



Derin öğrenme yöntemleri ve kelime yerleştirme modelleri kullanılarak Parkinson hastalığının duygu analiziyle değerlendirilmesi

The evaluation of Parkinson's disease with sentiment analysis using deep learning methods and word embedding models

Feyza ÇEVİK¹ , Zeynep Hilal KİLİMCİ^{2*} 

¹IQ Bender Teknoloji Limited Şirketi, İstanbul, Türkiye.

feyza.cevik@iqbender.com

²Bilişim Sistemleri Mühendisliği, Teknoloji Fakültesi, Kocaeli Üniversitesi, Kocaeli, Türkiye.

zeynep.kilimci@kocaeli.edu.tr

Geliş Tarihi/Received: 18.11.2019

Düzeltilme Tarihi/Revision: 23.04.2020

doi: 10.5505/pajes.2020.74429

Kabul Tarihi/Accepted: 23.05.2020

Araştırma Makalesi/Research Article

Öz

Parkinson hastalığı, hastanın yaşam kalitesini etkileyen, önemli sosyal ve ekonomik etkileri olan ve semptomların aşamalı görünümü nedeniyle erken teşhis edilmesi güç olan yaygın bir nörolojik hastalıktır. Parkinson hastalığının Twitter gibi sosyal medya platformlarında tartışılması, hastaların Parkinson hastalığının hem tanı hem de tedavi aşamasında birbirleriyle iletişim kurduğu bir platform sağlar. Bu çalışmanın amacı, derin öğrenme ve kelime yerleştirme modellerini kullanarak insanların Parkinson hastalığı ile ilgili duygu analizlerini değerlendirmek ve karşılaştırmaktır. Bildiğimiz kadarıyla, bu çalışma Parkinson hastalığını sosyal medya aracılığıyla kelime yerleştirme modelleri ve derin öğrenme algoritmaları kullanarak analiz etmek için yapılan ilk çalışmadır. Bu çalışmada, kelime yerleştirme modelleri olarak Word2Vec, GloVe ve FastText; Evrimsel Sinir Ağları (CNN'ler), Tekrarlayan Sinir Ağları (RNN'ler) ve Uzun Kısa Süreli Bellek Ağları (LSTM'ler) derin öğrenme teknikleri olarak harmanlanmış ve sınıflandırma amacıyla kullanılmıştır. Kelime yerleştirme modelleri ve derin öğrenme algoritmaları kullanılarak Parkinson hastalığı hakkında kullanıcı yorumlarının duygularını analiz etmek amacıyla kapsamlı deneyler İngilizce Twitter veri kümesi üzerinde gerçekleştirilmiştir. Deney sonuçlarında, Word2Vec kelime yerleştirme modelinin CNN derin öğrenme algoritmasıyla harmanlanması sonucu %75.12 doğruluk ile kayda değer bir sınıflandırma başarısı gözlemlenmiştir. Bu çalışma, hastaların gereksinimlerini anlamak için kelime yerleştirme modelleri ve derin öğrenme algoritmalarını kullanma etkinliğini ve Parkinson hastalarının ve yakınlarının duygularını sosyal medya aracılığı ile analiz ederek tedavi sürecine değerli bir katkı sağladığını göstermektedir.

Anahtar kelimeler: Parkinson hastalığı, Duygu analizi, Derin öğrenme modelleri, Kelime yerleştirme yöntemleri.

Abstract

Parkinson's disease is a common neurodegenerative neurological disorder, which affects the patient's quality of life, has significant social and economic effects, and is difficult to diagnose early due to the gradual appearance of symptoms. Examining the discussion of Parkinson's disease in social media platforms such as Twitter provides a platform where patients communicate each other in both diagnosis and treatment stage of the Parkinson's disease. The purpose of this work is to evaluate and compare the sentiment analysis of people about Parkinson's disease by using deep learning and word embedding models. To the best of our knowledge, this is the very first study to analyze Parkinson's disease through social media by using word embedding models and deep learning algorithms. In this study, Word2Vec, GloVe, and FastText as word embedding models and Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short Term Memory Networks (LSTMs) as deep learning techniques are blended and used for classification purpose. Extensive experiments are conducted to analyze the sentiments of user comments about Parkinson's disease using word embedding models and deep learning algorithms on English Twitter dataset. The remarkable classification success with 75.12% of accuracy is observed in the experiments through the result of blending Word2Vec as a word embedding model and CNN as a deep learning technique. This study demonstrates the effectiveness of using word embedding models and deep learning algorithms to understand patients' needs, and provides a valuable contribution to the treatment process by analyzing the feelings of Parkinson's patients and their relatives through social media.

Keywords: Parkinson's disease, Sentiment analysis, Deep learning models, Word embedding methods.

1 Giriş

Son yıllarda internet ve mobil teknolojinin gelişmesiyle birlikte, sosyal medya platformları hızla büyüdü. İnternet, hastaları ve akrabalarını hastalık süreci hakkında bilgilendirmek, görüş ve deneyimlerini paylaşmak ve tedavi seçenekleri hakkında ayrıntılı bilgi edinmek için önemli bir kaynak haline gelmiştir [1]. Twitter, bilgilerin olduğu gibi paylaşıldığı ve gerçek zamanlı olarak başkalarıyla bağlantı kurulabilen en popüler sosyal ağ hizmetidir [2],[3]. Aylık, yaklaşık 350 milyon aktif Twitter kullanıcısı ile Twitter, gerçek dünya hakkında bilgi toplamak için güçlü bir araçtır [4]. Twitter, aynı zamanda hastaların ve

hasta yakınlarının deneyimlerini paylaştığı ve duygularını anlamalarına olanak sağladığı güçlü bir sağlık veri kaynağıdır [5]-[7].

Sentiment analiz (Duygu/duyarlılık analizi), konuşma veya yazma gibi özne bir dil ögesinin belirli bir konuda pozitif veya negatif olarak kutuplaşmasıdır. Sınıflandırma her zaman pozitif veya negatif olmayabilir, ancak kullanılacak soruna veya verilere bağlı olarak çok pozitif, pozitif, nötr, negatif, çok negatif gibi sınıflarla detaylandırılabilir. Otomatik duygu analizi, öznel bilgi istatistiklerine ulaşmak isteyen bir birimin metinler aracılığıyla analiz yapmasını amaçlar.

*Yazışılan yazar/Corresponding author

Parkinson hastalığı (PD), merkezi sinir sisteminin dünya çapında 10 milyondan fazla insanı etkileyen, yaşam kalitesini, motor refleksini, konuşmasını, davranışını, zihinsel sürecini ve diğer hayati fonksiyonlarını etkileyen nörodejeneratif bir hastalıktır [8],[9]. Parkinson hastalarında titreme belirtileri, yavaş hareket, kas sertliği ve denge kaybı gibi birçok semptom vardır. Bu semptomlara ek olarak, zihinsel değişiklikler, konuşma bozuklukları, uyku bozuklukları ve depresyon belirtileri de sıklıkla gözlenir ve bu semptomlar zamanla kötüleşebilir. Konuşma bozuklukları hastaların yaklaşık yüzde 90'ında görülür. Ek olarak, bu hastalar düşük ses, donuk konuşma, konuşmaya başlamada zorluk, gürültüyü sürdürmede zorluk, telaffuzda hata ve akıcılığın azalması gibi konuşma bozukluğu belirtileri içerebilir [9]. Parkinson, tipik olarak 40-70 yaşları arasında görülür ve yaş ilerledikçe görülme sıklığı artış gösterir. Hastaların %75 kadarında ilk semptomlar, 60 yaşından sonra ortaya çıkmaktadır. Epidemiyolojik araştırmalarda yaş ve cinsiyet haricinde semptomların yanında diğer risk durumları arasında beyaz ırk, aile öyküsünde Parkinson hastalığı olması, kafa travması, karbonmonoksit, ağır metaller ve tarım ilaçları gibi çeşitli toksik maddelere maruz kalınması, kuyu suyu içimi ve kırsal kesimde yaşama gibi durumlar söz konusudur.

Derin öğrenme (DL) algoritmaları ve kelime yerleştirme (WE) modellerinin uygulanması son yıllarda görüntü işleme (Image Processing), görüntü sınıflandırma (Image Classification), nesne bulma (Object Detection), nesne takip etme (Object Tracking), doğal dil işleme (Natural Language Processing), stil transferi (Style Transferring) ve makine çevirisi gibi farklı araştırma alanlarında çok popülerdir. DL yöntemleri ve WE modelleri, geleneksel makine öğrenmesi algoritmalarına kıyasla daha iyi tahminler ve sonuçlar sunması nedeniyle araştırmacılar tarafından tercih edilmektedir [10]-[18]. Derin öğrenme modelleri temel olarak, derin sinir ağları aracılığıyla verilerin anlamlı bir şekilde temsil edilmesini sağlamak için karmaşık özellikleri minimal dış destekle eğiterek otomatik özellik çıkarımı sağlamak için kullanılır. Ayrıca, birçok alanda sınıflandırma görevlerinde derin öğrenme yöntemleri de kullanılmaktadır. Evrişimli sinir ağları (CNN'ler) [17],[18], tekrarlayan sinir ağları (RNN'ler) [10],[12],[14] uzun kısa süreli bellek ağları (LSTM'ler) [12],[18], derin inanç ağları (DBN'ler) derin öğrenme modelleri olarak ve Word2Vec [13],[15], GloVe [16], FastText [11] kelime yerleştirme modelleri olarak tanınmış mimarilerdir.

