Şekil 41' de yer almaktadır.



Şekil 39: Çiçek Ürünü KNN Algoritması Hata Tahmin Matrisi

KNN algoritması da olumlu sınıfı yüksek doğrulukla tahmin etmiştir. Gerçekte olumlu olan müşteri yorumlarının %98,41'i, model tarafından doğru bir biçimde olumlu olarak sınıflandırılmıştır. Bu oran, modelin olumlu geri bildirimleri başarılı bir şekilde tanıdığını göstermektedir. Ancak olumsuz yorumlara yönelik tahmin performansı düşüktür. Gerçekte olumsuz olan yorumların yalnızca %5,53'ü, model tarafından doğru biçimde olumsuz olarak belirlenmiştir. Geriye kalan %94,47'si hatalı şekilde olumlu olarak tahmin edilmiştir. Ayrıca, model gerçekte olumlu olan yorumların %1,59'unu olumsuz olarak sınıflandırmıştır.

SONUÇ

Araştırmaya Genel Bakış:

Dijitalleşmenin hız kazanmasıyla birlikte çevrimiçi alışveriş alışkanlıkları önemli ölçüde değişmiş, kullanıcı yorumları ise tüketici karar süreçlerinde belirleyici bir unsur haline gelmiştir. Web 2.0 teknolojilerinin sunduğu etkileşimli yapı sayesinde tüketiciler yalnızca bilgiye ulaşan değil, aynı zamanda bilgi üreten aktörler hâline gelmiştir. Bu çalışmada, özel günlerde çevrimiçi alışveriş yapan kullanıcıların yorumları makine öğrenmesi algoritmaları aracılığıyla analiz edilerek bu yorumların taşıdığı anlamın sınıflandırılması ve tahminlenmesi amaçlanmıştır.

Araştırma Sorusu 1:

Özel gün indirimleri kapsamında gerçekleştirilen çevrim içi alışverişlerin, tüketici yorumlarına yansıyan duygusal ifadeler aracılığıyla değerlendirme davranışları üzerindeki etkisi ortaya konabilir mi?

Elde edilen veriler, tüketici yorumlarının büyük oranda olumlu yönelimli olduğunu göstermektedir. Ancak kelime bulutu analizlerinde sıkça karşılaşılan “ama”, “fakat”, “ancak” gibi bağlaçlar, içerikte örtük bir memnuniyetsizlik ya da çekince barındırdığını işaret etmektedir. Bu tür bağlaçlar, metinde pozitif ifadelerle eşlik etse de kullanıcıların deneyimlerine dair tam anlamıyla net bir kanaat sunmadığını göstermektedir. Dolayısıyla yorumların yüzeysel olarak olumlu sınıflandırılması, içerdiği anlam katmanlarının gözden kaçmasına neden olabilir. Bu bulgu, tüketici davranışlarının yalnızca duygusal kutuplulukla değil; dilsel belirsizlikle de şekillenebildiğini ortaya koymuştur.

Araştırma Sorusu 2:

Farklı makine öğrenmesi algoritmaları, özel günlerde indirimli olarak satın alınan ürünlere ilişkin kullanıcı yorumlarını duygu analizi açısından sınıflandırmada ne düzeyde performans göstermektedir ve hangi algoritmalar bu açıdan öne çıkmaktadır?

Yedi farklı algoritmanın karşılaştırmalı analizinde, XGBoost, Gradyan Artırma ve Karar Ağacı modelleri yüksek sınıflandırma başarılarıyla öne çıkmıştır. XGBoost modeli, %95'in üzerinde doğruluk ve ROC-AUC skorları ile özellikle belirsiz ve kararsız duygu ifadelerini tanımada etkili olmuştur. Buna karşın Naive Bayes ve KNN algoritmaları, düşük F1 skorlarıyla özellikle olumsuz ya da örtük içerikleri sınıflamada yetersiz kalmıştır. Bu bulgular, dilin yapısal karmaşıklığını dikkate alan algoritmaların metin sınıflandırmasında daha başarılı olduğunu göstermektedir.

Araştırma Sorusu 3:

Farklı ürün kategorilerinde uygulanan makine öğrenmesi algoritmalarının duygu sınıflandırma performansları arasında; doğruluk, ROC-AUC ve F1 skoru gibi ölçütler temelinde anlamlı farklılıklar var mıdır ve bu farklılıklar hangi faktörlerden kaynaklanmaktadır?

