



Sosyal Medya Etkileşimlerinde Depresyonu Tanımlamak için Derin Öğrenme Tekniklerinin Kullanılması

Ordak İbrahim NOORULDEEN¹ ve Serkan SAVAŞ

How to cite: Nooruldeen, O. I., & Savaş, S. (2024). Sosyal medya etkileşimlerinde depresyonu tanımlamak için derin öğrenme tekniklerinin kullanılması. *Sinop Üniversitesi Fen Bilimleri Dergisi*, 9(2), 449-466. <https://doi.org/10.33484/sinopfbid.1456956>

Araştırma Makalesi

Sorumlu Yazar

Ordak İbrahim Nooruldeen
ordakibrahim@gmail.com

Yazarlara ait ORCID

O.I.N: 0009-0004-9477-1187
S.S: 0000-0003-3440-6271

Received: 22.03.2024

Accepted: 30.08.2024

Öz

Depresyon, en yaygın zihinsel sorunlardan biridir ve intiharların önemli bir nedenidir. Sosyal medya platformlarının kullanımının artması, kullanıcıların günlük dilini kullanarak ifade ettikleri cümleler üzerinden depresyonun erken teşhisine olanak sağlamıştır. Sosyal medya platformlarının bireylerin günlük hayatlarında merkezi bir rol oynamaya devam etmesiyle, bu platformları ruh sağlığı analizi için kullanma konusunda artan bir ilgi bulunmaktadır. Bu çalışmada, Twitter (günümüzde X) üzerinden depresyon sınıflandırması yapılmıştır. Bu çalışmanın amacı Twitter'dan alınan tweetler arasında depresif özellik taşıyan tweetleri tespit etmektir. Çalışmada, Çift Yönlü Uzun Kısa Süreli Bellek (Bi-LSTM) mimarisi kullanarak depresyon tahmini için yenilikçi bir model sunulmuştur. Bu model, tweetlerdeki dil özelliklerini kullanarak depresyonun daha doğru tespiti için uygun temizleme ve ön işleme tekniklerinden faydalanmaktadır. Çalışma için, Twitter API yoluyla elde edilen özel bir veri seti oluşturulmuş ve analizler bu veri seti üzerinde gerçekleştirilmiştir. Önerilen Bi-LSTM modeli, %97.22'lik bir doğruluk oranı elde ederek dikkate değer bir etkinlik göstermiştir. Elde edilen sonuçlar, Twitter kullanıcılarının duygularındaki depresyonla ilgili örüntüleri ayırt etmek için derin öğrenme tekniklerinin kullanılabilirliği ve etkinliğini göstermiştir. Bu araştırma, ruh sağlığı izlemede ileri düzey tahmin analitikleri için bir temel oluşturmakta ve depresyon tespit modellerinin doğruluğunu ve verimliliğini artırmada Bi-LSTM'in potansiyelini vurgulamaktadır.

Anahtar Kelimeler: Sosyal Medya, Depresyon Sınıflandırması, Twitter, Derin Öğrenme, Bi-LSTM

Using Deep Learning Techniques to Identify Depression in Social Media Interactions

Abstract

Depression is one of the most common mental problems and an important cause of suicide. The increased use of social media platforms has enabled early diagnosis of depression through the sentences expressed by users using everyday language. As social media platforms continue to play a central role in individuals' daily lives, there is a growing interest in using these platforms for mental health analysis. In this study, Twitter (nowadays X) was used to categorize depression. The aim of this study is to identify tweets with depressive characteristics among tweets retrieved from Twitter. We present an innovative model for depression prediction using a Bi-directional Long Short-Term Memory (Bi-LSTM) architecture. This model utilizes appropriate cleaning and preprocessing techniques for more accurate detection of depression using language features in tweets.

¹Çankırı Karatekin Üniversitesi,
Fen Bilimleri Enstitüsü,
Elektronik ve Bilgisayar
Mühendisliği Anabilim Dalı,
Çankırı, Türkiye.

² Kırıkkale Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Kırıkkale, Türkiye.	For the study, a special dataset obtained through Twitter API was created and analyses were performed on this dataset. The proposed Bi-LSTM model has shown remarkable effectiveness, achieving an accuracy rate of 97.22%. The results obtained demonstrated the feasibility and effectiveness of deep learning techniques for recognizing depression-related patterns in the emotions of Twitter users. This research provides a foundation for advanced predictive analytics in mental health monitoring and highlights the potential of Bi-LSTM in improving the accuracy and efficiency of depression detection models.
This work is licensed under a Creative Commons Attribution 4.0 International License	Keywords: Social Media, Depression Classification, Twitter, Deep Learning, Bi-LSTM

Giriş

Depresyon, küresel bir psikiyatrik bozukluk olarak, dünyanın sağlık problemlerinin önemli bir kısmını oluşturmakta ve nüfusun %4.4'ünü etkilemektedir [1]. Bu durum kişisel, ailesel ve toplumsal refahı derinden etkileyerek zor çocukluklar, istismar, bağımlılık ve toplumsal baskılar gibi çeşitli faktörlerle ortaya çıkabilir [2]. Bilinen tedavilere rağmen, ruh sağlığı sorunlarının şiddeti, özellikle depresyon hem bireyler hem de toplum için ciddi zorluklar yaratmaktadır [3, 4]. Özellikle ergenler için engelliliğe, kendine zarar vermeye ve hatta intihara yol açma potansiyeli göz önüne alındığında depresyonun ciddiyetini anlamak acil bir konu haline gelmektedir [5]. Günümüz dünyasında sosyal medya, görüşlerin ifade edilmesi ve duyguların paylaşılması için hayati öneme sahiptir. Twitter ve Facebook gibi büyük platformlar kullanıcıların duyguları üzerine değerli veriler sunmaktadır [4]. Bireylerde depresyon tespiti zor kalmakta, araştırmacıları erken tespit için sosyal medyanın potansiyelini keşfetmeye teşvik etmektedir [6]. Küresel nüfusun %57'den fazlası sosyal medya kullandığı için bu platformlar benzersiz bir fırsat sunmaktadır. Online tarama araçları, geleneksel yöntemleri geride bırakarak ruh sağlığı değerlendirmeleri için önemli hale gelmiştir [7, 8]. Özellikle Twitter'ın günlük büyük tweet hacmi düşünüldüğünde ruh sağlığı araştırmalarında etkili modellerin eğitimi için bu platform daha da değerli hale gelmektedir [9, 10]. Platformlar tarafından sunulan Uygulama Programlama Arayüzü (Application Programming Interface - API), düzenli ifadeler kullanarak sosyal medya platformlarından veri almayı kolaylaştırır [11 - 13]. Doğal Dil İşleme (Natural Language Processing - NLP) ve Duygu Analizi (Sentiment Analysis - SA), metinlerdeki duygusal yönleri değerlendirerek, sosyal medya içeriği aracılığıyla depresyon seviyelerine dair iç görüler sağlar [14, 15]. Dil kalıplarından yararlanarak Yapay Zeka (YZ) ve NLP tabanlı teknikler, sosyal medya kullanıcıları arasında depresyonun erken tespitinde araştırmacılara yardımcı olur [16]. Böylece makine öğrenimi ve derin öğrenme modelleri, özellikle depresyon olmak üzere ruh sağlığı bozuklukları hakkında değerli iç görüler ortaya çıkarma imkânı sunmaktadır. Bu çalışmada, ruh sağlığı bozukluklarıyla ilişkili çeşitli anahtar kelimeler kullanarak Twitter'dan tweetler alınmıştır. Çalışmanın amacı potansiyel olarak ruh sağlığı sorunları yaşayan veya depresyon deneyimleyen kullanıcıları tespit etmektir. Bu bağlamda sosyal medyadan empirik olarak desteklenen psikolojik iç görüleri gözlemlenebilir. Bu çalışma sosyal medya üzerinden depresyon tespitine yönelik bir temel oluşturmaktadır. Gelecek çalışmalarla birlikte ruh sağlığı

bozuklukları alanında duyguların ve hislerin daha kesin bir şekilde tanımlanmasına yol açabilir. Bu çalışmanın yapısı şöyle oluşturulmuştur. İkinci bölümde konu ile ilgili alan çalışmaları özetlenmiştir. Üçüncü bölümde çalışmada kullanılan materyal açıklanmış ve uygulanan metodoloji anlatılmıştır. Dördüncü bölümde çalışmadan elde edilen bulgular sunulmuş ve tartışılmıştır. Beşinci bölümde ise çalışma sonuçlandırılmıştır.

