Sosyal Medya Duygu Analizinde Word2Vec ve Doc2Vec Yöntemlerini, Logistic Regresyon ve LSTM Modelleriyle Değerlendirme

Evaluation of Word2Vec and Doc2Vec Methods in Social Media Sentiment Analysis with Logistic Regression and LSTM Models

*Özetçe*— Duygu analizi, bireylerin bir konu, hakkındaki görüşlerini ifade ettiği metinleri, doğal dil işleme ve yapay zeka aracılığıyla analiz ederek anlamlandırma çalışmalarını içerir. Bu alanda, özellikle sosyal medya ve diğer fikir paylaşım platformlarının kullanımının artmasıyla bu alanda üretilen veri miktarı zeta bayt seviyesine ulaşmış olması duygu analizi ve fikir madenciliği çalışmalarına olanak tanımıştır. Bu çalışmada, 2021 yılına ait Twitter’dan alınan etiketlenmiş tweet verileri kullanılarak, Word2Vec ve Doc2Vec vektörleşme yöntemleri üzerinden duygu analizi yapılmıştır. Ayrıca, lojistik regresyon ve bir derin öğrenme modeli olan LSTM'in her iki vektörleştirmeye dayalı duygu analizi model başarıları değerlendirilmiştir. Sosyal medya verilerinin sınıflandırılmasında eğitim algoritmalarının yanında vektörleşme yöntemlerinin kullanımı, sınıflandırmanın doğruluk oranının artışında etkili bir yöntem olarak öne çıkmıştır. Çalışmada, Doc2Vec ve lojistik regresyonun birlikte kullanımı, diğer deneylere oranla %26.15’lik bir sınıflandırma başarım oranı elde edilmiştir.

Anahtar Kelimeler—Makine Öğrenmesi, Duygu Analizi, Deep Learning, Doc2Vec, Word2Vec

*Abstract*—Sentiment analysis involves the analysis and interpretation of texts in which individuals express their opinions on a topic through natural language processing and artificial intelligence. In this field, especially with the increasing use of social media and other opinion sharing platforms, the amount of data produced in this field has reached the zeta byte level, which has enabled sentiment analysis and opinion mining studies. In this study, using tagged tweet data from Twitter for 2021, sentiment analysis was performed using Word2Vec and Doc2Vec vectorization methods. We also evaluate the performance of logistic regression and LSTM which is a deep learning model, in sentiment analysis based on both vectorization methods. The use of vectorization methods in addition to training algorithms in the classification of social media data has come to the fore as an effective method to increase the accuracy of classification. In the study, the combination of Doc2Vec and logistic regression yielded a classification success rate of 26.15% compared to other experiments.

Keywords—Machine Learning, Sentiment Analysis, Deep Learning, Doc2Vec, Word2Vec

# GİRİŞ

Son yıllarda, internet erişiminin yaygınlaşmasıyla birlikte internet, hayatımızın ayrılmaz bir unsuru olmuştur. Özellikle mobil cihazlara yönelik sosyal medya uygulamalarının yaygınlaşmasıyla, günlük yaşamımızın her anı sanal ortama taşınabilmektedir. İnsanlar, herhangi bir ürün veya hizmetle ilgili düşüncelerini hızlı ve kolay bir şekilde paylaşmalarına olanak tanıyan sosyal medya platformları, günümüzde oldukça popülerdir. İnternet kullanımının artmasıyla birlikte, sosyal platformlarda paylaşılan içeriklerin sayısında da artış gözlemlenmektedir[1].

Sosyal medya, kullanıcıların duygu ve düşüncelerini paylaştığı geniş bir platform olarak günümüz dijital dünyasında önemli bir yer tutmaktadır. Sosyal platformlarda paylaşılan metinler, duygu analizi, bilgi çıkarımı gibi araştırma alanları için geniş bir veri kaynağı sunmaktadır [2]. Bu platformlarda üretilen büyük veri miktarı, duygu analizi gibi metin madenciliği yöntemlerinin gelişimi için zengin bir kaynak sağlamaktadır. Duygu analizi, kullanıcıların metinlerindeki duygusal tonları belirleyerek işletmelere, pazarlamacılara ve araştırmacılara değerli bilgiler sunmaktadır[3].