Çalışmada incelenen farklı ürün kategorileri (örneğin tablet ve çam ağacı), kullanıcı yorumlarının içerik ve duygu yapısı bakımından belirgin farklılıklar göstermiştir. Bu farklılık, makine öğrenmesi algoritmalarının sınıflandırma başarısına doğrudan yansımıştır. Özellikle teknik özellikleri öne çıkan tablet ürünleriyle ilgili yorumlar daha açık, işlevsel ve net ifadeler içerdiğinden, algoritmalar bu kategori üzerinde daha yüksek doğruluk ve ROC-AUC değerleri üretmiştir. Karar Ağacı ve Gradyan Artırma algoritmaları bu tür ürünlerde %97'yi aşan doğruluk oranlarına ulaşmıştır.

Buna karşılık, çam ağacı gibi dekoratif ya da beklentilere daha öznel düzeyde hitap eden ürün kategorilerinde, yorumlar daha duygusal ve görelî ifadeler barındırmaktadır. Bu durum, yorumların duygu yöneliminin daha belirsiz olmasına ve dolayısıyla sınıflandırma hatalarının artmasına yol açmıştır.

Özellikle düşük frekanslı olumsuz ifadelerin bulunduğu yorumlarda bazı algoritmalar tutarsız sonuçlar üretmiştir.

Sonuç olarak, algoritmaların duygu sınıflandırma performansları ürün kategorisine bağlı olarak anlamlı farklılıklar göstermektedir.

Bu fark, yalnızca veri içeriğinden değil; aynı zamanda kullanıcıların ürün türüne bağlı olarak kullandığı dilin niteliğinden kaynaklanmaktadır. Bu nedenle duygu analizi uygulamalarında ürün türü, model seçimi açısından önemli bir değişken olarak değerlendirilmelidir.

Kullanıcı Yorumlarının Anlamsal Derinliği

Yorumlarda sıkça karşılaşılan “ama”, “fakat”, “değil”, “gibi” gibi bağlaç ifadeler, tüketicilerin yalnızca olumlu ya da olumsuz yargılar sunmakla kalmayıp deneyimlerini çok boyutlu olarak aktardığını göstermektedir. Bu da yorumların hem duygu hem içerik yönünden zengin olduğunu ve modellerin dilin bağlamsal yapısını doğru analiz etmesi gerektiğini ortaya koymaktadır. Ayrıca kelime bulutları (bkz. Şekil 23), olumlu ve olumsuz geri bildirimlerin birlikte değerlendirilerek ürün geliştirme süreçlerine veri temelli katkı sunduğunu göstermektedir.

Çalışmanın Kısıtları ve Gelecek Araştırmalar

Araştırmanın bazı sınırlamaları bulunmaktadır. Ürün kategorilerinin belirlenmesinde, yalnızca belirli bir zaman aralığında toplanan veriler ile sınırlı bir tüketici grubunun tercihleri dikkate alınmıştır. Bu durum, farklı dönemlerde ve çeşitli demografik özelliklere sahip tüketicilerin davranış kalıplarının yeterince temsil edilememesine neden olabilir. Ayrıca, veri toplama sürecinin yalnızca tek bir e-ticaret platformuyla sınırlı tutulması, elde edilen bulguların farklı çevrim içi alışveriş ortamlarına genellenebilirliğini azaltmaktadır. Çalışmada kullanılan veri seti, kapsam itibarıyla sınırlı sayıda öznitelik içermektedir ve ilgili literatürde sıklıkla incelenen bazı niteliksel değişkenler analiz sürecine dâhil edilmemiştir. Tüm bu faktörler, araştırmanın metodolojik kapsamını daraltan başlıca sınırlayıcı unsurlar arasında yer almaktadır.

Modelleme sürecinde, çevrimiçi kullanıcı yorumlarına yönelik duygu durumu sınıflandırması (olumlu/olumsuz) gerçekleştirilmiş; metinler üzerinde duygu etiketi çıkarımı başarılı şekilde yapılmıştır. Ancak, yorumların kelime sayısı ve görsel içerik (fotoğraf) içerip içermediği gibi öznitelikler modele dahil edilmemiştir.

Ayrıca, ileri düzey doğal dil işleme yöntemleri olan TF-IDF, Word2Vec, BERT gibi teknikler kullanılarak metinlerin anlamsal ve yapısal özellikleri analiz edilmemiştir. Bu nedenle, yorumların okunabilirlik düzeyleri, kelime tekrarları, bağlaç yapıları ya da sözcük birliktelikleri gibi daha derin metin analizi çıktıları bu çalışmanın kapsamı dışında kalmıştır.

