

## Türkçe Tweetler için Derin Özellik Çıkarımı Tabanlı Yeni Bir Duygu Sınıflandırma Modeli

Mehmet Umut SALUR<sup>1\*</sup>, İlhan AYDIN<sup>2</sup>

<sup>1</sup> Bilgisayar Teknolojisi Bölümü, Teknik Bilimler Meslek Yüksekokulu, Gaziantep İslam Bilim ve Teknoloji Üniversitesi, Gaziantep, Türkiye

<sup>2</sup> Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye

\*<sup>1</sup> mehmetumut.salur@gibtu.edu.tr, <sup>2</sup> iaydin@firat.edu.tr

(Geliş/Received: 17/06/2020);

Kabul/Accepted: 10/10/2021)

**Öz:** Sosyal medya uygulamalarının yaygın kullanımı insanların her dakika yeni veri üretmelerine neden olmuştur. Ses ve resim veri türlerine göre metin tabanlı verilerin boyutu daha hızlı artmaktadır. Metin tabanlı veriler, anlamlı kelimeler haricinde birçok içerik barındırabilmektedir. Metin işleme çalışmaları için bu içerikler gürültü olarak isimlendirilir ve metin ön işleme aşamasında bu içerikler veri kümelerinden çıkartılmaktadır. Özellikle Twitter veri kümeleri üzerinde yapılan duygu sınıflandırma çalışmalarında, veri kümeleri metin temsilleri oluşturulmadan önce yapılan ön işleme aşamasında URL, noktalama işareti ve emoji gibi içeriklerden arındırılmaktadır. Ancak Twitter için gürültü olarak nitelendirilen içerikler aslında bir bakıma kullanıcının duygu ve düşüncelerinin bir parçası niteliğindedir. Bu çalışmada, veri setinden çıkarılan gürültü verilerinden özellik çıkarımına odaklanılmıştır. Önerilen yöntem, gürültü olarak silinen içeriklerden manuel olarak çıkarılan özelliklerin ve derin öğrenme yöntemlerinden elde edilen derin özelliklerin birleştirilmesine dayanmaktadır. Önerilen model literatürde çalışılan Türkçe Twitter veri kümesi üzerinde gerçekleştirilmiştir. Yapılan deneylerle önerilen modelin sınıflandırma performansının önceki çalışmalardan daha iyi olduğu gösterilmiştir.

**Anahtar kelimeler:** Duygu Sınıflandırma, Özellik Çıkarımı, Derin Öğrenme, Twitter, Uzun Kısa Süreli Bellek.

### A Novel Sentiment Classification Model Based on Deep Feature Extraction for Turkish Tweets

**Abstract:** The widespread use of social media applications has caused people to generate new data every minute. The size of text-based data is increasing faster than audio and picture data types. Text-based data can contain many content other than meaningful words. In the text processing studies, these contents are named as noise and these contents are removed from data sets in the pre-processing phase. Especially in sentiment classification studies on Twitter data sets, data sets are purified from contents such as URLs, punctuations, and emojis at pre-processing phase before word embeddings are created. However, the content called noise in Twitter, might actually be the part of a user's sentiment or thought. In this study, we focused on feature extraction from the noise data that is removed from the dataset. The proposed method is based on the fusion of features extracted manually from contents deleted as noise and deep features obtained from deep learning methods. The proposed model was implemented on the Turkish Twitter dataset studied in the literature. It has been shown with the experiments that the classification performance of the proposed model is better than previous studies.

**Key words:** Sentiment Classification, Feature Extraction, Deep Learning, Twitter, Long Short Term Memory.

#### 1. Giriş

Günümüzde bilgi alışverişinin büyük bir kısmı Twitter, Facebook ve Instagram gibi sosyal medya uygulamaları üzerinden yapılmaktadır. Bu platformlardan Twitter aracılığıyla paylaşılan tweet'ler temelde bir ürün, etkinlik veya bir konuyla ilgili kamuoyunun görüşlerinden oluşmaktadır. Sosyal medya kullanıcıları tarafından paylaşılan bu içerikler büyük miktarda yapılandırılmamış veri içermektedir. Bu işlenmemiş veriler kar amacı güden şirketler, kurumlar ve kuruluşlardan kamuoyuna kadar geniş bir kitle için yüksek bir bilgi potansiyeline sahip olduğundan, keşfedilmesi gereken bir maden niteliğindedir. Bu verilerin işlenmesi ve bu verilerden anlamlı bilgilerin oluşturulması sağlık, pazarlama, ticaret, güvenlik, müşteri memnuniyeti gibi birçok alan için fırsat niteliği taşımaktadır. Paylaşılan her içerik kullanıcının bir konu, ürün, hizmet veya olay hakkındaki duygu, düşünce ve tutumunu temsil etmektedir. Duygu analizi, insanların yazdıkları metinlerden bu duygu, düşünce ve tutumlarını ortaya çıkarmayı hedeflemektedir [1]. Duygu sınıflandırma ise olumlu, olumsuz veya belirsiz gibi önceden belirlenen duygu kategorilerine göre metin verilerinin sınıflandırması alt işlemlerinden oluşmaktadır [2]. Twitter, 2020 yılı itibariyle 340 milyon aktif kullanıcıya sahip bir mikro blog sitesidir<sup>1</sup>.

\* Sorumlu yazar: [mehmetumut.salur@gibtu.edu.tr](mailto:mehmetumut.salur@gibtu.edu.tr). Yazarların ORCID Numarası: <sup>1</sup>0000-0003-0296-6266, <sup>2</sup>0000-0001-6880-4935

<sup>1</sup> <https://medyaakademi.com.tr/2020/02/03/2020-sosyal-medya-kullanici-sayilari/>

Tweet'lerden oluşan veri kümeleri üzerinde duygu analizi yapmak diğer veri kümelerine göre daha zordur [3]. Twitter'in karakter sınırlamasından dolayı metin içerikleri sıkıştırılmış anlama sahiptir. Tweet'ler karakter kısıtlamasından dolayı emoji ve özel kullanımlar içermektedir.

