



E-Ticaret Sitelerindeki Türkçe Ürün Yorumları Üzerine Makine Öğrenmesi Algoritmaları ile Duygu Analizi

Seda Kılıçer^{1,2*} , Ruya Samli² 

^{1*} İstanbul Beykent University, Computer Engineering Department, İstanbul, Türkiye

² İstanbul University -Cerrahpaşa, Computer Engineering Department, İstanbul, Türkiye

Özet

Bu çalışmada e-ticaret sitelerindeki elektronik cihazlar hakkında yapılan Türkçe 11.412 yorumdan oluşan veri seti üzerinde çeşitli makine öğrenmesi algoritmaları uygulanarak duygu analizi yapılmıştır. Bu veri setinden elde edilen duygu durumları pozitif, negatif, nötr olarak sınıflandırılmıştır. Kullanılan makine öğrenmesi algoritmaları K-En yakın komşu (IBK), K-Star, Naive Bayes (NB), Naive Bayes Multinomial (NBM), Lojistik Regresyon (LR), Destek Vektör Makinesi (DVM), Karar Ağacı (J48) ve Rastgele Orman (RO) şeklindedir. K-katmanlı çapraz doğrulama yöntemi ile makine öğrenmesi algoritmaları kullanılarak kesinlik, duyarlılık, F-skor, AUC kriterlerine göre elde edilen başarılar karşılaştırılmıştır. Pozitif, negative, nötr, yorumlar için ve ağırlıklı ortalamaya göre ayrı tablolarda elde edilen sonuçlar kıyaslanmıştır. Bu sonuçlara göre en yüksek AUC değeri NBM ve RO algoritmalar ile; ağırlıklı ortalama en yüksek AUC değeri de NBM algoritması ile elde edilmiştir. Büyük veri setleri ile nötr cümleler üzerine yapılan bir çalışma olarak elde edilen başarılar nötr cümleler üzerine daha fazla çalışma yapılmasını desteklemektedir. Ayrıca bu çalışmada 8 farklı algoritma kullanılarak algoritma seçiminin önemi gösterilmiştir.

Anahtar Kelimeler: Duygu Analizi, Makine öğrenmesi algoritmaları Naive Bayes Multinomial, F-skor, AUC.

Makale Bilgisi

Başvuru:

02/09/2023

Kabul:

22/10/2023

Sentiment Analysis with Machine Learning Algorithms on Turkish Product Reviews on E-Commerce Sites

Abstract

In this study, sentiment analysis was carried out by applying various machine learning algorithms on the data set consisting of 11,412 Turkish comments about electronic devices on e-commerce sites. Mood states obtained from this data set were classified as positive, negative and neutral. The machine learning algorithms used are K-Nearest neighbor (IBK), K-Star, Naive Bayes (NB), Naive Bayes Multinomial (NBM), Logistic Regression (LR), Support Vector Machine (DVM), Decision Tree (J48) and Random Forest (RO) form. The successes obtained according to precision, sensitivity, F-score, AUC criteria were compared using the K-layer cross validation method and machine learning algorithms. The results obtained in separate tables for positive, negative and neutral comments and according to the weighted average are compared. According to these results, the highest AUC value is with NBM and RO algorithms; The weighted average highest AUC value was also obtained with the NBM algorithm. The successes achieved as a study on neutral sentences with large datasets support further studies on neutral sentences. In addition, in this study, the importance of algorithm selection is shown by using 8 different algorithms.

* İletişim e-posta: sedakilicer@beykent.edu.tr

Keywords: Sentiment Analysis, Machine learning algorithms Naive Bayes Multinomial, F-measurement, AUC.

1 Giriş

Günümüzde insanlar yaşam kalitelerini arttırmak ve daha rahat bir yaşam sürdürmek için teknolojik gelişmelere günlük hayatlarında daha fazla yer vermektedir. Bu teknolojik gelişmeler içerisinde internetin gelişmesi ile birlikte, birçok işlemin yapılmasını sağlayan çevrimiçi platformlar büyük yer kaplamaktadır. Özellikle alışveriş konusunda insanlara büyük kolaylıklar sağlayan çevrimiçi platformlar sayesinde artık her türlü alışveriş bu şekilde yapılabilmektedir. Bu alışveriş platformlarında kişilerin farklı ürünleri satın alması ya da kiralaması, farklı ödeme yöntemleri ile ödeme işleminin gerçekleştirilmesi, çok farklı noktalara kargo hizmetinin verilmesi gibi sebeplerden ötürü platformlarda sunulan hizmetin güvenilir, hızlı, kolay vb özelliklere sahip olması gittikçe önem kazanmaktadır. Bu sebeple birçok platform, kullanıcılarının kendilerini değerlendirmelerine olanak vererek hizmet kalitelerini paylaşmaktadırlar. Böylece insanlar almak istedikleri ürün veya hizmet ile ilgili diğer insanların düşünce ve tecrübelerini kolaylıkla öğrenebilmektedir.

