Avrupa Bilim ve Teknoloji Dergisi
Sayı 27, S. 880-889, Kasım 2021

Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. 27, pp. 880-889, November 2021

Research Article

Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması

Esra Çelik[1*], Deniz Dal[2], Tolga Aydın[3]

[1*] Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye, (ORCID: 0000-0001-5333-6622), esra.celik@atauni.edu.tr
[2] Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye, (ORCID: 0000-0003-0120-4315), ddal@atauni.edu.tr
[3] Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye, (ORCID: 0000-0002-8971-3255), atolga@atauni.edu.tr

(İlk Geliş Tarihi 29 Mart 2021 ve Kabul Tarihi 21 Kasım 2021)

(**DOI:** 10.31590/ejosat.905259)

ATIF/REFERENCE: Çelik, E., Dal, D., & Aydın, T. (2021). Duygu Analizi İçin Veri Madenciliği Sınıflandırma Algoritmalarının Karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (27), 880-889.

Öz

Teknolojinin gelişmesi, internetin yaygınlaşması ve internet aracılığıyla bilgiye erişim kolaylığı insanların duygu ve düşüncelerini farklı iletişim araçlarını kullanarak paylaşmalarına imkân sağlamaktadır. Uyarlanabilir öğrenme ve karar verebilme gibi yeteneklerle donatılarak daha akıllı hale gelen söz konusu bu iletişim araçları, her geçen gün daha geniş kitlelere ulaşmaktadır. Bir zamanlar sadece ses iletimi için kullanılan bu araçlar şimdilerde insanların forum ve blog gibi sanal ortamlarda duygu ve düşüncelerini yazılı olarak paylaşmalarını mümkün kılmaktadır. Sanal ortamlar aracılığıyla yapılan bu yorumlar artık bir bilgi edinme kaynağı olarak görülmekte ve daha da önemlisi bu yorumlar bireylerin farklı konulara ilişkin düşüncelerinin analiz edilebilmelerini kolaylaştırdıkları için konu üzerinde çalışmalar yürüten araştırmacıların dikkatini fazlasıyla çekmektedir. Başka bir deyişle bu yorumlardan günümüzün popüler bir araştırma alanı olan duygu analizi için gerçek bir veri seti olarak faydalanılmaktadır. Bu çalışmada ürün, film ve restoran yorumlarını içeren farklı veri setlerinden faydalanılarak veri madenciliği sınıflandırma algoritmaları yardımıyla duygu analizi yapılmıştır. Bu amaçla Destek Vektör Makinesi, K-En Yakın Komşu, Naive Bayes, Karar Ağacı ve Rastgele Orman sınıflandırma algoritmalarından faydalanılmıştır. Veri boyutunu ve çeşitliliğini arttırmak amacıyla her biri içerisinde 500 olumlu, 500 olumsuz olmak üzere toplamda 1000 adet yorum içeren üç farklı veri seti birleştirilmiştir. Deneysel sonuçlar Destek Vektör Makinesi sınıflandırma algoritmasının duygu analizi noktasında diğer yöntemlere kıyasla daha başarılı olduğunu göstermiştir.

Anahtar Kelimeler: Duygu analizi, Veri madenciliği, Metin madenciliği, Makine öğrenmesi, K-en yakın komşu, Naive bayes, Destek vektör makinesi, Karar ağacı, Rastgele orman.

Comparison of Data Mining Classification Algorithms for Sentiment Analysis

Abstract

The development of technology, the spread of the Internet and the ease of access to the information through the Internet enable people to share their feelings and thoughts using different communication channels. These communication mediums, that have become smarter by being equipped with the skills such as the adaptive learning and decision making, are reaching a wider audience day by day. While these channels were once used only for the voice transmission, it nowadays enables people to share their feelings and thoughts in the virtual environments such as forums and blogs. The comments made through the virtual environments are now seen as a source of information, and more importantly, these comments attract the attention of researchers who are working on the subject, as they facilitate the analysis of individuals' opinions on different topics. In other words, these comments are used as a real data set for the sentiment analysis, that is one of the popular research areas. In this study, the sentiment analysis was carried out by means of the data mining classification algorithms applied on different data sets including the product, movie and restaurant reviews. For this purpose, Support Vector Machine, K-Nearest Neighbor, Naive Bayes, Decision Tree and Random Forest classification algorithms were taken into consideration. In order to increase the data size and its diversity, three different data sets, each containing 500 positive and 500 negative, 1000 comments in total, were combined. Experimental results showed that the Support Vector Machine classification algorithm is more successful than other methods in the sentiment analysis.

Keywords: Sentiment analysis, Data mining, Text mining, Machine learning, K-nearest neighbor, Naive bayes, Support vector machine, Decision tree, Random forest.

[*] Sorumlu Yazar: esra.celik@atauni.edu.tr

1. Giriş

Mobil cihaz teknolojilerinin gelişmesi, internetin yaygınlaşması ve internet aracılığıyla bilgiye erişim kolaylığı insanların duygu ve düşüncelerini farklı iletişim araçlarını kullanarak sesli, görüntülü ve yazılı formda paylaşabilmelerine imkân sağlamaktadır. Söz konusu her bir formun farklı platformlarda öne çıktığı, örneğin forum ve blog gibi sanal ortamlardaki paylaşım şeklinin genellikle yazılı olduğu bilinmektedir. Bu türden ortamlar aracılığıyla yapılan özellikle ürün, restoran ve film yorumları toplumdaki bireylerin fazlasıyla dikkatini çekmektedir çünkü bu yorumlar ürün incelemelerinde tüketicilerin ürünlerden, restoran incelemelerinde müşterilerin yemeklerden ve benzer şekilde film incelemelerinde de izleyicilerin izledikleri filmlerden memnun olup olmadıklarını ortaya çıkarmaktadır. Bu yorumlar diğer kullanıcıların tercihlerini etkileyebilen bir veri kaynağı olabilmelerinin yanında duygu analizi gibi araştırma alanlarında kullanılabilecek gerçek bir veri setini de oluşturmaktadır. Bu veri setlerinde yer alan yorumları analiz etme ihtiyacı ise her geçen gün artmaktadır. Söz konusu verinin hacmi göz önüne alındığında, bu analizin insan gücüyle ve manuel olarak yapılmasının pratik olmayacağı açıktır. Bu durum yorumların analizini hızlı bir şekilde gerçekleştirecek otomatik uygulamalara duyulan ihtiyacı ortaya koymaktadır. Bu nedenledir ki son yıllarda çeşitli makine öğrenmesi, yapay öğrenme ve veri madenciliği alanındaki algoritmalardan bu verileri kullanarak gizli örüntüleri çıkarma, analiz ve tahmin yapabilme gibi amaçlar için faydalanılmaktadır (Albayrak ve ark., 2017).

Yapay zekâ alanındaki gelişmeler doğal dil işleme ile geliştirilen uygulamaların da gün geçtikçe artmasına neden olmaktadır. Doğal dil işleme, eldeki mevcut verilerden bilgisayarların anlayabileceği anlamların çıkarılması işlemidir. Doğal dil işleme ile geliştirilen uygulamalar, manuel gerçekleştirilmesi zor olan analizler için bilgisayarlardan faydalanarak hızlı ve doğru sonuç alınmasına imkân vermektedir. Duygu analizi bu uygulamalardan biridir.