Duygu sınıflandırma için günümüzde makine öğrenmesi (MÖ) ve sözlük tabanlı yaklaşımlara göre derin öğrenme (DÖ) tabanlı yöntemler daha iyi sınıflandırma başarımları sağlamaktadır. MÖ temelli yaklaşımlarda metin veri kümelerinden kelime çantası (Bag of Words) (BoW) [4], TF-IDF (Term Frequency - Inverse Document Frequency) [5], sözcük etiketleme (Part of Speech) (PoS) [6] ve n-gram [7] gibi öznelik çıkarma yöntemleri sıklıkla tercih edilmektedir. Zhao ve Cao [3], önerdikleri MÖ tabanlı modelde n-gram, anlamsal ve polarite skoru özellikleri birleştirerek daha iyi sınıflandırma başarımları elde etmişlerdir. Çoban ve diğerleri [8], Türkçe tweet'lerden kelime çantası ve n-gram özellikleri çıkarıp, MÖ yöntemleriyle olumlu veya olumsuz olarak sınıflandırmışlardır. Siddiqua ve diğerleri [9], kural tabanlı bir sınıflandırıcıyı, denetimli sınıflandırıcıların çoğunluğuna dayalı bir oylama topluluğu ile birleştirmişlerdir. Chamansingh ve Hosein [10], her tweet için bir kelime skoru elde etmiş ve cümlenin skoruna göre tweet'leri duygu bakımından sınıflandırmışlardır. Jianqiang ve Xiaolin [11], altı farklı Twitter veri kümesi üzerinde, veri önışlemlerin sınıflandırma başarımına olan etkisini incelemişlerdir. Her veri seti için n-gram ve sözlük tabanlı özellikler elde etmiş ve MÖ yöntemleriyle sınıflandırmışlardır. Riyadh ve diğerleri [12], önışlemden geçirilen Twitter veri kümesinden unigram ve unigram ile sözcük etiketleme özelliklerini çıkarmışlardır. Özellikler Multinomial Naive Bayes (MNB) algoritmasıyla beş farklı sınıf (mutlu, üzgün, nefret, şaşırılmış ve belirsiz) doğrultusunda sınıflandırılmıştır. Tsapatsoulis ve Djouvas [13], Twitter paylaşımlarından manuel özellik çıkarmanın, klasik metin özellik çıkarma yöntemlerine göre performansını analiz etmişlerdir. Manuel özellik çıkarımının n-gram modellerine göre daha iyi sınıflandırma sağladığı görülmüştür. Kaur ve diğerleri [14], tweet'lerden n-gram yöntemiyle özellik çıkarımı yapmış ve tweet'leri K-En Yakın Komşu algoritmasıyla pozitif, negatif veya belirsiz olarak sınıflandırmışlardır. Permatasari ve diğerleri [15], Twitter'a özgü özellikler, sözdizimsel özellikler, PoS özellikleri, sözlük tabanlı özellikler ve BoW özelliklerini birleştirerek tweet'leri duygu bakımından sınıflandırmışlardır. Sadece BoW özellikleri kullanıldığında ise toplu özelliklere göre daha iyi başarımlar elde etmişlerdir. Aydın ve diğerleri [16], tweet'lere önışlem uyguladıktan sonra öznelik mühendisliği yaparak 10 adet özellik çıkarmışlardır. Elde edilen özellikler çoklu popülasyon tabanlı parçacık sürü optimizasyon yöntemiyle duygu bakımından sınıflandırılmıştır. Salur ve diğerleri [17], Türkiye'deki turizm merkezleri hakkında yapılan Twitter paylaşımlarından elle özellik çıkarımı yapmış olup, duygu bakımından sınıflandırıldıktan sonra her turizm merkezi hakkında oluşturulan duygu sonuçları mobil uygulama yardımıyla kullanıcılara sunulmuşlardır. Ahuja ve diğerleri [18], TF-IDF ve n-gram özellik çıkarma yöntemlerinin Twitter veri kümeleri üzerindeki etkisini araştırmışlardır. Yapılan sınıflandırmada TF-IDF yöntemiyle elde edilen özelliklerin n-gram özelliklerine göre daha iyi sonuç verdiği gözlemlenmiştir. Saad ve Yang [19], TF-IDF özellikleri üzerine bir dengeleme ve puanlama modeli oluşturularak, MÖ yöntemleriyle tweet'leri sınıflandırmışlardır. Karçioğlu ve Aydın [20], İngilizce ve Türkçe Twitter veri kümelerinden BoW ve Word2Vec yöntemleriyle çıkartılan özellikleri kullanarak tweet'leri MÖ yöntemleriyle sınıflandırmışlar. Çalışmalarında kelime köklerinin alınmasının duygu sınıflandırma başarımını üzerindeki etkisi araştırılmıştır. Sonuç olarak kök alma işlemi, hem Türkçe hem de İngilizce veri kümeleri için önerilmemiştir.

Günümüzde duygu analizi çalışmalarında MÖ yöntemlerine göre DÖ yöntemleri daha çok ilgi görmeye başlamıştır. DÖ tabanlı yaklaşımlarda metin formatındaki veriler Word2Vec [21], Glove [22], FastText [23] gibi yöntemlerle anlamsal ilişkiler barındıran sayısal vektörlerle temsil edildikten sonra DÖ algoritmalarına giriş olarak verilmektedir. DÖ yöntemleri olan Evrişimsel Sinir Ağı (Convolutional Neural Network) (CNN) ve Yinelemeli Sinir Ağı (Recurrent Neural Network) (RNN) algoritmalarının klasik MÖ algoritmalarına göre temel farklılıklarından biri, metin temsil yöntemleriyle elde edilen temsiller üzerinden özellik çıkarımlarını içsel yapılarıyla otomatikman yapmalarıdır. Salur ve Aydın [24], İngilizce Twitter paylaşımlarını hem MÖ yöntemleriyle hem de CNN yöntemiyle olumlu veya olumsuz olarak sınıflandırmışlardır. Çalışmanın sonucunda CNN'nin daha iyi başarımlar sağladığı görülmüştür. Meng ve diğerleri [25], LSTM (Long Short Term Memory) ve CNN yöntemlerinin birlikte kullanımına dayanan yeni bir dikkat mekanizması yardımıyla duygu analizi gerçekleştirmişlerdir. Kamiş ve Goularas [26], CNN ve LSTM ağlarına dayanan farklı DÖ yöntemleri kombinasyonlarını Twitter verileri üzerinde test etmişlerdir. Yapılan deneylerin sonucunda CNN ve LSTM ağlarının birlikte kullanımının tek başına kullanımlarına göre daha iyi performans gösterdiği gözlemlenmiştir. Salur ve Aydın [27], Türkçe Twitter verilerinin sayısal formata dönüştürülmesinde, iki farklı veri temsil yöntemi (FastText ve karakter temsilleri) ile iki farklı DÖ yönteminin (BiLSTM ve CNN) birlikte kullanımının sınıflandırma başarımını attırdığını göstermişlerdir. Muslim ve diğerleri [28], Twitter paylaşımlarını duygu bakımından hem LSTM hem de Çok Terimli Lojistik Regresyon (Multinomial Logistic Regression) yöntemiyle sınıflandırmıştır. LSTM'in daha iyi sınıflandırma performansı sağladığı görülmüştür. Kaladevi ve Thyagarajah'in önerdikleri CNN-LSTM tabanlı DÖ modeli [29], sıralı bir yapıda işlenen tek bir CNN katmanı ile iki LSTM

katmanından oluşmaktadır. CNN katmanından elde edilen öznitelikler LSTM ağına giriş olarak verilmiştir. Önerilen model klasik özellik çıkarma yöntemlerine göre daha iyi performans göstermiştir. Karakuş ve diğerleri, Türkçe film yorumlarının duygu bakımından sınıflandırılmasında DÖ yöntemlerinin (LSTM, BiLSTM ve CNN) performanslarını karşılaştırmışlardır [30]. Film yorumlarının sınıflandırılmasında en yüksek başarıyı CNN ve LSTM'i birlikte kullanan modelle elde etmişlerdir. Santur, duygu sınıflandırmada dengesiz veri kümelerinin sınıflandırma performansı üzerindeki olumsuz etkisini azaltmak için pozitif sınıfına ait örneklerde en sık geçen 20 kelimeyi çıkartma ilkesine dayanan DÖ modeli önermiştir [31]. Bu model Türkçe pozitif ve negatif kullanıcı yorumlarından oluşan dengesiz bir veri kümesi üzerinde uygulanmış ve sınıflandırma doğruluk değeri artırılmıştır.

Literatürdeki DÖ ve MÖ çalışmaları göz önüne alındığında; veri ön işleme aşamasında tweet'lerin içerisindeki bazı içerikler (URL, retweet, hashtag, rakamlar, noktalama işaretleri vb.) silinmektedir. DÖ modellerine giriş olarak verilen kelime temsil vektörlerinin (Word2Vec, FastText veya Glove ile elde edilen vektörler) oluşturulması aşamasından önce ön işleme aşamasında bahsedilen gürültülü içerikler temizlenmektedir. Diğer taraftan bu içerikler de aslında kullanıcının paylaşımının birer parçası olması gereği, bir noktada kullanıcının duygusuna katkı sağlaması beklenmektedir. Örneğin; sosyal medya uygulamalarında kullanıcılar yaptıkları bir paylaşımında vurgulamak istedikleri kelimeleri genellikle büyük harfle yazmaktadırlar. Metin ön işleme aşamasında genellikle tüm karakterler küçük harflere dönüştürüldüğünden bu anlamsal içerik kaybedilmektedir. Bu çalışmada bu anlamsal içeriğin de tweet'lerin duygusuna katkı sağlayacağı düşünüldüğünden önerilen modelle bu durumun etkisinin ortadan kaldırılması hedeflenmiştir. Önerilen modelde, DÖ için kelime temsilleri elde edildikten sonra, veri kümesinden özellik mühendisliği yardımıyla özellikler çıkartılmıştır. DÖ ile elde edilen özellikler ile elde edilen özelliklerin birleştirilmesiyle daha iyi sınıflandırma başarıyı elde edilmiştir. Bu çalışmanın katkıları aşağıda maddeler halinde verilmiştir:

- Veri ön işleme aşamasında silinen gürültülü verilerden öznitelikler çıkarılarak gürültülü verilerin duygu analizi açısından sınıflandırma performansına katkıda bulunmasını sağlamıştır.
- Türkçe Twitter veri kümeleri için yeni bir sınıflandırma yaklaşımı sunulmuştur.
- Türkçe tweet'lerden manuel olarak çıkarılabilen özellik kümeleri sunulmuştur.
- Önerilen modelin performansını doğrulamak için birçok deney yapılmış ve sonuçlar literatürdeki yöntemlerle karşılaştırılmıştır.