Teknolojik gelişmelerin çevrimiçi ortamlardaki veri miktarının hızla artmasını sağladığı bilinmektedir. Bu veri yığının içerisinde insanlar için faydalı bilginin manuel olarak bulunabilmesi çok zordur. Hızlı bir şekilde artan verilerden gerekli olan bilginin de hızlı ve güvenilir bir şekilde elde edilebilmesine ihtiyaç duyulmaktadır. Bu amaçla çeşitli yöntemler kullanılmaktadır. Yukarıda ifade edilen ve çevrimiçi alışveriş sitelerinde (kısaca e-ticaret siteleri olarak da ifade edilebilir) yer alan müşteri yorumları da diğer çevrimiçi platformlardakine benzer şekilde; bu sitelerdeki veri miktarının hızla artmasını sağlamaktadır ve dolayısıyla bu sitelerde de verilerin doğru bir şekilde analiz edilmesine ihtiyaç bulunmaktadır.

Verilerin işlenerek insanlar için faydalı bilgiye dönüştürülmesini sağlayan alanlardan biri de duygu analizidir. Duygu analizi, herhangi bir veri seti üzerinden kişilerin hissettiklerinin tespit edilmesi ile gerçekleştirilir ve herhangi bir ürün veya hizmet ile ilgili deneyimlerin paylaşılmasıyla elde edilen verilerin hızla analiz edilerek insanlar

için faydalı bilgilere dönüştürülmesi işlemi olarak ifade edilebilir [1]. Duygu analizi ile yazılı metinlerdeki genel duygu durumu belirlenmektedir [2]. Bu sayede insanların hızlı bir şekilde değerlendirme yaparak karar verme süreçlerinin kolaylaştırılması hedeflenmektedir [3].

Literatürde duygu analizi ile bir cümledeki duygu durumunun belirlenmesi için sözlük tabanlı yaklaşım, makine öğrenmesi yaklaşımı ve hibrit yaklaşım olmak üzere 3 farklı yaklaşım kullanılmaktadır.

Sözlük tabanlı yaklaşım, kelimeler ve kelimelere atanan duygu puanlarından oluşturulan sözlük kullanılarak duygu durumlarının tahmin edildiği yöntemlere verilen isimdir [4]. Sözlük tabanlı yaklaşımlar kullanılarak duygu analizi yapılırken herkese açık olarak oluşturulan sözlükler kullanılabilir. SentiTurkNet duygu analizi sözlüğü Türkçe kelimelerden oluşturulan ilk duygu kutbu sözlüğüdür ve birçok çalışmada kullanılmıştır [5].

Makine öğrenmesi yaklaşımı, eğitim veri seti ve test veri seti olarak ayrılan veri setleri ile çeşitli makine öğrenmesi algoritmalarının kullanılarak duygu durumlarının tahmin edildiği yöntemlerdir. Makine öğrenmesi yaklaşımında, diğer alanlardaki makine öğrenme çalışmalarına benzer şekilde sınıflandırma işleminin başarılı bir şekilde yapılması önemlidir ve bu açıdan verilere çeşitli sınıflandırma algoritmaları uygulanmaktadır.

Hibrit yaklaşım ise sözlük tabanlı yaklaşım ve makine öğrenmesi yaklaşımının birlikte uygulanarak duygu durumlarının tahmin edildiği yöntemlere verilen isimdir [6].

Bu çalışmada e-ticaret sitelerindeki Türkçe ürün yorumları üzerine makine öğrenmesi algoritmaları ile duygu durumlarının tahmin edilmesi işlemi gerçekleştirilmiştir.

Bu çalışmanın organizasyonu aşağıdaki gibidir;

Literatür Taraması başlığı taşıyan 2. bölümde; benzer algoritmaları kullanan literatürdeki çalışmalar özetlenmiştir. Materyal ve metot başlığı taşıyan 3. Bölümde kullanılan veri seti, kullanılan makine öğrenmesi algoritmaları, kullanılan başarı kriterlerinden bahsedilmiştir. Bulgular başlığı

taşıyan 4. Bölümde çalışma sonucunda elde edilen bulgular değerlendirilmiştir. Sonuç başlığı taşıyan 5. Bölümde çalışma sonucunda elde edilen sonuçlardan bahsedilmiştir.

2 Literatür Taraması

Farklı dillerde yazılan metinlerde yapılan duygu analizi çalışmalarında kullanılan veri setleri, yöntemler ve elde edilen başarıların karşılaştırılabilmesi için bu alanda yapılan “Farklı dillerdeki metinler üzerinde yapılan duygu analizi çalışmalarının incelenmesi” çalışması incelenmiştir.

Çalışma sonucunda Türkçe olarak yazılan metinlerde elde edilen başarı oranlarının diğer dillerde yazılan metinlerde elde edilen başarı oranlarına göre daha düşük olduğu gözlenmiştir [7]. Türkçe metinlerin incelenmesi, birçok dile göre farklı zorluklar içermekte olduğundan Türkçe metinler ile ilgili yapılan her çalışma literatüre bir katkı anlamına gelmektedir. Tablo 1’de 2013-2023 yılları arasında literatürdeki Türkçe metinlerdeki duygu analizi çalışmalarında makine öğrenmesi yaklaşımı kullanılarak elde edilen başarı oranları gösterilmiştir.