Makine öğrenmesi ve derin öğrenme algoritmaları, sosyal medya verilerini sınıflamak için etkili araçlar olarak öne çıkmaktadır [4]. Bu çalışma, 2021 yılına ait Twitter’dan alınan verileri kullanarak Word2Vec ve Doc2Vec vektörleme yöntemlerinin duygu analizi modelinde etkinliğini, Lojistik Regresyon (LR) ve bir derin öğrenme modeli olan Long Short-Term Memory (LSTM) algoritmasının duygu analizindeki performansını karşılaştırmayı amaçlamaktadır.

Çalışmada, etiketlenmiş tweet verileri kullanılarak Word2Vec ve Doc2Vec yöntemleriyle oluşturulan vektörler, LR modeli ile eğitilmiş ve sonuçlar değerlendirilmiştir. Ayrıca, bu vektörleme yöntemleri kullanılarak oluşturulan derin öğrenme modeli de incelenmiştir. Bu karşılaştırmalar, sosyal medya verilerinin analizinde en uygun yöntem ve algoritmaları belirlemeye yönelik önemli bulgular sunmaktadır.

# GENEL BİLGİLER VE LİTERATÜR TARAMASI

## DUYGU ANALİZİ

Duygu analizi, insanların internet platformlarında paylaştıkları görüş ve duyguların, analiz edilmesini sağlayan bir araştırma alanıdır [5]. Günümüzde, internetin yaygınlaşmasıyla birlikte insanlar forumlarda, bloglarda, sosyal medya ve içerik paylaşım servislerinde düşüncelerini açıkça ifade etmektedirler. Bu durum, duygu analizini doğal dil işleme, istatistik ve programlama gibi çeşitli teknikler ile gerçekleştiren güncel bir araştırma alanı haline getirmiştir.

Duygu analizinin esas amacı, bireylerin durumları hakkındaki düşüncelerini anlamak ve açığa çıkarmaktır. Görüşler ve duygular, karar verme sürecini etkileyen unsurlar olup, duygu analizi yöntemi ise teknoloji ve yeniliklere ilişkin olumlu veya olumsuz yorumları araştırır. Özellikle günümüzde bu konular ön plandadır ve çeşitli çalışmaların temel amaçlarını oluşturur [5].

Rekabet halindeki firmalar, geleneksel pazarlama yöntemlerinin yetersiz olduğunu fark etmişler ve sosyal medyayı gündemlerine almışlardır. Sosyal medya, birçok sektörün yönelimini belirleyen değerli bir veri kaynağıdır. Sosyal medyada paylaşılan düşünceler, ürün ve hizmetlerin başarısını göstermekte ve dolaylı olarak da ekonomiyi etkilemektedir [6].

Twitter gibi sosyal medya platformları, kullanıcıların çeşitli konularla ilgili düşüncelerini paylaşabildikleri popüler platformlardan biridir. Bu platformlardaki mesajlar, çeşitli konular hakkındaki kullanıcı görüşleri ve hislerini barındırır. Bu nedenle, bu çalışmada Twitter gibi bir sosyal medya platformundan elde edilen mesajlar, duygu analizi sınıflandırması amacıyla kullanılmaktadır [7].

Duygu analizi, sentiment analizi adıyla da bilinir ve bir varlığa yönelik insanların tutumlarını, düşüncelerini ve duygularını bilgisayar bilimleri yardımıyla belirlemeyi amaçlar. Bu analiz, üç farklı sonuç elde etmeyi amaçlar: pozitif, negatif ve nötr [8] .