Bu çalışmanın ikinci bölümünde kullanılan yöntemlere dair temel bilgiler verilmiştir. Üçüncü bölümde yapılan çalışmanın metodolojisi sunulmuştur. Dördüncü bölümde ise önerilen modelin performansını gösteren deneysel çalışmalar ve sonuçları paylaşılmıştır. Çalışmanın son bölümünde ise çalışmada elde edilen sonuçlar özet olarak verilmiştir.

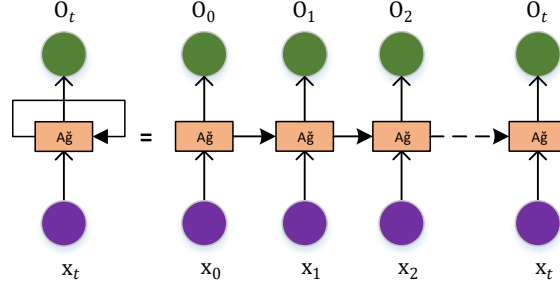
## 2. Temel Bilgiler

### 2.1. Yinelemeli sinir ağları

Yinelemeli sinir ağları klasik ileri beslemeli yapay sinir ağlarına göre dinamik bir yapıya sahiptirler. Bu ağlar sınıflandırmanın her bir adımında içsel bir durum bilgisi taşırlar. Bu ağdaki durum bilgisi, yüksek ve alt katman düğümler arasındaki dairesel bağlantılar ve kendi kendine geribildirim bağlantılar sayesinde oluşmaktadır. Bu bağlantılar, verinin önceki işlem adımından geçerli işlem adımlarına yayılmasını sağlamaktadır. Bunun sonucunda ise RNN'ler zaman serisi olaylarının bir hafızasını oluşturmaktadır [32]. Yani RNN için, herhangi bir düğümdeki  $(t + 1)$  anındaki çıkış değeri,  $t$  anındaki giriş değerinden doğrudan etkilenmektedir. RNN'nin zaman eksenini boyunca açılmış hali Şekil 1'de verilmiştir. RNN'ler zaman serileri verileri ve ardışıl veri tiplerinde etkili sınıflandırma performansı gösterdiğinden doğal dil işleme problemlerinde sıklıkla tercih edilmektedir [33].

RNN'lerdeki gradyan yok olma (vanishing gradient problem) veya gradyan uçması (exploding gradient problem) problemleri için farklı ağ tipleri geliştirilmiştir. Bu ağ tiplerinin başında ve en çok tercih edilenler arasında Uzun Kısa Süreli Bellek (Long Short Term Memory) (LSTM) ve Kapılı Yinelemeli Hücre (Gated Recurrent Unit) (GRU) gelmektedir. LSTM giriş olarak verilen veriyi baştan sona olacak şekilde ağ üzerinde geçirmektedir. Bu şekilde zaman eksenini boyunca tek yönlü ilişkiler ve bağılıklar ortaya çıkarılmaktadır. Metin türündeki veriler için bu ilişkiler her zaman tek yönlü değildir. Bunun için de çift yönlü LSTM (BiLSTM) ağ tipi geliştirilmiştir. BiLSTM hem baştan sona, hem de sondan başa kelimeler arasındaki ilişkileri ortaya çıkarmaktadır. GRU ise, LSTM'in yapısındaki bir kapının kaldırılması sonucu oluşan ağ tipidir. LSTM mimarisinde giriş (input), çıkış (output) ve unutma (forget) kapıları varken GRU mimarisinde sıfırlama (reset) ve güncelleme (update) kapıları bulunmaktadır. Zaman serisi boyunca birbiriyle ilişkili veri kümelerinde gösterdiği yüksek

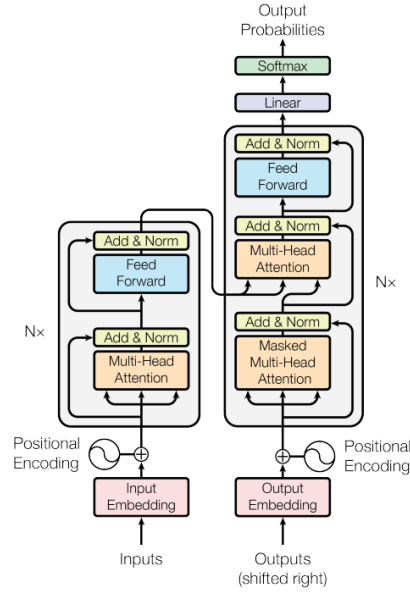
performanslarından dolayı bu çalışma kapsamında da tweet'lerin sınıflandırılmasında DÖ yöntemlerinden LSTM, BiLSTM ve GRU ağları tercih edilmiştir [34].



Şekil 1. RNN'nin zaman eksenini boyunca açılmış halinin temsili görüntüsü.

## 2.1. BERT dil modeli

Günümüzde birçok dil modeli kodlayıcı (encoder) - kod çözücü (decoder) mimarisi üzerine inşa edilmiştir. Kodlayıcı giriş olarak sembollerden oluşan (örneğin bir ziyaretçi yorumu) girdiyi, çıktı olarak başka bir girdiye eşlemektedir. Kod çözücü ise kodlayıcıdan aldığı semboller dizisi için her seferinde bir elemanı olan bir dizi üretmektedir. Kodlayıcı-kod çözücü mimarisi diller arasında çeviri, metinlerin sınıflandırılması ve duygu analizinde sıklıkla kullanılmaktadır. RNN modelleri kodlayıcı-kod çözücü mimarisini kullanmaktadır. RNN modelleri uzun bir cümlenin kelimeleri arasındaki ilişkiyi korumada yetersiz kalırken, dikkat mekanizmaları (attention mechanism) bu problemi büyük oranda ortadan kaldırmaktadır. Dikkat mekanizması kodlayıcıdaki bütün bilgiyi sabit uzunlukta bir vektörler temsil ettiğinden kelimeler arasındaki anlamsal ilişkinin korunmasını sağlamaktadır. Dikkat mekanizmasından sonra ortaya çıkan Transformer mimarisi ise dikkat mekanizmasının yanı sıra paralel hesaplama yapma yeteneği ve çok başlı dikkat mekanizması (Multi-Headed Attention) yapısıyla metin sınıflandırmada büyük performans göstermiştir [35]. Transformer mimarisi şema görünümü Şekil 2'de verilmiştir. Transformer modelinde birden fazla kodlayıcı ve kod çözücünün ardışık olarak konumlandırılmıştır. Bu mimarideki çok başlı dikkat mekanizması bir kelimenin metin içerisinde farklı yerlerdeki anlamlarına farklı bakış açılarıyla bakılmasını sağlamaktadır. BERT (Bidirectional Encoder Representations from Transformers) modeli ise veriyi iki yönlü işleyen bir transformer modelidir. 2018 yılında Google tarafından büyük bir işlem gücüyle ve veri kümesiyle eğitilmiştir [36].