Tablo 1. Türkçe metinlerde yapılan duygu analizi çalışmaları

Ref	Veri Seti	Algoritmalar	En İyi Algoritma	En İyi Sonuçlar
[8]	Twitter: Pozitif, negatif ve nötr etiketli 2 veri seti	NB, RO, IBk, DVM, J48	DVM	Başarı •A veri seti: %65, •B veri seti: %65,3
[9]	Twitter: Pozitif, negatif ve nötr etiketli 2 veri seti dengeli: 824 veri dengesiz: 2.000 veri	NB, J48, IBk, RO, DVM	DVM	Doğruluk •Dengeli veri seti: %72,33 •Dengesiz veri seti: %66,40
[10]	Twitter: Apple, Samsung ve LG markaları ile ilgili 500'er tweet verisi	NB, DT, IBK, RO, DVM	NB	Doğruluk •Apple veri seti: %53,30 •Samsung veri seti: %56,53 •LG veri seti: %65,63
[11]	Twitter: 14.777adet tweet verisi	DVM, IBK, NB, NBM	NBM	Doğruluk •%66
[12]	Twitter: Pozitif ve negatif dengeli 10.600 tweet verisi	NB, DVM, LR	NB ve LR	Doğruluk •%77
[13]	3 farklı cep telefonu markasına ait pozitif ve negatif dengeli 1.000 yorum verisi	NB, DVM, IBK	DVM ve NB	Doğruluk •Pozitif tahminleme: (DVM) %74 •Negatif tahminleme: (NB) %83
[14]	Kitap yorumu: 700 pozitif, 700 negatif yorumdan oluşan dengeli veri seti	NB, DVM, LR, Çok Katmanlı Algılayıcı (MLP)	MLP	Sınıflandırma başarısı •%89
[15]	Yemek Sepeti: 676.000 verilik veri seti	DVM, IBK, NB, LR, Yapay Sinir Ağları (YSA)	YSA	Doğruluk •%86
[16]	Pozitif, negatif ve nötr etiketli 3.135 verilik veri seti	DVM, RO, XGBoost, IBK	DVM	Doğruluk •%89
[17]	Twitter: 16.000 tane pozitif ve 16.000 tane negatif etiketli Türkçe tweet verisi	NB, IBK, DVM, LR, J48	LR	F-skoru %71

3 Materyal ve Metot

Bu bölümde çalışmada kullanılan veri seti, veri seti üzerinde yapılan ön işlemler, kullanılan

3.1 Veri Seti

Bu çalışmada Müjdat Çabuk tarafından Kaggle sitesinde paylaşılan Türkçe yorumlardan oluşan veri seti kullanılmıştır [18]. Bu veri seti çeşitli e-ticaret sitelerindeki elektronik cihazlar hakkında yapılmış 15.170 müşteri ürün yorumundan oluşmaktadır. Türkçe cümlelerle duygu durumlarının tahmin edilebilmesi çalışmalarında duygu durumları farklı şekillerde kategori edilmektedir. Bazı çalışmalarda pozitif ve negatif duygu durumu incelenirken bazılarında pozitif, negatif ve nötr olarak incelenmektedir.

- Pozitif yorum, içerisinde güzel düşünceler barındıran ve genelde ürünün iyi ve beğenilen yönlerinin ifade edildiği yorumlardır.
- Negatif yorum, içerisinde kötü düşünceler barındıran ve genelde ürünün kötü ve

algoritmalar hakkında bilgi verilmiştir. Ayrıca çalışma sonucunda elde edilen sonuçların kıyaslanmasında kullanılan başarı ölçütleri tanıtılmıştır.

beğenilmeyen yönlerinin ifade edildiği yorumlardır.

- Nötr yorum, bu cümleler hakkında net bir duygu tanımlaması yapılamayan yorumlardır. Bu yorumlar iyi ve kötü duyguyu aynı anda barındırmaktadır. Ürün hakkında iyi ve kötü yorumlar aynı yorum içerisinde belirtilmektedir. Bu sebeple bu yorumlar hakkında net bir duygu durumu belirlenmemektedir ve bu yorumların duygu durumu nötr olarak ifade edilmektedir.

Bu veri setinde duygu durumlarını göstermek için 0,1,2 rakamları kullanılmıştır. Bu rakamlardan 1 pozitif, 0 negatif, ve 2 nötr duygu ifade eden cümleleri belirtmek için kullanılmıştır. Tablo 2’de bahsedilen veri seti ile ilgili örneklerden bir kısım paylaşılmıştır.