Duygu analizi özünde bir metin işleme sürecidir ve bir metnin duygusal olarak ifade etmek istediği sınıfı belirlemeyi amaçlamaktadır (Seker, 2016). Başka bir deyişle bir bilimsel çıkarım tekniği olan duygu analizi, çok miktarda metinsel içerikten oluşan veri kümelerinde ifade edilmek istenen duygunun ortaya çıkarılmasını sağlayan bir metin madenciliği çalışmasıdır. Öte yandan metin madenciliği metni bir veri kaynağı olarak kabul eden ve metin üzerinden yapısallaştırılmış veri elde etmeyi amaçlayan bir veri madenciliği çalışması olarak tanımlanmaktadır (Kılınç ve ark., 2016) Duygu analizi alanında farklı teknikleri içeren çalışmalara literatürde sıkça rastlanmaktadır. Bu çalışmaların birçoğu duygu analizi için veri madenciliği, sınıflandırma ve makine öğrenmesi tekniklerinin etkisini incelemiştir (Albayrak ve ark., 2017) (Topaçan, 2016) (Nalçakan ve ark., 2015) (Kaynar ve ark., 2016) (Pang ve ark., 2002) (Zhang ve ark., 2014) (Khan ve ark., 2016). Duygu analizinin ilk uygulamaları metinlerin duygusal olarak olumlu veya olumsuz şeklinde iki gruba ayrılmaya çalışıldığı duygusal kutupsallık üzerine olmuştur (Seker, 2016). Bu amaçla sosyal medyada tartışılan sosyal bir konunun verilerine ait duygu analizleri yapılmıştır ve sosyal medya aracılığıyla üretilen metinlerin istatistiksel yöntemler kullanılarak analiz edilmesi sağlanmıştır (Albayrak ve ark., 2017). Bu kapsamda metin yazarlarının bir konuya karşı düşüncesinin olumlu, olumsuz ya da tarafsız sınıflandırılmasıyla ağı oluşturan topluluğun genel eğiliminin tahmin edilmesi de sağlanmıştır (Topaçan, 2016). Bir başka çalışmada bir sınıf dışı eğitim faaliyeti olan bir projenin, öğrencilerin tutumlarına ve duygularına etkisini belirlemek amacıyla duygu analizi gerçekleştirilmiştir (Demir ve Yılmaz, 2018). Nalçakan ve ark., (2015) referanslı çalışmada Twitter üzerinden Samsung, Apple ve LG markaları için yapılan yorumlar için bir makine öğrenmesi algoritmasıyla iyi, kötü ve duygu belirtmeyen şeklinde bir geribildirim elde edilmiştir. Nitelik seçim yöntemlerinin Türkçe Twitter verileri üzerindeki duygu analizi işleminin performansına etkisi Parlar ve ark., (2017) referanslı çalışma kapsamında değerlendirilmiştir. Burcu ve Şimşek (2018) referansıyla verilen çalışma ile bir televizyon kanalına ait seyirci görüşlerinin sosyal medya üzerinden derlenip kanal için faydalı bilgi elde edilmesi ve işletmeye değer katması amacıyla bir karar destek unsuru oluşturulmuştur. Bir diğer çalışmada film yorumlarının içeriğine göre farklı sınıflandırma algoritmaları kullanılarak duygu ve düşünce analizi yapılmıştır (Kaynar ve ark., 2016). Makine öğrenmesi ile duygu analizi ilk defa 2002 yılında sinema filmlerinin yorumlarını pozitif ve negatif olarak sınıflandırmak amacıyla kullanılmıştır (Pang ve ark., 2002). Sadece pozitif/negatif şeklindeki bir ikili sınıflandırma yerine sinirli, mutlu, mutsuz ve suçluluk hissi gibi birden fazla sınıfa göre duygu sınıflandırması Balahur ve ark., (2012) referanslı çalışmaya konu olmuştur. Zhang ve ark., (2014) referanslı çalışmada dolaylı anlatımla ifade edilen duyguların tespit edilebilmesi için AppleStore'daki mobil uygulama yorumları üzerinde denetimli makine öğrenmesi yöntemleri ve bu yöntemlerin farklı parametrelere göre performansları karşılaştırılmıştır. Sinema yorumları, ürün yorumları gibi farklı veri setleri üzerinde denetimli makine öğrenmesi yöntemi ve destek vektör makinesi sınıflandırıcısı ile gerçekleştirilen bir duygu analizi Khan et al. (2016) referanslı çalışma kapsamında yürütülmüştür. Literatürdeki çalışmalardan anlaşılacağı üzere duygu analizinde genellikle bir veri kümesi içerisindeki ifadelerin sınıflandırılarak olumlu veya olumsuz olup olmadığı sorgulanmaktadır. Bu sorgulama insanların seçimlerini ve herhangi bir konu hakkındaki karar verme sürecini kolaylaştırmaktadır. Öte yandan bu aşamada çok sayıda yorum içeren veri setlerinin analizinin manuel bir şekilde gerçeklenmesi zorluğu karşımıza çıkmaktadır. Bu nedenledir ki duygu analizinin otomatik bir şekilde ve hızlı olarak yapılabilmesi veri madenciliği ve makine öğrenmesi alanları için önemli bir çalışma konusu olmuştur.

Bu çalışmada duygu içeren metinlerle oluşturulmuş büyük veri kümelerinin doğru bir şekilde analiz edilmesi için veri madenciliği ile sınıflandırma yapılmıştır. Bu amaçla *K-En Yakın Komşu (KNN)*, *Naive Bayes (NB)*, *Destek Vektör Makinesi (DVM)*, *Karar Ağacı (KA)* ve *Rastgele Orman (RO)* makine öğrenmesi algoritmalarından faydalanılmıştır ve bu algoritmalar Python programlama dili ve *scikit-learn* kütüphanesi ile gerçeklenmiştir. Söz konusu algoritmaların etkinliği literatürde yaygın olarak kullanılan *Amazon*, *Imdb* ve *Yelp* Anonim, (2020) isimli üç ünlü veri seti ve bu veri setleri birleştirilerek elde edilen dördüncü bir yeni veri seti ile test edilmiştir. Deneysel sonuçlar Destek Vektör Makinesi sınıflandırma algoritmasının duygu analizi noktasında diğer yöntemlere kıyasla daha başarılı olduğunu göstermiştir.

3. Bölüm'de daha detaylı olarak bahsedileceği üzere birleştirme sonrası veri kapasitesi ve veri çeşitliliği artan yeni veri seti sayesinde ilgili makine öğrenmesi algoritmalarının doğruluk ve hassaslık gibi başarım ölçütlerinde önemli derecede

bir iyileşme gözlenmiştir. Söz konusu bu yeni veri setinin performans üzerine etkisinin incelenmesi çalışmamızı literatürdeki benzer çalışmalardan farklı kılan noktalardan birisidir. Bir diğeri ise çalışma kapsamında değerlendirmeye alınan sınıflandırma algoritmalarının çeşitliliğidir.

Bu çalışmanın geri kalan bölümü şu şekilde organize edilmiştir. 2. Bölüm'de materyal ve yönteme yer verilmiştir. 3. Bölüm'de deney sonuçları detaylı grafikler ve tablolar eşliğinde analiz edilmiştir. Son olarak 4. Bölüm ile çalışma sonuçlandırılmıştır.

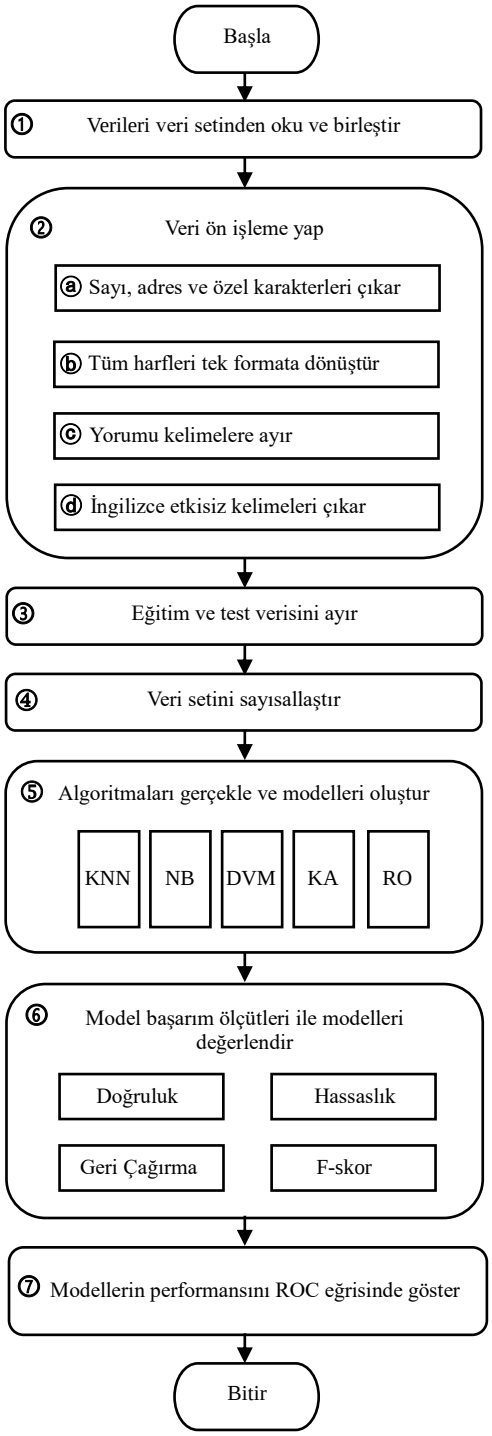
2. Materyal ve Metot

Veri madenciliği sınıflandırma algoritmalarının duygu analizi açısından karşılaştırılmasına yönelik hazırlanan bu çalışmaya ait tasarım akışı diyagramı Şekil 1 ile verilmiştir. Bu diyagramda yer alan her bir adım ise takip eden alt bölümlerde detaylandırılmıştır.