İlgili Çalışmalar

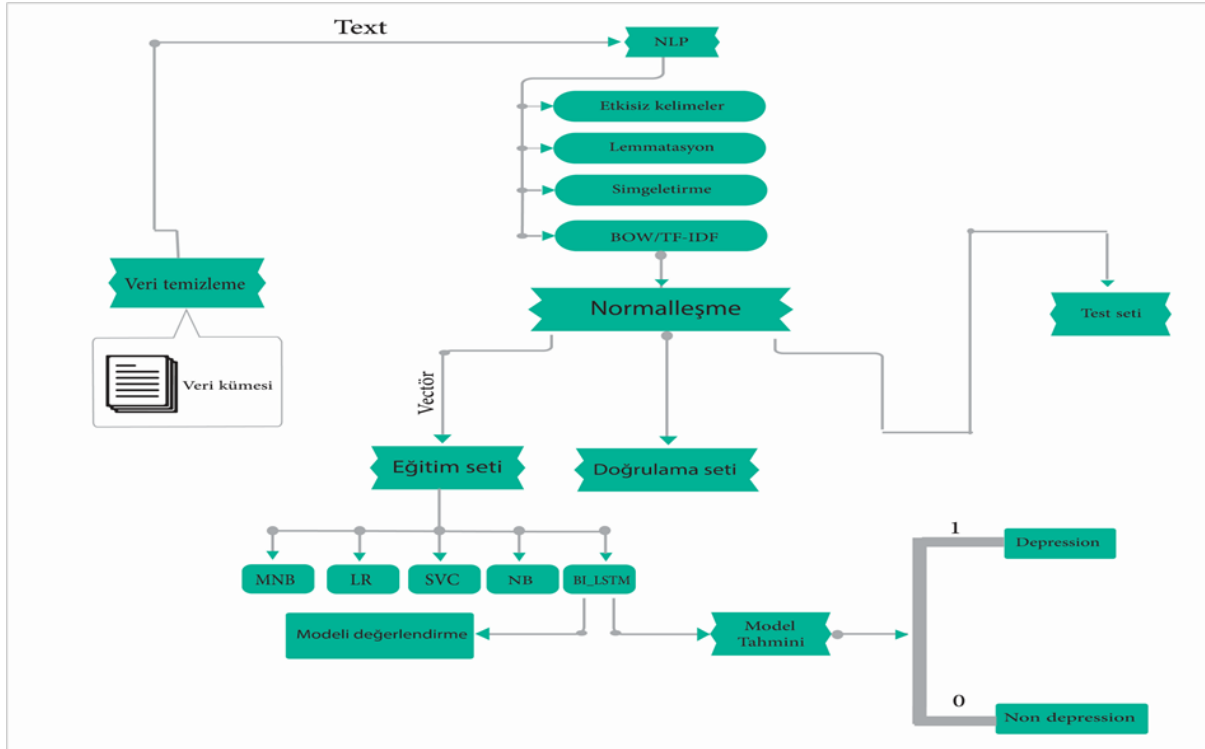
Sosyal medya metin verilerinin sınıflandırılması için çeşitli makine öğrenimi algoritmaları ve istatistiksel teknikler kullanılmaktadır. Geleneksel çalışmalar, depresyon ile sosyal medya kullanımı arasında bir ilişki olduğunu göstermektedir [17]. Bu amaçla literatürde farklı çalışmalar gerçekleştirilmiştir. Costello ve ark. [17], internet platformlarında kullanılan dil temelinde psikolojik özelliklerin çevrimiçi davranışlarda haritalanabileceğini belirtmişlerdir. Eichstaedt ve ark. [18], Facebook durum güncellemelerinde makine öğrenimi uygulayarak %69'luk bir depresyon tahmini elde etmişlerdir. Ahmad ve ark. [19] ise sosyal medyada %93 doğruluk oranıyla depresyonu tanımlamak için derin öğrenme kullanmışlardır. Bir başka çalışmada Priva ve ark. [20] tarafından beş makine öğrenimi algoritması kullanılmış ve Naif Bayes (Naive Bayes - NB) algoritmasının en yüksek doğruluğa sahip olduğu belirtilmiştir. Mori ve Haruno [21] tarafından Twitter verilerinde dört farklı makine öğrenimi algoritmasıyla 24 kişisel veriye ait 239 özellik analiz edilmiştir. Çalışmada dört tür sosyal ağ hizmeti bilgisine ait 23 alt ölçeğin toplu olarak tahmin edebileceği belirtilmiştir [21]. Tao ve ark. [22] tarafından gerçekleştirilen çalışmada depresif içerik tahminine odaklanılarak %90 F1 puanı elde edilmiştir. Bu çalışmalar, sosyal medya aracılığıyla depresyonu karakterize etmede çeşitli makine öğrenimi yaklaşımlarının başarılı sonuçlar üretebileceğini göstermektedir [22]. Guntuku ve arkadaşlarının [23] yaklaşık 400 milyon tweet toplayarak yalnızlıkla ilgili terimler üzerine gerçekleştirdiği çalışmada yalnızlık seviyeleri tahmin edilmiştir. Makine öğrenimi sınıflandırıcıları ve NLP tekniklerini kullanarak Rastgele Orman (Random Forest - RF) modeliyle 0.86'lık bir AUC (Area Under the Curve) elde etmişlerdir [23]. Smys ve Raj [24], bir Twitter veri seti kullanarak hastalık riskini erken teşhis yoluyla azaltmayı hedeflemişlerdir. Araştırmacılar çeşitli makine öğrenimi algoritmalarını kullanmış ve bireysel sınıflandırıcılardan daha üstün performans gösteren, duyguları tanıma ve tahmin doğruluğunda daha yüksek başarı elde eden hibrit bir yaklaşım önermişlerdir [24]. Hiraga [25] ise araştırmasında, Japon bloglarında ruhsal bozuklukları sınıflandırmak için geleneksel makine öğrenimi algoritmalarını kullanarak NB ile %95'lik en yüksek doğruluk oranını elde etmiştir. Wu ve ark. [26], Weibo kullanıcılarında tükenmişlik hissini tahmin etmek için makine öğrenimi algoritmalarını kullanmışlardır. Araştırmacılar yüksek geri çağırma, hassasiyet ve F1 puanı ile Aşırı Gradyan Artırma (eXtreme Gradient Boosting - XGBoost) modelinin tükenmişlik hislerini tespit etmedeki başarısını vurgulamışlardır [26]. Bir başka çalışmada sosyal medya metinlerinden doğum sonrası depresyonu tahmin etmek için makine öğrenimi teknikleri uygulanmıştır [14]. Çalışmada, Çok Katmanlı Algılayıcı (Multilayer Perceptron -