Tablo 2. Veri Seti Örneği

Metin	Durum
evet anlatıldığı gibi	1
Daha öncede almıştım bu cihazdan ense ve sakal tüketmek için on numara sıfıra yakın alıyor	1
Ürün gayet başarılı sakal kesmede başlık sayısı biraz daha fazla olabilirdi.Hem 0 a yakın alıyor. hem de kirli sakal için 3 numara başlık ideal.	1
Daha öncede aynısını almıştım çok güzel ve kaliteli bir ürün.	1
Erkek kuaförüyüm ense ve sıfır sakal traşı için uygun bir ürün	2
ürün gerçekten çok güzel	1
Ürün beklediğimden güzel çıktı gayet kullanışlı tarklarda bir iki numara saha olsa daha iyidi ama her şey harika	1
tavsiye edebileceğim çok güzel bir makina	2
ürün geldiğinde şarjı vardı. ilk lullanım öncesi 10 saat kadar şarjda kaldı. yaklaşık 30 dk kadar kesintisiz kullandım vebhala şarjı iyi seviyedeydi. 30 dk kullanım sürecinde ısınmadı. ben saç traşı oldum. herhangi bir takılma-çekme vs sorun yaşamadım (saçlarım seyrel ve ince telli. uzunluğu 3-4 mm kopardı, taraksız traş oldum) bu performansı devam ederse mükemmel bir ürün diyebilirim.	1
Ürün sadece aldığım gün güzel bir şekilde kesti. sonrasında saatlerce uğraşıp sadece sakalımı toplayabildim. Allah kahretsin başka bı makine alacam ama bu sefer powertec mi ? ASLA	0

Bu veri seti incelendiğinde veri setindeki cümlelerin e-ticaret sitelerinden çekildiği gibi (yazım yanlışları vb dahil olmak üzere) paylaşıldığı gözlenmiştir. Bu veri seti ilk olarak önışlem

aşamasından geçirilmiştir. Aşağıda belirtilen ön işlem adımları sırasıyla yapılmıştır.

- Veri setinde yer alan İngilizce yorumlar silinmiştir.
- Tekrarlı veriler silinmiştir.

- @, &#, ?,
 ifadeleri silinmiştir.
- Kelimelerin yazımı sırasında kullanılan gereksiz harf tekrarları silinmiştir.
- Metin içerisindeki ç, ğ, ı, ö, ş, ü harfleri sırasıyla c,g,i,o,s,u harfleri ile değiştirilmiştir.

Ön işlemlerden sonra veri setinde ürün yorumları ile ilgili 11.412 yorum bulunmaktadır. Bu yorumların belirlenen duygu durumlarına göre dağılımları Tablo 3'te detaylı olarak gösterilmiştir.

Tablo 3. Duygu durumlarına göre dağılımları

Duygu Durumu	Yorum Sayısı
Pozitif Yorum Sayısı	4.244
Negatif Yorum Sayısı	4.235
Nötr Yorum Sayısı	2.933
Toplam Yorum Sayısı	11.412

Bu çalışmada yorumların duygu durumlarının tahmin edilebilmesi için WEKA isimli yazılım kullanılmış ve bu yazılım ile makine öğrenmesi algoritmaları olan IBk, K-Star, NB, NBM, Lojistik, DVM, J48 ve RO algoritmaları kullanılarak elde edilen sonuçlar incelenmiştir.

Türkçe yorumlar ile duygu analizi uygulaması yapılırken ilk olarak tüm veriler eğitim verisi ve test verisi olarak ayrılmıştır. Eğitim ve test verilerinin belirlenmesi sırasında tüm verilerin %70'i eğitim verisi ve %30'u test verisi olarak ayrılmıştır. Eğitim veri seti belirlenen her duygu durumuna (pozitif, negatif, nötr) ait eşit veri olacak şekilde oluşturulmuştur. Test veri seti ise tüm veri setindeki verilerden eğitim veri seti için ayrılan verilerin çıkarılmasıyla elde edilmiştir. Eğitim veri setindeki cümlelerin duygu durumları sisteme öğretilerek daha sonra test veri setinde yer alan cümlelerin duygu durumlarının tahmin edilmesi sağlanmıştır. Eğitim veri seti her yorumun yanına belirlenen duygu durumlarından hangi duygu durumuna ait olduğu bilgisi eklenerek

Bu çalışma sonucunda K-Star algoritmasının hem eğitim hem de uygulama sonucunda elde edilen doğruluk oranının çok düşük olduğu gözlenmiştir. Bu durum K-Star algoritması ile sadece yorumlardan oluşan bir veri seti kullanılarak başarılı sonuçlar elde edilemeyeceği düşünülmektedir. Ayrıca eğitim ve test veri setleri

- Emojiler silinmiştir.

oluşturulmuştur. Eğitim için hazırlanan eğitim veri seti uygun formata dönüştürülerek WEKA (<https://sourceforge.net/projects/weka/>) yazılımına yüklenmiştir. Daha sonra makine öğrenmesi algoritmaları sırasıyla seçilerek, seçilen algoritma ile eğitim işlemleri yapılmıştır. Tablo 5'te kullanılan algoritmalar ve eğitim sonucunda elde edilen başarı oranları karşılaştırmalı olarak gösterilmiştir.