2.1. Veri Seti

Bu çalışma kapsamında literatürde yaygın olarak kullanılan Amazon, Imdb ve Yelp isimli web sitelerinden elde edilen üç farklı veri seti kullanılmıştır. Her bir veri seti içerisinde 500 olumlu ve 500 olumsuz olmak üzere toplam 1000 yorum ve her bir yoruma ait birer etiket bulunmaktadır. Etiket adı verilen bileşen ilgili yorumun olumlu veya olumsuz olduğunu ifade eden sayısal bir değerdir. Veri setlerinde olumlu yorum 1 ve olumsuz yorum ise 0 etiketi ile işaretlenmiştir. Amazon isimli veri seti içerisinde amazon.com'da cep telefonları ve aksesuarları kategorisinde satılan ürünler için yapılan yorumlar ve etiketler mevcuttur. Imdb isimli veri seti içerisinde imdb.com'da yayınlanan filmlere ait yorumlara ve etiketlere, Yelp isimli veri setinde ise yelp.com'daki restoran incelemelerini içeren yorumlara ve etiketlere yer verilmiştir (Kotzias ve ark., 2015). Veri setlerindeki yorumlar sınıflandırma yaklaşımlarını değerlendirmek için kullanılmaktadır.

Bilimsel araştırmalarda kullanılan modellerin değerlendirilmesi noktasında veri setlerinin büyüklüğü ve çeşitliliği oldukça önemlidir (Onan ve Korukoğlu, 2016). Bu nedenle bu çalışmada çeşitliliği sağlamak ve veri setini büyütmek amacıyla tasarım akışının ① numaralı adımında yukarıda bahsedilen üç farklı veri seti birleştirilmiştir. Böylelikle toplamda 3000 yorumu ve bu yorumlara ait 3000 etiketi içeren tek bir veri seti ile modellerin doğruluğu değerlendirilmiştir. Bu aşamada veri çeşitliliğini sağlamak amacıyla eğitim ve test kümeleri rastgele oluşturulmuştur. Başka bir deyişle algoritmanın eğitildiği küme içerisinde hem ürün hem film ve hem de restoran yorumları mevcuttur. Benzer bir durum test kümesi için de geçerlidir.



Okuyucuya bir fikir vermesi açısından bu çalışmada kullanılan veri setlerinin her birinden alınan olumlu ve olumsuz yorum örnekleri ve bunların etiketleriyle Tablo 1 hazırlanmıştır. Yorumlar İngilizce olduğu için parantez içerisinde bu yorumların Türkçe karşılıklarına da Tablo 1'de yer verilmiştir.

Tablo 1. Örnek yorumlar

Veri Seti	Yorum	Etiket
Amazon	It has all the features I want (İstediğim tüm özelliklere sahip)	1
	Returned 8 hours later. (8 saat sonra geri dönüş yapıldı.)	0
Imdb	Excellent short film. (Mükemmel kısa film.)	1
	It was so BORING! (O çok sıkıcıydı!)	0
Yelp	Great food and awesome service! (İyi yemek ve müthiş hizmet!)	1
	Service sucks. (Servis berbattı.)	0

2.2. Veri Ön İşleme

Dil bilgisi kuralları her dil için farklı olduğundan doğal dil işleme süreçleri de dilden dile değişiklik göstermektedir. Doğal dil işleme sürecinde bilgisayar kelimenin kökünü ayrı, kelimelerin dizilimini ayrı, cümlenin ve tüm metnin anlamını ayrı ayrı inceleyerek anlatılmak istenileni öğrenmekte ve bir anlam çıkarmaktadır. Örneğin, İngilizce metinlerin doğal dil işleme süreci Türkçe'den farklıdır. İngilizce, Türkçe gibi sondan eklemeli bir dil değildir, bu nedenle kelime kökünün yapısal incelemesi basittir. Bilgisayar İngilizce kelimelerin diziliminde özne, yüklem ve nesne sıralamasını dikkate almaktadır (bkz. Tablo 1). Öte yandan bilgisayarın cümlede anlatılandan bir anlam çıkarması için cümlenin daha küçük anlamlı birimlere ayrılması gerekmektedir. Bu durum doğal dil işlemede veri ön işleme olarak nitelendirilmektedir.

Bilimsel araştırmalar için toplanan verilerin genellikle herhangi bir ön işlemden geçirilmeden, başka bir ifadeyle yapılandırılmadan veri setlerine dahil edildiği bilinmektedir (Uçkan ve ark., 2019). Oysa ki veri setindeki verilerin kalitesi ile o verileri kullanan algoritmalarla elde edilen sonuçların doğruluğu arasında bir korelasyon bulunmaktadır. Daha da önemlisi bu şekilde hiçbir işlemden geçmemiş veriler sadece sonuçları değil analiz sürecini de olumsuz etkilemektedir. Bu nedenle tasarımın ② numaralı adımında her bir veri seti önce gereksiz içeriklerden arındırma ve yazım hatalarının düzeltilmesi gibi bir takım veri önişleme sürecinden geçirilmiştir. Bu amaçla ⓐ ile ifade edilen alt adımda veri setlerindeki her bir yorumun içeriği kontrol edilerek e-mail adresleri, sayılar, IP adresleri ve özel karakterler *replace* fonksiyonu kullanılarak yorumdan çıkarılmıştır. Daha sonra verileri uygun bir formata dönüştürmek amacıyla ⓑ ile ifade edilen adımda *lower* fonksiyonu yardımıyla tüm harfler küçük harfe çevrilmiştir. Öte yandan yorumları anlamlı birimlere, başka bir ifadeyle kelimelere ayırmak için ⓒ adımında Python'a ait *nltk.tokenize* kütüphanesinde yer alan *word_tokenize* metodundan faydalanılmıştır. Yorumların tamamı İngilizcedir ve yorumlardan elde edilen kelimelerin bir kısmı etkisiz kelimelerdir (stop words). Bu kelimelerin analize bir katkısı olmadığı için ⓓ ile belirtilen adımda Python'a ait *nltk.corpus* kütüphanesinde yer alan *stopwords* metodundan faydalanılarak etkisiz kelimeler yorumlardan arındırılmış ve tüm bu ön işlemeler sonucunda veriler uygun formata dönüştürülmüştür. Örneğin, Amazon isimli veri setinden alınan "*$50 Down the drain.*" yorumu için ön işleme şu şekilde gerçekleştirilmiştir. İlk olarak yorum sayılardan ve özel karakterlerden arındırılmış, daha sonra "*down the drain*" şeklinde tamamen küçük harflerle temsil edilmiş ve son olarak etkisiz kelimeler ayrıştırılarak "*drain*" şeklini almıştır.

2.3. Veri Setini Bölme

Bu çalışma kapsamında dikkate alınan sınıflandırma algoritmalarının tamamında her bir veri setini etkin bir şekilde analiz edebilmek amacıyla tasarımın ③ numaralı adımında verinin %75'lik rastgele kısmı eğitim, %25'lik rastgele kısmı ise modelin testi için ayrılmıştır. Duygu analizi için kullanılan algoritmaların tamamında model oluşturulurken veri setinin eğitim için ayrılan %75'lik kısmından yararlanılmıştır. Model oluşturma işleminden sonra %25'lik test verisi kullanılarak ilgili model üzerinde tahminleme yapılmış ve modelin performansı değerlendirilmiştir.