MLP) algoritmasının depresif içeriği tanımlamada ve doğum sonrası depresyon içeriğini yüksek doğrulukla tahmin etmede Destek Vektör Makinesi (Support Vector Machine - SVM) ve Lojistik Regresyon'dan (Logistic Regression - LR) daha üstün performans gösterdiği belirtilmiştir [14]. Orabi ve ark. [27], Twitter kullanıcılarında zihinsel hastalık belirtilerini tespit etmek için derin sinir ağları kullanmışlardır. Çalışmada dört sinir ağı modeli değerlendirilmiş ve Evrişimli Sinir Ağı (Convolutional Neural Network - CNN) ile maksimum %87,957'lik en yüksek doğruluğa ulaşılmıştır. CNN tabanlı modellerin Yinelemeli Sinir Ağları (Recurrent Neural Network – RNN) tabanlı modellere göre daha iyi performans gösterdiği görülmüştür [27]. Farklı bir platform olarak Reddit kullanıcılarında erken depresyon tespiti için Shah ve ark. [28], derin öğrenme tabanlı kelime yerleştirme tekniklerini kullanmış ve Word2VecEmbed ile Meta özelliklerin kombinasyonu ile depresyon tespitinde üstün performans elde etmişlerdir [28]. Shen ve ark. [29] tarafından gerçekleştirilen çalışmada Twitter'da depresif kullanıcıları tespit etmek için çok modlu sözlük öğrenme yöntemi önerilmiş ve minimum tanım uzunluğu ilkesiyle %85'lik en yüksek F1 puanına ulaşılmıştır. De Choudhury ve ark. [30], sosyal medyada depresyon tespiti için SVM algoritmasıyla, etiketlenmiş eğitim verilerinden ayırt edici özellikler kullanarak %70 sınıflandırma doğruluğu elde etmişlerdir. Sosyal ağlarda kullanıcı tarafından üretilen içerik analizi yoluyla ruh sağlığı seviyelerini ayırt etme potansiyelini inceleyen bir çalışmada Aldarwish ve Ahmad [31], NB yöntemi kullanarak %92'lik bir doğruluk oranı elde etmişlerdir. Guntuku ve ark. [32] tarafından gerçekleştirilen çalışmada sosyal medya aracılığıyla depresyonun tahmin edilmesi için çeşitli yöntemler karşılaştırılmış ve standart verilerin sosyal medya verileriyle entegre edilmesinin gerekliliği vurgulanmıştır. Çalışmada 0.70 ile 0.91 arasında değişen AUC değerleri, sosyal medyanın geleneksel yöntemlere kıyasla ruhsal hastalıkları tahmin etme potansiyelini göstermiştir [32]. Ahmad ve ark. [33] tarafından gerçekleştirilen çalışmada Twitter'dan elde edilen veri seti kullanılarak çevrimiçi kullanıcıların depresyon belirtileri gösterip göstermediğini belirlemek için makine öğrenmesi ve derin öğrenme modelleri kullanılmıştır. Çift Yönlü Uzun Kısa Süreli Bellek (Bidirectional Long Short-Term Memory - Bi-LSTM) modeli, %93.5 doğruluk oranıyla diğer modellerden daha iyi bir performans sergilemiştir [33]. Syms ve Raj [24], zihinsel sağlık sorunları ve intihar risklerinden korunmaya yönelik erken tahmin sistemi geliştirmek için bir hibrit yöntem önermişlerdir. Shah ve arkadaşlarının [28] gerçekleştirdiği çalışmada, kullanıcıların metin gönderilerini analiz ederek depresyonu tespit edebilen bir model önerilmiştir. Bu model, eğitim verileri kullanılarak derin öğrenme algoritmalarıyla eğitilmiş ve performansı, Reddit'in CLEF eRisk 2017'de yayınlanan 'Depresyonun Erken Tespiti' adlı pilot çalışmasının test verileriyle değerlendirilmiştir. Çalışmada Bi-LSTM modeli, çeşitli kelime gömme teknikleri ve meta veri özellikleri kullanarak başarılı sonuçlar elde etmiştir [28]. Figuerêdo ve ark. [34] tarafından gerçekleştirilen çalışmada depresyonun erken tespitine yönelik CNN tabanlı bir yöntem önerilmiştir. Önerilen yöntemin potansiyel olarak depresyon belirtileri gösteren kullanıcıları tespit edebildiği ve birçok temel metrik açısından diğer yöntemlere eşit veya daha iyi performans gösterdiği belirtilmiştir [36]. Marriwala ve ark. [35] tarafından depresyon tespiti için, hastanın yanıtlarının metinsel

ve işitsel özelliklerini birleştiren hibrit bir model önerilmiştir. Depresyonlu hastaların davranışsal özelliklerini incelemek için DAIC-WoZ veri tabanı kullanılmıştır. Önerilen yöntem üç bileşenden oluşmaktadır. İlk bileşen yalnızca metin özellikleriyle eğitilen bir metinsel CNN modelidir. İkinci bileşen yalnızca ses özellikleriyle eğitilen bir işitsel CNN modelidir. Üçüncü bileşen ise ses ve metinsel özelliklerin bir kombinasyonundan oluşan hibrit modeldir. Bu modellerde LSTM algoritmaları kullanılmıştır. Önerilen çalışmada, Bi-LSTM modeli olarak adlandırılan geliştirilmiş bir LSTM versiyonu da kullanılmıştır. Bi-LSTM'nin %88 doğruluk oranı ile diğer modellere göre daha iyi öğrenme performansı sergilediği gözlenmiştir [36].

Materyal ve Metodoloji

Bu çalışmada, tweetlerin depresyonu işaret edip etmediğini belirlemek için Twitter kullanıcı yorumlarından veri toplama ve analiz etme konusunda kapsamlı bir yöntem kullanılmıştır. Twitter'dan veri toplama ve depresyon belirtisi gösteren bireylerin tweetlerini diğerlerinden ayırarak iki ayrı dosyada yerleştirme işlemi gerçekleştirilmiştir. Sonrasında elde edilen bu veriler analiz edilerek çalışmada oluşturulan modelin performansı değerlendirilmiştir. Çalışmanın blok diyagramı Şekil 1'de gösterilmiştir.

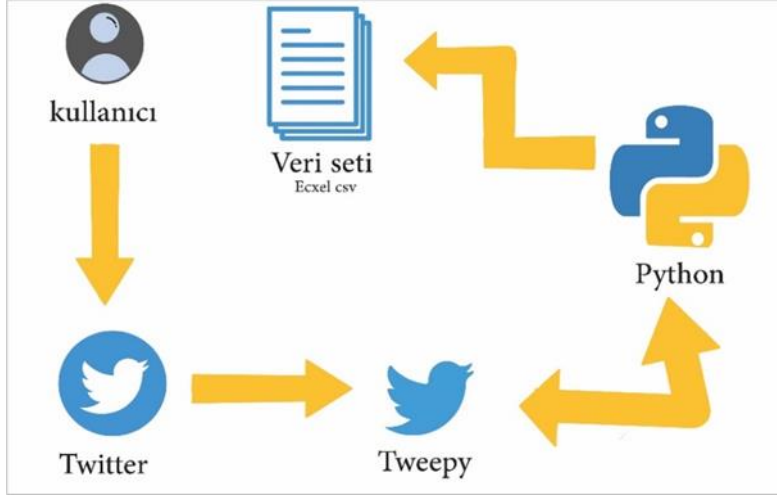


Şekil 1. Veriler üzerinde ön işleme ve özellik çıkarımı adımları

Veri Seti

Bu çalışmada kullanıcılardan veri toplamak için Twitter'ın API kullanılmıştır. API, geliştiricilerin manuel işlemlere gerek kalmadan büyük miktarda kamuya açık veriyle programatik olarak etkileşimde bulunmalarını sağlar. API erişimi, bir geliştirici hesabı oluşturarak araştırma çalışmasını detaylandırma

ve geliştirici anlaşması ile sağlanmıştır. Sonuç olarak, API 'anahtar' ve 'erişim belirteci' kullanımı için yetki alınmıştır. Veri alımını kolaylaştırmak için bir Python kütüphanesi olan Tweepy kullanılmıştır. Şekil 2, veri setini edinmek için kullanılan metodolojiyi göstermektedir.



Şekil 2. Veri setini edinmek için kullanılan metodoloji

Çalışmada gerçekleştirilen analiz hem tweetlerdeki sözcük seçimlerine hem de ifade edilen depresif belirtilerin türlerine dayanmaktadır. Twitter API yardımıyla toplamda 100,033 veri çekilmiştir. Bu veriler içerisinde %60 oranında eğitim verisi, %20 oranında doğrulama verisi ve %20 oranında ise test verisi rastgele seçilerek alt kümeler oluşturulmuştur. Bu tweetlerde ifade edilen duygular iki sınıfa ayrılmıştır. İlk sınıf 'depresyon', 'kötü', 'kendimi kötü hissediyorum', 'gergin', 'intihar', 'şizofren', 'yorgun', 'stres', 'anksiyete', 'çaresiz' ve 'kriz' gibi terimleri içermektedir. İkinci sınıf ise 'mutlu', 'iyi', 'memnun', 'neşeli', 'daha iyi', 'umut', 'güzel' ve 'heyecanlı' gibi pozitif duyguları kapsamaktadır. Veriler daha sonra analiz için Excel CSV formatına dönüştürülmüştür. Analiz için tweet veri seti gruplara ayrılmıştır. Her grup; 'No', 'Tarih', 'tweet_id', 'Metin', 'KullanıcıAdı', 'Ad', 'Konum', 'takipçi_sayısı', 'Dil' ve 'Etiket' sütunlarını içermektedir. Son adım, bu tweetleri filtreleyerek depresif belirtilerle en alakalı olanları belirlemektir.