Bu çalışmada, Parkinson hastaları için kişisel bir bakım planının geliştirilmesini kolaylaştırmayı ve bireysel ve örgütsel kullanıcıların duygu analizini kullanarak, artan bir Parkinson hasta popülasyonuna kişisel sağlık bilgileri için değerli bir tıbbi kaynak sağlamayı öneriyoruz. Bu amaçla, Parkinson hastalığı ile ilgili tweet'lerin anlamlarını ve bağlamlarını zenginleştirmek için Word2Vec, GloVe ve FastText adlı üç farklı kelime yerleştirme modeli kullanılmaktadır. Ayrıca, bireysel ve örgütsel kullanıcıların tweetlerini yorumlamak ve pozitif, negatif ve tarafsız gibi tweetleri sınıflandırarak Parkinson hastalığı algısını anlamak için derin öğrenme modelleri kullanılmaktadır. Bildiğimiz kadarıyla, bu, derin öğrenme ve kelime yerleştirme temelli modelleri kullanarak Parkinson hastalığına ilişkin duyguları analiz etmeye yönelik ilk girişimdir. Deney sonuçları, Word2Vec kelime yerleştirme modelinin derin öğrenme algoritmasıyla harmanlanması sonucu %75.12 doğruluk ile kayda değer bir sınıflandırma başarısı göstermektedir. Makalenin geri kalanı şu şekilde düzenlenmiştir: Bölüm 2, Parkinson hastalığı ile ilgili yapılan

çalışmaların bir özetini sunar. Sistem modeli ve sonuçlar Bölüm 3 ve 4'te verilmektedir.

2 Literatür çalışmaları

Bu bölüm, Parkinson Hastalığı (PD) hakkındaki çalışmaların literatür taramasının kısa bir özetini sunar.

Bir çalışmada yazarlar, Probabilistic Neural Network (PNN) sınıflandırıcısı, Parkinson hastalığının tanısı için kullanılır ve %80.92'lik test başarısı elde eder. Bu çalışmada kullanılan veri seti 23'ü Parkinson hastalığı olan 31 denekten alınan çeşitli biyomedikal ses ölçümlerinden oluşmaktadır [19]. Vouloimos ve diğ. Parkinson hastalarından (42 kişi) oluşan ve hasta başına yaklaşık 200 ses kaydı içeren Parkinson veri kümesinin sınıflandırılmasına odaklanmışlardır. Destek vektör makinesi, önerilen modelin sınıflandırıcısı olarak kullanılır. Öte yandan, filtre seçiminde ilgisiz ölçüm özellikleri seçmek için kullanılır. Deney sonuçları, önerilen modelin sınıflandırma performansının %91.40 başarı oranına ulaştığını göstermektedir [20]. Başka bir çalışmada, konvansiyonel önyükleme ya da bir-bir-dışa doğrulama validasyon yöntemi ve karşılıklı olarak bilgiye dayalı bir nitelik seçimi metodu (maksimum alaka düzeyi minimum yedeklilik - MRMR) tarafından toplanan özellikleri kullanır. PD veri seti, 195 denekten oluşan 32 kişinin (24 PD ve 8 sağlıklı) ses kayıtlarından meydana gelmiştir ve %92.9 doğruluk performansı rapor edilmiştir [21]. Başka bir çalışmada, Parkinson hastalarını ayırt etmek için sinir ağları, veri madenciliği sinir ağları modeli (DMneural), regresyon ve karar ağacı olmak üzere dört bağımsız sınıflandırma modeli uygulanmıştır. Sinirsel ağ tabanlı sınıflandırıcı %92.9 doğruluk sonucuyla diğerlerinden daha iyi performans gösterir [22]. Çağlar ve diğ. hem sınıflandırma görevi hem de özellik seçimi için insan biyomedikal sesini içeren Parkinson hastalığı veri kümesi örneklerinde dilsel olarak güçlü adaptif sinir bulanık sınıflandırıcı (ANFC + LH) kullanılmaktadırlar. ANFC+LH kullanımının 94.72 sınıflandırma performansı gösterdiği rapor edilmiştir [23].

Başka bir çalışmada, Parkinson hastalığının sınıflandırılması, bulanık C-ortalama kümelemesi temelinde özellik ağırlıklandırma yöntemi kullanılarak önerilmiştir. PD veri seti, Kaliforniya Üniversitesi Irvine makine öğrenim veri tabanından elde edilmiştir. Veri kümesinin içeriği, 31'i erkek, 23'ü kadın PD tanısı alan 195 uzun vadeli sesli fondu oluşmaktadır. Ağırlıklandırılmış Parkinson hasta veri kümesi için k-NN sınıflandırma algoritması kullanılarak farklı k değerleri için %97.93'lük bir başarı oranı elde edilmiştir [24]. Diğer bir çalışmada, dört tıbbi veri seti, dermatoloji, Pima-Hint diyabeti, meme kanseri ve Parkinson veri kümesi kullanılarak, benzerlik sınıflandırıcısı (SC) ve bulanık entropi (FE) özellik seçim yöntemleriyle %85.03 sınıflandırma performansı elde edilmiştir [25]. Diğer bir çalışmada yazarlar, yapay bağışıklık sistemi (AIS) ve klon niteliği seçim yaklaşımlarını kullanarak Parkinson hastalığının tanınmasına odaklanılmaktadırlar. Bu amaçla, veri kümesi, yukarıda belirtilen çalışmaların çoğunda olduğu gibi Oxford Üniversitesi'nde, Ulusal Ses ve Konuşma Merkezi, Denver, Colorado ile iş birliği içinde oluşturulmuştur. Yazarlar AIS kullanımının %92.70 sınıflandırma başarısı sağladığını bildirmektedirler [26]. Başka bir çalışma, Parkinson hastalığının tanısı için özellik seçim yöntemlerinin vokal ölçümlerden karşılaştırılmasına odaklanılmaktadır. Bu amaçla, SVM sınıflandırıcısı altı farklı özellik seçim yöntemiyle harmanlanmıştır. Parkinson hastalığı veri kümesi ve birleşik Parkinson hastalığı veri kümesi, iki

Parkinson hastalığı veri kümesi grubu, Parkinson hastalığının tanısında biyomedikal ses parametreleri kullanılarak test edilmiştir. Deney sonuçları, özylenelemeli özellikli elemeli (SVM-RFE) özellikli seçim yöntemine dayanan destek vektör makinesi yöntemlerinin, önerilen sistemin sınıflandırma başarısını diğerlerine göre %95.13 doğruluk değeriyle artırdığını göstermiştir [27].

Başka bir çalışmada [28] yazarlar, Parkinson hastalığında klinik ölçekler ve makine öğrenme modelleri kullanarak yeni ve geliştirilmiş aşama tahmini önermişlerdir. Lojistik regresyon (OLR), destek vektör makinesi (SVM), Adaptif Boost (AdaBoost) ve Random Under sampling Boosting (RUSBoost) tabanlı sınıflandırıcılar gibi makine öğrenme tekniklerini kullanarak PD'nin aşamasını ve ciddiyetini tahmin etmek için tahmin modelleri geliştirmişlerdir. Ek olarak, PD veri kümesindeki özelliklerin önemi rastgele ormanlar kullanılarak belirlenmiştir. AdaBoost tabanlı topluluk modeli %97.46 ile en yüksek sınıflandırma sonucunu sunmuştur. Başka bir çalışmada [29], yazarlar hasta anketi ve prediktif modelleme yoluyla Parkinson hastalığının erken saptanmasına odaklanmışlardır. Bu amaçla, hareket bozukluğu toplum-birleşik Parkinson hastalığı derecelendirme ölçeğinden (MDS-UPDRS) gelen hasta anketi veri kümesi olarak kullanılmıştır. Kişiyi sağlıklı, normal ve erken olarak sınıflandırabilecek modelleri tahmin etmek için, lojistik regresyon, destek vektör makinesi, ağaçları artırma, rastgele ormanlar gibi yaygın olarak kullanılan makine öğrenme tekniklerini kullanmışlardır. Bu tekniklerin, erken PD sınıflamasında ROC eğrisi altında (her ikisinde > %95) yüksek doğruluk ve performans gösterdiğini bildirmişlerdir. Bu çalışma ile, bu tahmin modellerinin, bir anketin maddelerini makine öğrenmesi yoluyla birleştirerek uzmanların teşhis sürecine yardımcı olma potansiyeline sahip olduğu sonucuna varılmıştır.