2.4. Veri Setini Sayısallaştırma

Değerlendirmeye alınan algoritmaların tamamında matematiksel hesaplamaların yapılabilmesi için ön işleme sürecinden geçirilerek temizlenen verilerin ④ numaralı tasarım adımında sayısal değerlere dönüştürülmesi ve her yorumun bir vektörle temsil edilmesi gerekmektedir. Bu amaçla bir doğal dil işleme süreci devreye girmekte ve Terim Frekansı–Ters Metin Frekansı (TF-IDF) yönteminden faydalanılmaktadır. TF-IDF yorumlardaki kelimelerin ön işlemeden geçirilmiş veri setindeki önem derecesini belirleyen bir yöntemdir. Bu yöntem algoritmalara girdi olarak verilmek üzere kelimelerin miktarını dikkate alarak ilgili kelimelerin sayısal ağırlığını hesaplamaktadır (Güran ve Kınık, 2021). Bu çalışmada TF-IDF'yi kullanmak için Python'a ait *sklearn.feature_extraction.text* kütüphanesinde tanımlı *TfidfVectorizer* metodundan faydalanılmış ve bu sayede makine öğrenmesi algoritmalarının önem dereceleri belli olan kelimeleri sayısal olarak kullanabilmeleri sağlanmıştır.

TF-IDF'nin çalışma mantığı Amazon isimli veri setinden alınan ve üzerlerinde ön işleme yapılmış "*good product good seller*" ve "*great sound service*" yorumları kullanılarak Tablo 2'de örneklendirilmiştir. Bu işlem için ❶ numaralı adımda her bir kelimeye ait TF değerlerini hesaplamak amacıyla yorumlardaki kelime sayılarını içeren vektör temsilleri oluşturulmuştur (örneğin *good* kelimesi iki kez kullanıldığı için vektör temsilinde değeri 2'dir). ❷ numaralı adımda her bir yorumdaki toplam kelime sayısı belirlenmiştir. ❸ numaralı adımda vektör temsilindeki her bir değer ilgili yorumdaki toplam kelime sayısına bölünmüş ve normalleştirilmiş TF değerleri elde edilmiştir (örneğin good kelimesinin normalleştirilmiş TF değeri $2/4 = 0.5$'dir). IDF değerini hesaplamak için ❹ numaralı adımda, kullanılan toplam yorum sayısı her bir kelimenin geçtiği yorum sayısına bölünerek logaritması alınmıştır (örneğin *good* kelimesinin IDF değeri $\log(2/1) = 0.3$'tür). Son olarak ❺ numaralı adımda kelimelerin ağırlıklarını hesaplamak için her bir kelimeye ait TF ve IDF değerleri çarpılmıştır (örneğin *good* kelimesinin ağırlığı $0.5 * 0.3 = 0.15$'dir).

Tablo 2. TF-IDF hesaplama

TF-IDF hesaplama adımı	Yorum	
	good product good seller	*great sound service*
❶	[2 1 1]	[1 1 1]
❷	4	3
❸	[0.5 0.25 0.25]	[0.3 0.3 0.3]
❹	[0.3 0.3 0.3]	[0.3 0.3 0.3]
❺	[0.15 0.07 0.07]	[0.09 0.09 0.09]

2.5. Veri Madenciliği ve Sınıflandırma

Yapay zekâ teknolojisi son yıllarda hızlı bir gelişme göstermiştir. Yapay zekânın uyarlanabilir öğrenme ve karar verme yetenekleri sayesinde daha akıllı hale gelen dijital cihazların sayısında görülen artış da dikkat çekmektedir. Bu tür cihazların sayısındaki artış ve bu cihazlar ile yapılan işlemlerin dijital ortamda kayıt altına alınması neticesinde artık büyük veri olarak adlandırılan bir veri ortaya çıkmıştır. Önceleri görülen veri kıtlığı, yerini aşırı bolluğa bırakmış ve bilgiye erişim endişesinin yerini artık erişilebilen miktarla başa çıkma endişesi almıştır (Şentürk, 2006). Biriken bu büyük veri, insanları bu verilerden faydalı bilgiler çıkarmaya yöneltmiş, böylece bu verinin değersiz kalmasının önüne geçilmiştir. Bir zamanlar herhangi bir anlam ifade etmeyen verileri, farklı tekniklerle işleyerek ve analiz ederek anlaşılabilir kılan işlemlerin bütününe veri madenciliği denilmektedir.

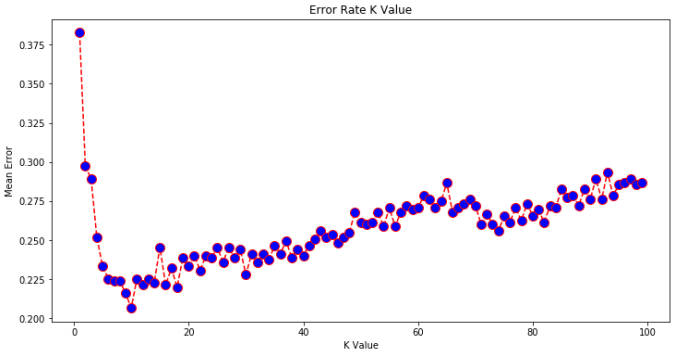
Veri madenciliği, karar verilmesi gereken durumlarda kümeleme ve birliktelik kuralı gibi yöntemleri kullanırken, tahmin etmeye dayalı durumlarda sınıflandırma ve veri kümesindeki değişkenler arasındaki ilişkileri inceleyen regresyon üzerine yoğunlaşmaktadır. Bu çalışmada bir keşif ve öngörü metodu ve aynı zamanda veri madenciliğinin en popüler yöntemlerinden birisi olan sınıflandırmadan faydalanılmıştır. Sınıflandırma, önceden belirlenmiş sınıfları kullanarak sonuçları bilinmeyen verilere ait sınıfların tahmin edilmesini sağlamaktadır. Bu amaçla başlangıçta üzerinde çalışılacak veri seti belirlenen bir yüzdeliğe göre eğitim ve test verisi olarak iki gruba ayrılmaktadır. Daha sonra bu veri setinin eğitim için ayrılan parçası kullanılarak sınıflandırma kuralının oluşturulması sağlanmaktadır. Son aşamaya gelindiğinde ise hazırlanan sınıflandırma kurallarıyla sonuçları bilinmeyen verilere ait sınıfların tahmin edilmesi gerçekleştirilmektedir. Öte yandan sınıflandırma kuralının oluşturulması aşamasında bir veri madenciliği yöntemi olan makine öğrenmesi algoritmalarından da faydalanılmaktadır. Veri madenciliği ve makine öğrenmesi alanları birbirleriyle önemli ölçüde örtüşmektedir. Ortak yönleri oldukça fazla olan bu iki alan gerekli durumlarda destek vererek birbirlerini tamamlamaktadır. Bu sayede günümüzde veri madenciliği ve makine öğrenmesi teknikleri farklı sektörler tarafından tercih edilmektedir (Orakcı ve ark., 2019) (Coşkun ve Baykal, 2011).

Makine öğrenmesi tekniklerinin öğrenme sürecinde uygulanmasının en temel adımları model oluşturma ve değerlendirmedir. Model oluşturma adımında, farklı öğrenme modellerinden veri setine en uygun olanları kullanılarak öğrenme işlemi gerçekleştirilmektedir. Makine öğrenmesinde kullanılan öğrenme modelleri denetimli ve denetimsiz olmak üzere iki gruba ayrılmaktadır. Denetimli öğrenmede girişleri ve çıkışları içeren bir veri kümesi algoritmaya girdi olarak verilmekte, algoritma ise bu girişlere ve çıkışlara nasıl ulaşılacağını belirleyen bir yöntem bulmaktadır. Denetimsiz öğrenmede ise var olan girdilere ait bir çıktı (etiket) bulunmadığından (sınıflar önceden belli olmadığından) algoritma mevcut verileri analiz ederek ilişkileri kendisi belirlemektedir. Bu çalışmada duygu analizini etkili bir şekilde gerçekleştirmek için tasarımın ⑤ numaralı adımında veri madenciliğinin sınıflandırma için kullanılan K-En Yakın Komşu, Naive Bayes, Destek Vektör Makinesi, Karar Ağacı ve Rastgele Orman denetimli öğrenme algoritmalarından faydalanılmıştır.