Sentiment analizi gerçekleştirilirken, önce etiketlenmiş verilerden oluşan bir eğitim veri seti oluşturulur ve sonrasında, çeşitli veri madenciliği yöntemleriyle veriler ön işleme sürecinden geçerek sınıflandırma için uygun hale getirilir. Bu süreçte metinlerdeki yazım hataları, noktalama işaretleri ve gereksiz kelimeler gibi sorunlar tespit edilerek temizlenir. Sonrasında, sınıflandırma işlemi için makine öğrenmesi algoritmaları kullanılarak modeller eğitilir ve test edilir [9].

1. İLGİLİ ÇALIŞMALAR

Araştırmalar, toplumsal düşünce ve duyguları incelemenin yanı sıra ürünlerin piyasadaki tepkilerini ölçme, siyasi seçimleri tahmin etme ve borsa gibi sosyoekonomik faktörleri anlama amacı güden firmalar için duygu analizinin önemli bir araç olduğunu göstermektedir [10]. Örneğin, Tansa Trisna Astono Putri ve ekibi Asya’daki bir ülkenin büyük şehirlerindeki tweetleri Destek Vektör Makinesi (SVM) yöntemini kullanarak analiz ederek olumlu, olumsuz ve tarafsız tweetleri sınıflandırmışlardır [11]. Benzer şekilde, Xia ve ekibi ise Ensemble Learning yönteminin duygu analizindeki etkinliğini araştırmışlar ve farklı özellik kümelerini ve sınıflandırma algoritmalarını birleştirerek daha güçlü sınıflandırıcılar geliştirmeyi amaçlamışlardır [12]. Malhar ve Ram ise Twitter verilerini sınıflandırmak için denetimli yöntemleri kullanmışlar ve SVM'nin diğer yöntemlere göre daha yüksek performans gösterdiğini belirtmişlerdir. Bu çalışmada, Temel Bileşen Analizini (PCA) kullanarak hibrit özellik seçimini optimize etmişlerdir [13]. Sharma ve diğerleri ise metin belgelerinin polaritesine göre gözetimsiz bir belge tabanlı duyarlılık analizi yapan bir sistem önermişlerdir. Bu sistem belgeleri pozitif, negatif ve tarafsız olarak sınıflandırarak duyarlılık kelimelerini belge koleksiyonlarından tespit etmekte ve karar vericilere yardımcı olacak özet raporlar sunmaktadır [14]. Son olarak, Po-Wei Liang ve ekibi sosyal medya mesajlarının duygularını otomatik olarak araştıran "opinion miner (görüş madencisi)" adlı bir sistem geliştirmiş ve tweetleri Naïve Bayes sınıflandırıcısı ile görüş ve görüş dışı olarak sınıflandırmışlardır [15].

# materyal ve yöntemler

## VERİ TOPLAMA

Çalışma için gerekli olan 2021 yılına ait tweet verileri, bir sosyal medya platformu olan Twitter'dan toplanmıştır. Bu veriler, belirli anahtar kelimeler kullanılarak çekilmiş ve duygu analizi için uygun hale getirilmiştir. Tablo 1’de toplanan tweetler etiketlenmiş veri seti gösterilmektedir [16].

Tablo 1: Duygu Analizi Veri Seti Dağılım Tablosu

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Veri Setleri** | **Total Veri Sayısı** | **Pozitif Veri Sayısı** | **Negatif Veri Sayısı** | **Nötr Veri Sayısı** |
| **Etiketli Veri Seti** | 12960 | 1607 | 2904 | 8449 |

## VERİ ÖN İŞLEME

Toplanan veriler üzerinde çeşitli ön işleme adımları uygulanmıştır.

Veri Temizleme Fonksiyonları: Veri temizleme işlemleri, metinlerin gereksiz karakterlerden arındırılması ve tutarlılığın sağlanması için uygulanır. Bu adımda HTML etiketlerinin kaldırılması, URL'lerin temizlenmesi, kullanıcı mentionlarının ve noktalama işaretlerinin kaldırılması gibi işlemler yapılır. Ayrıca, metinlerin küçük harfe dönüştürülmesi ve emojilerin temizlenmesi de bu adımda gerçekleştirilir.