Şekil 2. Transformer ağ mimarisi [35].

### 3. Metodoloji

Bu bölümde çalışmada kullanılan veri kümesi, veri kümesi üzerinde gerçekleştirilen önışlemler, veri kümesinin sayısal vektörlerle temsil edilmesi için kullanılan veri temsil yöntemi ile yapılan özellik mühendisliğinden bahsedilmektedir. Bu bölümün son kısmında ise bu çalışmada sunulan yeni duygu sınıflandırma modelinin yapısı ve çalışma yöntemi anlatılmıştır.

#### 3.1. Veri kümesi

Bu çalışmada kullanılan veri kümesi Twitter ortamında paylaşılan genel kullanıcı içeriklerinden oluşmaktadır. Bu veri kümesi, Çoban ve diğerleri [8] tarafından tweet'in içerdiği emojilere göre pozitif veya negatif olarak etiketlenmiştir. Veri kümesi 6.887 adet pozitif ve 7.890 adet negatif olmak üzere toplamda 14.777 adet tweet'ten oluşmaktadır. Veri kümesi sınıfsal dağılım açısından eşit sayıda örnek içermemektedir.

#### 3.2. Veri önışleme

İnternet ortamında oluşturulan metin tabanlı içerikler her zaman yazım dilinin kurallarına riayet edilerek yazılmamaktadır. Kullanıcının kişisel deneyimi doğrultusunda metin içerisinde anlamlı veya anlamsız birçok içerik olabilmektedir. Twitter tabanlı duygu sınıflandırma çalışmaları için bu içerikler, URL bilgisi, hashtag bilgisi, rakam bilgisi, kelimeler arası boşluk, noktalama işaretleri vb. kullanımlar olarak sıralanabilir. Twitter veri kümeleri için bu içeriklerin temizlenmesinin sınıflandırma başarımını arttırdığı görülmüştür [37]. Ayrıca DÖ modelleri içinde de temel önışlemlerin sınıflandırma başarımını arttırdığı belirtilmiştir [38]. Bu doğrultuda veri kümesi DÖ algoritmalarına giriş olarak verilmenden önce bir dizi önışlemden geçirilmiştir. Veri kümesinin önışlem adımlarından önceki ve sonraki kelime bulutları Şekil 3'te verilmiştir.



Şekil 3. Veri kümesinin kelime bulutu gösterimi: a-) Veri kümesinin önışlem öncesi kelime bulutu gösterimi b-) Veri kümesinin önışlem sonrası kelime bulutu gösterimi.

Veri kümesi üzerinde önışlemler gerçekleştirmedeki amaç, veri kümesinden daha anlamlı kelime temsil vektörleri elde etmektir. Ayrıca Türkçe tweet'lerin sınıflandırılmasında kelime gövdeleme işleminin sınıflandırma performansını arttırmadığı tespit edildiğinden [20], bu çalışmada da kelimelerde gövdeleme işlemi yapılmamıştır. Bu çalışma kapsamında DÖ yöntemlerine giriş olarak verilecek vektörlerin elde edilmesinden önce veri kümesi üzerinde gerçekleştirilen önışlemler aşağıdaki gibi sıralanabilir:

- **URL Bilgisi Silme:** Tweet'ler içerisinde bulunan URL bilgisi, paylaşımın duygu fikrine katkı sağlamadığından tüm link bilgileri yalnızca "URL" kelimesiyle değiştirilmiştir.
- **Boşluk Silme:** Tweet'ler içerisinde başta ve sonda bulunan fazla boşluklar silinmiştir. Ayrıca kelimeler arasındaki boşluk sayısı bire düşürülmüştür.
- **Noktalama İşareti Silme:** Tweet'ler içerisinde geçen tüm noktalama işaretleri silinmiştir.
- **Kullanıcı Adı Bilgisi Silme:** Tweet'ler içerisinde geçen kullanıcı adı bilgisi silinmiştir.
- **Anlamsız Kelime Silme:** Tweet'ler içerisinde geçen kelimelerdeki yazım yanlışları Zembek [39] yardımıyla düzeltilmiştir. Kelime görünümünde yazılmış anlamsız karakter dizilimleri tespit edilmiş ve silinmiştir.
- **Rakam Silme:** Tweet'ler içerisinde geçen rakamlar silinmiştir.
- **Tek Karakter Silme:** Tweet'ler içerisinde geçen tek karakterlik içerikler silinmiştir.

- **Kelime Düzeltme:** “MERhabaaaa” gibi karakter tekrarı içeren kelimeler, Zemberek yardımıyla Türkçe dil bilgisi kuralına göre düzeltilmiştir.

### 3.3. Veri temsillerinin oluşturulması

Bu bölümde veri kümesinin sayısal olarak temsil edilmesinde kullanılan teknikler ayrıntılı bir şekilde ele alınmıştır. Bu çalışmada veri kümesi önemli temsil yöntemlerinden hem FastText ile sayısallaştırılmıştır hem de veri kümesinden manuel olarak 12 özellik olarak dahil edilmemiştir.

#### 3.3.1. Elle özellik çıkarımı

FastText yardımıyla veri kümesinden iyi temsil vektörleri elde etmek için veri önışlem aşamasında uygulanan operasyonlarla kullanıcının paylaşımındaki bazı içerikler silinmiştir. Bu içerikler duygu bakımından anlamsal bir değere sahip olmadığı düşünülse de, aslında kullanıcının duygu ve düşüncesinin bir parçası niteliğindedir. Bu bakış açısıyla, veri kümesi önışlemlerden geçirilmeden önce sözdizimsel, anlamsal ve Twitter’a özgü kullanımlardan oluşan üç kategori altında 12 adet elle özellik çıkarımı (EÖÇ) yapılmıştır. Çıkarılan bu özellikler ve açıklamaları Tablo 1’de verilmiştir. Anlamsal özelliklerden pozitif ve negatif emoji sayıları [40]’daki emoji skorlarına göre belirlenmiştir. Ayrıca veri kümesinin etiketlenmesinde [8]’de referans alınan emoji türleri (olumlu- “:-)”, “:””, “=)”, “:D” ve olumsuz- “:-(”, “:(”, “=(”, “;(” ) özellik olarak eklenmemiştir.

**Tablo 1.** Veri kümesinden elle çıkartılan sözdizimsel, anlamsal ve Twitter özgü özellikler ve açıklamaları.

Özellik Türü	Özellik No	Özellik Adı	Özellik Açıklaması
Sözdizimsel	Öz-1	Toplam Karakter Sayısı	Bir tweet içerisindeki toplam karakter sayısını ifade eder.
	Öz-2	Büyük Harf Sayısı	Bir tweet içerisindeki büyük harflerin sayısını ifade eder.
	Öz-3	Küçük Harf Sayısı	Bir tweet içerisindeki küçük harflerin sayısını ifade eder.
	Öz-4	Rakam Sayısı	Bir tweet içerisindeki rakamların sayısını ifade eder.
	Öz-5	Noktalama İşareti Sayısı	Bir tweet içerisindeki noktalama işaretlerinin sayısını ifade eder.
	Öz-6	URL Sayısı	Bir tweet içerisindeki link bilgilerinin sayısını ifade eder.
Twitter	Öz-7	Bahsetme (mention) Sayısı	Bir tweet içerisinden etiketlenen yani bahsedilen kullanıcı sayısını ifade eder.
	Öz-8	Hashtag (#) Sayısı	Bir tweet içerisinden geçen hashtag kullanımlarının (Örneğin, #pazartesi gibi.) sayısını ifade eder.
Anlamsal	Öz-9	Pozitif Emoji Sayısı	Bir tweet içerisinde geçen pozitif emoji sayısını ifade eder.
	Öz-10	Negatif Emoji Sayısı	Bir tweet içerisinde geçen negatif emoji sayısını ifade eder.
	Öz-11	Pozitif Ünlem Sayısı	Bir tweet içerisinde geçen pozitif ünlem sayısını ifade eder. Örneğin; “ohhhh”
	Öz-12	Negatif Ünlem Sayısı	Bir tweet içerisinde geçen negatif emoji sayısını ifade eder. Örneğin; “uffff”