2.6. K-En Yakın Komşu

K-En Yakın Komşu (KNN) ile sınıflandırma işlemi nesnelerin birbirleri arasındaki yakınlık ilişkilerine göre yapılmaktadır. Geliştirme kolaylığı avantajına sahip KNN algoritmasının yüksek miktarda bellek alanına gereksinim duyması, veri seti ve boyutu arttıkça işlem yükünün ve maliyetin önemli ölçüde yükselmesi, performansın K komşu sayısı gibi parametre ve özelliklere bağlı olarak etkilenmesi gibi birtakım dezavantajlara sahip olduğu da bilinmektedir (Liu ve Zhang, 2012).

KNN ile sınıflandırma işleminde ilk olarak komşu sayısını ifade eden K değerine bakılarak eleman sayısı belirlenmektedir. Algoritma yeni bir veri ile karşılaştığında K'ya olan mesafeleri hesaplamakta, sıralamakta ve en küçük uzaklığa bağlı olarak yeni değeri en yakın komşuların bulunduğu kümeye eklemektedir. Mesafe hesaplama işleminde ise Öklid uzaklığı sıkça tercih edilmektedir (Aydın, 2018). Bu süreçte K değeri 1 olduğu zaman sadece en yakın komşunun bulunduğu sınıfa atama yapılmakta, K değeri örnek sayısına yaklaştığında ise veri setinde yer alan tüm veriler dikkate alınmaktadır. Algoritma bir eğitim verisi içermekte ve her yeni değer için bu süreç tekrar etmektedir. Bu nedenle eğitim kümesinin büyük olması ve K değerinin uygun seçilmesi KNN açısından çok önemlidir. Bu çalışmada da ilk olarak ilgili veri setine en uygun komşuluk değerini (K) bulmak için belirli aralıktaki K değerlerine karşılık gelen hata oranları tespit edilmiştir. Şekil 2'de yer alan 1 ile 100 arasındaki tüm K değerlerine ait test verisinin tahmin edilen değerleri için ortalama hata grafiği hazırlanmıştır. Grafikten de anlaşılacağı üzere ilgili veri seti için ortalama hatanın en az olduğu 10 değeri en uygun komşuluk değeri olarak tespit edilmiştir. Ayrıca mesafe hesaplama işleminde Öklid uzaklığı kullanılmıştır. Son olarak Python'a ait *sklearn.neighbors* kütüphanesinin *KNeighborsClassifier* metodundan faydalanılmış ve bu metoda tespit edilen en yakın komşuluk değeri parametre verilerek bir model oluşturulmuştur.



Şekil 2. K değeri hata oranı

2.7. Naive Bayes

Naive Bayes (NB), sınıflandırma yapmak amacıyla kullanılan olasılıksal bir makine öğrenmesi algoritmasıdır (Aydın, 2018). NB algoritması sınıflandırma işleminde değerleri birbirinden bağımsız olarak ele almaktadır. Sınıfların ve örnek verilerin hangi sınıflara ait olduğu bellidir. Algoritma bir eleman için her durumun olasılığını hesaplamakta ve olasılık değeri en yüksek olana göre sınıflandırma yapmaktadır. Test kümesindeki bir verinin eğitim kümesinde bir karşılığı mevcut değilse o veri için olasılık değeri olarak 0 atanmakta ve tahmin yapılamamaktadır. Bu durum literatürde Sıfır Frekans (Zero Frequency) olarak bilinmektedir (Wu, 2013). Bu çalışmada NB

sınıflandırma algoritması için Python'a ait *sklearn.naive_bayes* kütüphanesinin *GaussianNB* metodundan faydalanılarak bir model oluşturulmuştur.

2.8. Destek Vektör Makinesi

Destek Vektör Makinesi (DVM), sınıflandırma amacıyla kendisinden faydalanılan oldukça etkili ve basit öğrenme teorisine dayalı bir algoritmadır (Taşçı ve Şamlı, 2020). Temel olarak iki sınıfa ait verileri birbirinden en uygun şekilde ayırmak için kullanılmaktadır. Bunun için bir düzlemde bulunan iki grup arasında bir sınır çizilmesi gerekmektedir. Algoritma bu sınırın nasıl çizileceğini belirlemektedir. DVM sınıflandırma problemini kareli optimizasyon problemine dönüştürüp çözmektedir. Bu dönüşüm ile öğrenme aşamasında işlem sayısı azalmakta ve diğer algoritmalara göre daha hızlı çözüme ulaşılmaktadır (Osowski ve ark., 2004). Öte yandan algoritma çekirdek (*kernel*) fonksiyonları sayesinde doğrusal (*linear*) olarak ayrıştırılabilen sınıfların belirlenmesinde de sıkça tercih edilmektedir. Bu çalışmada DVM algoritması için Python'a ait *sklearn.svm* kütüphanesinin *SVC* metodundan doğrusal olarak faydalanılarak bir model oluşturulmuştur.

2.9. Karar Ağacı

Karar Ağacı (KA), karar ve yaprak düğümlerinden oluşan ve ağaç yapısında model oluşturan bir sınıflandırma algoritmasıdır. KA algoritması oldukça büyük veri kümelerini karar verme kurallarına göre küçük parçalara bölerek işlem yapmaktadır. Karar ağacı farklı türden verilerden oluşabilmektedir ve karar ağaçlarıyla sınıflandırma yapmak için CART (Classification and Regression Trees) algoritmasından sıklıkla faydalanılmaktadır. Bu algoritma ile kök düğümünden başlanarak her bir düğüm iki yaprağa ayrılmakta ve karar ağacı yapısında ikili dallanmalar oluşmaktadır. Öte yandan bir karar ağacında düğüme bağlı her yaprak bir sınıf değerini göstermekte ve hangi dalın en iyi seçim olacağına karar vermek için yapraklarda olumlu ve olumsuz sınıflarının sayısını kontrol etmek yeterli olmaktadır. Yaprakta olumlu ve olumsuz tek bir sınıf oluştuğunda ve daha fazla bölünmeye gerek duyulmadığında aşağı yönde hareket eden bölme işlemi sonlandırılmaktadır (Aksu ve Doğan, 2019). Bu çalışmada CART karar ağacı algoritması için Python'a ait *sklearn.tree* kütüphanesinin *DecisionTreeClassifier* metodundan faydalanılarak bir model oluşturulmuştur.

2.10. Rastgele Orman

Rastgele Orman (RO), veri setinin eğitim için ayrılmış parçasındaki verilerden birden fazla karar ağacı oluşturan bir sınıflandırma algoritmasıdır. Literatürde topluluk öğrenme yöntemi olarak da adlandırılan rastgele orman, çok sayıda karar ağacına ait sınıflandırma sonuçlarından yararlanarak verilen test girdisinin sınıfına çoğunluk oyu ile karar vermektedir (Kalaycı, 2018). Algoritma, ilk başta eğitim verisini kullanarak çok sayıda karar ağacı oluşturmaktadır. Daha sonra veri setinin test için ayrılmış parçasını sınıflandırmak için test verilerini her bir ağaca yerleştirmektedir. Son aşamada her bir ağaçtan elde ettiği sınıflandırmayı değerlendirerek en yüksek değere sahip olanı seçmektedir. Bu çalışmada eğitim verisinden faydalanarak ilgili veri seti için en uygun ağaç sayısı değeri 100 olarak tespit edilmiştir. Ayrıca RO için Python'a ait sklearn.ensemble kütüphanesinin RandomForestClassifier metodundan faydalanılarak bir model oluşturulmuştur.

2.11. Model Başarım Ölçütleri

Değerlendirme sürecinde (⑥) sınıflandırma algoritmalarının performansını hesaplamak amacıyla literatürde sık kullanılan bir takım model başarım ölçütünden faydalanılmıştır (Kaynar ve ark., 2016) (Varol ve İşeri, 2019). Bu ölçütlerin formülasyonunda kullanılan parametreler şu şekilde tanımlanmaktadır:

TP (True Positive, Doğru Pozitif): Olumlu olan ve aynı zamanda sınıflandırıcı tarafından olumlu kabul edilen yorumların sayısıdır.