Ön İşleme

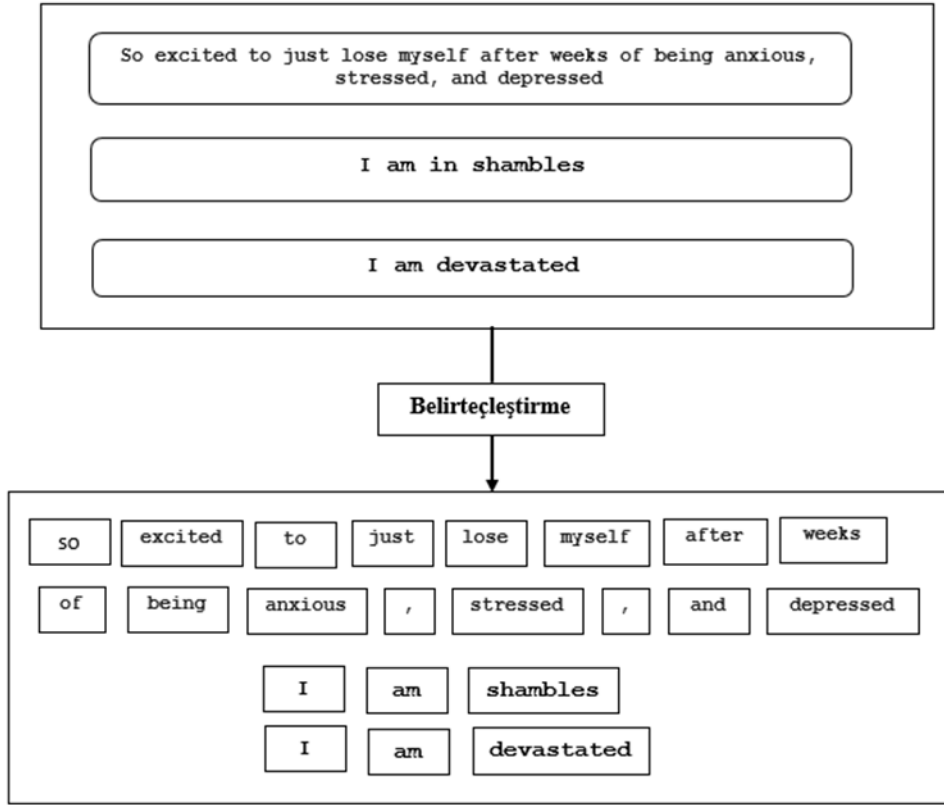
Veri madenciliği sürecinde veri ön işleme, ham metin verilerini temizleme ve ardından yapılacak işlemler veya analizler için verileri hazırlama için kritik bir adımdır. Bu süreç, tokenizasyon, etkisiz kelimeleri (stop words) kaldırma, kök bulma (stemming) ve kökten türetme (lemmatization) gibi birkaç önemli adımı içermektedir. Bu adımlar, metin verilerinin boyutunu azaltmaya ve NLP görevlerinde doğruluğu artırmaya yardımcı olmaktadır. Metin verileri doğası gereği yapılandırılmamış olup sıklıkla yazım hataları, dilbilgisi hataları ve standart olmayan biçimlendirme gibi gürültü içerir. Bu gürültüler de verilerin işlenmesini zorlaştırır. Bu çalışmadaki ön işleme aşamaları verideki gürültüyü hafifletmeyi ve böylece veri analizini kolaylaştırmayı hedeflemektedir. Kullanılan veri ön işleme yöntemleri arasında tweetleri temizleme, dengesizlikleri giderme, retweetleri eleme ve analiz için kullanıcı tarafından

üretileen içerik ve kullanıcı profil verilerinin kullanılabilirliğini sağlama yer almaktadır. Özellikle kullanıcı tanımlayıcıları (örn. @kullanıcıadı), etiketler, bağlantılar, harf olmayan karakterler, semboller, sayılar, boş dizgiler, NaN değerleri içeren satırlar ve çift girişler gibi fazlalık unsurların kaldırılması hedeflenmiştir. Tweet içindeki bağlantılar, tahmin amaçları için alakasız oldukları ve hesaplama karmaşıklığını artırdıkları için hariç tutulmuştur. Ayrıca, analizi çarpıtmamak için tweetlerden tarihler, saatler, sayılar ve etiketler kaldırılmıştır. Emojiler, fazla boşluklar ve etkisiz kelimeler de çıkarılmıştır. Doğal dil aracı (Natural Language Toolkit - NLTK) paketi, metinlerden etkisiz kelimeleri çıkarmada önemli bir rol oynamıştır. Porter Stemmer gibi teknikler kullanılarak kelime köklerine indirgemek ve ekleri çıkarmak amacıyla kök bulma işlemi gerçekleştirilmiştir. Şekil 3, ön işlemenin ardından temizlenmiş depresyon tweetlerini göstermektedir.

	edited_tweet	Original_Tweet
0	thread thinking bsp lust misery melancholy sp...	rt @itsamenanni: thread thinking bsp + lust mi...
1	taekookau college student jeongguk mob bos ta...	rt @mykingjk: ã°ã,ã¼ã¼ã¼/#taekookau college stud...
2	time avvabothundhi ante inactive hour active ...	rt @roaringrebels_: time 11 avvabothundhi ante...
3	almost new tesla getting towed stopped inters...	rt @mrinalsri: my almost new tesla getting tow...
4	disappointed know woman try commit suicide suc...	@tragicbirdapp she'll disappointed know women ...

Şekil 3. Ön işlemenin ardından temizlenmiş depresyon tweetlerini göstermektedir

Makine öğrenimi algoritmaları ilk başta metin verilerini anlamsal olarak yorumlama yeteneğinden doğal olarak yoksundur. Bu açığı kapatmak için kelimeleri sayısal değerlere eşlemek gereklidir. Bu işlem özellik çıkarma yoluyla gerçekleştirilir. Bu sayısal eşleme metinlerin anlamlarını etkili bir şekilde kapsüllemeyi amaçlar. Bu bağlamda metin verilerini sayısal özelliklere dönüştüren iki yaygın teknik araştırılmıştır. Bunlar, kelime torbası modelinden türeyen CountVectorizer ve Terim Frekansı-Ters Belge Frekansı (Term Frequency-Inverse Document Frequency - TF-IDF) yöntemleridir. CountVectorizer, metin analizinde kullanılan bir yöntemdir ve bir metin belgeleri koleksiyonunu kelime (token) sayılarının bir matrisine dönüştürür. Temelde metni bireysel kelimelere ayırır ve her bir belgedeki oluşum sayılarını sayar. Bu işlem metin verilerini makine öğrenimi algoritmalarının yorumlayabileceği sayısal bir formata dönüştürür. Ayrıca veri işlemede veri normalleştirme tekniği kullanılmıştır. Normalizasyon, bir veri kümesindeki değerleri, değerlerin aralığındaki farkları bozmadan veya bilgi kaybetmeden ortak bir ölçüğe ayarlamak için kullanılan bir tekniktir. Veri setindeki her bir satır, metnin kelimelere ayrılmasıyla belirteçleştirilmiştir (tokenizasyon). Bu işlem NLP yöntemleri için ön işlemenin temel adımudur. Şekil 4'te veri setindeki metinleri kelimelere ayırma işlemi gösterilmektedir.



Şekil 4. Veri setindeki metinleri kelimelere ayırma işlemi

Kelime bulutu, bu çalışmada tweetlerden elde edilen ve ön işlemeden geçirilen depresyon içeren ve depresyon içermeyen veri setlerini temsil etmektedir. Şekil 5'te, ön işleme sonrası oluşturulan kelime bulutu görülmektedir.



Şekil 5. Depresyon içeren ve içermeyen kelime bulutları

Kullanılan Modeller

Çalışmada analiz işlemlerini gerçekleştirmek makine öğrenimi modellerinden faydalanılmıştır. NB, Lojistik Regresyon (LR), Çok Terimli Naive Bayes (MultinomialNB), Destek Vektör Sınıflandırıcısı (Support Vector Classifier - SVC) ve Bi-LSTM gibi modeller performansları açısından karşılaştırılmıştır.