Test veri seti ise her yorumun yanına duygu durumu bilgisinin kullanılan algoritma tarafından tahmin edilebilmesi için boş bir sütun eklenerek oluşturulmuştur. Eğitim işlemi tamamlandıktan sonra kaydedilen eğitim dosyası ile birlikte hazırlanan test veri seti sisteme yüklenerek yorumların duygu durumları tahmin edilmiştir. Eğitim dosyası sisteme yüklendiğinde test veri seti için de aynı algoritma otomatik olarak seçili gelmektedir. Bu işlem sonucunda her yorumun duygu durumu tahmini sırasıyla çıktı olarak elde edilmiştir.

Test veri setindeki manuel olarak etiketlemiş yorumlar ile uygulama sonucunda elde edilen tahmini duygu durumları karşılaştırılmıştır. Manuel olarak etiketlenen duygu durumu ile uygulama tarafından tahmin edilen duygu durumu aynı ise sonuç doğru olarak kabul edilmiştir. Tablo 4'te kullanılan algoritmalara göre elde edilen doğruluk oranları karşılaştırmalı olarak gösterilmiştir.

Tablo 4. Doğruluk Oranları

Algoritma	Eğitim Veri Seti Doğruluk Oranı	Uygulama Doğruluk Oranı
IBk	%99	%56
K-Star	%33	%7
NB	%63	%65
NBM	%73	%71
Lojistik	%86	%65
DVM	%83	%68
J48	%80	%65
RO	%99	%71

arasında elde edilen doğruluk oranları farkının da incelenmesi gerektiği ve farklı başarı kriterlerine göre değerlendirilmesi gerektiği düşünülmüştür.

3.2 Algoritmalar

Bu çalışmada Türkçe metinlerde duygu durumları analizinde makine öğrenmesi algoritmalarından

IBk, K-Star, NB, NBM, Lojistik, DVM, J48, RO algoritmaları kullanılmıştır. Algoritmalar aşağıda kısaca açıklanmıştır.

IBK algoritması: Sınıflandırılan verilerin incelenerek yeni sınıflandırılacak verinin en yakın komşularının incelenerek doğru sınıfın tahmin edilmesine dayanan bir algoritmadır [13].

K-Star algoritması: Entropi uzaklığının hesaplanarak sınıflandırma işleminde kullanıldığı bir algoritmadır [19].

NB algoritması: Sisteme sunulan öğretilmiş verileri kullanarak yeni sunulan veri setinin sınıflarının tahmin edilmesinde NB algoritması olasılığı yüksek olan bir algoritmadır [13].

NBM algoritması: Metin dokümanlarındaki kelimelerin frekansına göre sınıflandırma tahmininde bulunan bir algoritmadır. Metin sınıflandırma problemlerinde tercih edilmektedir [20].

Lojistik algoritması: Bağımlı ve bağımsız değişkenler arasındaki ilişkiyi matematiksel olarak bularak tahminde bulunmaya dayanan bir algoritmadır [6].

DVM algoritması: İstatistiksel öğrenme yöntemine dayanan bir algoritmadır. Eğitim veri seti kullanılarak öğrenme işlemi tamamlandıktan sonra test veri seti üzerinde tahmin yapılmasına dayalı bir algoritmadır [13].

J48 algoritması: Bir karar algoritmasıdır. Düğüm ve dallardan oluşan bir yapı oluşturularak sınıfların tahmin edilmesine dayalı bir algoritmadır [21].

RO algoritması: Karar ağaçları oluşturularak başarılı bir sınıflandırma yapmayı hedefleyen bir algoritmadır. Farklı karar ağaçları farklı eğitim verileri kullanılarak eğitilir ve daha sonra sonuçlar birleştirilir. Test verisi ise bu sınıflandırma ağaçlarının tahminleri kullanılarak tahmin edilmeye çalışılır [20]

3.3 Başarı Kriterleri

Duygu analizi ile ilgili çalışmalar incelendiğinde genellikle; doğruluk, kesinlik, duyarlılık, F-skor ve AUC değerlerinin başarı kriterleri olarak kullanıldığı görüldüğünden [8-17] bu çalışmada da aynı kriterler ele alınmış ve k-katmanlı çapraz doğrulama yöntemi kullanılarak algoritmalar bu kriterlere göre karşılaştırılmıştır.

K-katmanlı çapraz doğrulama aşırı öğrenme veya eksik öğrenme olup olmadığını tespit etmek

amacıyla kullanılmaktadır. Veri seti k eşit parçaya bölünerek her parça eğitim ve test için kullanılır. Burada k değeri 10 olarak seçilmiştir.

Kullanılan modelin başarısı sınıfı doğru belirlenen yorum sayısı ve sınıfı yanlış belirlenen yorum sayısı karşılaştırılarak belirlenmektedir [22]. Çapraz doğrulama yöntemi kullanılarak ise seçilen algoritmaya göre elde edilen doğruluk oranı, hassaslık, geri çağırma, F-skor, alıcı işletim karakteristiği (Receiver Operating Characteristic-ROC) eğrisinin altındaki alan (Area Under Curve - AUC) gibi kriterlere göre elde edilen sonuçlar karşılaştırılır. Bu değerlerin hesaplanabilmesi için karmaşıklık matrisindeki değerler kullanılmaktadır.

2 sınıf için oluşturulan karmaşıklık matrisi Tablo 5'te görülmektedir.