Ayrıca, kullanıcıların yorumlara verdiği tepkiler (beğeni, faydalı bulma gibi) de veri setine entegre edilmemiştir. Bu tür etkileşimlerin dahil edilmesi, gelecekteki çalışmalarda yorumların etki gücünü ve karar destek sistemlerindeki önemini daha sağlıklı ölçmeye olanak sağlayabilir.

Gelecek araştırmalarda, farklı yaş gruplarının yorumlama alışkanlıkları, farklı özel günlerin etkileri ve çok daha geniş ürün kategorilerinin dahil edilmesiyle modelin kapsamı genişletilebilir. Aynı zamanda, metinlerin okunabilirlik analizi, duygu yoğunluğu ölçümü, pozitif/negatif sözcük oranları ve kullanıcı tepkilerinin analizi gibi unsurların sınıflandırma algoritmalarına entegre edilmesi, daha kapsamlı ve doğru sonuçlara ulaşılmasını mümkün kılacaktır.

KAYNAKÇA

- Arı, O. (2022). *Tüketici Yorumlarının Fayda Düzeyinin Tahminlenmesine Yönelik Bir Araştırma: Makine Öğrenmesi Algoritmalarının Karşılaştırılması*. Yüksek Lisans Tezi, Sakarya Üniversitesi SBE, Sakarya.
- Arslan, E. (2020, May 19). *Makine Öğrenmesi — KNN (K-Nearest Neighbors) Algoritması Nedir?* 01 15, 2024 tarihinde <https://arslanev.medium.com/makine-%C3%B6%C4%9Frenmesi-knn-k-nearest-neighbors-algoritmas%C4%B1-bdfb688d7c5f> adresinden alındı
- Ashish Kumar Shakya, G. P. (2023, November 30). Reinforcement learning algorithms: A brief survey. *Expert Systems with Applications*, 231 .
- Ayaz, Y. Y. (2023). *Reklam Genomu Pratikte ve Teoride Reklam*. Eğitim Yayınevi.
- Aydemir, M. (2011). *SEO: Search Engine Optimization*. İstanbul: Koblab Yayın Dağıtım.
- Barış, Y. O. (2003). *Tüketici Davranışları*. İstanbul: MadiaCat Akademi.
- Berners-Lee, T. (1998/05/07, May 7). *The World Wide Web: A very short personal history*. <https://www.w3.org/People/Berners-Lee/ShortHistory.html> adresinden alındı
- Can, A. (2019). *SPSS ile Bilimsel Araştırma Sürecinde Nicel Veri Analizi* (7 b.). Ankara: Pegem Akademi.
- Choudhury, N. (2014). World Wide Web and Its Journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Information Technologies*, 5(6), s. 8096-8100.
- Cömert, G. (2016). *Tüketicilerin İnternet Reklamlarına Yönelik İnançlar ve Tutumları Arasındaki İlişki ve Reklamları İzleme Davranışları Üzerine Etkisi*. Yüksek Lisans Tezi, Beykent Üniversitesi SBE., İstanbul.
- Dokoohaki, R. K. (2011). Diversifying Product Review Rankings: Getting the Full Picture. *IEEE/WIC/ACM Uluslararası Web Zekası ve Akıllı Aracı Teknolojisi Konferansları Bildirileri*, 1, s. 138-145. <https://doi.org/10.1109/WI-IAT.2011.33> adresinden alındı
- Ersöz, B. (2020). Yeni Nesil Web Paradigması: Web 4.0. *Bilgisayar Bilimleri ve Teknolojileri Dergisi*, 1(2), s. 58-65.