TN (True Negative, Doğru Negatif): Olumsuz olan ve aynı zamanda sınıflandırıcı tarafından olumsuz kabul edilen yorumların sayısıdır.

FP (False Positive, Yanlış Pozitif): Olumsuz olan ancak sınıflandırıcı tarafından olumlu kabul edilen yorumların sayısıdır.

FN (False Negative, Yanlış Negatif): Olumlu olan ancak sınıflandırıcı tarafından olumsuz kabul edilen yorumların sayısıdır.

Bu parametreler kullanılarak hesaplanan model başarım ölçütlerinin tanımı ve formülasyonu ise aşağıda yer almaktadır:

Doğruluk (Accuracy): Doğru tahminlerin tüm tahminlere oranıdır. Doğruluk ölçütü (1) ile verilen formül kullanılarak hesaplanmaktadır.

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

Hassaslık (Precision): Doğru pozitif tahminlerin sayısının doğru ve yanlış pozitif tahminlerin sayılarının toplamına oranıdır. Hassaslık ölçütü (2) ile verilen formül kullanılarak hesaplanmaktadır.

$$Hassaslık = \frac{TP}{TP+FP} \quad \textbf{(2)}$$

Geri Çağırma (Recall): Doğru pozitif tahminlerin sayısının doğru pozitif ve yanlış negatif tahminlerin sayılarının toplamına oranıdır. Geri çağırma ölçütü (3) ile verilen formül kullanılarak hesaplanmaktadır.

$$Geri\ Çağırma = \frac{TP}{TP+FN} \quad \textbf{(3)}$$

F-skor (F-score): Hassaslığın ve geri çağırmanın harmonik ortalamasıdır. F-skor ölçütü (4) ile verilen formül kullanılarak hesaplanmaktadır.

$$F-skor = 2 \times \frac{precision \times recall}{precision+recall} \quad \textbf{(4)}$$

2.12. Test Ortamı

Bu çalışmada değerlendirmeye alınan sınıflandırma algoritmaları Tablo 3'de yer alan konfigürasyonlara sahip bir dizüstü bilgisayarda gerçeklenmiştir ve çalıştırılmıştır.

Tablo 3. Konfigürasyon

Sistem	Özellik
İşlemci (Processor)	Intel® Core™ i7-4870HQ Processor, 6 MB Cache, 2.5 GHz (Launch Date: Q3'14)
Bellek (Memory)	16 GB, 1600 MHz DDR3
İşletim Sistemi (Operating System)	Part 1: macOS High Sierra Part 2: Windows 10+Ubuntu 18.04 (WSL, Windows Subsystem for Linux)

3. Analiz ve Deney Sonuçları

Duygu analizi için veri madenciliği sınıflandırma algoritmalarının karşılaştırılması üzerine hazırlanan bu çalışmada materyal ve yöntem bölümünde detaylandırılan veri seti ve algoritmalar kullanılarak ürün, film ve restoran yorumlarına dayalı analizler yapılmıştır. Yine bu çalışma kapsamında duygu analizi üzerine yapılmış literatürdeki diğer çalışmaların aksine daha fazla model ile kıyaslama gerçekleştirilmiştir.

Bu çalışmada K-En Yakın Komşu, Naive Bayes, Destek Vektör Makinesi, Karar Ağacı ve Rastgele Orman sınıflandırma algoritmalarından yararlanılmıştır. Bu algoritmaların her biri *Anaconda* ortamında Python programlama dili ve *scikit-learn* kütüphanesi ile gerçeklenmiş ve içerisinde 3000 farklı verinin yer aldığı kapsamlı bir veri seti ile test edilmiştir. Test amacıyla veri setinin %25'lik kısmından faydalanılmış, %75'lik kısmı ise eğitim için kullanılmıştır.

Tablo 4, sadece Amazon isimli veri setinde yer alan 1000 adet ürün yorumu kullanılarak gerçeklenen sınıflandırma algoritmalarının dört farklı model başarım ölçütüyle performanslarının karşılaştırıldığı tablodur. Bu tablo incelendiğinde en iyi sonuçlara %79.6 doğru sınıflandırma oranıyla DVM ve %75.6 doğru sınıflandırma oranıyla KNN sınıflandırma algoritmaları ile ulaşıldığı ve NB algoritmasının diğer yaklaşımlara kıyasla daha düşük bir performans sergilediği görülmektedir.

Tablo 4. Amazon veri seti model başarım ölçütleri

Algoritma	Doğruluk	Hassaslık	Geri Çağırma	F-skor
DVM	0.796	0.794	0.795	0.794
KNN	0.756	0.760	0.762	0.756
RO	0.740	0.740	0.731	0.733
KA	0.720	0.717	0.715	0.716
NB	0.712	0.720	0.720	0.712

Tablo 5, sadece Imdb isimli veri setinde yer alan 1000 adet film yorumu kullanılarak gerçeklenen sınıflandırma algoritmalarının dört farklı model başarım ölçütüyle performanslarının karşılaştırıldığı tablodur. Bu tablo incelendiğinde en iyi sonuçlara %80 doğru sınıflandırma oranıyla DVM ve %75.6 doğru sınıflandırma oranıyla KNN sınıflandırma algoritmaları ile ulaşıldığı ve KA algoritmasının diğer yaklaşımlara kıyasla daha düşük bir performans sergilediği görülmektedir.

Tablo 5. Imdb veri seti model başarım ölçütleri

Algoritma	Doğruluk	Hassaslık	Geri Çağırma	F-skor
DVM	0.800	0.796	0.794	0.795
KNN	0.756	0.758	0.765	0.752
RO	0.744	0.748	0.740	0.742
NB	0.676	0.679	0.679	0.676
KA	0.664	0.660	0.658	0.660

Tablo 6, sadece Yelp isimli veri setinde yer alan 1000 adet restoran yorumu kullanılarak gerçeklenen sınıflandırma algoritmalarının dört farklı model başarım ölçütüyle performanslarının karşılaştırıldığı tablodur. Bu tablo incelendiğinde en iyi sonuçlara %79.2 doğru sınıflandırma oranıyla DVM ve %77.2 doğru sınıflandırma oranıyla KNN sınıflandırma algoritmaları ile ulaşıldığı ve NB algoritmasının diğer yaklaşımlara kıyasla daha düşük bir performans sergilediği görülmektedir.

Tablo 6. Yelp veri seti model başarım ölçütleri

Algoritma	Doğruluk	Hassaslık	Geri Çağırma	F-skor
DVM	0.792	0.801	0.792	0.790
KNN	0.772	0.776	0.772	0.771
RO	0.756	0.781	0.756	0.751
KA	0.724	0.725	0.724	0.724
NB	0.696	0.699	0.696	0.695

Tablo 7 ise Amazon, Imdb ve Yelp isimli veri setlerinin birleştirilmesiyle elde edilen ve 3000 yorumdan oluşan yeni veri seti kullanılarak gerçeklenen sınıflandırma algoritmalarının dört farklı model başarım ölçütüyle performanslarının karşılaştırıldığı tablodur. Bu tablo incelendiğinde en iyi sonuçlara %85 doğru sınıflandırma oranıyla DVM ve %81.3 doğru sınıflandırma oranıyla KNN sınıflandırma algoritmaları ile ulaşıldığı görülmektedir.

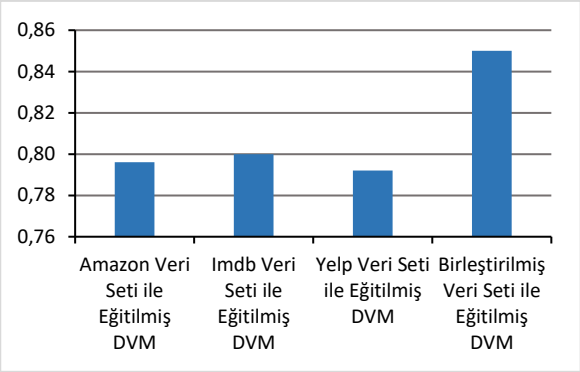
Tablo 7. Birleştirilmiş veri seti model başarım ölçütleri

Algoritma	Doğruluk	Hassaslık	Geri Çağırma	F-skor
DVM	0.850	0.851	0.850	0.850
KNN	0.813	0.818	0.814	0.813
RO	0.770	0.771	0.770	0.770
KA	0.723	0.738	0.725	0.720
NB	0.703	0.712	0.702	0.699

Bu durum DVM'nin büyük ve veri çeşitliliği fazla veri setleri için KNN'ye kıyasla daha etkili olduğuna işaret etmektedir. Öte yandan RO'nun barındırdığı birden fazla karar ağacına ait sınıflandırma sonuçlarını kullanarak en ideal sonuca ulaşması nedeniyle tek bir ağaca sahip klasik KA algoritmasına göre %93.8 oranında (RO ve KA'nın doğruluk ölçütüne göre kıyaslanması sonucu ulaşılan oran) daha başarılı bir performans sergilediği anlaşılmaktadır. Bu durumun veri büyüklüğünün ve

çeşitliliğinin bir sonucu olduğu değerlendirilmektedir. Öte yandan NB algoritmasının diğer yaklaşımlara kıyasla daha düşük bir performans sergilediği anlaşılmaktadır.