a) Bi-LSTM Modeli

Bi-LSTM özellikle NLP görevlerinde kullanılan temel LSTM modelinin gelişmiş bir varyantıdır. Veriyi tek bir yönde (ileri veya geri) işleyen standart LSTM algoritmasının aksine Bi-LSTM, veriyi eş zamanlı olarak her iki yönde de işler. Bu çift yönlü işleme modelin veri dizisinin hem geçmiş (geri) hem de gelecek (ileri) durumlarından bilgi toplamasına olanak tanır. Bi-LSTM algoritmasının temel avantajı geleneksel LSTM algoritmasına göre bağlamı daha iyi anlayabilmesidir. Algoritma, tek yönlü bir yaklaşımla kaçırılacak olan verilerdeki desenleri ve bağımlılıkları etkili bir şekilde öğrenebilir. Potansiyel bir dezavantajı ise hesaplama karmaşıklığıdır. İki yönden veri işlemek daha fazla hesaplama kaynağı ve zaman gerektirir, bu da modeli standart LSTM algoritmasına kıyasla daha kaynak yoğun hale getirir. Tasarımı özellikle sıralı girdi verilerinden öğrenmek ve hem ileri hem de geri yönde bağımlılıkları yakalamak için özelleştirilmiştir. Önerilen Bi-LSTM modelinin mimarisi, depresyon belirtisi içeren ve içermeyen tweetlerin sınıflandırılması amacıyla tasarlanmıştır. Bu mimarinin temel bileşeni, giriş metnini bir dizi belirteç indeksine dönüştüren metin vektörleştirme katmanıdır. Müteakip aşamada, bu belirteçler, her kelimeye eğitilebilir bir vektör atayan gömme katmanına iletilir. Yeterli eğitim süreci sonrasında bu vektörler, semantik açıdan benzer kelimelerin benzer vektörel temsillere sahip olmasını sağlayacak şekilde optimize edilir. Elde edilen veriler, ardından bu dizileri işleyen ve nihai olarak tek bir logit değerine indirgeyerek sınıflandırma çıktısı üreten Çift Yönlü LSTM katmanlarına aktarılır. Depresyon sınıflandırması bağlamında, her bir tweet "normal" veya "depresif" olmak üzere ikili bir sınıflandırmaya tabi tutulmuştur. Bu amaçla geliştirilen Bi-LSTM modeli, Çizelge 1'de detaylandırılan parametreler kullanılarak uygulanmıştır.

Tablo 1. Bi-LSTM modelinin parametre ayarları

Model Parametresi	Değer
Num hidden units	97
Input size	1
Num classes	2
Loss	Categorical_crossentropy
Optimizer	Adam
metrics	Accuracy
Epoches	17
Validation split	0.2
Verbose	1

b) Naive Bayes Modeli

Bayes teoremi, NB algoritması yoluyla doğal dil işlemede de kullanılan matematiksel bir çerçeve olup bir dilin "kurallarına" ve "türlerine" dayanan bir sınıflandırma teorisidir. Bayes teoreminin formülü Denklem (1)'de gösterilmiştir.

$$P(A | B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (1)$$

c) Lojistik Regresyon Modeli

LR, iki tür veriyi analiz etmek için kullanılan istatistiksel bir yöntemdir. Verileri analiz etmek için belirli bir etiketli veri kategorisi kullanır. Ayrıca karar alma süreçlerinde çeşitli kaynaklardan gelen verileri analiz etme ve şirketlerde en iyi eylem planını tahmin etmek gibi durumlarda da kullanılır. LR, birden fazla veri türünü analiz etmenin karmaşıklıklarını ele almak için önemli bir araçtır [36].

d) Destek Vektör Sınıflandırıcı Modeli

SVC, SVM algoritmasının bir varyantı olarak çeşitli sınıflandırma görevleri için kullanılan bir makine öğrenimi algoritmasıdır. Büyük ve karmaşık veri kümelerinden öğrenmek için son derece etkilidir. Bu algoritma veri setlerinden veri analizi yapmak, verileri daha anlamlı bir formata dönüştürmek ve en son bilgilere dayalı yeni veri üretmek gibi birçok uygulamada kullanılabilir [37].

e) Çok Terimli Naive Bayes Modeli

MultinomialNB, çeşitli veri türleri arasındaki benzerlikleri ve farklılıkları belirleme ve analiz etmeye odaklanan bir NB algoritmasıdır. Bir makine öğrenimi yaklaşımı olarak MultinomialNB çok yönlüdür ve metin, e-postalar, SMS ve daha fazlası gibi geniş bir veri yelpazesine uygulanabilir. Bayes teoremini kullanarak en yaygın veri türlerini ve aralarındaki ilişkileri belirler.

Değerlendirme Metrikleri

Çeşitli ölçümler kullanılarak kapsamlı bir değerlendirme yapılmadan bir makine öğrenimi modelini yeni verilere uygulamak sorunlara ve yanlış tahminlere yol açabilir [38]. Modellerin tahmin performansını iyileştirmek için doğruluk, kesinlik, hatırlama ve F1 puanı gibi metrikler önemlidir. Karışıklık matrisi yoluyla elde edilen doğru pozitif, yanlış pozitif, doğru negatif ve yanlış negatif sonuçları, modellerin performansı ile ilgili öngörüler sağlar. Değerlendirme için kullanılan metrikler, Denklemler (2), (3), (4) ve (5)'te verilmiştir.

$$\text{Doğruluk} = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

$$\text{Kesinlik} = \frac{TP}{TP+FP} \quad (3)$$

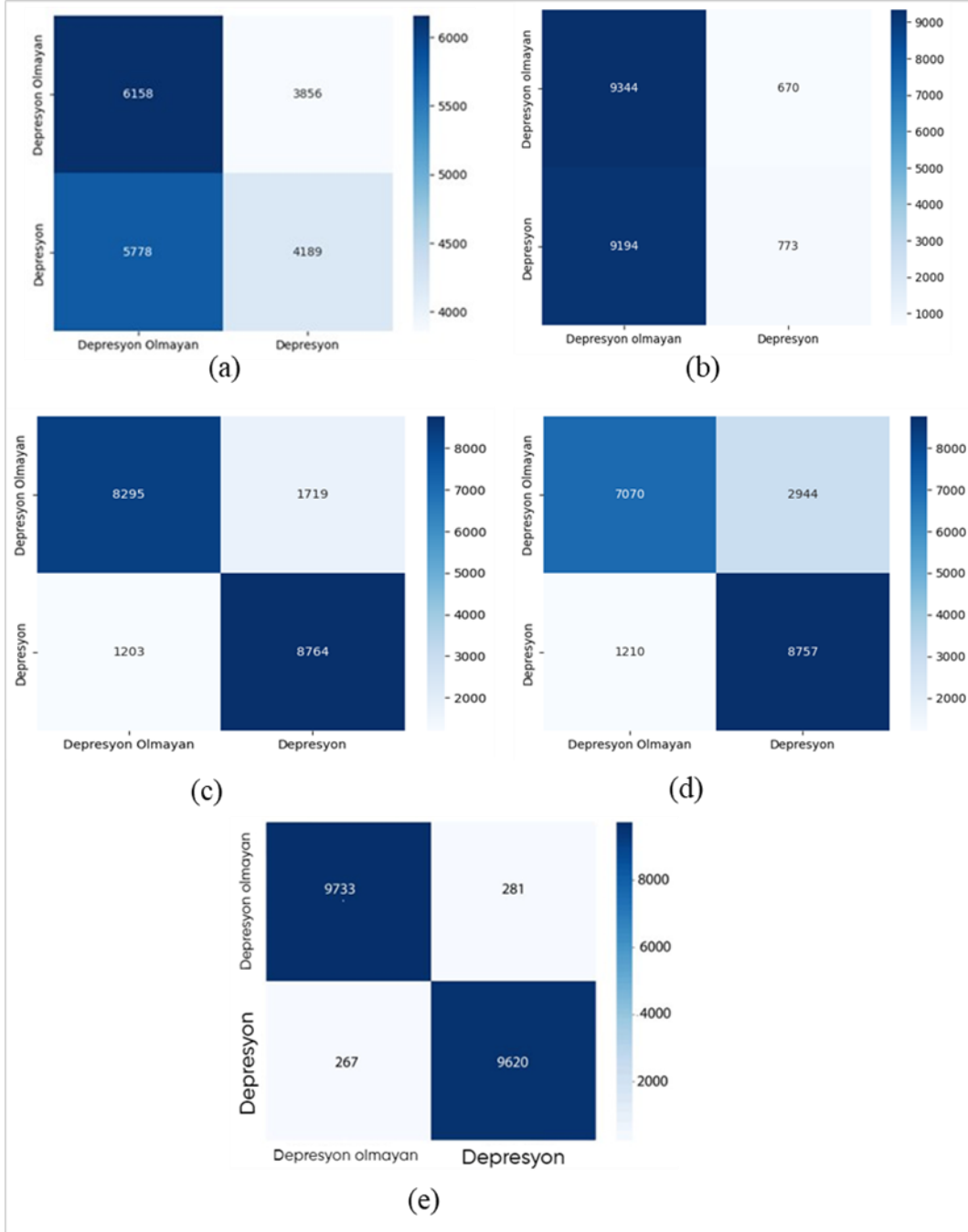
$$\text{Hatırlama} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 \text{ puanı} = 2 \times \frac{\text{Kesinlik} \times \text{Hatırlama}}{\text{Kesinlik} + \text{Hatırlama}} \quad (5)$$

Denklemlerde doğru pozitif için TP, yanlış pozitif için FP, doğru negatif için TN ve yanlış negatif içinse FN kısaltmaları kullanılmıştır.