Tablo 5. Karmaşıklık Matrisi

		Gerçek Değer	
		Doğru	Yanlış
Tahmin Değeri	Doğru	Doğru Pozitif (True Positive - TP)	Yanlış Pozitif (False Positive - FP)
	Yanlış	Yanlış Negatif (False Negative- FN)	Doğru Negatif (True Negative - TN)

Doğruluk Oranı

Doğru sınıflandırılmış veri sayısının toplam veri sayısına oranı olarak hesaplanır [23]. Doğruluk formülü denklem (1)'de verilmiştir.

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Kesinlik

Doğru pozitif örnek sayısının tüm pozitif örnek sayısına oranıdır [22]. Kesinlik formülü denklem (2)'de verilmiştir.

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (2)$$

Duyarlılık

Doğru tahmin edilen pozitif örneklerin sayısının, toplam doğru örnek sayısına oranıdır [19]. Duyarlılık formülü denklem (3)'te verilmiştir.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (3)$$

F-skor

Kesinlik ve duyarlılık sonuçlarının harmonik ortalaması olarak ifade edilmektedir [24]. F-skor formülü denklem (4)'te verilmiştir.

$$F - skor = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (4)$$

AUC (Eğri Altında Kalan Alan)

ROC eğrisi sınıflandırmadaki performansı gösteren bir diğer kriterdir. Olasılık eğrisi olarak ifade edilmektedir. Bu eğrinin altında kalan alan ise AUC olarak ifade edilmektedir [25]. AUC değeri 1'e ne kadar yakın ise başarı oranının o kadar yüksek olduğu anlamı çıkarılır.

4 Bulgular

Çeşitli e-ticaret sitelerindeki elektronik cihazlar hakkında yapılmış 11.412 müşteri ürün yorumundan oluşan veri seti WEKA yazılımı kullanılarak değerlendirilmiştir. Bu işlemler esnasında k-katmanlı çaprazlama kullanılmış ve makine öğrenmesi algoritmaları duyarlılık, kesinlik, F-skor ve AUC başarı kriterleri değerlerine göre karşılaştırılmıştır.

Türkçe metinler üzerinde farklı algoritmalar ile yapılan duygu analizi uygulamaları sonucunda başarı kriterleri olarak belirlenen kesinlik, duyarlılık, F-skor, AUC değerleri elde edilmiştir.

Pozitif, negatif, nötr yorumlar için ve ağırlıklı ortalama sonucunda uygulanan algoritmalar ile elde edilen başarı oranları Tablo6-Tablo 9'da detaylı olarak gösterilmiştir.

Tablo 6. Pozitif duygu sınıfı sonuçları

POZİTİF DUYGU SINIFI				
	Kesinlik	Duyarlılık	F-skor	AUC
IBk	0,544	0,829	0,657	0,800
K-Star	-	0,000	-	0,500
NB	0,741	0,664	0,701	0,827
NBM	0,765	0,840	0,801	0,903
LR	0,737	0,783	0,759	0,870
DVM	0,767	0,808	0,787	0,864
J48	0,702	0,766	0,732	0,813
RO	0,727	0,843	0,781	0,898

Tablo 7. Negatif duygu sınıfı sonuçları

NEGATİF DUYGU SINIFI				
	Kesinlik	Duyarlılık	F-skor	AUC
IBk	0,724	0,488	0,583	0,796
K-Star	-	0,000	-	0,500
NB	0,691	0,711	0,520	0,758
NBM	0,757	0,676	0,714	0,858
LR	0,687	0,690	0,689	0,826
DVM	0,713	0,724	0,718	0,822
J48	0,656	0,711	0,683	0,781
RO	0,688	0,814	0,746	0,883

Tablo 8. Nötr duygu sınıfı sonuçları

NÖTR DUYGU SINIFI				
	Kesinlik	Duyarlılık	F-skor	AUC
IBk	0,381	0,271	0,317	0,595
K-Star	0,257	1,000	0,409	0,257
NB	0,453	0,503	0,476	0,716
NBM	0,515	0,522	0,518	0,745
LR	0,462	0,417	0,438	0,682
DVM	0,482	0,434	0,457	0,629
J48	0,422	0,315	0,361	0,600
RO	0,545	0,275	0,365	0,721

Tablo 9. Ağırlıklı ortalama sonuçları

AĞIRLIKLI ORTALAMA				
	Kesinlik	Duyarlılık	F-skor	AUC
IBk	0,569	0,559	0,542	0,746
K-Star	-	0,257	-	0,500
NB	0,648	0,640	0,643	0,802
NBM	0,698	0,697	0,696	0,846
LR	0,648	0,655	0,651	0,805
DVM	0,674	0,681	0,677	0,788
J48	0,613	0,630	0,618	0,747
RO	0,666	0,686	0,661	0,847

K-katmanlı çapraz doğrulama yöntemi ile farklı makine öğrenmesi algoritmaları kullanılarak Türkçe metinlerde yapılan duygu analizi uygulamaları sonucunda elde edilen doğruluk oranları Tablo 10'da detaylı olarak gösterilmiştir.