Gürültü Temizleme: Metinlerdeki gereksiz karakterler, bağlaçlar ve özel semboller temizlenmiştir.

Tokenizasyon: Metinler, kelime ve cümle düzeyinde parçalarına ayrılmıştır.

Kelime Köklerine Ayırma (Stemming): Kelimeler NLTK kütüphanesindeki SnowballStemmer kullanılarak köklerine ayrılıp daha anlamlı hale getirilmiştir.

Stopword’lerin Kaldırılması : Anlam taşımayan sık kullanılan kelimeler (stopwords) metinlerden çıkarılmıştır. Standart stopword listelerinin yetersiz kaldığı durumlar göz önünde bulundurularak, çalışma kapsamında özel bir stopword listesi oluşturulmuş ve kullanılmıştır.

Label Encoding: Duygu etiketleri sayısal değerlere dönüştürülmüştür.

Bu adımlar, verileri gereksiz karakterlerden arındırarak ve daha tutarlı hale getirerek, duygu analizinde daha iyi sonuçlar elde etmek amacıyla yapılmıştır [12].

1. VEKTÖRLEME YÖNTEMLERİ

Veri ön işleme adımlarından sonra, tweet verileri Word2Vec ve Doc2Vec vektörleme yöntemleri kullanılarak vektörleştirilmiştir.

1. *Word2Vec :* Word2Vec, Mikolov ve diğerleri tarafından önerilen en popüler kelime temsil yöntemidir [17,18]. Bu teknik, sinir ağları aracılığıyla bir eğitim setindeki kelimelerin vektör temsillerini öğrenir. Anlamları benzer olan kelimeler, gömülü vektör uzayında yakın vektörlerle eşleştirilir.

Word2Vec, kelimelerin dağıtılmış bir temsilini sağlamak için Mikolov ve arkadaşları (2013) tarafından geliştirilen Word2Vec, CBOW ve Skip-gram isimli iki model mimarisi sunar. CBOW modeli, bir kelimenin belirli bir bağlamda tahmin edilmesi için kullanılırken, Skip-gram modeli bir kelimenin kökünden sınıfını tahmin etmek için kullanılır.

Bu yöntem, büyük miktarda metin verisi üzerinde eğitilerek kelimeler arasındaki anlamsal ilişkileri yakalar . Etiketlenmiş tweet veri seti üzerinde Word2Vec modeli eğitilmiştir [17,18].

1. *Doc2Vec :* Le ve Mikolov (2014) tarafından geliştirilen Doc2Vec, cümle veya belgelerin anlamlarını vektörlere dönüştüren bir yöntemdir. Word2Vec'in bir uzantısı olarak, sadece kelimeleri değil, aynı zamanda kelimelerin bağlamını ve cümlelerin ya da belgelerin anlamını temsil eden vektörler oluşturur. Doc2Vec, Paragraf Vektör-Distributed Memory (PV-DM) ve Paragraf Vektör-Distributed Bag of Words (PV-DBOW) olmak üzere iki ana model içerir. PV-DM modeli, cümle veya paragraf vektörlerini, kelime vektörleri ile birlikte öğrenirken, PV-DBOW modeli sadece cümle veya paragraf vektörlerini öğrenir. Bu yöntem, metinlerin bağlamsal ilişkilerini daha iyi anlamak için kullanılmıştır [19-21].

Doc2Vec yöntemi, temelde Word2Vec yönteminden türetilmiştir. Word2Vec algoritması, kelimeleri vektörler halinde uzaysal bir düzleme dönüştürürken, Doc2Vec cümleler veya paragraflar için benzer işlemleri yapar. Bu sebeple bazı kaynaklarda Paragraph2Vec olarak da anılır. Doc2Vec, Word2Vec'in dokümanlara uyarlanmış bir versiyonudur [19-21]. Word2Vec'de eğitim sürecinin amacı, ortalama olasılığı maksimize etmektir.