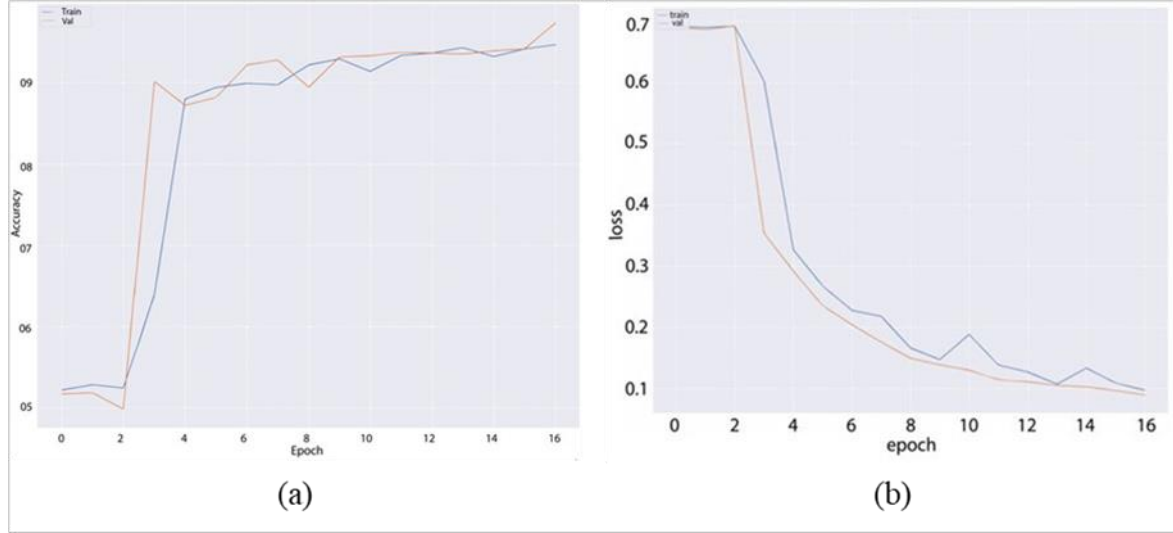
Bulgular ve Tartışma

Bu bölümde her modelin performansı ve sonuçları ilgili değerlendirmeler gerçekleştirilmiştir. Şekil 6'da modellerin karışıklık matrisi sonuçları gösterilmiştir.



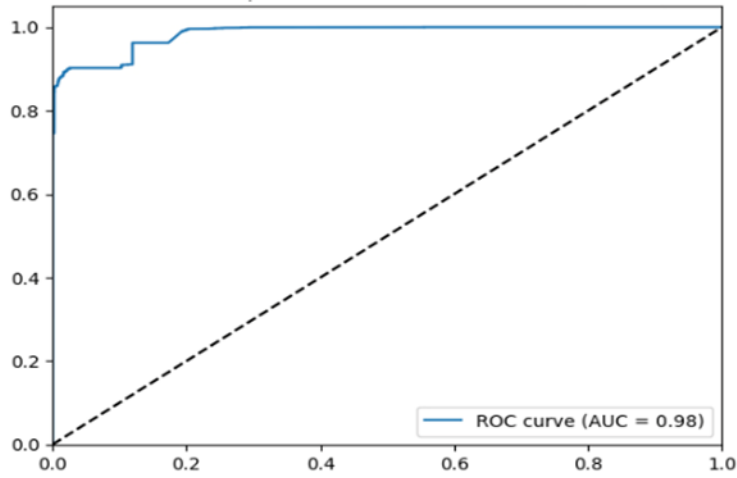
Şekil 6. Modellerin karışıklık matrisleri (a) MultinomialNB, (b) NB, (c) LR, (d) SVC (e) Bi-LSTM

Çalışmada kullanılan modeller içerisinde en başarılı model olarak karışıklık matrisinde ön plana çıkan (doğru pozitif sayısı 9620 ve doğru negatif sayısı 9733) Bi-LSTM modeline ait doğruluk ve kayıp grafikleri Şekil 7'de gösterilmiştir.



Şekil 7. Bi-LSTM modelinin doğruluk ve kayıp grafikleri

Bir sınıflandırıcının değeri ne kadar yüksekse, pozitif ve negatif kategoriler arasında ayırma yeteneği o kadar güçlüdür. Şekil 8’de Bi-LSTM modeline ait Alıcı İşletim Karakteristiği (Receiver Operating Characteristic - ROC) eğrisinin AUC değeri gösterilmiştir. Sonuç olarak önerilen Bi-LSTM modelinin veri kümesinin pozitif sınıfını etkili bir şekilde sınıflandırdığı sonucuna varılabilir.



Şekil 8. ROC grafiği

Metin sınıflandırmasında genellikle metin verilerinden elde edilen sayısal özellikler için minimum-maksimum normalizasyon tercih edilir. Bu teknik, özellikleri (örneğin kelime sayıları veya TF-IDF değerleri) belirli bir aralıkta (genellikle [0, 1]) standartlaştırarak, model eğitimi verimliliğini ve performansını artırabilir. Modelin normalizasyon olmadan eğitilmesi ise özelliklerin farklı ölçeklerde ve dağılımlarda olmasına bağlı zorluklarla karşılaşılmasına neden olabilir. Ayrıca modelin yeni veriye etkili şekilde genelleme yapma yeteneğini olumsuz etkileyebilir. Tablo 2, metin sınıflandırmasında min-max normalleştirme ve normalleştirme olmadan model eğitimi konusunda iki farklı yaklaşımı temsil etmektedir.

Tablo 2. Bi-LSTM modellerinin karşılaştırması

Model	Doğruluk	Kesinlik	Hatırlama	F1 Puanı
Normalizasyon ile eğitilen Bi-LSTM modeli	%97.22	%98	%97	%97
Normalizasyon olmadan eğitilen Bi-LSTM modeli	%91	%97	%84	%90

Çalışmadaki Bi-LSTM modeli, yüksek doğruluk oranlarıyla (pozitif 9620 ve negatif 9733) dikkat çekerek tweetlerde depresyonu tanımlamada oldukça etkili olmuştur. Öte yandan NB modeli, depresif olmayan birçok tweeti doğru bir şekilde tanımlamasına rağmen, depresif tweetlerde önemli sayıda yanlış sınıflandırma gerçekleştirmiştir. Bu, bağlamsal duyarlılık eksikliğine veya daha yaygın sınıflara karşı bir önyargıya işaret etmektedir. LR modeli, depresif tweetler için makul sayıda doğru pozitiflerle daha dengeli bir yaklaşım sergilese de yüksek yanlış sınıflandırma oranları göstermiştir. Bu, depresif ve depresif olmayan tweetler arasında orta düzeyde bir ayırım yeteneğine işaret etmektedir. SVC modeli, depresif tweetleri yüksek sayıda doğru pozitiflerle tanımlamada iyi bir performans gösterirken, her iki sınıf için de önemli yanlış sınıflandırma sorunları göstermiştir. Bu, NB modeline benzer şekilde tweetlerde kullanılan nüanslı dili kavramada zorluk çektiğine işaret etmektedir. Bu nedenle, metin içindeki bağlamı anlama yeteneği ve üstün doğruluğu ile Bi-LSTM modeli bu uygulama için en uygun seçim olarak öne çıkmaktadır. Buna karşılık, NB, LR ve SVC modelleri, belirli açılardan yararlı olmalarına rağmen, özellikle sosyal medya içeriğinden depresyon tespit etme gibi uygulamalarda dilin inceliklerinin kritik olduğu doğal dil işleme ve duygu analizi alanındaki karmaşıklıklarla başa çıkmada sınırlılıklar göstermektedir. Tablo 3'te beş modelin doğruluk, kesinlik, hatırlama ve F1 puanları karşılaştırılmıştır. Bu karşılaştırmaya göre Bi-LSTM modeli, Twitter veri setini 'depresif' ve 'depresif olmayan' kategorilere sınıflandırmada diğer yaklaşımlardan daha başarılı performans göstermektedir.