Oscar ve diğ. [30], çalışmamıza benzer şekilde Twitter ortamından elde edilen verilerle çalışmaktadırlar. Oscar ve diğ. çalışmalarında sadece atılan tweetlerin polaritesini ölçümlerken, bizim çalışmamız hem atılan tweetlerdeki duygu analizini yapmakta hem de atılan tweetler üzerinde doğruluk ölçüm metriğiyle sınıflandırma yapıp gerçekte ne kadar doğruluk ile duygunun saptandığını belirlemektedir. Ayrıca, Oscar ve diğ. çalışmamızdan farklı olarak Parkinson hastalığına değil de Alzheimer hastalığı üzerine atılan yorumları analiz etmişlerdir. Literatürdeki birçok çalışma İngilizce metinlerdeki duygu analizini saptarken İnan ve diğ. [31], geliştirdikleri model ile Türkçe metinler üzerindeki duygu polaritesini bulmayı hedeflemişlerdir. Metinlerdeki gizlenen duygunun saptanması için çok çeşitli duygu analizi araçları olsa da bu araçların Türkçe dili için duygu polaritesini doğru bir biçimde saptaması söz konusu değildir. Bunun için kullanıcının ekstra olarak platforma önceden eğitilmiş bir Türkçe veri kümesi sunması gereklidir. İnan ve diğ. geliştirdikleri modelde Türkçe veri kümesi kullanılacak çalışmalarda, bu zorunluluğu ortadan kaldırmışlar.

Bir başka çalışmada [32] yazarlar, derin öğrenme yaklaşımlarını kullanarak Türkçe metinlerden anlamsal çıkarımlarda bulunmuşlardır. Türkçe metinleri toplamak için çeşitli kategorilerde ürün satışı yapan bir internet sitesini kullanmışlardır. Bu siteden toplanan metinleri analiz ederek pozitif ve negatif anlam içeren kelimelerin vektör temsillerini öğrenerek Word2Vec modelinde eğitmişlerdir. Eğitilmiş olan kelime kümeleri kullanılarak Rastgele Orman (Random Forest-RF) modeli geliştirilmiştir. Klasik tekniklere ek olarak, derin

öğrenme yaklaşımlarından LSTM ve CNN modelleri duygu sınıflandırma için kullanılmıştır. Sonuç olarak CNN modeline göre LSTM modelinin %94.21 daha iyi sonuç ürettiği görülmüştür.

Krizhevsky ve diğ. [33], yapay sinir ağı modelinin cümle ve belge modellemesine dikkat çekmişlerdir. Çalışmalarında CNN ve RNN isimli iki temel modeli doğal dilleri anlamada kullanmışlardır. Bu iki mimarının güçlü yönlerini birleştirmişlerdir. Mimariler duygu sınıflandırması ve soru sınıflandırması görevleri üzerinde değerlendirilmiştir. Deneysel sonuçlar, iki modelin birlikte kullanıldığı yeni model olan C-LSTM'in daha iyi bir performans gösterdiğini ortaya koymuştur. Liu ve diğ. [34], 16 farklı metin veri kümesi kullanarak sınıflandırmayı LSTM modeli ile gerçekleştirmişlerdir.

Çalışmamız, yukarıda bahsedilen literatür çalışmalarından farklıdır; bunun sebebi Parkinson hastalığının algı ve duygu analizini, derin öğrenme algoritmaları ve kelime yerleştirme modelleri kullanarak sosyal medya platformu aracılığıyla saptamaya yönelik ilk girişim olmasıdır. Literatür araştırmalarından farklı olarak, Parkinson hastaları için kişisel bir bakım planının geliştirilmesini kolaylaştırmayı ve sosyal medyadan gelen metin verilerini kullanarak, artan bir Parkinson hasta popülasyonuna kişisel sağlık bilgileri için değerli bir tıbbi kaynak sağlamayı önermekteyiz.

3 Sistem modeli

Yöntemler, kullanılan araçlar ve önerilen model bu bölümde sunulmaktadır.

3.1 Veri toplama ve önerilen çerçeve

Bu çalışmada, Parkinson hastalığı ile ilgili gönderilen kullanıcı yorumlarının duygu analizlerini değerlendirerek Parkinson hastalığının saptanmasına odaklandık. Bu amaçla, Selenium tarayıcısı, Parkinson hastalığı ile ilgili olarak aşağıdaki anahtar kelimelerden oluşan kullanıcı yorumlarını toplamak için kullanılmaktadır: "Parkinson", "ParkinsonDisease", "ParkinsonsCure", "ParkinsonsTreatment", "ParkinsonDiagnosis", "ParkinsonSymptom". Twitter' da bireysel ve organizasyonel olmak üzere iki ana kullanıcı hesabı vardır. Parkinson hastalarının ve akrabalarının Parkinson hastalığı hakkındaki görüş ve deneyimlerini anlamak ve yorumlamak için Twitter' da tweet adı verilen bireysel ve kurumsal kullanıcı yorumları toplanmaktadır. Bireysel hesaplar hem Parkinson hastalarına hem de onların yakınlarına ait olabilirken, organizasyon hesapları nörologları, haber kaynaklarını ve hasta olmayanları içerebilir. Bu çalışmada, bazı bireysel hesapların korunan tweetleri nedeniyle halka açık bireysel hesaplar analiz edilmiştir. Toplamda 11043 tweet hem bireysel hem de kurumsal hesaplardan indirilmiştir. Selenium kullanarak, Twitter API' sinin izin verdiği sınır sorunu hakkında endişelenmeden istenildiği kadar tweet toplanabilmiştir. Bu çalışmada, denetimli makine öğrenme stratejisine odaklanılmıştır. Her kullanıcının Parkinson hastalığına karşı tutumunu belirlemek için kullanıcılardan gönderilen her tweeti pozitif, negatif veya nötr olarak etiketleme ihtiyacı vardır. Bu çalışmada, her yoruma ait duygunun belirlenmesi için TextBlob platformu [35] içinde yer alan Naive Bayes sınıflandırıcısı kullanılmıştır. Böylece, her tweet için pozitif, negatif veya nötr olarak sınıf olasılığını oluşturulmuştur. TextBlob' ın yanı sıra pozitif, negatif ve nötr duygu polaritesini belirlemek amacıyla ücretsiz olarak sunulan VADER [40], MonkeyLearn [41] gibi çeşitli duygu analizi araçları bulunmaktadır. Bu çalışmada

kullanılan etiketlenmemiş veriler, TextBlob, VADER, ve MonkeyLearn araçlarının uygulama programlama arayüzü (API) ile Python programlama dilinde etiketlenmiş ve sınıflandırma performansları kıyaslanmıştır. Kullanıcı yorumlarını anlamada TextBlob platformu ile ortalama %79.13 sınıflandırma başarısı sunulurken VADER ve MonkeyLearn ortamlarından sırasıyla %80.37 ve %78.25 doğruluk sonuçları elde edilmiştir. Her bir platformdan farklı doğruluk sonuçlarının elde edilmesinin sebebi farklı sınıflandırıcıların ve ön işleme yöntemlerinin kullanılıyor olmasındandır. Tablo 1'de pozitif, negatif veya nötr olarak etiketlenen yorumlardan birkaç örnek sunulmaktadır.

Tablo 1. Ön işleme modellerinin kelime yerleştirme yöntemlerine göre doğruluk sonuçları.

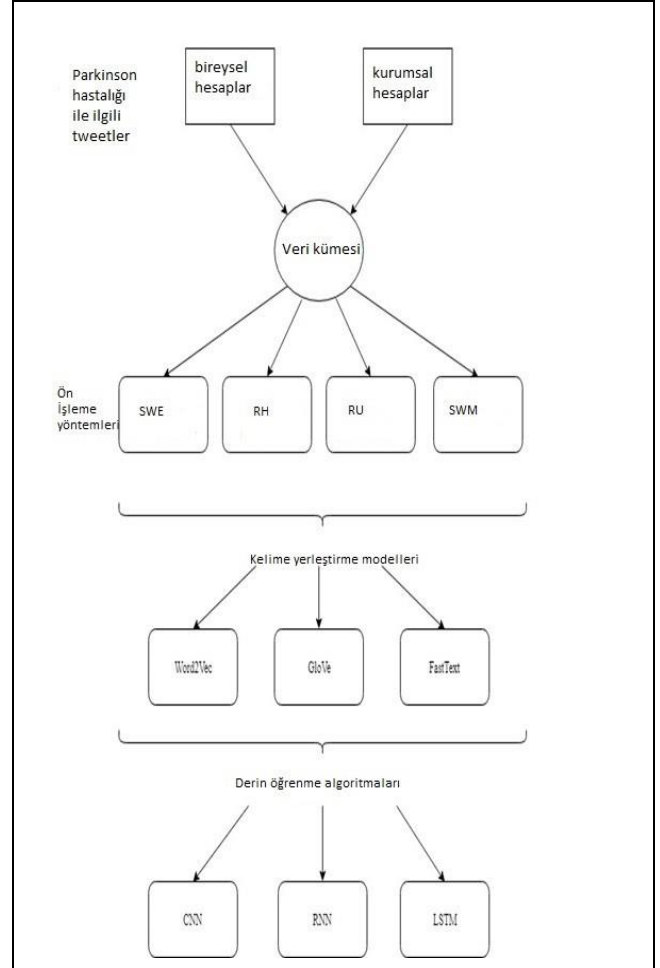
Table 1. Accuracy results of preprocessing models in terms of word embedding methods.