- Fabian Pedregosa, G. V. (2011, Jun). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, s. 2825–2830.
- Fanni Agmeka, R. N. (2019). The Influence of Discount Framing towards Brand Reputation and Brand Image on Purchase Intention and Actual Behaviour in e-commerce. *The Fifth Information Systems International Conference 2019* (s. 851–858). Endonezya: Procedia Computer Science.
- Fu Liu, H. W. (2023, October). The influence of online review dispersion on consumers' purchase intention: The moderating role of dialectical thinking. *Journal of Business Research*.
- Gabriel Almeida Lucas, G. L. (2023, March–April). From e-commerce to m-commerce: An analysis of the user's experience with different access platforms. *Electronic Commerce Research and Applications*.
- GUPTA, L. T. (2014). *Amerika Birleşik Devletleri Patent No. US 8,738.436 B2*. .
- Gürbüz, S., & Şahin, F. (2015). *Sosyal Bilimlerde Araştırma Yöntemleri*. Seçkin Yayıncılık.
- Harker, P. T., & Vargas, L. G. (1987). The Theory of Ratio Scale Estimation: Saaty's Analytic Hierarchy Process. *Management Science*, s. 1367-1509. <https://doi.org/10.1287/mnsc.33.11.1383> adresinden alındı
- Hülya Kurgun, O. A. ve Aktaş, E. (2018, June). What does Web 4.0 Promise for Tourism Ecosystem? A Qualitative Research on. *Journal of Tourism and Hospitality Management*, 6(1), s. 55-65.
- Kaynar, O., Tuna, M. F., Görmez, Y., & Deveci, M. A. (2017). Makine Öğrenmesi Yöntemleriyle Müşteri Kaybı Analizi. *C.Ü. İktisadi ve İdari Bilimler Dergisi*, 18(1).
- Koz, K. A., & Işık, U. (2022). İnternet Gazeteciliği İçin Kılavuz: Google Trendlerde Haber. *Uluslararası Medeniyet Çalışmaları Dergisi*, VII(1), s. 24-48.
- Kütük, C. D. (2010). Anlamsal Web (Web 3.0) ve Ontolojilerine Genel Bir Bakış. *istanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, s. 97-108.
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), s. 436-44.
- Lee, B. P. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in information retrieval*, 2(1-2), s. 1-135.
- Li Li, L. Y. (2023, August). Influence of online E-commerce interaction on consumer satisfaction based on big data algorithm. *Heliyon*.

- Liu, M. H. (2004). Mining and Summarizing Customer Reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, (s. 168–177).
- MacLennan, J., Tang, Z., & Crivat, B. (2005). *Data Mining with Microsoft SQL Server 2005*. Wiley.
- Önder DEMİR, A. I. (2019). Türkçe Metinlerde Sözlük Tabanlı Yaklaşımla Duygu Analizi. *Int. Per. of Recent Tech. in App.*, 2, s. 58-66.
- Önder DEMİR, A. I. (2019). Türkçe Metinlerde Sözlük Tabanlı Yaklaşımla Duygu Analizi. *Int. Per. of Recent Tech. in App. Eng*, 2, s. 58-66.
- Özbek, F. C. (2021). *Arama Motoru (Google) Reklamlarının Veri Madenciliği Yöntemleri İle İncelenmesi: Bir E- ticaret Sitesi Örneği*. Yüksek Lisans Tezi, Bandırma.
- Özden, A. T. (2022, May). World From 1.0 To 5.0: Web, Marketing, Industry And Society 1.0'dan 5.0'a Dünya: Web, Pazarlama, Endüstri Ve Toplum. *Journal Of Business In The Digital Age*(5), s. 29-44.
- Page, D. (2016). *Stop Using Google Trends*. 12 21, 2023 tarihinde <https://medium.com/@dannypage/stop-using-google-trends-a5014dd32588> adresinden alındı
- Plutchik, R. (2001). The Nature of Emotions. *American Scientist*, 89(4), s. 344-350.
- Prabhu, T. N. (2020). *Google Trends API for Python*. 12 21, 2023 tarihinde <https://towardsdatascience.com/google-trends-api-for-python-a84bc25db88f> adresinden alındı
- Purdon, E. B. (2006). Introducing Web 2.0: social networking and socialbookmarking for health librarians. *Journal of the Canadian Health Libraries*, s. 65-67.
- Rahm, E., & Do, H. H. (2000). Data Cleaning: Problems and Current Approaches. *IEEE Data Engineering Bulletin*, 23(4).
- Remzi Altunışık, Ş. Ö. (2006). *Pazarlamaya Giriş*. Adapazarı: Sakarya Yayıncılık.
- Ruohola, E. (2021). *Structuring a Modern Web Service for Users and Search Engines*. Yüksek Lisans Tezi, University of Turku. https://www.utupub.fi/bitstream/handle/10024/152595/Ruohola_Eero_opinnayte.pdf?sequence=1 adresinden alındı
- Samuel, A. L. (1959). Some Studies in Machine Learning. *Originally published in IBM Journal*, 3(3), s. 210-229.
- Sarsılmaz, A. E. (2019). *İletişim Çalışmaları*. Sakarya.