Tablo 10. Doğru sınıflandırma yüzdeleri

Kullanılan Algoritma	Doğru Sınıflandırma Yüzdeleri
IBK	%55,9
K-Star	%25,7
NB	%63,9
NBM	%69,7
LR	%65,4
DVM	%68,0
J48	%62,9
RO	%68,6

5 Sonuç

Duygu analizi çok önemli bir konu olmasına rağmen literatürde Türkçe metinler üzerine yapılan duygu çalışmalarının diğer dillerde yapılan çalışmalara göre daha az olduğu ve başarı oranlarının da çok düşük olduğu gözlenmiştir. Literatürdeki çalışmalar incelendiğinde genel olarak pozitif ve negatif duygu durumları ile incelemeler yapıldığı görülmüştür. Nötr duygu durumunun da incelendiği çalışmalarda ise veri sayısının çok daha az olduğu görülmüştür. Bu çalışmada hem 11.412 yorumdan oluşan büyük bir veri seti kullanılmış hem de 3 duygu durumuna göre duygu analizi yapılmıştır. Bu nedenle bu çalışma literatürdeki diğer çalışmalardan ayrılmaktadır. Bu çalışmada farklı e-ticaret sitelerindeki elektronik cihazlar ile ilgili yapılan Türkçe müşteri yorumları incelenmiştir. Kaggle sitesinden alınan 15.170 yorumluk veri seti üzerinde veri temizleme ön işlemleri gerçekleştirilmiş; oluşan 11.412 yorumluk veri seti çalışmanın ana veri seti olarak kullanılmıştır.

Türkçe yorumlar üzerine duygu analizi uygulanmasında makine öğrenmesi yaklaşımı kullanılmıştır. Makine öğrenmesi yaklaşımı uygulanırken IBK, K-Star, NB, NBM, Lojistik, DVM, J48, RO algoritmaları kullanılmıştır. Bu çalışmada 8 farklı algoritma kullanılarak veri seti ile arasındaki ilişkiler incelenmiştir. Farklı algoritmaların sonuçları üzerindeki etkisi karşılaştırılmıştır.

Çalışmada kullanılan veri seti %70 eğitim %30 test veri seti olarak ayrılarak algoritmaların öğrenmesi için gerekli olan işlemlerin yapılması sağlanmıştır. Bu sayede, ilk olarak algoritmalar ile ilgili eğitim ve test sürecindeki doğruluk sonuçları elde edilmiş ve duygu analizi çalışması için bu algoritmaların uygunluğu ile ilgili bir izlenim oluşması sağlanmıştır.

Elde edilen sonuçlar incelendiğinde, en yüksek AUC, doğru sınıflandırma ve F-skor değerlerinin NBM ve RO algoritmaları ile elde edildiği gözlenmiştir. Bu başarı kriterleri aşağıdaki gibi elde edilmiştir.

NBM:

- AUC: %84
- Doğru Sınıflandırma: %69
- F-skor: %69

RO:

- AUC: %84
- Doğru Sınıflandırma: %68
- F-skor: %66

Bu çalışmada pozitif, negatif ve nötr duygu durumu üzerine yapılan çalışmalara kıyasla çok daha büyük bir veri seti kullanılmıştır. Elde edilen sonuçlar kıyaslandığında ise literatürdeki diğer çalışmalara oranla bu kadar büyük bir veri seti ile 3 farklı duygu durumu kullanılarak yüksek bir doğruluk oranı elde edilmiştir. Bu durum nötr olarak değerlendirilen cümleler üzerinde de daha fazla çalışılması gerektiğini göstermiştir.

Kullanılan 8 farklı algoritma ile elde edilen sonuçlar karşılaştırıldığında ise hem AUC, F-skor değerlerine göre hem de doğru sınıflandırma yüzdelerine göre aralarında büyük farkların olduğu gözlenmiştir. Bu durum kullanılan veri seti ile seçilen algoritma arasındaki uyumun ne kadar önemli olduğunu göstermiştir. Literatürdeki çalışmalar incelendiğinde bu kadar çeşitli algoritma ile uygulamanın pek olmadığı ve genel olarak benzer algoritmaların kullanıldığı gözlenmiştir. Bu çalışmanın uygulamalardaki algoritma çeşitliliğinin önemini gösterilmesi açısından da literatür için önemli bir çalışma olduğu düşünülmektedir.

Kaynaklar

- [1] Şahinaslan, Ö., Dalyan, H., Şahinaslan, E. 2022. Naive Bayes Sınıflandırıcısı Kullanılarak Youtube Verileri Üzerinden Çok Dilli Duygu Analizi. Bilişim Teknolojileri Dergisi, 15(2), 221-229.