Tablo 3. Önerilen modeller arasındaki sonuçların karşılaştırılması

Model	Doğruluk	Kesinlik	Hatırlama	F1 Puanı
Bi-LSTM	%97.22	%98	%97	%97
NB	%51	%50	%92	%65
LR	%85	%88	%83	%85
SVC	%79	%86	%71	%77
MultinomialNB	%52	%52	%61	%56

Bi-LSTM modeli, tüm metriklerde üstün performans sergileyerek %97.22 doğruluk, %98 kesinlik, %97 hatırlama ve %97 F1 puanı ile öne çıkmaktadır. Bu, ilgili durumları doğru bir şekilde tanımlamada yüksek güvenilirliğini ve kesinlik ile hatırlama arasında dengeli bir tutum sergilediğini göstermektedir. Diğer yandan, NB modeli, %51 doğruluk ve %50 kesinlik ile performansı düşükken, %92 ile üstün bir hatırlama göstermektedir. Bu, %65'lik F1 puanı ile yüksek hatırlama ve düşük kesinlik arasındaki tutarsızlığı göstermektedir. LR modeli, %85 doğruluk, %88 kesinlik, %83 hatırlama ve %85 F1 puanı ile orta derecede bir performans göstermektedir. Model, kesinlik ve F1 puanı konusunda dengeli bir yetenek sergilemektedir. SVC, %79 doğruluk, %86 kesinlik, %71 hatırlama ve %77 F1 puanı ile düşük bir performans sergilemiştir. Benzer şekilde MultinomialNB modeli %52 doğruluk, %52 kesinlik, %61

hatırlama ve %56 F1 puanı ile nispeten daha düşük performans sergilemiştir. Bu sonuçlar, Bi-LSTM'nin yüksek doğruluk ve kesinlik ile hatırlama arasında dengeli bir yaklaşım gerektiren görevler için en uygun model olduğunu göstermektedir. Diğer modellerin, özellikle NB'nin, yüksek geri çağırma öncelikli senaryolarda, artan yanlış pozitiflerin potansiyeli göz önünde bulundurularak daha uygun olabileceğini göstermektedir. Pratik uygulamalardaki model seçimi, bu nedenle eldeki görevin özel gereksinimlerine bağlı olacaktır. Çalışma yöntemi, sosyal medyada zihinsel sağlık sorunları olan kullanıcıları belirlemede etkili bir araç olma potansiyeline sahiptir ve araştırmada yeni bir dönemin habercisi olabilir. Ayrıca, ruhsal bozuklukları tanımak için özelleştirilmiş NLP yöntemleri kullanılması, kullanıcıların gönderilerine dayanarak potansiyel ruh sağlığı sorunlarını tespit etmeyi kolaylaştırabilir. Bu yaklaşım, kolayca erişilebilir sosyal medya verilerinden yararlanmaktadır. Bu tür analizler için kullanıcıların hesaplarına erişmeden önce onaylarının alınması gerektiğini belirtmek önemlidir. Onay alındıktan sonra bu doğrulanmış modeller, çevrimiçi platformlardaki kullanıcıların gönderilerine dayanarak çeşitli zihinsel bozuklukların olasılığını ortaya çıkarabilir. Sonuç olarak bu çalışma, sosyal medyada zihinsel hastalıkları tespit etmenin gelecekte önemli bir araştırma alanı olabileceğini önermektedir.

Sonuç

Bu çalışmada derin sinir ağı modeli olan Bi-LSTM kullanılarak Twitter gönderilerinden depresyonlu tweetleri tespit etmeye yönelik bir model önerilmiştir. Önerilen Bi-LSTM modeli, metin verilerindeki sıralı bağımlılıkları etkili bir şekilde yakalayabilir ve kullanabilir. Bu özellik, kelimelerin veya ifadelerin sırasının metnin anlamını, duygusunu veya sınıflandırmasını önemli ölçüde etkilediği görevler için son derece önemlidir. Bi-LSTM, bilgileri hem ileri hem de geri yönde aynı anda işleyerek metin içindeki bağlamı ve nüansları daha iyi anlayabilir. Bi-LSTM'deki çift yönlülük, modelin tahmin yaparken hem önceki hem de sonraki kelimeleri dikkate alabilmesine olanak tanır. Bu yetenek, bir cümle içindeki duygu değişimlerini anlamak veya sosyal medya gönderilerindeki depresyonun inceliklerini belirlemek gibi cümle bağlamının kritik bir rol oynadığı görevler için değerlidir. Son olarak çalışmada tweetler normal veya depresif olarak sınıflandırılmıştır. Farklı makine öğrenimi modelleriyle yapılan deneylerde modellerin performansları Twitter API aracılığıyla elde edilen veri seti üzerinde değerlendirilmiştir. Önerilen model depresyon tespitinde iyi performans göstermiştir. Bi-LSTM modelinin F1 puanı (%97), hatırlama (%91), doğruluk (%97.22) ve kesinlik (%98) metriklerinde diğer yaklaşımlara göre daha iyi sonuçlar verdiği gözlemlenmiştir. Bu araştırmanın temel bulgusu, makine öğrenimi modellerinin Twitter kullanıcılarında depresif belirtileri tespit etme ve anlamada önemli bir etkiye sahip olduğudur. Bu içgörü kritiktir çünkü makine öğrenimi ve Bi-LSTM gibi tekniklerin, teorik araştırmaların yanı sıra pratik uygulamalarda da özellikle depresyonla başa çıkan bireylerin yaşam kalitesini artırmada potansiyeli olduğunu vurgulamaktadır. Çalışmanın sonuçları önemlidir ve ruh sağlığı izleme ve müdahalesinde ileri analitiklerin kullanımı için umut verici bir gelecek sunmaktadır. Bu çalışma aynı zamanda

profesyonellerin depresyonlu kişileri analiz etmek için ilk verilere ulaşmalarına yardımcı olabilir. Gelecekte, bir kullanıcının sosyal ağının ve etkileşimlerinin onun depresif durumunu nasıl etkileyebileceği dikkate alınarak bu çalışma genişletilebilir ve faktörlerin etkisini ölçmenin daha güvenilir bir yolunun bulunması sağlanabilir.

Teşekkür "Araştırma süreci boyunca değerli katkıları, önerileri ve desteği için Dr. Öğr. Üyesi Serkan SAVAŞ'a teşekkür ederim."

Fon/Finansman Bilgileri Bu çalışma için herhangi bir kurum ve/veya kuruluştan destek alınmamıştır.

Etik Kurul Onayı ve İzinler Çalışma, etik kurul izni veya herhangi bir özel izin gerektirmemektedir.

Çıkar çatışmaları/Çatışan Çıkarlar Yazarlar arasında herhangi bir çıkar çatışması yoktur

Yazarların Katkısı *Yazarların Katkısı* 1. yazar %60 oranında, 2. yazar %40 oranında katkı sağlamıştır

Kaynaklar

- [1] Uddin, M. Z., Dysthe, K. K., Følstad, A., & Brandtzaeg, P. B. (2022). Deep learning for prediction of depressive symptoms in a large textual dataset. *Neural Computing and Applications*, 34(1), 721–744. <https://doi.org/10.1007/s00521-021-06426-4>
- [2] Oquendo, M. A., Ellis, S. P., Greenwald, S., Malone, K. M., Weissman, M. M., & Mann, J. J. (2001). Ethnic and sex differences in suicide rates relative to major depression in the United States. *American Journal of Psychiatry*, 158(10), 1652–1658. <https://doi.org/10.1176/appi.ajp.158.10.1652>
- [3] Zafar, A., & Chitnis, S. (2020). Survey of depression detection using social networking sites via data mining. *IEEE Xplore*, 88-93. <https://doi.org/10.1109/Confluence47617.2020.9058189>
- [4] Martínez-Castaño, R., Pichel, J. C., & Losada, D. E. (2020). A big data platform for real time analysis of signs of depression in social media. *International Journal of Environmental Research and Public Health*, 17(13), 4752. <https://doi.org/10.3390/ijerph17134752>
- [5] Patel, V., Ramasundarahettige, C., Vijayakumar, L., Thakur, J., Gajalakshmi, V., Gururaj, G., Suraweera, W., & Jha, P. (2012). Suicide mortality in India: A nationally representative survey. *The Lancet*, 379(9834), 2343–2351. [https://doi.org/10.1016/S0140-6736\(12\)60606-0](https://doi.org/10.1016/S0140-6736(12)60606-0)
- [6] Lin, C., Hu, P., Su, H., Li, S., Mei, J., Zhou, J., & Leung, H. (2020, June 8-11). *Sensemood: depression detection on social media [Conference Presentation]*. International Conference on Multimedia Retrieval, Dublin Ireland. <https://doi.org/10.1145/3372278.3391932>
- [7] Conway, M. & O'Connor, D. (2016). Social media, big data, and mental health: Current advances and ethical implications. *Current Opinion in Psychology*, 9, 77–82. <https://doi.org/10.1016/j.copsyc.2016.01.004>
- [8] Ebert, D. D., Harrer, M., Apolinário-Hagen, J., & Baumeister, H. (2019). Digital Interventions for Mental Disorders: Key Features, Efficacy, and Potential for Artificial Intelligence Applications. In: Kim, Y. K. (ed), *Frontiers in Psychiatry*, (pp. 583–627). Springer.