#### 3.3.2. FastText ile kelime temsillerinin oluşturulması

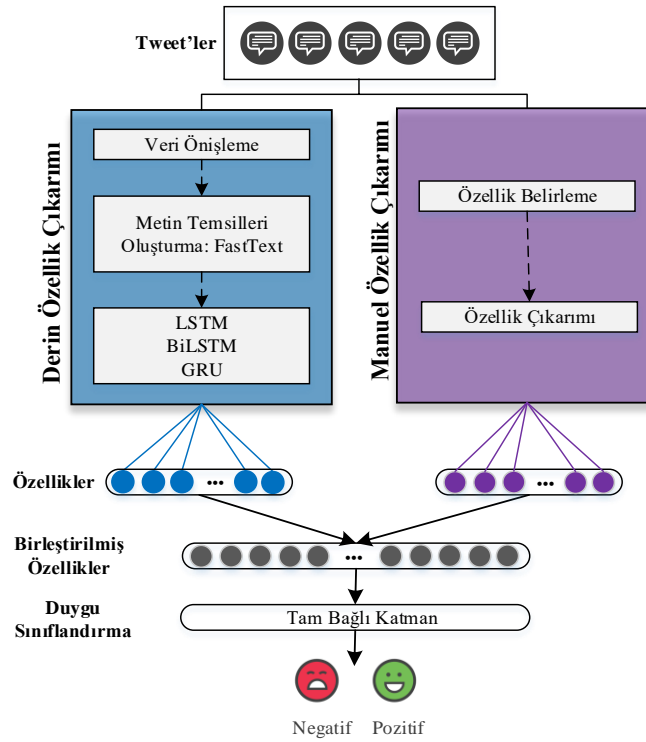
Metin tabanlı duygu sınıflandırma problemlerinin en önemli alt aşamalarından biri metin içerisindeki kelimeler arasındaki anlamsal ilişki ve bağlantıları, kelimeleri sayısal formatta temsil ederken de korumaktır. Bu aşamada kullanılan yöntemin başarısı, doğrudan sınıflandırma başarımı üzerinde olumlu etki oluşturduğundan, bu aşama ayrıca iyi özellik çıkarmanın da göstergesi niteliğindedir. Günümüzde metin tabanlı verilerin anlamsal bilgiye sahip sayısal vektör olarak temsil edilmesinde Word2Vec, Glove, FastText gibi yöntemler sıklıkla tercih edilmektedir.

Glove, giriş olarak aldığı kelimelerden, sayısal vektörleri üreten ve bağlamdaki kelimelerin birlikte bulunma istatistiklerine dayanan gözetimsiz bir öğrenme yöntemidir. Word2Vec ise, giriş olarak kelimeleri alan ve çıkış olarak çok boyutlu vektör üreten bir yapay sinir ağı modelidir. Word2Vec temelde, kelime vektör uzayında benzer kelimeleri birlikte gruplandırmaktadır. Bunu iki farklı yöntemle gerçekleştirmektedir. Bunlardan birincisi vektörü oluşturulacak kelimenin öncesinde ve sonrasında kullanılan kelimelerin vektörlerinden elde etmek yani bağlamdan kelime vektörünü (Continuous Bag of Words) (CBoW) elde etmektir. İkinci yöntemde ise, bağlamda geçen bir kelimeden (Skip-gram) yeni kelime vektörlerini oluşturma şeklindedir. Word2Vec yönteminin en büyük eksikliği, bağlamda olmayan bir kelime için vektör oluşturma aşamasında rastgele sayısal bir vektör oluşturmaktır. Diğer taraftan FastText yöntemi, vektör oluşturma aşamasında n-gram yaklaşımını kullandığından Word2Vec

yönteminin bu eksikliğini gidermektedir. Bu nedenle bu çalışma kapsamında veri kümesinin sayısal temsillerinin oluşturulmasında FastText yöntemi tercih edilmiştir. Veri kümesi önışlemlerden geçirildikten sonra, Türkçe Wikipedia metinleri üzerinde eğitilen FastText<sup>2</sup> modeli yardımıyla 300 boyutlu sayısal vektörler olarak temsil edilmiştir.

### 3.4. Önerilen model

Twitter veri kümeleri üzerinde yapılan duygu sınıflandırma çalışmaları incelendiğinde, iyi veri temsilleri elde etmek için veri normalleştirme önışlemleri önerilmektedir. Bu önışlemlerle veri içerisinde sözdizimsel yapıların büyük bir kısmı silinmektedir. Bu çalışmada önerilen yöntemle, silinen bu içerikleri de duygu bakımından anlamsal değerlere dönüştürmek için EÖÇ yapılmış olup, elde edilen özellikler ile DÖ yöntemlerinden elde edilen özelliklerin birleştirilerek kullanılmasının sınıflandırma başarımına olan katkısı araştırılmıştır.



Şekil 4. Önerilen derin özellik çıkarımı tabanlı modelin yapısı.

Önerilen model, FastText kelime temsil yöntemi ve RNN'lerden elde edilen derin özelliklerle, veri önışlem aşamasından önce veri kümesi üzerinde elle çıkartılan 12 özelliğin birleştirilmesi ilkesine dayanmaktadır. Önerilen modelin şema gösterimi Şekil 4'te verilmiştir. Önerilen model iki temel kısımdan oluşmaktadır. Birinci kısımda, FastText kelime temsil yöntemi ile metin formatındaki veriler sayısal vektörlerle temsil edildikten sonra, LSTM, BiLSTM ve GRU yardımıyla veri kümesinden derin özellikler çıkartılmıştır. İkinci kısımda ise, veri kümesinden önceki bölümde verilen 12 özellik çıkartılmıştır. Her iki kısımdan çıkartılan özellikler birleştirilerek, sınıflandırması amacıyla tam bağlı katmana iletilmiştir. Elle çıkartılan özellikler, RNN modelleri (LSTM, BiLSTM, GRU) yardımıyla çıkarılan derin özellikler ve her iki grup özelliklerin birleşimi olan özelliklerin sınıflandırma başarımı üzerindeki etkisini görmek için üç girişli ve üç çıkışlı DÖ modeli tasarlanmıştır. Modelin geliştirilmesi için TensorFlow<sup>3</sup> DÖ yazılım çatısı ile Keras<sup>4</sup> yazılım çatısının fonksiyonel API'si (Application Programming Interface) kullanılmıştır.

<sup>2</sup> <https://fasttext.cc/docs/en/crawl-vectors.html>

<sup>3</sup> <https://www.tensorflow.org>

<sup>4</sup> <https://keras.io/api>

#### 4. Deneysel Sonuçlar ve Tartışma

Bu bölümde öncelikle önerilen modelin başarımını ölçmek için kullanılan performans değerlendirme metrikleri verilmiştir. Sonrasında bu metrikler kullanılarak önerilen modelin sınıflandırma sonuçları verilmiş ve sınıflandırma sonuçları tartışılmıştır.

##### 4.1. Performans metrikleri

Sınıflandırma problemlerinde, modellerin başarımları genellikle karmaşıklık matrisi (Confusion Matrix) ve bu matris yardımıyla hesaplanan ölçütlerle değerlendirilmektedir. Bu çalışmada da önerilen modelin sınıflandırma performansının değerlendirilmesi ve elde edilen sınıflandırma sonuçlarının önceki çalışmalarla mukayese edilmesi amacıyla karmaşıklık matrisi ve ölçütleri kullanılmıştır. İki sınıflı bir sınıflandırıcı modelinin sınıflandırma sonuçları için karmaşıklık matrisi Tablo 2’de verilmiştir.