Word2Vec algoritmasında, Continuous Bag of Words (CBoW) ve Skip-Gram (SG) olmak üzere iki farklı yöntem bulunmaktadır. Bu yöntemler, Doc2Vec için yapılandırılarak Distributed Memory (DM) ve Distributed Bag of Words (DBoW) adlarını almıştır [22].

1. MODEL EĞİTİMİ

*1)Lojistik Regresyon (LR)*: Bağımlı değişkenlerin kategorik olarak olasılığının tahmini için kullanılan bir istatistiksel algoritmadır [23]. LR modeli, bağımlı değişkenlerin olasılığı üzerinden sınıflama yapmak için önemli özellikleri belirleyerek data analizi yapar [24].

LR, eğitim veri noktalarının çok boyutlu uzayda her birinden minimum mesafeye sahip bir grafik gösterimi hesaplamak için çeşitli lojistik fonksiyonları kullanır [25].

Denklem 1’de P(y=1∣x), belirli bir giriş, x için çıktının 1 sınıfına ait olma olasılığı, w giriş özelliklerinin ağırlık vektörü, b bias terimi, e Euler sayısıdır (yaklaşık olarak 2.71828). Bu fonksiyon, sonsuz sayıda değer alabilen bir değişkenin değerini 0 ile 1 arasında bir değere dönüştürür. Bu özelliği, sınıflandırma problemlerindeki olasılık tahminini yapmak için ideal kılar.

(1)

Modelin eğitim aşamasında, giriş değişkenlerinin ağırlıkları ve bias terimi, eğitim veri setine en iyi şekilde uyacak şekilde ayarlanır. Ardından, yeni bir veri noktasının sınıfını belirlemek için bu ağırlıklar ve bias terimi kullanılır.

LR'un sınıflandırma yetenekleri genellikle basit ve yorumlanabilir modeller üretmesinden gelir. Bu özellikleri, özellikle tıp, biyoloji ve sosyal bilimler gibi alanlarda, modelin karar süreçlerinin anlaşılması ve yorumlanması gereken durumlarda önemlidir [25].

*2)LSTM* : Derin öğrenme, son yıllarda büyük ilgi gören bir konudur ve doğrusal olmayan işlem birimlerinin katmanlar halinde kullanıldığı bir makine öğrenme türüdür. Bu katmanlar sıralı olarak çalışır ve her katman, bir önceki katmanın çıktısını alır. Derin öğrenme algoritmaları, yapay sinir ağlarının karmaşık formları olarak düşünülebilir. Tek katmanlı yapay sinir ağlarından farklı olarak, doğrusal olmayan problemleri çözme yeteneğine sahiptir. Derin öğrenme yaklaşımında gizli katmanlar bulunur [26].

LSTM, zaman serileri gibi aralıklı bağlantılara sahip bir derin öğrenme algoritmasıdır. Duygu analizi, metin sınıflandırma, konuşma tanımlama ve dil çevirisi gibi çeşitli alanlarda kullanılmaktadır. LSTM'nin yapısı, girdi, çıktı ve unutma katmanlarına sahip özyinelemeli ağ yapısından farklıdır. LSTM, klasik sinir ağlarından farklı olarak,  hangi bilgilerin tutulacağına, geçmiş bilgiye bakarak karar verebilmektedir. Bu karar, örneğin bir ortamın model için gerekli olup olmadığına bağlı olarak alınır [27].

LSTM algoritması, Recurrent Neural Networks (RNN) yapısında ortaya çıkan türev problemlerine çözüm olarak geliştirilmiştir [27]. Girdi ve çıktı katmanlarının yanı sıra, verinin unutulup unutulmayacağına karar veren unutma katmanları da içerir. Bu katmanlar, bir ortamdan başka bir ortama geçildiğinde eski ortamın model için gerekli olup olmadığına karar verir ve gereksiz bilgileri modelden kaldırır.