Tweet	Etiket
1) Please don't give unsolicited and inane advice and instead just say "that sucks" or say nothing..instead of "my grandpa had Parkinson's, but my grandma gave him fish oil everyday, so he had no symptoms....REALLY? No symptoms? Then he didn't have Parkinson's!	Negatif
2) On WorldParkinsonsDay discover how a clinical study led by EUfunded TreatER project is administering a promising Parkinson's drug to the first human patients.	Pozitif
3) A new test agent could lead to early diagnosis of prion diseases, Alzheimer's and Parkinson's for the first time.	Pozitif
4) 20 years ago no one would have believed that the gut was an issue in Parkinson, let alone something deserving a special issue!	Negatif
5) Chapter examines alpha-syn's activity as an antimicrobial peptide Right pointing backhand index and proposes a model for how #infectious agents, in conjunction w/ inflammatory, environmental, and genetic facilitators, may result in transfer of alpha-syn from the periphery to the brain in Parkinson's/MSA.	Nötr

Sosyal medya platformlarından her bir kullanıcıdan toplanan ham veri kümesi oldukça kirlidir. Bu nedenle, farklı ön işleme teknikleri uygulamasına ihtiyaç vardır. Bu çalışmada, bağlaçların kaldırılması (stop-word elimination), hashtaglerin kaldırılması (removing hashtags), URLlerin kaldırılması (removing URLs), ve kök bulma (stemming) teknikleri uygulanmaktadır [3].

Twitter'ın karakter kısıtlaması nedeniyle kullanıcıların duygularını yeterince ifade edemedikleri bir gerçektir. Bu sorunu gidermek için Word2Vec, GloVe ve FastText gibi yerleştirme modellerine odaklanılmıştır. Bu şekilde, her yorum kelime yerleştirme modelleri kullanılarak anlam, içerik ve söz dizimi açısından zenginleştirilmiştir. Bu yöntemlerle, Twitter'da fikirlerini ifade etmedeki sınır sorunu kullanıcı duygularını anlamak için bir sorun olmaktan çıkmıştır. Kelime yerleştirme modellerinden çıkan kelime vektörleri, derin öğrenme algoritmalarına girdi olarak verilmiştir. Bu aşamadan sonra, geleneksel makine öğrenme algoritmaları kullanmak yerine, evrişimli sinir ağları (CNN), tekrarlayan sinir ağları (RNN) ve uzun kısa süreli bellek ağları (LSTM) gibi üç farklı derin

öğrenme mimarisi sınıflandırma amacıyla kullanılmaktadır. Önerilen sistemin akış şeması Şekil 1'de verilmiştir. Şekil 1'de yer alan ön işleme yöntemleri, kelime yerleştirme modelleri ve derin öğrenme algoritmaları için şu kısaltmalar kullanılmıştır: SWE: Bağlaçları kaldırma, RH: Hashtagleri kaldırma, RU: URL'leri kaldırma, STM: Kök bulma, AOT: Bunların hepsi, CNN: Evrişimli sinir ağı, RNN: Tekrarlayan sinir ağı, LSTM: Uzun kısa süreli hafıza ağı.



Şekil 1. Önerilen sistemin akış şeması.

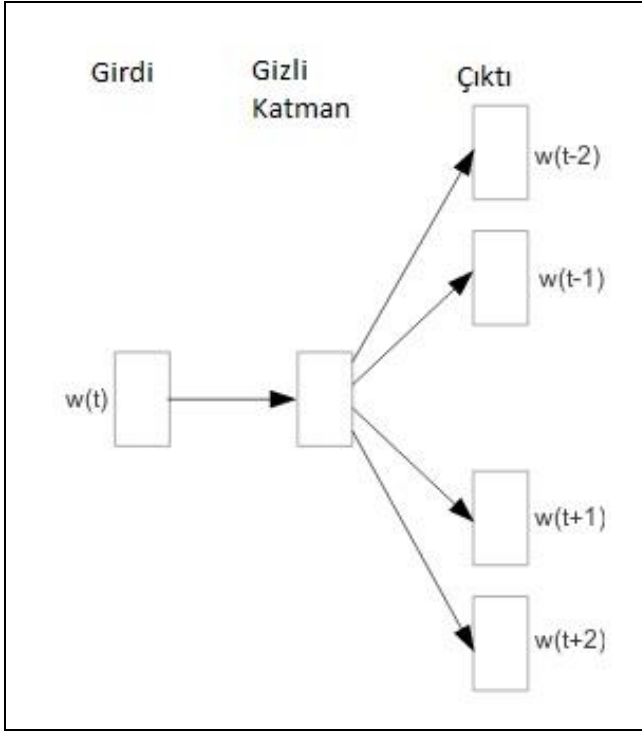
Figure 1. Flow chart of the proposed system.

3.2 Kelime yerleştirme modelleri

3.2.1 Word2Vec

Word2Vec, metni işleyen iki katmanlı bir sinir ağıdır. Word2Vec derin bir sinir ağı olmasa da metni derin ağların anlayabileceği sayısal bir forma dönüştürür. Kelime vektörlerini oluştururken pencere genişliği, yerleştirme boyutu gibi hiper parametreler bulunmaktadır. Pencere genişliği hedef kelimenin sağında ve solunda kaç kelime olması gerektiğini belirtirken, yerleştirme boyutu ise her bir kelimenin kaç boyutlu vektör olarak tanımlanacağını belirtir. Bu durum gizli katmandaki nöron sayısına karşılık gelmektedir.

Mikolov ve arkadaşları [13],[15] Skip-gram ve Sürekli Sözcük Torbası (CBOW) modelleri olmak üzere iki farklı model önermektedir. Şekil 2'de Skip-gram modeli mimarisi verilmiştir. Skip-gram modelleri, bir cümle veya belgedeki bir kelimenin etrafındaki kelimeleri tahmin eder [12].



Şekil 2. Skip-gram model mimarisi.

Figure 2. Architecture of the Skip-gram model.

Girdi ve çıktıları olasılık olarak benzetip anlamsal olarak en uygun biçimde temsil etmek amaçlanır. Örneğin, 'ben bugün çok fazla çalıştım' cümlesi için, 'çok' kelimesi girdi olarak kullanılır ve pencere boyutu 5 olarak ayarlandığında, 'ben', 'bugün', 'fazla' ve 'çalıştım' kelimeleri çıktı olarak kullanılır. Matematiksel olarak ifade edildiğinde, $w_1, w_2, w_3, \dots, w_T$ kelime dizisi aşağıdaki denklemi sağlar. Burada, j verilen bir kelime için öngörülme çalışılan komşu kelime sayısının iki katıdır ve eğitilecek bağlamın büyüklüğünü temsil etmektedir. J arttırıldığında doğruluk oranları daha yüksek vektörler elde edilmesi beklenmektedir, ancak eğitim süresi uzamaktadır. Skip-gram modelinin amacı, ortalama doküman olasılığını en üst düzeye çıkarmaktır. Bu çalışmada Skip-gram modeline odaklanılmıştır. Denklem 1, Skip-gram modelini ifade eder. Burada w_T , orta kelimeyi temsil ederken eğitim için verilen kelimeler dizisi $w_1, w_2, w_3, \dots, w_T$ şeklindedir. Ayrıca, n daha fazla veya daha az çevreleyen kelimeleri içerecek şekilde ayarlanabilen eğitim bağlamının boyutunu belirtir.

Mikolov ve arkadaşları [13],[15] Skip-gram ve Sürekli Sözcük Torbası (CBOW) modelleri olmak üzere iki farklı model önermektedir. Şekil 2'de Skip-gram modeli mimarisi verilmiştir. Skip-gram modelleri, bir cümle veya belgedeki bir kelimenin etrafındaki kelimeleri tahmin eder [12]. Girdi ve çıktıları olasılık olarak benzetip anlamsal olarak en uygun biçimde temsil etmek amaçlanır. Örneğin, 'ben bugün çok fazla çalıştım' cümlesi için, 'çok' kelimesi girdi olarak kullanılır ve pencere boyutu 5 olarak ayarlandığında, 'ben', 'bugün', 'fazla' ve 'çalıştım' kelimeleri çıktı olarak kullanılır. Matematiksel olarak ifade edildiğinde, $w_1, w_2, w_3, \dots, w_T$ kelime dizisi aşağıdaki denklemi sağlar. Burada, j verilen bir kelime için öngörülme çalışılan komşu kelime sayısının iki katıdır ve eğitilecek bağlamın büyüklüğünü temsil etmektedir. J arttırıldığında doğruluk oranları daha yüksek vektörler elde edilmesi beklenmektedir, ancak eğitim süresi uzamaktadır. Skip-gram

modelinin amacı, ortalama doküman olasılığını en üst düzeye çıkarmaktır. Bu çalışmada Skip-gram modeline odaklanılmıştır. Denklem 1, Skip-gram modelini ifade eder. Burada w_T , orta kelimeyi temsil ederken eğitim için verilen kelimeler dizisi $w_1, w_2, w_3, \dots, w_T$ şeklindedir. Ayrıca, n daha fazla veya daha az çevreleyen kelimeleri içerecek şekilde ayarlanabilen eğitim bağlamının boyutunu belirtir.

$$J_0 = \frac{1}{T} \sum_{t=1}^T \sum_{-j \leq 0 \leq j \neq 0} \log p(w_{t+j} | w_t) \quad (1)$$

3.2.2 GloVe

GloVe, kelime temsili için global vektörler, önerilen bir başka kelime yerleştirme modelidir. İstatistikleri kontrol etmeyen algoritmalara dayanır. Bir korpustaki kelime oluşumlarının istatistikleri, kelime temsillerini öğrenmek için tüm denetlenmemiş yöntemlerin kullanabileceği temel bilgi kaynağıdır. Kelime vektörleri bu anlamı temsil edebilir.