**Tablo 2.** İki sınıflı bir sınıflandırma problemi için karmaşıklık matrisi

		Tahmin Edilen Sınıf Etiket	
		Pozitif	Negatif
Gerçek Sınıf Etiketi	Pozitif	DP	YN
	Negatif	YP	DN

Tablo 2’de verilen karmaşıklık matrisindeki DP (Doğru Pozitif), YP (Yanlış Pozitif), YN (Yanlış Negatif) ve DN (Doğru Negatif) kısaltmaları aşağıdaki anlamlara gelmektedir:

- **DP:** Gerçek sınıf etiketinin pozitif ve tahmin edilen sınıf etiketinin pozitif olduğu örneklerin sayısı
- **YP:** Gerçek sınıf etiketinin negatif ve tahmin edilen sınıf etiketinin pozitif olduğu örneklerin sayısı
- **YN:** Gerçek sınıf etiketinin pozitif ve tahmin edilen sınıf etiketinin negatif olduğu örneklerin sayısı
- **DN:** Gerçek sınıf etiketinin negatif ve tahmin edilen sınıf etiketinin negatif olduğu örneklerin sayısı

Tablo 2’de verilen karmaşıklık matrisine göre Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall) ve F1 ölçümü hesaplanmaktadır. Doğruluk değeri Denklem 1’e, Kesinlik Denklem 2’ye, Duyarlılık Denklem 3’e ve F1 değeri ise Denklem 4’e göre hesaplanmaktadır. Kappa değeri ise genellikle dengesiz sınıf dağılımına sahip veri kümeleri için referans alınan bir başarımlar metriğidir. Kappa değeri Denklem 5 ve Denklem 6 yardımıyla hesaplanmaktadır. Denklem 5’te Kappa değeri için gerekli olan beklenen uzlaşma (expected agreement) ( $P_e$ ) değeri hesaplanmaktadır. Ayrıca dengesiz dağılıma sahip veri kümeleri için önem ifade eden bir diğer metrik olan ve ROC (receiver operating characteristic curve) eğrisi ile hesaplanan eğri altında kalan alanı ifade eden AUC (Area Under Curve) değeri de hesaplanmıştır.

$$\text{Doğruluk} = \frac{DP + DN}{DP + DN + YP + YN} \quad (1)$$

$$\text{Kesinlik} = \frac{DP}{DP + YP} \quad (2)$$

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (3)$$

$$F_1 = \frac{2 * \text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4)$$



$$P_e = \frac{(DN + YP) * (DN + YN) + (YN + DP) * (YP + DP)}{(DP + DN + YP + YN) * (DP + DN + YP + YN)} \quad (5)$$

$$Kappa = \frac{Doğruluk - P_e}{1 - P_e} \quad (6)$$

#### 4.2. Sınıflandırma sonuçları

Veri kümesi üzerinde yapılan önceki çalışmalarda 10 kat çaprazlama tekniği kullanılmasına rağmen, bu çalışmada veri kümesinin %90'ı eğitim ve %10'u test amacıyla kullanılmıştır. Her model için 10 deney gerçekleştirilmiş olup, ortalama başarımlar metrikleri kayda geçirilmiştir. Önerilen modelde LSTM, BiLSTM ve GRU için birim sayısı 64, seyreltme (dropout) oranı 0.7, tekrarlı seyreltme (recurrent dropout) oranı 0.7, optimizasyon fonksiyonu 'adam' ve aktivasyon fonksiyonu olarak da sigmoid kullanılmıştır.

Yapılan deneylerde LSTM, BiLSTM ve GRU'nun derin özellik çıkarımındaki başarımlarını test etmek için Model-A (LSTM), Model-B (BiLSTM) ve Model-C (GRU) olmak üzere üç ayrı model oluşturulmuştur. Model-A, Model-B ve Model-C'nin başarımlar metrikleri Tablo 3'te verilmiştir. Her üç modelde de elde edilen özellikler ile derin özelliklerin birleşiminden oluşan modellerin başarımları daha yüksek olduğu görülmektedir. LSTM, BiLSTM ve GRU arasından da en yüksek doğruluk sınıflandırma başarımları % 75,71 olarak BiLSTM ile elde edilmiştir. Sadece derin özellik çıkarımında da BiLSTM (%74,68) GRU (%74,32) ve LSTM (74,33)'e göre daha iyi doğruluk değeriyle performans göstermiştir. Tablo 3'teki sonuçlara göre DÖ'den elde edilen özellikler ile EÖÇ sonucunda elde edilen özelliklerin her üç model için de doğruluk değeri yaklaşık %1 ile %2 arasında arttırdığı görülmüştür. Ancak özellik birleştirmesinden sonra duyarlılık değeri düştüğünden F1 ölçütü değerinin artışı sınırlı kalmıştır.

**Tablo 3.** RNN modellerinin ve BERT modelinin sınıflandırma başarımlar metrikleri.

		Başarımlar Metrikleri					
	Özellik Türü	Doğruluk	Kesinlik	Duyarlılık	F1	Kappa	ROC-AUC
Model-A	EÖÇ	0.654570	0.650137	0.577497	0.609345	0.302776	0.650678
	LSTM	0.743331	0.769972	0.656840	0.704357	0.481097	0.739023
	EÖÇ + LSTM	0.753486	0.797972	0.639818	0.707688	0.499762	0.747397
Model-B	EÖÇ	0.661408	0.648530	0.584145	0.611975	0.313154	0.655471
	BiLSTM	0.746851	0.752564	0.677385	0.710698	0.486354	0.741238
	EÖÇ + BiLSTM	<b>0.757142</b>	0.784286	<b>0.653795</b>	<b>0.712064</b>	<b>0.504839</b>	<b>0.749311</b>
Model-C	EÖÇ	0.659174	0.644400	0.595417	0.616405	0.310701	0.654636
	GRU	0.743263	0.780974	0.626742	0.692594	0.476945	0.735500
	EÖÇ + GRU	0.748679	<b>0.800400</b>	0.613487	0.692753	0.486523	0.739566
Model-BERT	BERT	0.753482	0.753385	0.753482	0.753428	0.535731	0.753518

Bu çalışmada kullanılan veri kümesi son teknoloji yöntemlerden olan ilgi (attention) mekanizması tabanlı transformers modeli olan BERT modeliyle de sınıflandırılmıştır. BERT tabanlı sınıflandırma için Model-BERT oluşturulmuştur. BERT modeli için önceden eğitilmiş kelime vektörleri kullanılmıştır. Sınıflandırma için kullanılan önceden eğitilmiş BERT Base Türkçe modeli Hugging Face<sup>5</sup> kütüphanesinden alınmıştır. BERT Base modeli 12 kodlayıcı katmanına sahiptir ve eğitim aşamasında 110 milyon parametre ayarlaması yapılmaktadır. Model-BERT'in son katmanında ince ayar (fine-tuning) yapılarak, önceden eğitilmiş modelin bu çalışmada kullanılan veri kümesine göre parametre ayarlaması yapılmıştır. BERT modellerinin eğitimi büyük bir GPU hesaplama gerektirdiğinden veri kümelerinin sınıflandırması aşamasında Google Colab<sup>6</sup> kullanılmıştır. Tablo 3'te

<sup>5</sup> <https://huggingface.co/dbmdz>

<sup>6</sup> <https://colab.research.google.com/notebooks/intro.ipynb#recent=true>

verilen BERT modeli sınıflandırma sonuçlarına bakıldığında doğruluk değeri bakımından Model-B (%75,71) ile yakın sonuçlar elde edilmiştir. F1 skoru açısından BERT modelinin skoru diğer üç modele göre daha yüksek elde edilmiştir. BERT modeli birçok ilgi mekanizması barındırdığından genel olarak LSTM, BiLSTM ve GRU'ya göre yüksek sınıflandırma performansı sağlamıştır.