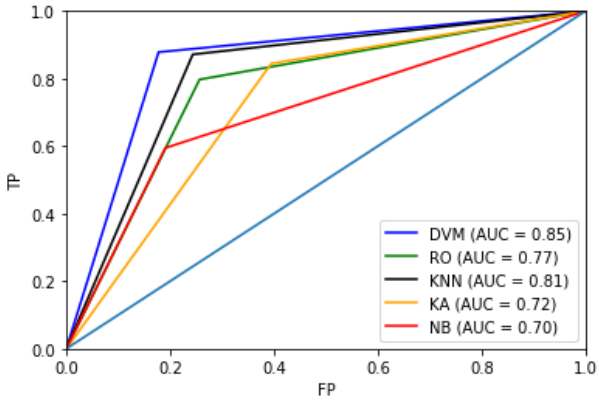
Tüm algoritmalar arasında en iyi performansı gösteren DVM algoritmasının, kullanılan her bir veri seti ile ayrı ayrı ve tüm veri setlerinin birleşiminden oluşan tek bir veri seti ile eğitilmesi sonucu ulaşılan doğruluk değerleri Şekil 3 ile verilen grafikte yer almaktadır. Bu grafik incelendiğinde DVM algoritmasının performansının dördüncü veri setinin heterojen yapısına rağmen arttığı anlaşılmaktadır. Aynı algoritmanın homojen verilerden oluşan diğer 3 veri seti ile ayrı ayrı eğitilmesi durumunda performansının azaldığı ise bu grafikten çıkarılabilecek bir başka sonuçtur.



Şekil 3. DVM algoritmasının doğruluk grafiği

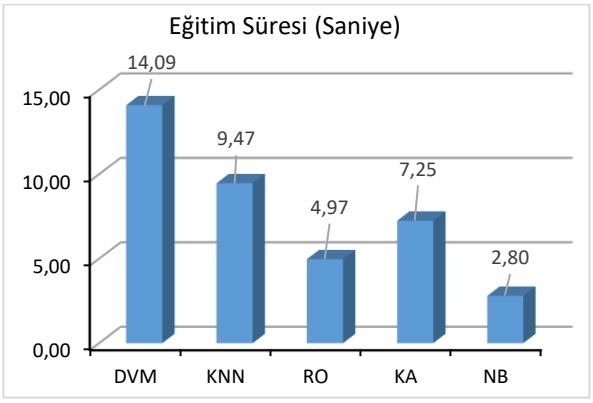
Çalışmada elde edilen sonuçların daha iyi değerlendirilebilmesi amacıyla tasarımın ⑦ numaralı adımında algoritmaların performansları günümüzde yaygın olarak kullanılan AUC-ROC eğrisi ile görselleştirilmiştir. ROC eğrisinin altındaki alanı ifade eden AUC'nin büyüklüğü modelin başarısının bir ölçüsü olarak kabul edilmektedir. Görselde yer alan eğrinin sol üst köşesine yaklaşıldıkça sonuçların doğruluğu artmaktadır.

Şekil 4'de yer alan ROC eğrisi incelendiğinde DVM'nin sol üst köşeye en yakın algoritma olduğu ve tüm algoritmaların doğruluğunun %50'nin üzerinde olduğu görülmektedir.



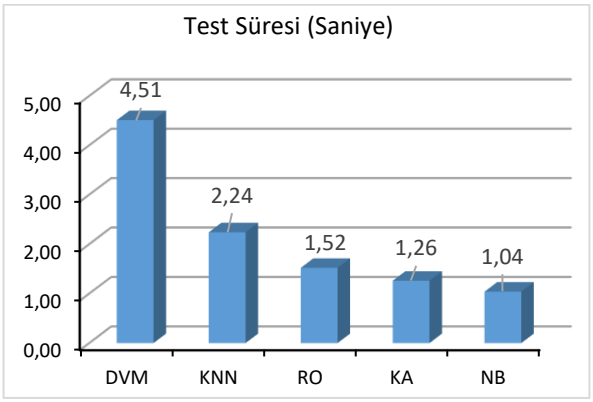
Şekil 4. ROC eğrisi

Şekil 5'de bu çalışma kapsamında performansı analiz edilen beş algoritmanın her birine ait eğitim süreleri yer almaktadır. Bu şekle göre duygu analizinde eğitim sürecini en hızlı tamamlayan algoritma NB'dir. Veri setinde yer alan işlemlerin ve özniteliklerin sayısının fazlalığı DVM algoritmasının eğitim sürecini diğer algoritmalara kıyasla oldukça olumsuz etkilemiştir. Öte yandan en AUROC skorunda en yüksek algoritmaya en yakın olan KNN, modeli eğitmek için DVM'nin yarısından daha fazla bir süreye ihtiyaç duymuştur.



Şekil 5. Algoritmaların eğitim süreleri

Şekil 6'da ise test süreleri görülmektedir. Algoritmaların test sürelerinin eğitim sürelerine hemen hemen paralel bir trend izlediği anlaşılmaktadır. Başka bir deyişle test süreci en uzun süren algoritmalar DVM ve KNN iken, bu süreci en hızlı tamamlayan algoritma yine NB olmuştur.



Şekil 6. Algoritmaların test süreleri

3.1. Literatürdeki Benzer Çalışmalarla Performans Karşılaştırması

Bu alt bölümde, doğruluk kriteri açısından bu çalışmanın en iyi algoritması olan DVM, literatürde aynı veri setlerini kullanan çalışmalardaki farklı algoritmalar ile karşılaştırılmıştır.

İlk çalışmada (Bari ve Saatcioglu (2018)), beş farklı algoritma tüm veri setlerine (Amazon, Yelp, Imdb) uygulanmıştır. Tablo 8, DVM modelimizin doğruluk kriteri açısından ilgili çalışmadaki tüm algoritmalara kıyasla daha başarılı olduğunu göstermektedir.

İkinci çalışmada (Rathee ve ark., (2018)), dokuz farklı algoritma tüm veri setlerine (Amazon, Yelp, Imdb) uygulanmıştır. Tablo 8, DVM modelimizin doğruluk kriteri açısından ilgili çalışmadaki tüm algoritmalara kıyasla daha başarılı olduğunu göstermektedir.

Üçüncü çalışmada (Wei (2021)), dört farklı algoritma Imdb ve Yelp veri setine geri kalan dört algoritma ise sadece Yelp veri setine uygulanmıştır (Amazon veri setine ait doğruluk değerleri bu çalışmada paylaşılmamıştır). Tablo 8, DVM modelimizin doğruluk kriteri açısından ilgili çalışmadaki tüm algoritmalara kıyasla daha başarılı olduğunu göstermektedir.