Skip-gram ve CBOW gibi modeller anlamsal bilgileri yakalar, ancak birlikte çalışabilirlik için istatistik kullanmazlar. Matris ayrıştırma yöntemleri bu istatistikleri kullanmasına rağmen, anlamsal ilişkileri yakalayamazlar. Bu kelime yerleştirme modeli, problemi olasılık istatistiklerini kullanarak yeni bir amaç fonksiyonu yaratarak çözmeyi amaçlar. GloVe modeli kelime yerleştirme işlemlerini elde etmek için aşağıda yer alan Denklem 2' deki işlevi kullanır [16]. Burada, X kelime birliktelik matrisini, X_{ij} 'in her elemanı i kelimesinin j bağlam kelimesinde kaç kez görüldüğünü göstermektedir.

$$J_0 = \sum_{i,j=1}^V f(X_{ij})(w_i^T + b_i + \bar{b}_j - \log X_{ij})^2 \quad (2)$$

3.2.3 FastText

FastText, Word2Vec modelinin bir uzantısı olan başka bir kelime yerleştirme modeli yöntemidir. FastText, son birkaç on yılda doğal dil işleme ve makine öğrenen toplulukların getirdiği en başarılı konseptlerden bazılarını birleştirir. Bunlar, cümleleri kelime torbası ve n-gram torbasıyla temsil etmenin yanı sıra alt kelime bilgisini kullanma ve gizli bir temsil yoluyla sınıflar arasında bilgi paylaşımını içerir. Ayrıca hesaplamayı hızlandırmak için sınıfların dengesiz dağılımından faydalanan hierarchical softmax kullanılır. Örneğin, $n=3$ ile "artificial" kelimesi seçilirse, FastText aşamaları <ar, art, rti, tif, ifi, fic, ici, ial, al>, kelimenin başlangıcı ve bitişidir. Bir kelime veya harften ne kadar olduğunu anlamamızı sağlar. Çok sıklıkla geçmeyen kelimelerin n-gramlarının ortaya çıkma olasılığı zayıf olduğu için bu kelimeler daha doğru bir biçimde temsil edilir. FastText, değiştirilmiş bir Skip gram kaybı fonksiyonuna sahip Word2Vec için önerilen negatif örnekleme Skip gram modelini kullanır. $G_w \subset \{1, \dots, G\}$ a kelimesinde görünen n-gram kümesi olsun, kelimenin puanı n-gramın vektör aşamalarının toplamıyla Denklem 3'deki FastText genel denklemi elde edilir [12]. Burada, w kelimesi, kelimenin kendisi için bir vektörden oluşan tüm n-gram zg' nin toplamı olarak temsil edilir. Ayrıca v ile sözlüğü, c ise gizli bir katman çıktısı bağlam kelimesi vektörünün ağırlıklı ortalamasıdır.

$$s(w, c) = \sum_{g \in G_w} z_g^T v_c \quad (3)$$

3.3 Derin öğrenme teknikleri

Bu çalışmada, Evrişimli Sinir Ağları (CNN'ler), Tekrarlayan Sinir Ağları (RNN) ve Uzun Kısa Süreli Bellek ağları (LSTM) gibi yaygın olarak kullanılan üç derin öğrenme algoritmasına odaklanıyoruz.

3.3.1 Evrişimli sinir ağları (CNN)

Evrişimli sinir ağları (CNN'ler) bir tür çok katmanlı algılayıcıdır (MLP) [17]. İleriye dönük bir sinir ağı olan CNN, hayvanların görsel merkezinden ilham almıştır. Buradaki matematiksel evrişim süreci, bir nöronun uyarıcı alanından uyarıcılara tepkisi olarak düşünülebilir. Tüm CNN katmanları tamamen birbirine bağlanacak şekilde tasarlanmıştır ve her bir evrişim filtresi öğrenilecek özellikleri oluşturur. Bu katlamalı katmanlar, havuzlama yoluyla hem ebatta hem de eğitim süresinde optimizasyon sağlar.

Bu katmanlar;

- Evrişimsel Katman-Özellikleri belirlemek amacıyla kullanılır,
- Non-Linearity Katman-Sisteme lineer olmayanlığın (non-linearity) belirtilmesi için kullanılır,
- Pooling (Downsampling) Katman-Ağırlık sayısını düşürür ve uygunluğu kontrol eder,
- Flattening Katman-Klasik sinir ağı için verileri hazır hale getirir,
- Fully-Connected Katman-Sınıflamada işleme alınan standart sinir ağıdır.

Temel olarak, CNN, sınıflandırma probleminin çözüm yolu için standart sinir ağı kullanır ve bilgileri tespit etmek ve bazı özellik durumlarını belirlemek için diğer katmanları kullanır.

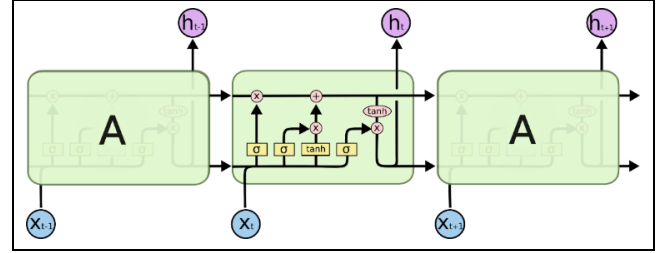
3.3.2 Tekrarlayan sinir ağları (RNN)

Tekrarlayan sinir ağı (RNN), önceki adımdan gelen çıkışın mevcut aşamaya girdi olarak beslendiği bir tür sinir ağıdır [36]. RNN, kelimeleri hatırlama gereği nedeniyle önerilmektedir. Bu sorun gizli katman yardımı ile çözülür [14],[37]. RNN'in en önemli özelliği, bir dizi hakkında bazı bilgileri hatırlayan gizli durumdur (secret state). RNN, hesaplananlarla ilgili tüm bilgileri hatırlayan bir belleğe sahiptir. RNN, diğer sinir ağlarından farklı olarak, parametrelerin karmaşıklığını azaltır. Çıktıyı üretmek için tüm girdilerde veya gizli katmanlarda aynı görevi gerçekleştirir. Her giriş için aynı parametreleri kullanmak, parametrelerin karmaşıklığını azaltır.

3.3.3 Uzun kısa süreli bellek ağları (LSTM)

Uzun Kısa Süreli Bellek ağları genellikle "LSTM" olarak adlandırılır ve uzun vadeli bağımlılıkları öğrenebilen özel bir RNN türüdür. Başlangıç noktası, derin sinir ağlarını eğitirken, geri yayılım algoritmasını kullanarak üssel hata büyümesi sorununa bir çözüm sağlamaktır. Bu sorunun temel nedeni, aktivasyon fonksiyonu tarafından üretilen değerlerin sürekli olarak -1, 1 aralığında olması, böylece bu değerlerin geri yayılma algoritmasına verilmesi ve sıfır ile çarpılmasıdır. Bu problemten kaçınmak ve karmaşık yapılar için daha iyi öğrenme algoritmaları tasarlamak için geliştirilen LSTM, uzun vadeli bağımlılıkların ve uzun vadeli bilgilerin hatırlanması gereken problemlerde iyi sonuçlar vermektedir. RNN hücrelerine ayrıca bellek eşlik eder. Her adımda, öğrenilen hücrelerin hangilerinin atılması gerektiğine ve hangilerinin güncelleneceğine karar verilir. Nöral makine çevirisi için Google tarafından başarıyla kullanılan bir yapıdır. Bu

çalışmada, daha önce kullanılan kelimelerin anlamlarını öğrenmek ve bu anlamlara dayalı tahminler üretmek için kullanılmıştır [12],[18]. Şekil 3'te LSTM ağı mimarisi gösterilmiştir.



Şekil 3. LSTM ağı mimarisi.

Figure 3. Architecture of LSTM network.

4 Sonuçlar

Bu çalışmada, Parkinson hastalığı ile ilgili kullanıcı yorumlarının duygularını kelime yerleştirme modelleri ve derin öğrenme algoritmaları kullanarak analiz etmek için kapsamlı deneyler yapılmıştır. Her modelin sınıflandırma performansını ve çalışmalarımızın katkısını göstermek için deneylerde bir değerlendirme ölçütü olarak doğruluk metriği kullanılmıştır. Önceki literatür çalışmalarına benzer şekilde [37],[38] verilerin eğitimi için veri kümesinin %80' i ve test için ise %20' si kullanılır. 100 boyutlu belgeleri göstermek için Google Haberler veri kümesinden Word2Vec Skip-gram modeli ile önceden eğitilmiş vektörler kullanılmıştır.

Bölüm 3.1'de bahsedilen ön işleme yöntemleri için SWE, RH, RU, STM ve AOT ve derin öğrenme algoritmaları için CNN, RNN ve LSTM kısaltmaları kullanılmıştır. Elde edilen en iyi doğruluk sonuçları kalın harflerle belirtilmiştir. İlk aşamada, ön işleme yöntemlerinin kelime yerleştirme modelleri üzerindeki etkisi gösterilmiştir.

Tablo 2'de, ön işleme yöntemlerinin %80 eğitim kümesinde kelime yerleştirme modelleri üzerindeki etkisi yukarıdaki gibi gösterilmiştir.