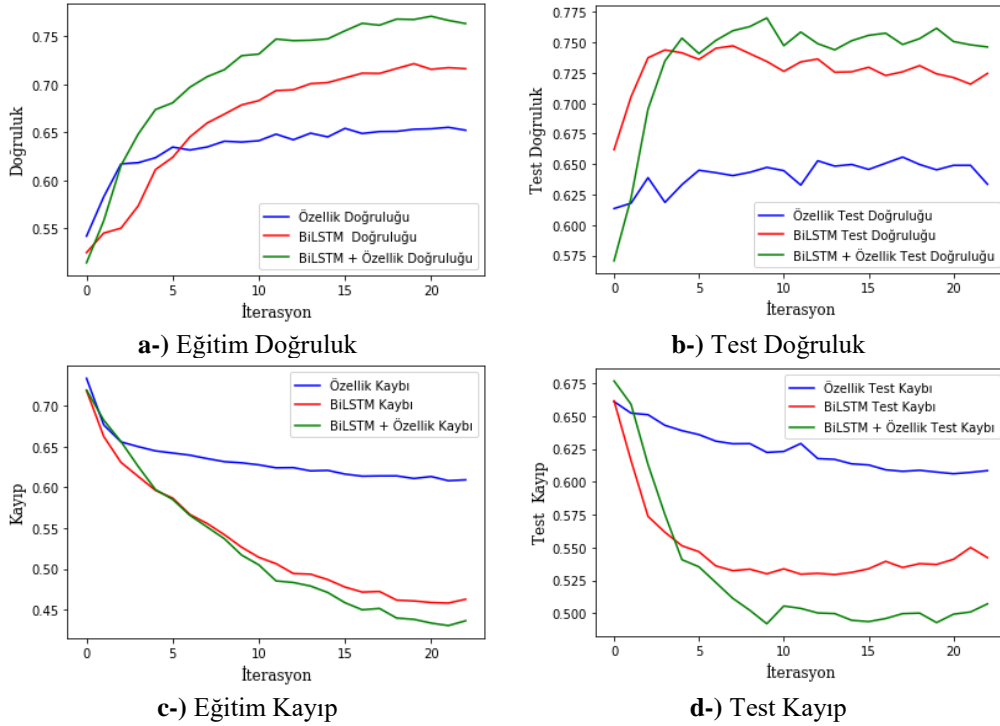
Bu çalışmada kullanılan veri kümesi pozitif ve negatif örnek sayısı bakımından dengesiz bir dağılıma sahiptir. Dengesiz veri kümelerinde başarımların metriği olarak genellikle Kappa değeri referans alınmaktadır. Bu çalışmada da her model için Kappa değeri hesaplanmıştır. Kappa değerine bakıldığında BiLSTM tabanlı olan Model-B'nin daha iyi performans gösterdiği görülmektedir. ROC-AUC ve F1 metriğine bakıldığında ise benzer şekilde de Model-B'nin başarımlarını Model-A ve Model-C'ye göre daha iyi olduğu görülmektedir. Öte yandan, her üç modelde de özellik birleştirme sonrasında duyarlılık değerinin yalnız BiLSTM, GRU, LSTM'e göre düştüğü görülmektedir.

Önerilen modelin performansı aynı veri kümesi üzerinde daha önce gerçekleştirilen çalışmalarla kıyaslandığında, daha yüksek sınıflandırma performansı elde edilerek, önerilen modelin etkinliği gösterilmiştir. Veri kümesi üzerinde gerçekleştirilen önceki çalışmaların özellik çıkarım yöntemleri, sınıflandırma yöntemleri ve sınıflandırma başarımları Tablo 4'te verilmiştir. Sınıflandırma başarımları incelendiğinde önerilen modelin sınıflandırma başarımları diğer çalışmalardan daha iyi olduğu görülmektedir.

**Tablo 4.** Twitter veri kümesi üzerinde önerilen model ile önceki çalışmaların başarımlarını sonuçları.

Referans	Özellik Çıkarım Yöntemi	Sınıflandırma Yöntemi	Doğruluk (%)
[20]	BoW özellikleri	Destek Vektör Makineleri	65,62
[8]	N-gram özellikleri	MNB	66,06
<b>Önerilen Model (Model-B)</b>	<b>EÖÇ + Derin özellikler (FastText + BiLSTM)</b>	<b>Tam Bağlı Yapay Sinir Ağı</b>	<b>75,71</b>

Özelliklerin birleştirilmesinden sonra en yüksek başarımların elde edildiği Model-B'nin her iterasyondaki performansı izlenmiştir. Şekil 5'te Model-B'nin bir kez çalıştırılması sırasındaki elde çıkarılan özellikler, derin özellikler ve iki özelliğin birleşimi olan önerilen modelin doğruluk ve kayıp değerleri görülmektedir.



**Şekil 5.** Önerilen modelin eğitim ve test aşamalarındaki doğruluk ve kayıp değerleri eğrileri.

Şekil 5'teki grafiklerden de görüldüğü üzere elle çıkartılan özellikler (mavi eğri) ile BiLSTM (kırmızı eğri) ile elle çıkartılan özelliklerin birleşimi (yeşil eğri) daha iyi sınıflandırma performansı göstermiştir. Her üç model için de 50 iterasyona kadar eğitim yapılması hedeflenmiş olmasına rağmen modeller genellikle eğitimlerini 20 ile 30 iterasyon arasında tamamlamışlardır.

Önerilen yöntemin duygu sınıflandırma açısından avantaj ve dezavantajları bulunmaktadır. Önerilen modelin avantajları arasında gürültü olarak nitelendirilen içeriklerin duygu sınıflandırması için tekrardan kullanılması ve sınıflandırma başarımına katkısı gelmektedir. Bir başka avantaj ise önerilen model Türkçe haricindeki birçok dil için uygulanabilir bir yapıda olmasıdır. Diğer taraftan her dil için manuel özelliklerin analiz edilmesi ve belirlenmesi önerilen modelin zayıf yönleri arasındadır. Ayrıca manuel özelliklerin veri kümesinden elde edilmesi süreci ek olarak çaba gerektirmektedir.

## 5. Sonuç

Twitter gibi anlık paylaşım olanağı sunan sosyal medya uygulamaları üzerinden kullanıcıların duygu ve düşüncelerini belirlemek, birçok alan için önemli fırsatlar sunması açısından önemlidir. Gelişen teknolojik imkânlar ve yöntemler sayesinde günümüzde bu kullanıcı içerikleri üzerinde duygu analizi çalışmaları gerçekleştirme aşamasında DÖ yöntemleri sıklıkla tercih edilmektedir. DÖ modelleri için kelime temsil yöntemleri yardımıyla giriş verisi oluşturulurken, iyi kelime temsil vektörleri elde etmek amacıyla veri ön işlemleri gerçekleştirilmektedir. Bu ön işlemlerde kullanıcının paylaşımında bulunan bazı içerikler silinmektedir. Bu çalışmada önerilen modelle, silinen bu içeriklerden EÖÇ yapılarak, bu özellikler ile derin model yardımıyla çıkartılan özelliklerin birleştirilmesinin daha iyi sınıflandırma performansı sağladığı gösterilmiştir. Veri kümesi üzerinde üç farklı RNN modeli (Model-A, Model-B, Model-C) çalıştırılmıştır. Bu modellerden en yüksek doğruluk değeri BiLSTM tabanlı olan Model-B ile %75,71 olarak elde edilmiştir. DÖ'den elde edilen özellikler ile EÖÇ sonucunda elde edilen özelliklerin her üç model için de doğruluk değeri yaklaşık %1 ile %2 arasında artırdığı görülmüştür.

Önerilen model ile veri kümesi üzerinde yapılan önceki çalışmaların sınıflandırma performansları kıyaslandığında önerilen modelin daha iyi performans sağladığı görülmektedir. Önceki çalışmalardan elde edilen en yüksek sınıflandırma doğruluk değerine göre %10 daha iyi başarımlar elde edilmiştir. Gelecek çalışmalarda önerilen modelin farklı veri temsil yöntemleri üzerindeki etkisi araştırılması planlanmaktadır.