- Sarvjeet Kaur Chatrath, G. B. (2022, September 28). Handling consumer vulnerability in e-commerce product images using machine learning. *Heliyon*.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, s. 85-117.
- Schuff, S. M. (2010, March). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *In MIS Quarterly: Management Information*, 34(1), s. 185-200. <https://www.jstor.org/stable/20721420> adresinden alındı
- Sefer Gümüő, S. S. (2014). *İnternet Reklamlarının Tüketicinin Satın Alma Davranışlarına Etkileri*. İstanbul: Hiperlink Yayınları.
- Selvi, A. (2020). *Bilecik İlinde İlköğretimden Liseye Geçiş Sınavlarında Makine Öğrenmesi Yöntemleri ile Öğrenci Başarısının Tahmini*. Bilecik Şeyh Edebali Üniversitesi. Bilecik: Fen Bilimleri Enstitüsü.
- Semila Fernandes, R. P. (2022, September). Measuring the impact of online reviews on consumer purchase decisions – A scale development study. *Journal of Retailing and Consumer Services*.
- Sevli, O. (2019, Ağustos). Göğüs Kanseri Teşhisinde Farklı Makine Öğrenmesi Tekniklerinin Performans Karşılaştırması. *Avrupa Bilim ve Teknoloji Dergisi*(16), s. 176-185.
- Skiena, S. S. (2008). *The Algorithm Design Manual* (2 b.). Springer.
- Solomon, M. R. (2009). *Consumer Behavior: Buying, Having and Being* (8 b.). Pearson/Prentice Hall, Upper Saddle River, N.J. .
- Şenol ÇELİK, N.K.(2022). Veri Madenciliği ve Makine Öğrenmesi ile Farklı Alanlarda Uygulamalar
Holistence Publications.
- Taeyong Kim, S. H. (2022, November). Text analysis of online customer reviews for products in the FCB quadrants: Procedure, outcomes, and implications. *Journal of Business Research*, 150, s. 676-689.
- Tosun, N. (2004). Kullanılan Teknikler Açısından İnternet Reklamları. *Selçuk Üniversitesi İletişim Fakültesi Akademik Dergisi*, 3(2), s. 159-167.
- Turan, T., Kemaloğlu, N., & Küçükşille, E. (2020). Hukuk'ta Yapay Zeka: Çalışmalar ve Gelecek Öngörüler. *Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 11(2), s. 246-255.
- Uzun , Y., Bilban, M., & Arıkan, H. (2018). Tarım ve Kırsal Kalkınmada Yapay Zeka Kullanımı. *VI. KOP Bölgesel Kalkınma Sempozyumu- -KOPBKS*. Konya.

- Ünal, S. (2022). Blok Zincir Tabanlı Arama Motorları ve Günlük Yaşamda Kullanımı. *Yeni Medya Elektronik Dergi*, 6(3), s. 241-255.
- Yaman, F. (2013, Haziran). Tüketici Davranışı Olarak Özgeciliğin İncelenmesi. *Tüketici ve Tüketim Araştırmaları Dergisi*, 5(1), s. 79-92.
- Yaman, Y. (2021). Müşteri Yolculuğu Analitiğini Makine Öğrenmesi Modeliyle Tahminleme. *YBS Ansiklopedi*, 9(1).
- Yılmazel, B. (2019). *Algoritmalar ve Programlama*. Eskişehir: Anadolu Üniversitesi.
- Yuming Li, J. C. (2023, November). Mixed emotion extraction analysis and visualisation of social media text. *Data & Knowledge Engineering*, 148.
- Zülfiye Erdoğan, Ö. H. (tarih yok). *Öğrenci Başarısının Bileşik Makine Öğrenme Teknikleri Kullanılarak Tahmini*.

<https://support.google.com/> "Çevrimiçi" Erişim Tarihi: 20/12/2023

<https://gs.statcounter.com/search-engine-market-share#monthly-200901-202312>

"Çevrimiçi" Erişim Tarihi: 21/12/2023

<https://trends.google.com/trends/explore?geo=TR&hl=tr> "Çevrimiçi" Erişim Tarihi: 21/12/2023

https://tr.wikipedia.org/wiki/Soyut_makine "Çevrimiçi" Erişim Tarihi: 08/01/2024