Tablo 2. Ön işleme modellerinin kelime yerleştirme yöntemlerine göre doğruluk sonuçları.

Table 2. Accuracy results by word embedding methods of preprocessing models.

Ön işleme Modelleri	Word2Vec	GloVe	FastText
SWE	81.23	80.87	78.98
RH	83.75	81.51	80.72
RU	83.94	82.25	81.86
STM	78.36	76.79	75.69
AOT	82.48	80.59	78.83

RU, önerilen sistemin başarısını artırarak, %83.75 doğruluk başarısı ile rekabet eden RH'ye kıyasla %83.94 doğrulukla üstün sınıflandırma performansı sergiler. Bunu %82.48 ile AOT, %81.23 ile SWE ve %78.36 ile STM izlemektedir. Yerleştirme modelleri arasında Word2Vec, ön işleme yöntemi RU olarak ayarlandığında, sınıflandırma başarısını %83.94 doğrulukla dikkat çekici bir şekilde göstermektedir. Ayrıca, Word2Vec diğer kelime yerleştirme modellerinden daha iyi performans gösterirken, FastText kelime yerleştirme modellerinde %81.86 ile en düşük sınıflandırma performansına sahiptir. Dahası, Word2Vec, FastText'in başarısı göz önüne alındığında yaklaşık %2 oranında bir iyileşme

sağlarken, GloVe FastText'in sınıflandırma performansına kıyasla yaklaşık %1 oranında bir geliştirme sağlar.

RU'nun üstün başarısına ek olarak, STM sınıflandırmadaki en kötü performansı sergilemektedir. Tablo 2'nin sonucu olarak ön işleme modellerinin sınıflandırma doğruluklarının sırası genellikle şöyledir: RU> RH> AOT> SWE> STM. Bu çalışmada RU, sınıflandırma üzerindeki başarılı etkisinden dolayı önerilen sistemin ön işleme modeli olarak seçilmiştir.

Tablo 3'te doküman boyutuna göre Word2vec modelinin derin öğrenme modelleriyle harmanlandığı durumda derin öğrenme modellerinin sınıflandırma performansı değerlendirilmiştir. CNN, tüm doküman sayılarında en iyi performansı gösterirken, doküman sayısı 7500 değerinde %75.12 doğruluk oranı ile en iyi derin öğrenme modeli olduğu gösterilmiştir. CNN, RNN'ye kıyasla yaklaşık %2 daha iyi bir başarıya sahiptir, ancak ikisi de birbirlerine çok yakın sınıflandırma performansları sergiler. Dahası, LSTM tüm doküman sayısı boyutlarında diğer derin öğrenme modellerine göre yaklaşık %20 daha başarısız olduğu görülüyor. Diğer doküman sayısı boyutlarında, CNN ve RNN arasındaki doğruluk farkı, %2 ile %3 arasında değişirken, CNN, doğruluk yüzdesi bakımından LSTM'ye kıyasla minimum %20, maksimum %24 artış gösterir. Tüm derin öğrenme teknikleri için en yüksek doğruluk oranı doküman sayısı 7500 boyutunda elde edilir. Özet olarak, derin öğrenme algoritmalarının sınıflandırma başarısı tüm doküman sayısı boyutlarında CNN> RNN> LSTM şeklinde sıralanır.

Tablo 3. Doküman boyutuna göre Word2Vec modelinin derin öğrenme modelleriyle harmanlandığı durumdaki doğruluk sonuçları.

Table 3. Accuracy results in a situation where the Word2Vec model is blended with deep learning models based on document size.

Doküman sayısı	CNN	RNN	LSTM
100	69.01	65.45	48.12
500	70.98	68.53	49.56
1000	70.59	69.29	50.37
2500	72.01	70.21	50.78
5000	73.39	71.87	51.89
7500	75.12	73.90	52.20
10000	74.15	72.54	51.68

Önerilen hibrid modelimizin Word2Vec için Skip-gram modeli ile önceden eğitilmiş vektörlerle kullanıldığından bahsetmiştik. Skip-gram modeli yerine önerilen hibrid modelin Word2Vec' in CBOW modeliyle gerçekleştirilmesi durumunda, sistem performansı Tablo 3'te yer alan sonuçlarla kıyaslandığında doküman sayısının 7500 olduğu durumda CNN modeli için yaklaşık %3, RNN algoritması için yaklaşık %5, LSTM metodu için yaklaşık olarak %1 oranında bir performans kaybı gözlenmiştir. Buradan hareketle, önerilen modelde, Word2Vec kelime yerleştirme yöntemi için Skip-gram yöntemi uygulanmıştır.

Tablo 4'te doküman boyutuna göre GloVe modelinin derin öğrenme modelleriyle harmanlandığı durumda derin öğrenme modellerinin sınıflandırma performansı değerlendirilmiştir. Tablo 4'te en iyi sınıflandırma doğruluğunun Tablo 3'te olduğu gibi 7500 doküman boyutunda olduğu açıkça görülmektedir. Tüm derin öğrenme teknikleri için 7500 doküman sayısından 10000 doküman sayısına, sınıflandırma başarısında yaklaşık %2 azalma vardır. Daha düşük doküman sayılarında, sınıflandırma performansındaki düşüş, doküman sayısı 7500

ve 10000 aralığına kıyasla keskin bir şekilde gözlenir. Daha geniş bir bakış açısıyla, RNN doküman sayısı 7500'de diğerlerine kıyasla %74.46 doğrulukla üstün sınıflandırma performansı sergiler. Bunu %69.13 doğrulukla CNN ve %51.94 doğruluk sonucu ile LSTM takip eder.

Tablo 4. Doküman boyutuna göre GloVe modelinin derin öğrenme modelleriyle harmanlandığı durumdaki doğruluk sonuçları.

Table 4. Accuracy results in a situation where the GloVe model is blended with deep learning models based on document size.

Doküman sayısı	CNN	RNN	LSTM
100	62.56	68.41	46.95
500	63.21	69.48	47.12
1000	65.93	71.32	47.71
2500	67.45	72.59	49.07
5000	68.01	74.08	50.23
7500	69.13	74.46	51.94
10000	67.91	72.11	50.09

Tablo 5'te doküman boyutuna göre FastText modelinin derin öğrenme modelleriyle harmanlandığı durumda derin öğrenme modellerinin sınıflandırma performansı değerlendirilmiştir. RNN, tüm doküman sayılarında en iyi performansı gösterir. Doküman sayısı 7500 iken %73.34 doğruluk oranı ile RNN' in en iyi derin öğrenme modeli olduğu gösterilmiştir. RNN, CNN'e kıyasla yaklaşık %2 daha iyi bir başarıya sahiptir, ancak ikisi de birbirlerine çok yakın sınıflandırma performansları sergiler. Ayrıca, Tablo 3'e benzer şekilde, LSTM' in tüm doküman sayısı boyutlarında diğer derin öğrenme modellerine göre yaklaşık %20 daha başarısız olduğu gözlemlenmiştir. RNN, doküman sayısı boyutu 7500 olarak ayarlandığında %73.34 doğruluk başarısı ile doküman sayısı 5000 olarak ayarlandığında elde edilen %71.93 doğruluk sonucuna kıyasla üstün sınıflandırma performansı sergiler. Derin öğrenme algoritmalarının sınıflandırma başarısı tüm doküman sayılarında RNN> CNN> LSTM şeklinde sıralanır.

Tablo 5. Doküman boyutuna göre FastText modelinin derin öğrenme modelleriyle harmanlandığı durumdaki doğruluk sonuçları.

Table 5. Accuracy results in a situation where the FastText model is blended with deep learning models based on document size.

Doküman sayısı	CNN	RNN	LSTM
100	64.35	66.48	44.77
500	65.92	68.32	46.13
1000	68.63	69.04	47.75
2500	69.22	69.85	49.81
5000	70.47	71.93	50.66
7500	71.84	73.34	51.94
10000	70.34	71.27	50.82

Bu çalışmada, Parkinson hastalığına ilişkin son çalışmaların aksine, kullanıcıların sosyal medya platformlarındaki duyguları analiz edilerek Parkinson hastalığının tespitine odaklanıldı. Kullanıcıların, Parkinson hastalığı hakkındaki görüşlerini anlamak ve analiz etmek için kelime yerleştirme modelleri ve derin öğrenme yöntemleri harmanlanmış ve sınıflandırma amacıyla kullanılmıştır. Bu amaçla CNN, RNN ve LSTM derin öğrenme algoritmaları olarak kullanılmakta, Word2Vec, GloVe ve FastText ise kelime yerleştirme modelleri olarak değerlendirilmektedir. Dahası, önerilen modelin sınıflandırma

performansını arttırmak için ön işleme yöntemleri olarak SWE, RH, RU ve STM üzerinde duruyoruz. Ayrıca, doküman sayılarının etkisi de sınıflandırma performanslarına göre incelenmiştir. Sonuç olarak, ön işleme aşamasında CNN ile Word2Vec+RU kombinasyonunun kullanılması, kullanıcıların Parkinson hastalığına karşı duyarlılıklarını belirlemek amacıyla doküman sayısı 7500 için kayda değer bir sınıflandırma başarısı gerçekleştirir.