Tablo 8. DVM modelinin literatürdeki diğer ilgili çalışmalarla karşılaştırılması

Algoritma	Amazon Doğruluk	Imdb Doğruluk	Yelp Doğruluk
TB Polarity	0.747	0.740	0.761
TB Subjectivity	0.695	0.628	0.686
OF Polarity	0.640	0.667	0.657
OF Subjectivity	0.525	0.551	0.544
Stanford NLP	0.758	0.784	0.774
Bari ve Saatcioglu (2018)			
Logistic Regression	0.752	0.740	0.756
K-Nearest Neighbors	0.652	0.616	0.656
Support Vector Machine Classifier	0.476	0.484	0.448
Decision Tree Classifier	0.748	0.676	0.740
Random Forest Classifier	0.752	0.760	0.760
AdaBoost Classifier	0.744	0.708	0.728
Gaussian Naive Bayes	0.676	0.724	0.648
Bagging (Random Forest)	0.708	0.720	0.756
Bagging (Ada Boost)	0.764	0.680	0.728
Rathee ve ark., (2018)			
UPNN		0.435	0.608
HUAPA		0.550	0.686
NSC+LA		0.487	0.630
NSC+UPA	-	0.533	0.667
VistaNet			0.619
VS-CNN(Itemoriented)		-	0.620
VS-CNN(Useroriented)			0.649
Wei (2021)			
Bu Çalışmanın DVM Uygulaması	**0.796**	**0.800**	**0.792**

4. Sonuç

Bu çalışmada duygu analizi için beş farklı veri madenciliği sınıflandırma algoritmasının performansı karşılaştırılmıştır. Tüm algoritmaların kullanılan her bir veri seti ile ayrı ayrı çalıştırılmasına kıyasla tüm veri setlerinin birleşiminden oluşan tek bir veri seti ile eğitilmesinin performans artışına neden olduğu gözlemlenmiştir. Bu durum veri çeşitliliğinin ve büyüklüğünün önemini bir kez daha gözler önüne sermiştir. Öte yandan DVM'nin birleştirilmiş veri setinin büyüklüğüne ve çeşitliliğine rağmen %85'lik bir doğruluk oranıyla diğer algoritmalara kıyasla en etkili analiz aracı olduğu ortaya çıkmıştır. Bu amaçla dört farklı başarım ölçütünden faydalanılmıştır ve bu ölçütlerden biri ROC eğrisi ile görselleştirilmiştir.

Gelecekte yapılacak çalışmalarda veri seti çeşitliliğinin ve boyutunun daha da arttırılması, ayrıca derin öğrenmeye dayalı sınıflandırma algoritmaları kullanılarak duygu analizinin gerçekleştirilmesi planlanmaktadır.

5. Referanslar

Aksu, G., & Dogan, N. (2019). Comparison of Decision Trees Used in Data Mining= Veri madenciliginde kullanilan karar agaçlarinin karsilastirilmasi. *Pegem Journal of Education and Instruction*, *9*(4), 1183-1208.

Albayrak, M., Topal, K., Altıntaş, V. (2017). Sosyal Medya Üzerinde Veri Analizi: Twitter. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *22*(Kayfor 15 Özel Sayısı), 1991-1998.

Anonim (2020). Internet: UCI ML Repository Sentiment Analysis Dataset, http://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences, Son Erişim Tarihi: 28.03.2021.

Aydın, C. (2018). Makine öğrenmesi algoritmaları kullanılarak itfaiye istasyonu ihtiyacının sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (14), 169-175.

Balahur, A., Hermida, J. M., Montoyo, A. (2012). Detecting implicit expressions of emotion in text: A comparative analysis. *Decision Support Systems*, 53(4), 742-753.

Bari, A., & Saatcioglu, G. (2018, August). Emotion artificial intelligence derived from ensemble learning. In *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* (pp. 1763-1770). IEEE.

Burcu, A. K. I. N., Şimşek, U. T. G. (2018). Sosyal Medya Analitiği İle Değer Yaratma: Duygu Analizi İle Geleceğe Yönelim. *Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 5(3), 797-811.

Coşkun, C., & Baykal, A. (2011). Veri madenciliğinde sınıflandırma algoritmalarının bir örnek üzerinde karşılaştırılması. *Akademik Bilişim*, *2011*, 1-8.

Demir, C. G., Yılmaz, H. (2018). Sınıf dışı eğitim faaliyetlerinin öğrencilerin bilim ve teknolojiye yönelik tutumlarına etkisi ve duygu analizi. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, *7*(5), 101-116.

Güran, A., & Kınık, D. (2021). TF-IDF ve Doc2Vec Tabanlı Türkçe Metin Sınıflandırma Sisteminin Başarım Değerinin Ardışık Kelime Grubu Tespiti ile Arttırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (21), 323-332.

Kalaycı, T. E. (2018). Kimlik hırsızı web sitelerinin sınıflandırılması için makine öğrenmesi yöntemlerinin karşılaştırılması. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *24*(5), 870-878.

Kaynar, O., Yıldız, M., Görmez, Y., & Albayrak, A. (2016). td öğrenmesi yöntemleri ile Duygu Analizi. In *International Artificial Intelligence and Data Processing Symposium (IDAP'16)* (pp. 17-18).

Khan, F. H., Qamar, U., & Bashir, S. (2016). eSAP: A decision support framework for enhanced sentiment analysis and polarity classification. *Information Sciences*, *367*, 862-873.

Kotzias, D., Denil, M., De Freitas, N., & Smyth, P. (2015, August). From group to individual labels using deep features. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 597-606).

Kılınç, D., Borandağ, E., Yücalar, F., Tunali, V., Şimşek, M., & Özçift, A. (2016). KNN algoritması ve r dili ile metin madenciliği kullanılarak bilimsel makale tasnifi. *Marmara Fen Bilimleri Dergisi*, *28*(3), 89-94.

Liu, H., & Zhang, S. (2012). Noisy data elimination using mutual k-nearest neighbor for classification mining. *Journal of Systems and Software*, *85*(5), 1067-1074.

Nalçakan, Y., Bayramoğlu, Ş. S., & Tuna, S. (2015). *Sosyal Medya Verileri Üzerinde Yapay Öğrenme ile Duygu Analizi Çalışması*. Technical Report.

Onan, A., & Korukoğlu, S. (2016). Makine öğrenmesi yöntemlerinin görüş madenciliğinde kullanılması üzerine bir

literatür araştırması. *Pamukkale University Journal of Engineering Sciences*, *22*(2).

Orakcı, M., Ciylan, B., Kök, İ., & Sevri, M. (2019). Suç Analizinde Veri Madenciliği Teknikleri Ve Makine Öğrenmesi Algoritmalarının Kullanılması.

Osowski, S., Siwek, K., & Markiewicz, T. (2004, June). Mlp and svm networks-a comparative study. In *Proceedings of the 6th Nordic Signal Processing Symposium, 2004. NORSIG 2004.* (pp. 37-40). IEEE.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.

Parlar, T., Saraç, E., & Özel, S. A. (2017, May). Comparison of feature selection methods for sentiment analysis on Turkish Twitter data. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Rathee, N., Joshi, N., & Kaur, J. (2018, June). Sentiment analysis using machine learning techniques on python. In 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 779-785). IEEE.

Seker, S. E. (2016). Duygu Analizi (Sentimental Analysis). *YBS Ansiklopedi*, *3*(3), 21-36.

Şentürk, A. (2006). *Veri madenciliği: kavram ve teknikler*. Ekin Yayınevi.

Taşçı, M. E., & Şamlı, R. (2020). Veri Madenciliği İle Kalp Hastalığı Teşhisi. *Avrupa Bilim ve Teknoloji Dergisi*, 88-95.

Topaçan, Ü. (2016). Sosyal medya paylaşımlarında duygu analizi: makine öğrenimi yaklaşımı üzerine bir araştırma. Doktora Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 1-216.

Uçkan, T., Cengiz, H. A. R. K., Seyyarer, E., & Karcı, A. Ağırlıklandırılmış Çizgelerde Tf-Idf ve Eigen Ayrışımı Kullanarak Metin Sınıflandırma. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, *8*(4), 1349-1362.

Varol, A. B., & İşeri, İ. (2019). Lenf Kanserine İlişkin Patoloji Görüntülerinin Makine Öğrenimi Yöntemleri ile Sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, 404-410.

Zhang, L., Hua, K., Wang, H., Qian, G., & Zhang, L. (2014). Sentiment analysis on reviews of mobile users. *Procedia Computer Science*, *34*, 458-465.

Wei, Y. (2021, January). A Survey of Sentiment Analysis Based on Product Review. In 2021 2nd International Conference on Computing and Data Science (CDS) (pp. 57-63). IEEE.

Wu, J., Cai, Z., & Zhu, X. (2013, August). Self-adaptive probability estimation for naive bayes classification. In *The 2013 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.