- [2] Mandloi, L., Patel, R. 2020. Twitter Sentiments Analysis Using Machine Learning Methods, International Conference for Emerging Technology, 1-5.
- [3] Alaca, M., Güran, A. 2022. Covid-19 Salgını Sürecinde Duygu Skorlarının Ve Teknik İndikatörlerin Kullanılmasıyla Bist 100 Endeksi Trend Tahmini. Bilişim Teknolojileri Dergisi, 15(4), 379-388.
- [4] Demir Ö, Chawai, A.I.B., Doğan, B. 2019. Türkçe Metinlerde Sözlük Tabanlı Yaklaşımla Duygu Analizi Ve Görselleştirme. International Periodical Of Recent Technologies In Applied Engineering, 1(2), 58-66.
- [5] Dehkharghani, R., Saygin, Y., Yanikoglu, B., Oflazer, K. 2016. SentiTurkNet: a Turkish polarity lexicon for sentiment analysis. Language Resources and Evaluation, 50, 667-685.
- [6] Yadav, K, Patel, M. 2021. Sentiment Analysis on Restaurant Review using Hybrid Approach. International Journal for Research in Applied Science & Engineering Technology, 9(5), 1999-2006.
- [7] Kılıcer, S., Samli, R. 2022. Farklı dillerdeki metinler üzerinde yapılan duygu analizi çalışmalarının incelenmesi. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 13(3), 493-509.
- [8] Çetin M, Amasyalı M.F. 2013. Eğitici ve Geleneksel Terim Ağırlıklandırma Yöntemleriyle Duygu Analizi. Signal Processing and Communications Applications Conference.
- [9] Nizam H, Akın S.S. 2014. Sosyal Medyada Makine Öğrenmesi ile Duygu Analizinde Dengeli Ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması. Türkiye'de İnternet Konferansı, 1(6), 873-883.
- [10] Nalçakan, Y. Bayramoğlu, Ş. S. Tuna, S. 2015. Sosyal Medya Verileri Üzerinde Yapay Öğrenme ile Duygu Analizi Çalışması. Technical Report.
- [11] Çoban Ö, Özyet B, Tümöklü Özyer G. 2015. Türkçe Twitter Mesajlarının Duygu Analizi. Signal Processing and Communications Applications Conference.
- [12] Aytuğ O, 2017. Twitter mesajları üzerinde makine öğrenmesi yöntemlerine dayalı duygu analizi. Yönetim Bilişim Sistemleri Dergisi, 3(2), 1- 14.
- [13] Karamanlı E., 2019. Makine Öğrenmesi Algoritmaları Kullanarak, Metin Madenciliği Ve Duygu Analizi İle Müşteri Deneyiminin Geliştirilmesi. İstanbul Üniversitesi Yüksek Lisans Tezi.
- [14] Tuzcu, S. 2020. Çevrimiçi Kullanıcı Yorumlarının Duygu Analizi ile Sınıflandırılması. Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 1(2), 1-5.
- [15] Aktaş, Ö., Coskuner, B., Soner, İ. 2021. Turkish Sentiment Analysis Using Machine Learning Methods: Application on Online Food Order Site Reviews. Journal of Artificial Intelligence and Data Science, 1(1) 1-10.
- [16] Çılgın, C., Gökçen, H., Gökşen, Y. 2023. Sentiment analysis of public sensitivity to COVID-19 vaccines on Twitter by majority voting classifier-based machine learning. Journal of the Faculty of Engineering and Architecture of Gazi University, 38(2), 1093-1104.
- [17] Kumaş, E. 2021. Türkçe Twitter Verilerinden Duygu Analizi Yapılırken Sınıflandırıcıların Karşılaştırılması. Eskişehir Türk Dünyası uygulama ve Araştırma Merkezi Bilişim Dergisi, 2(2), 1-5.
- [18] <https://www.kaggle.com/datasets/mujdatcabuk/eticaret-urun-yorumlari>
- [19] Şenel, F. A., Saygın, R. R., Saygın, M., & Öztürk, Ö. 2021. Makine Öğrenmesi Algoritmaları Kullanılarak Vücut Analizi ile Uyku Apnesi Teşhisi. Uyku Bülteni, 2(1), 6-10.
- [20] Shirzadova, S., & Uysal, A. K. 2022. Türkçe YouTube Yorumları Üzerinde Spam Filtreleme. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 10(4), 1793-1810.
- [21] B. Oralhan, 2020. Veri madenciliği yaklaşımı ile telekomünikasyon sektöründe arıza giderme analizi. Business & Management Studies: An International Journal, 8,(1), 1026-1043.
- [22] Başer B.Ö, Yangın, M., Sarıdaş E.S. 2021. Makine öğrenmesi teknikleriyle diyabet hastalığının sınıflandırılması. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 25(1), 112-120.
- [23] Bilgin, M. 2017. Gerçek veri setlerinde klasik makine öğrenmesi yöntemlerinin performans analizi. Akademik Bilişim Konferansı, 2(9), 683.
- [24] Çelik, E., Deniz, D.A.L., Aydın, T. 2021. Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması. Avrupa Bilim ve Teknoloji Dergisi, (27), 880-889.
- [25] Göçkün Ö, Onan A., 2021. Amazon ürün değerlendirmeleri üzerinde derin öğrenme/makine öğrenmesi tabanlı duygu analizi yapılması, Avrupa Bilim ve Teknoloji Dergisi, 24,445-448.