1. PERFORMANS METRİKLERİ

Modellerin performansı, doğruluk oranları ve diğer metrikler kullanılarak değerlendirilmiştir:

1. *Confusion Matrix* :Confusion matrix, bir sınıflandırma algoritmasının performansını değerlendirmek için kullanılan temel bir araçtır. Her bir sınıfın gerçek ve tahmini değerlerini içerir. Bu matris, True Positive (TP), True Negative (TN), False Positive (FP) ve False Negative (FN) değerlerini gösterir. Denklik 2’de örnek bir confusion matrix gösterilmektedir [28].

(2)

1. *Accuracy* :Accuracy, doğru tahmin edilen toplam örnek sayısının toplam örnek sayısına oranını ifade eder. Denklik 3’teki formüle göre hesaplanır [28].

(3)

1. *Precision*: Pozitif olarak tahmin edilen örneklerin gerçekte pozitif olma olasılığını ifade eder ve Denklik 4’teki formüle göre hesaplanır [28].

(4)

1. *Recall*: Gerçekte pozitif olan örneklerin ne kadarının doğru bir şekilde pozitif olarak tahmin edildiğini gösterir ve Denklik 5’teki formüle göre hesaplanır [28].

(5)

1. *F1-Score:* F1-Score, precision (kesinlik) ve recall (duyarlılık) değerlerinin harmonik ortalamasıdır. Özellikle dengesiz sınıflandırma durumlarında kullanışlıdır ve Denklik 6’daki formüle göre hesaplanır [28].

F1-Score=2 (6)

# UYGULAMA VE BULGULAR

Çalışmada, duygu analizi modeli oluşturmak için öncelikle klasik Natural Language Processing (NLP) işlemleri kullanılarak LR algoritması denenmiştir. Bu ilk aşamada, duygu analizi modellerinde derin öğrenme algoritmalarının daha iyi sonuç verebileceği düşünülerek model bu kez LSTM sinir ağı ile eğitilmiştir. Ancak, modellerin doğruluk sonuçları beklendiği gibi olmamış; LR için doğruluk oranı %70 iken, LSTM için %67 olarak gerçekleşmiştir (Tablo 2). Böylelikle çalışmanın ilk odak noktası, duygu analizi modelinde makine öğrenmesi algoritmaları ve derin öğrenme algoritmalarının başarısını karşılaştırılması olmuştur.

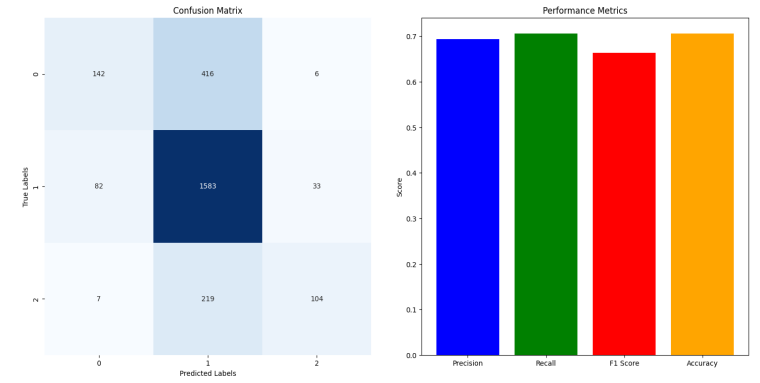
Çalışmanın ikinci odak noktası, Word2Vec ve Doc2Vec vektörleme yöntemlerinin karşılaştırılmasıdır. İlk aşamada, etiketlenmiş tweet veri seti üzerinde Word2Vec vektörleme yöntemi uygulanmıştır. Daha sonra, elde edilen vektörler kullanılarak ayrı ayrı LR ve LSTM sinir ağı modeli eğitilmiştir. Bu aşama, Word2Vec vektörleme yönteminin başarısını değerlendirmek adına önemli bir adım olmuştur. Word2Vec ve LR kombinasyonu %81 doğruluk oranına ulaşırken, Word2Vec ve LSTM kombinasyonu %77 doğruluk oranı sağlamıştır. Bu sonuçlar, doğruluk oranlarını artırmış olsa da, derin öğrenme algoritması yeterli performansı gösterememiştir.