## Teşekkür

Bu çalışma Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TUBITAK) tarafından 119E655 numaralı proje ile desteklenmiştir.

## Kaynaklar

- [1] Yue L, Chen W, Li X, Zuo W, Yin M. A survey of sentiment analysis in social media. *Knowledge and Information Systems*, 2019; 60(2): 617–663.
- [2] Singh NK, Tomar DS, Sangaiah AK. Sentiment analysis: a review and comparative analysis over social media. *J. Ambient Intell. Hum. Comput*, 2020; 11(1): 97–117.
- [3] Zhao J, Cao X. Combining semantic and prior polarity for boosting twitter sentiment analysis. 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), December 19-21 2015; Chengdu, China. 832–837.
- [4] Wallach HM. Topic Modeling : Beyond Bag-of-Words. 23rd International Conference on Machine learning, June 25-29, 2006; Pennsylvania, USA. 977–984.
- [5] Ramos J. Using Tf-Idf to Determine Word Relevance in Document Queries. *Proceedings of the first instructional conference on machine learning*, December 3-8, 2003; Piscataway, NJ USA. 29–48.
- [6] Gimpel K, Schneider N, O'Connor B, Das D, Mills D, Eisenstein J, Heilman M, Yogatama D, Flanigan J, Smith NA. Part-of-speech tagging for twitter: Annotation, features, and experiments. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 2011; Stroudsburg, PA, USA. 42–47.
- [7] Cavnar WBC, Trenkle JM. N-Gram-Based Text Categorization. *Proceedings of SDAIR-94, 3rd annual symposium on document analysis and information retrieval*, April 11-13 1994; Las Vegas, Nevada.
- [8] Çoban Ö, Özyer B, Özyer GT. Sentiment analysis for Turkish Twitter feeds. 2015 23rd Signal Processing and Communications Applications Conference (SIU), May 16-19 2015; Malatya, Turkey. 2388–2391.
- [9] Siddiqua UA, Ahsan T, Chy AN. Combining a rule-based classifier with ensemble of feature sets and machine learning techniques for sentiment analysis on microblog. 19th International Conference on Computer and Information Technology (ICIT), December 18-20, 2016; Dhaka, Bangladesh. 304–309.
- [10] Chamansingh N, Hosein P. Efficient sentiment classification of Twitter feeds. 3rd International Conference on

- Knowledge Engineering and Applications (ICKEA), September 28-30 2016; Singapore. 78–82.
- [11] Jianqiang Z, Xiaolin G. Comparison research on text pre-processing methods on twitter sentiment analysis. *IEEE Access*, 2017; 5: 2870–2879.
- [12] Riyadh AZ, Alvi N, Talukder KH. Exploring human emotion via Twitter. 20th International Conference of Computer and Information Technology, December 22-24 2017; Dhaka, Bangladesh. 1–5.
- [13] Tsapatsoulis N, Djouvas C. Feature extraction for tweet classification: Do the humans perform better?. 12th International Workshop on Semantic and Social Media Adaptation and Personalization, July 9-10 2017; Bratislava, Slovakia. 53–58.
- [14] Kaur S, Sikka G, Awasthi LK. Sentiment Analysis Approach Based on N-gram and KNN Classifier. 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), December 15-17, 2018; Jalandhar, India. 13–16.
- [15] Permatasari PI, Fauzi MA, Adikara PP, Sari EDL. Twitter Sentiment Analysis of Movie Reviews using Ensemble Features Based Naïve Bayes. 3rd International Conference on Sustainable Information Engineering and Technology, November 10-12 2018; Malang, Indonesia. 92–95.
- [16] Aydın İ, Salur MU, Başkaya F. Duygu Analizi için Çoklu Populasyon Tabanlı Parçacık Sürü Optimizasyonu. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 2018; 11(1): 52–64.
- [17] Salur MU, Aydın I, Alghrsi SA. SmartSenti: A Twitter-Based Sentiment Analysis System for the Smart Tourism in Turkey. 2019 International Conference on Artificial Intelligence and Data Processing Symposium (IDAP), September 21–22 2019; Malatya, Turkey. 1–5.
- [18] Ahuja R, Chug A, Kohli S, Gupta S, Ahuja P. The impact of features extraction on the sentiment analysis. *Procedia Comput. Sci*, 2019; 152: 341–348.
- [19] Saad SE, Yang J. Twitter Sentiment Analysis Based on Ordinal Regression. *IEEE Access* 2019, 7: 163677–163685.
- [20] Karcioğlu AA, Aydın T. Sentiment Analysis of Turkish and English Twitter Feeds Using Word2Vec Model. 27th Signal Processing and Communications Applications Conference (SIU), April 24-26 2019; Sivas, Turkey. 1–4.
- [21] Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed Representations of Words and Phrases and their Compositionality. *Advances in neural information processing systems*, 2013; 3111–3119.
- [22] Pennington J, Socher R, Manning CD. GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, October 25-29 2014; Doha, Qatar. 1532–1543.
- [23] Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 2017; 5: 135–146.
- [24] Salur MU, Aydın I. Derin Öğrenme Tabanlı Duygu Sınıflandırma. 26. Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU), 2-5 Mayıs 2018; İzmir, Türkiye. 1-4.
- [25] Meng W, Wei Y, Liu P, Zhu Z, Yin H. Aspect Based Sentiment Analysis with Feature Enhanced Attention CNN-BiLSTM. *IEEE Access*, 2019; 7: 167240–167249.
- [26] Kamiş S, Goularas D. Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data. 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications, August 26-28 2019; İstanbul, Turkey. 12–17.
- [27] Salur MU, Aydın I. A Novel Hybrid Deep Learning Model for Sentiment Classification. *IEEE Access* 2020; 8: 58080–58093.
- [28] Muslim A, Mutiara AB, Refianti R, Karyati CM, Setiawan G. Comparison of accuracy between long short-term memory-deep learning and multinomial logistic regression-machine learning in sentiment analysis on twitter. *Int. J. Adv. Comput. Sci. Appl.* 2020; 11(2): 747–754.
- [29] Kaladevi P, Thyagarajah K. Integrated CNN- and LSTM-DNN-based sentiment analysis over big social data for opinion mining. *Behav Inform Technol* 2019; 1–9.
- [30] Karakuş BA, Talo M, Hallaç İR, Aydın G. Evaluating deep learning models for sentiment classification. *Concurrency and Computation: Practice and Experience*, 2018; 30(21): e4783.
- [31] Santur Y. Derin Öğrenme ve Aşağı Örnekleme Yaklaşımları Kullanılarak Duygu Sınıflandırma Performansının İyileştirilmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 2020; 32(2): 561–570.
- [32] Staudemeyer RC, Morris ER. Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv Prepr. arXiv1909.09586*. 2019; 1–42.
- [33] Chollet F. *Deep learning with Python*. Shelter Island: Manning Publications, 2017.
- [34] Dang NC, Moreno-García MN, De la Prieta F. Sentiment analysis based on deep learning: A comparative study. *Electronics*, 2020; 9(3): 1-29.
- [35] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez a, Kaiser L, Polosukhin I. Attention is all you need. *Advances in neural information processing systems*, 2017. 5998–6008.
- [36] Devlin J, Chang M, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 2018. 4171–4186.
- [37] Zhao J. Pre-processing boosting twitter sentiment analysis. 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), December 19-21 2015; Chengdu, China; 748–753.
- [38] Salur MU, Aydın I. The Impact of Preprocessing on Classification Performance in Convolutional Neural Networks for Turkish Text. 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), September 28-30

- 2018; Malatya, Turkey. 1-4.
- [39] Akın AA, Akın MD. Zemberek, An Open Source Nlp Framework for Turkic Languages, *Structure* 2007; 10: 1–5.
- [40] Novak PK, Smailović J, Sluban B, Mozetič I. Sentiment of emojis. *PLoS One*, 2015; 10(12): 1–22.