Bu çalışma, [39]'daki çalışmanın genişletilmiş halidir. [39]'da yazarlar derin öğrenme algoritmalarının ve kelime yerleştirme modellerinin performansını birbirinden bağımsız olarak farklı eğitim yüzdelerinde değerlendirmektedirler. Bu çalışmamızda, eğitim yüzdesi %80 olarak ayarlanmış olup kelime yerleştirme modelleriyle derin öğrenme modelleri harmanlanmıştır. Dolayısıyla, her iki çalışmada da önerilen modeller birbirinden farklıdır. Ayrıca [39]'daki çalışmada sadece farklı eğitim yüzdelerinde ve ön işleme yöntemlerinde derin öğrenme algoritmalarının ve kelime yerleştirme modellerinin ayrı ayrı performansları incelenirken bu çalışmada, önerilen hibrid modellerin doküman sayıları üzerindeki başarısı da sunulmuştur. Dahası, bu çalışmadaki ön işleme yöntemleri sadece kelime yerleştirme modelleri üzerinden değerlendirilmiş olup [39]'daki çalışmanın kapsamı amacıyla ön işleme yöntemlerinin her modele olan ilişkisi gösterilmiştir. Önceki çalışmamızda [39], deney sonuçlarının karşılaştırılması açısından %80 eğitim verisi dikkate alındığında kelime yerleştirme modellerinden Word2Vec, %76.64 doğruluk sonucu sunmakta olup bu çalışmada, hibridlenen modellerden Word2Vec kelime yerleştirme modelinin en iyi sonucu CNN algoritmasıyla birleştirilerek ve %75.12 doğruluk sonucu sunarak elde edildiği gözlenmiştir. Aradaki fark, ön işleme yöntemlerinden sadece bir tanesinin kullanılmasından kaynaklı olabilir. Ayrıca [39]'daki çalışmadan en iyi sınıflandırma performansı sistem tek başına modellendiğinde LSTM algoritmasına ait iken bu çalışmada, hibridlenecek kelime yerleştirme modeline göre değişkenlik göstermekte olup en kötü performans, LSTM derin öğrenme algoritması ile elde edilmiştir.

Bildiğimiz kadarıyla, bu; sözcük yerleştirme modelleri ve derin öğrenme algoritmaları kullanarak Parkinson hastalığının algısını analiz eden ilk çalışmadır. Bu şekilde, kullanıcıların Parkinson hastalığına ilişkin duygu analizi, Parkinson hastaları ve yakınları için hem erken tanı hem de tedavi sürecinde yeni bir bakış açısı sağlayabilir. Ayrıca, bu çalışma, Parkinson hastalığı ile ilgili tweetlerden elde edilen analizlerin, Parkinson hastalarının bakım sürecinde nelere ihtiyaç duyduğunu ve Parkinson hastalarının ve yakınlarının sağlık sorunları hakkında birbirleriyle nasıl iletişim kurduğunu ortaya çıkaran değerli bir sağlık hizmeti kaynağı olduğunu göstermektedir. Gelecekteki bir çalışma olarak, Parkinson hastalığının erken tespitini güçlendirerek hem metin içerikli hem de sesli kayıt verilerini içeren bir hibrit modeli geliştirmeyi planlıyoruz.

5 Conclusions

In this study, extensive experiments are conducted to analyze the sentiments of user comments related to Parkinson's disease using word embedding models and deep learning algorithms. The accuracy metric is used as an evaluation criterion in experiments to demonstrate the classification performance of each model and the contribution of our work. Similar to previous literature studies [37],[38], 80% of the data set is used for training and remaining part is for testing. Pre-trained

vectors are used with the Word2Vec Skip-gram model from the Google News data set to show 100-dimensional documents.

The abbreviations SWE, RH, RU, STM and AOT are used for the preprocessing methods mentioned in Section 3.1, and CNN, RNN and LSTM are employed for deep learning algorithms. The best accuracy results obtained are indicated in bold. In the first stage, the effect of preprocessing methods on word embedding models is shown.

In Table 2, the effect of preprocessing methods on word embedding models in 80% training set is shown as above. RU exhibits superior classification performance with 83.94% accuracy compared to the competing RH with 83.75% accuracy success, increasing the success of the proposed system. It is followed by AOT with 82.48%, SWE with 81.23%, and STM with 78.36%. Among word embedding models, Word2Vec remarkably demonstrates classification success with 83.94% accuracy when the preprocessing method is set to RU. In addition, Word2Vec performs better than other word embedding models, while FastText has the lowest classification performance with 81.86% in word embedding models. Moreover, Word2Vec provides an improvement of about 2% given the success of FastText, while GloVe provides an improvement of about 1% compared to classification performance of FastText. In addition to the outstanding success of RU, STM is the worst performer in the classification. As a result of Table 2, the order of classification accuracy of preprocessing models is usually as follows: RU> RH> AOT> SWE> STM. In this study, RU is chosen as the preprocessing model of the proposed system due to its successful effect on classification.

In Table 3, the classification performance of the deep learning models is evaluated when the Word2Vec model is blended with the deep learning models based on the document size. CNN exhibits the best performance in all document sizes, while the best deep learning model CNN with an accuracy rate of 75.12% when document size is set to 7500. CNN has about 2% better success compared to RNN, but both of them exhibit very close classification performances. Moreover, LSTM appears to fail about 20% more than other deep learning models in all document sizes. In other document sizes, the accuracy difference between CNN and RNN ranges from 2% to 3%, while CNN increases by a minimum of 20% and a maximum of 24% compared to LSTM in terms of accuracy percentage. The highest accuracy rate for all deep learning techniques is achieved in the size of 7500 documents. In summary, the classification success of deep learning algorithms is ranked as CNN> RNN> LSTM across all document sizes. We mention that the proposed hybrid model is used with vectors pre-trained with the Skip-gram model for Word2Vec. If the proposed hybrid model instead of the Skip-gram model is implemented with the CBOW model of Word2Vec, the loss of system performance is approximately 3% for the CNN model, approximately 5% for the RNN algorithm, and approximately 1% for the LSTM method when the number of documents is 7500 compared to the results in Table 3. Based on this, the Skip-gram method is applied to the Word2Vec word embedding method in the proposed model.

In Table 4, the classification performance of deep learning models is evaluated when the GloVe model is blended with deep learning models based on the document size. In Table 4, it is clearly observed that the best classification accuracy is obtained when the document size is adjusted as 7500, as in

Table 3. From the number of 7500 documents to the number of 10000 documents for all deep learning techniques, there is about a 2% reduction in classification success. At lower document numbers, the decrease in classification performance is observed sharply compared to the 7500 and 10000 document count range. More broadly, the RNN performs superior classification performance at 7500 document size with an accuracy of 74.46% compared to the others. It is followed by CNN with 69.13% of accuracy and LSTM with 51.94% of accuracy.

In Table 5, the classification performance of deep learning models is evaluated when the FastText model is blended with deep learning models based on the document size. RNN exhibits the best classification success in all document sizes. RNN has been shown to be the best deep learning model with an accuracy of 73.34%, while the number of documents is 7500. RNN has about 2% better success compared to CNN, but both of them demonstrate very close classification performances. Furthermore, similar to Table 3, it is observed that LSTM failed approximately 20% more in all document sizes than the other deep learning models. RNN has superior classification performance compared to 73.34% accuracy success when the number of documents is set to 7500 and 71.93% accuracy result when the number of documents is set to 5000. Classification success of deep learning algorithms is sorted in all document numbers as RNN> CNN> LSTM.

In this study, unlike recent studies on Parkinson's disease, the focus is on detecting Parkinson's disease by analyzing users' emotions on social media platforms. Word embedding models and deep learning methods are collated and used for classification purposes to understand and analyze users' views on Parkinson's disease. For this purpose, CNN, RNN and LSTM are used as deep learning algorithms, while Word2Vec, GloVe and FastText are evaluated as word embedding models. Moreover, we also focus on SWE, RH, RU and STM as preprocessing methods to improve the classification performance of the proposed model. In addition, the effect of document sizes is also studied according to classification performance. As a result, the use of a combination of CNN and Word2Vec+RU during the pre-processing phase achieves a remarkable classification success at 7500 document sizes to determine the sensitivity of users to Parkinson's disease.

This study is an extended version of the study in [39]. In [39], the authors evaluate the performance of deep learning algorithms and word embedding models at different training percentages independently of each other. In this study, the training percentage is set as 80% and word embedding models and deep learning models are blended. Therefore, the models proposed in both studies are different from each other. In addition, the study in [39] examined only the separate performance of deep learning algorithms and word embedding models in different training set percentages and preprocessing methods, while the success of the proposed hybrid models on the number of documents is also presented in this study. Moreover, the preprocessing methods in this study are evaluated only on word embedding models and the relationship of preprocessing methods to each model is shown for the purpose of the scope of the study in [39]. In our previous study [39], Word2Vec from word embedding models offers 76.64% accuracy result when 80% training data is taken into account in terms of comparing the results of the experiment. In this study, it is observed that the best result of the Word2Vec model from the hybridized models is achieved by combining it

with the CNN algorithm and delivering a 75.12% of accuracy result. The difference may be due to the use of only one of the preprocessing methods. In addition, the best classification performance from the study in [39] belongs to the LSTM algorithm when the system is modeled alone, while in this study, it varies according to the hybrid word embedding model, and the worst performance is achieved by the LSTM deep learning algorithm. To our knowledge, this is the first study to analyze the perception of Parkinson's disease using word embedding models and deep learning algorithms. In this way, users' analysis of feelings about Parkinson's disease can provide a new perspective for Parkinson's patients and their relatives in both the early diagnosis and treatment process. Furthermore, this study shows that analyses from tweets about Parkinson's disease are a valuable source of health care, revealing what Parkinson's patients need in the care process and how Parkinson's patients and their relatives communicate with each other about their health problems.