Çalışmanın son odak noktasında ise vektörleşme yöntemi karşılaştırması amacıyla modellere bu kez Word2Vec yerine Doc2Vec vektörleme yöntemi tercih edilmiştir. Bu yöntem, kelimelerin bağlamını daha iyi anlamak için önceki denemeden farklı bir perspektife sahiptir. Etiketlenmiş veri seti üzerinde Doc2Vec ile elde edilen vektörler, Word2Vec yönteminde yapıldığı gibi LR ve LSTM modeli eğitilerek uygulanmış ve başarılar karşılaştırılmıştır. Ancak, dikkat çekici bir şekilde, her iki yöntem de benzer doğruluk oranlarına ulaşmıştır. Doc2Vec ve LR kombinasyonu %83 doğruluk oranına ulaşırken, Doc2Vec ve LSTM kombinasyonu %65 doğruluk oranı sağlamıştır (Tablo 2).

Yapılan denemeler sonucunda, elde edilen sonuçların veri yetersizliğinden dolayı derin öğrenme modellerinin tam potansiyelini gösteremediği şeklinde özetlenmiştir. Bu durum, vektörleme yöntemlerinin seçiminde doğruluk oranı yerine potansiyel avantajları göz önünde bulundurmanın önemini vurgulamıştır. Doc2Vec'in avantajları, kelime bağlamını daha iyi anlama ve duygusal içeriği etkili bir şekilde temsil etme yeteneği üzerinde odaklanmıştır. Performans metrikleriyle yapılan analiz, her iki yöntemin benzer doğruluk oranlarına sahip olmasına rağmen, Doc2Vec'in duygu analizi görevinde daha etkili sonuçlar elde etme eğiliminde olduğunu göstermiştir.

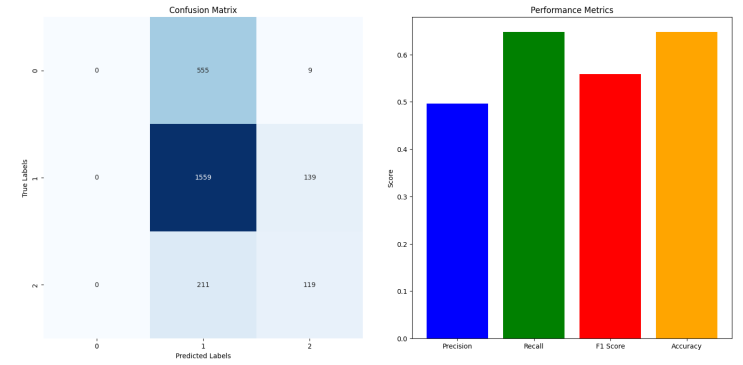
1. *Sayısal Değerlendirme Sonuçları* : Görsel 1, klasik

NLP işlemleriyle eğitilmiş LR modelinin performansını göstermektedir. Model, pozitif ve negatif sınıflar arasında yüksek doğruluk oranları sergilemiştir. Pozitif sınıfta (0), doğru tahmin oranı yüksekken, yanlış negatif sayısı düşüktür. Negatif sınıfta da doğru tahmin oranı yüksektir, ancak yanlış negatif ve pozitif sayıları %36 daha fazladır. Genel doğruluk oranı %70 olarak belirlenmiştir ve LR modeli duygu analizi görevinde etkili sonuçlar vermektedir. Performans metrikleri, modelin dengeli ve güvenilir bir performans sergilediğini göstermektedir.