- [9] Javed, A. R., Sarwar, M. U., Beg, M. O., Asim, M., Baker, T., & Tawfik, H. (2020). A collaborative healthcare framework for shared healthcare plan with ambient intelligence. *Human-Centric Computing and Information Sciences*, 10, 1-21. <https://doi.org/10.1186/s13673-020-00245-7>
- [10] Kale, S. S. (2015). *Tracking mental disorders across Twitter users* [Doctoral Dissertation, University of Georgia].
- [11] Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying Mental Health Signals in Twitter. In Resnik, P., Resnik, R., and Mitchell, M.(eds), *Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, (pp. 51-60). Association for Computational Linguistics.
- [12] Losada, D. E., & Crestani, F. (2016). A test collection for research on depression and language use. *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, 28-39. https://doi.org/10.1007/978-3-319-44564-9_3
- [13] Rissola, E. A., Aliannejadi, M., & Crestani, F. (2020, April 14–17). Beyond Modelling: Understanding Mental Disorders in Online Social Media. In J. M. Jose., Yilmaz, E., Magalhães, J., Castells, P., Ferro, N., Silva, M.J., Martins, F. (Eds.), *Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR*, (pp. 296-310). Springer International Publishing. https://doi.org/10.1007/978-3-030-45439-5_20
- [14] Fatima, I., Abbasi, B. U. D., Khan, S., Al-Saeed, M., Ahmad, H. F., & Mumtaz, R. (2019). Prediction of postpartum depression using machine learning techniques from social media text. *Expert Systems*, 36(4), e12409. <https://doi.org/10.1111/exsy.12409>
- [15] Suman, S. K., Shalu, H., Agrawal, L. A., Agrawal, A., & Kadiwala, J. (2020). A novel sentiment analysis engine for preliminary depression status estimation on social media. *arXiv:2011.14280*. <https://doi.org/10.1111/exsy.12409>
- [16] Zehra, W., Javed, A. R., Jalil, Z., Khan, H. U., & Gadekallu, T. R. (2021). Cross corpus multi-lingual speech emotion recognition using ensemble learning. *Complex & Intelligent Systems*, 7(4), 1-10. <https://doi.org/10.1525/collabra.18731>
- [17] Costello, C., Srivastava, S., Rejaie, R., & Zalewski, M. (2021). Predicting mental health from followed accounts on Twitter. *Collabra: Psychology*, 7(1), 18731. <https://doi.org/10.1525/collabra.18731>
- [18] Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preotiuc-Pietro, D., & Schwartz, H. (2018). Facebook language predicts depression in medical records. *National Academy of Sciences*, 115(44), 11203-11208. <https://doi.org/10.1073/pnas.1802331115>
- [19] Ahmad, H., Asghar, M. Z., Alotaibi, F. M., & Hameed, I. A. (2020). Applying deep learning technique for depression classification in social media text. *Journal of Medical Imaging and Health Informatics*, 10(10), 2446- 2451. <https://doi.org/10.1166/jmihi.2020.3169>
- [20] Priya, A., Garg, S., & Tigga, N. P. (2020). Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Computer Science*, 167, 1258-1267. <https://doi.org/10.1016/j.procs.2020.03.442>
- [21] Mori, K., & Haruno, M. (2021). Differential ability of network and natural language information on social media to predict interpersonal and mental health traits. *Journal of Personality*, 89(2), 228-243. <https://doi.org/10.1111/jopy.12578>

- [22] Tao, X., Zhou, X., Zhang, J., & Yong, J. (2016, December 12-15). *Sentiment analysis for depression detection on social networks* [Conference presentation]. *12th International Conference, ADMA Australia*. https://doi.org/10.1007/978-3-319-49586-6_59
- [23] Guntuku, S. C., Schneider, R., Pelullo, A., Young, J., Wong, V., Ungar, L., & Merchant, R. (2019). Studying expressions of loneliness in individuals using twitter: an observational study. *BMJ Open*, 9(11). <https://doi.org/10.1136/bmjopen-2019-030355>
- [24] Smys, D. S., & Raj, D. J. S. (2021). Analysis of deep learning techniques for early detection of depression on social media network-a comparative study. *Journal of Trends in Computer Science and Smart Technology*, 3(1), 24-39. <https://doi.org/10.36548/jtcsst.2021.1.003>
- [25] Hiraga, M. (2017). Predicting depression for japanese blog text. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics- Student Research Workshop*, 107–113 <https://doi.org/10.18653/v1/P17-3018>
- [26] Wu, J., Ma, J., Wang, Y., & Wang, J. (2021). Understanding and predicting the burst of burnout via social media. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 1-27. <https://doi.org/10.1145/3434174>
- [27] Orabi, A. H., Buddhitha, P., Orabi, M. H., & Inkpen, D. (2018). Deep learning for depression detection of twitter users. *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, 88-97. <https://doi.org/10.18653/v1/W18-0609>
- [28] Shah, F. M., Ahmed, F., Joy, S. K. S., Ahmed, S., Sadek, S., Shil, R., & Kabir, M. H. (2020). Early depression detection from social network using deep learning techniques, *IEEE Region 10 Symposium (TENSYMP)*, 823-826. <https://doi.org/10.1109/TENSYMP50017.2020.9231008>
- [29] Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., & Zhu, W. (2017). Depression detection via harvesting social media: A multimodal dictionary learning solution. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*, 3838-3844. <https://doi.org/10.24963/ijcai.2017/536>
- [30] De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1), 128-137. <https://doi.org/10.1609/icwsm.v7i1.14432>
- [31] Aldarwish, M. M., & Ahmad, H. F. (2017). Predicting depression levels using social media posts. *2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*. 277-280. <https://doi.org/10.1109/ISADS.2017.41>
- [32] Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18, 43-49. <https://doi.org/10.1016/j.cobeha.2017.07.005>
- [33] Ahmad, H., Asghar, M. Z., Alotaibi, F. M., & Hameed, I. A. (2020). Applying deep learning technique for depression classification in social media text. *Journal of Medical Imaging and Health Informatics*, 10(10), 2446-2451. <https://doi.org/10.1166/jmih.2020.3169>
- [34] Figuerêdo, José Solenir L., Ana Lúcia L. M. Maia, & Rodrigo Tripodi Calumby. (2022). Early depression detection in social media based on deep learning and underlying emotions. *Online Social Networks and Media*, 31, 100225. <https://doi.org/10.1016/j.osnem.2022.100225>

- [35] Marriwala, Nikhil, & Deepti Chaudhary. (2023). A hybrid model for depression detection using deep learning. *Measurement: Sensors*, 25, 100587. <https://doi.org/10.1016/j.measen.2022.100587>
- [36] Yapıcı, M. M. (2022). Lojistik Regresyon Modeli. In Savaş, S. & Buyrukoğlu, S., (Eds). *Teori ve Uygulamada Makine Öğrenmesi* (pp 37-67). Ankara: Nobel Akademik Yayıncılık Eğitim Danışmanlık TİC. LTD. ŞTİ.
- [37] Karakış, R. (2022). Destek Vektör Makinesi In Savaş, S. & Buyrukoğlu, S., (Eds). *Teori ve Uygulamada Makine Öğrenmesi* (pp 93-118). Ankara: Nobel Akademik Yayıncılık Eğitim Danışmanlık TİC. LTD. ŞTİ.
- [38] Schwartz, H. A., Eichstaedt, J., Kern, M., Park, G., Sap, M., Stillwell, D., & Ungar, L. (2014). Towards assessing changes in degree of depression through facebook. In *Resnik, P., Resnik, R., and Mitchell, M.(eds), Proceedings of the workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality.* (pp 118–125) Association for Computational Linguistics. <https://doi.org/10.3115/v1/W14-3214>