As a future study, we plan to develop a hybrid model that includes both text and voice recording data, strengthening early detection of Parkinson's disease.

6 Yazar katkı beyanı

Gerçekleştirilen çalışmada Feyza Çevik kaynak ve malzemelerin temini, veri toplama, analizlerin gerçekleştirilmesi, literatür taraması, yazım denetimi ve içerik açısından makalenin kontrol edilmesi başlıklarında; Zeynep Hilal Kilimci fikrin oluşturulması, tasarımın yapılması, sonuçların incelenmesi ve değerlendirilmesi, yazım denetimi ve içerik açısından makalenin kontrol edilmesi başlıklarında katkı sunmuşlardır.

7 Etik kurul onayı ve çıkar çatışması beyanı

Hazırlanan makalede etik kurul izni alınmasına gerek yoktur.

Hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

8 Kaynaklar

- [1] Eckler P, Worsowicz G, Rayburn J. "Social media and healthcare: an overview". *Physical Medicine and Rehabilitation*, 2(11), 1046-1050, 2010.
- [2] Prieto VM, Matos S, Alvarez M, Cacheda F, Oliveira JL. "Twitter: A good place to detect health conditions". *Public Library of Science One*, 9(1), 1-11, 2014.
- [3] Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. "Twitter as a tool for health research: A systematic review". *American Journal of Public Health*, 107(1), 1-8, 2017.
- [4] Neiger BL, Thackeray R, Burton SH, Thackeray CR, Reese JH. "Use of Twitter among local health departments: An analysis of information sharing, engagement, and action". *Journal of Medical Internet Research*, 15(8), 1-10, 2013.
- [5] Beykikhoshk A, Arandjelović O, Phung D, Venkatesh S, Caelli T. "Data-mining Twitter and the autism spectrum disorder: A pilot study". *IEEE/ACM 2014 International Conference on Advances in Social Networks Analysis and Mining*, Beijing, China, 17-20 August 2014.
- [6] Beykikhoshk A, Arandjelović O, Phung D, Venkatesh S. "Overcoming data scarcity of Twitter: Using tweets as bootstrap with application to autism-related topic content analysis". *IEEE/ACM 2015 International Conference on Advances in Social Networks Analysis and Mining*, Paris, France, 25-28 August 2015.

- [7] Chew C, Eysenbach G. "Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak". *Public Library of Science One*, 5(11), 1-13, 2010.
- [8] DeRijk MD, Tzourio C, Breteler MM, Dartigues JF, Amaducci L, Lopez-Pousa S, Rocca WD. "Prevalence of parkinsonism and Parkinson's disease in Europe: The Europarkinson collaborative study. European community concerted action on the epidemiology of Parkinson's disease". *Journal of Neurology, Neurosurgery & Psychiatry*, 62(1), 10-15, 1997.
- [9] Fahn S. "Description of Parkinson's disease as a clinical syndrome". *Annals of the New York Academy of Sciences*, 991, 1-14, 2003.
- [10] Elman JL. "Finding structure in time". *Cognitive Science*, 14(2), 179-211, 1990.
- [11] Joulin A, Grave E, Bojanowski P, Douze M, Jégou H, Mikolov T. "Fasttext. zip: Compressing text classification models". *ICLR 2017 International Conference on Learning Representations*, Toulon, France, 24-26 April, 2017.
- [12] Kilimci ZH, Akyokuş S. "Deep learning-and word embedding based heterogeneous classifier ensembles for text classification". *Complexity*, 2018, 106-116, 2018.
- [13] Le Q, Mikolov T. "Distributed representations of sentences and documents". *International Conference on Machine Learning*, Beijing, China, 21-26 June, 2014.
- [14] Lipton ZC, Berkowitz J, Elkan C. "A critical review of recurrent neural networks for sequence learning". <https://arxiv.org/abs/1506.00019> (15.04.2019).
- [15] Mikolov T, Chen K, Corrado G, Dean J. "Efficient estimation of word representations in vector space". *International Conference on Learning Representations Workshop*, Scottsdale, Arizona, 2-4 May 2013.
- [16] Pennington J, Socher R, Manning C. "Glove: Global vectors for word representation". In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, Doha, Qatar, 25-29 October 2014.
- [17] Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E. "Deep learning for computer vision: A brief review". *Computational Intelligence and Neuroscience*, 2018, 130-143, 2018.
- [18] Zhang L, Wang S, Liu B. "Deep learning for sentiment analysis: A survey". *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), 1-25, 2018.
- [19] Ene M. "Neural network-based approach to discriminate healthy people from those with Parkinson's disease". *Annals of the University of Craiova-Mathematics and Computer Science Series*, 35, 112-116, 2008.
- [20] Little M, McSharry P, Hunter E, Spielman J, Ramig L. "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease". *Nature Precedings*, 1(1), 1-27, 2008.
- [21] Sakar CO, Kursun O. "Teliagnosis of Parkinson's disease using measurements of dysphonia". *Journal of Medical Systems*, 34(4), 591-599, 2010.
- [22] Das R. "A comparison of multiple classification methods for diagnosis of Parkinson disease". *Expert Systems with Applications*, 37(2), 1568-1572, 2010.
- [23] Çağlar MF, Çetişli B, Toprak İB. "Automatic recognition of Parkinson's disease from sustained phonation tests using ANN and adaptive neuro-fuzzy classifier". *Mühendislik Bilimleri ve Tasarım Dergisi*, 1(2), 59-64, 2010.
- [24] Polat K. "Classification of Parkinson's disease using feature weighting method on the basis of fuzzy c-means clustering". *International Journal of Systems Science*, 43(4), 597-609, 2012.
- [25] Luukka P. "Feature selection using fuzzy entropy measures with similarity classifier". *Expert Systems with Applications*, 38(4), 4600-4607, 2011.
- [26] Kihel BK, Benyettou M. "Parkinson's disease recognition using artificial immune system". *Journal of Software Engineering and Applications*, 4(7), 391-395, 2011.
- [27] Eskidere Ö. "A Comparison of feature selection methods for diagnosis of Parkinson's disease from vocal measurements". *Sigma*, 30, 402-414, 2012.
- [28] Prashanth R, Roy SD. "Novel and improved stage estimation in Parkinson's disease using clinical scales and machine learning". *Neurocomputing*, 305, 78-103, 2018.
- [29] Prashanth R, Roy SD. "Early detection of Parkinson's disease through patient questionnaire and predictive modelling". *International Journal of Medical Informatics*, 119, 75-87, 2018.
- [30] Oscar N, Fox PA, Croucher R, Wernick R, Keune J, Hooker K. "Machine learning, sentiment analysis, and tweets: An examination of Alzheimer's disease stigma on Twitter". *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 72(5), 742-751, 2017.
- [31] Inan E, Soygazi F, Mostafapour V. "TurkiS: A Turkish sentiment analyzer using domain-specific automatic labelled dataset". *International Journal of Intelligent Systems and Applications in Engineering*, 7(2), 99-103, 2019.
- [32] Pervan N, Keleş Y. Derin Öğrenme Yaklaşımları Kullanarak Türkçe Metinlerden Anlamsal Çıkarım Yapma. Yüksek Lisans Tezi, Ankara Üniversitesi, Ankara, Türkiye, 2019.
- [33] Krizhevsky A, Sutskever I, Hinton GE. "Imagenet classification with deep convolutional neural networks". In *Advances in Neural Information Processing Systems*, Nevada, USA, 2-6 December 2012.
- [34] Liu P, Qiu X, Huang X. "Adversarial multi-task learning for text classification". *Association for Computational Linguistics*, Vancouver, Canada, 30 July- 4 August 2017.
- [35] Loria S. "Textblob documentation". <https://buildmedia.readthedocs.org/media/pdf/textblob/latest/textblob.pdf> (21.05.2019).
- [36] Elman JL. "Finding structure in time". *Cognitive Science*, 14(2), 179-211, 1990.
- [37] Kilimci ZH, Akyokuş S, Omurca SI. "The effectiveness of homogenous ensemble classifiers for Turkish and English texts". *IEEE 2016 International Symposium on INnovations in Intelligent Systems and Applications*, Sinaia, Romania, 2-5 August 2016.
- [38] Schneider KM. "On word frequency information and negative evidence in Naive Bayes text classification". *International Conference on Natural Language Processing*, Alicante, Spain, 20-22 October 2004.
- [39] Çevik F, Kilimci ZH. "Analysis of Parkinson's disease using deep learning and word embedding models". *International Symposium on Innovative Technologies in Engineering and Science*, Şanlıurfa, Turkey, 22-24 November 2019.

[40] Hutto CJ, Gilbert E. "Vader: A parsimonious rule-based model for sentiment analysis of social media text". *Eighth International AAAI Conference on Weblogs and Social Media*, 27-29 May 2014.

[41] MonkeyLearn API Reference.
<https://monkeylearn.com/docs/article/api-reference/>
(15.04.2020).