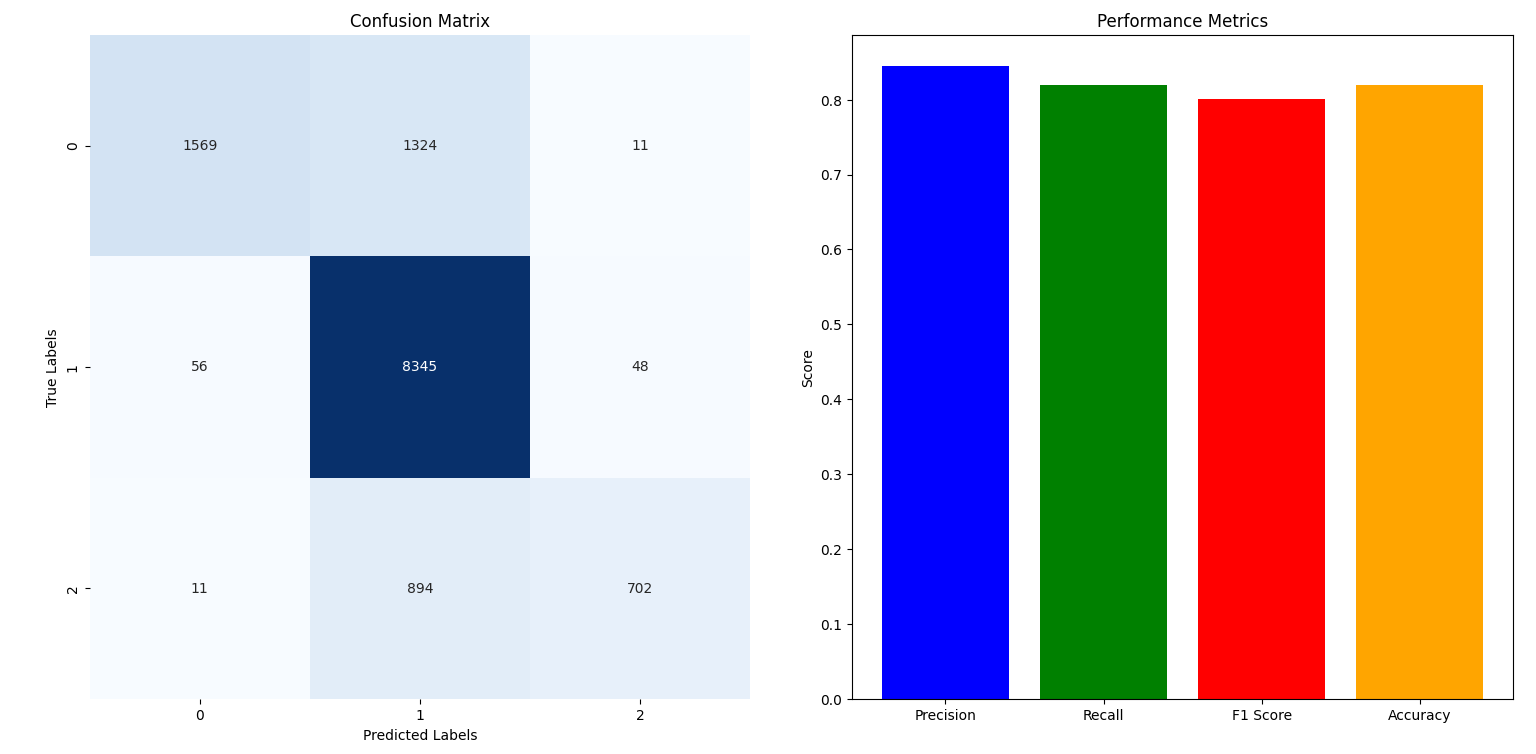
*Görsel 1: LR Modeli Karmaşıklık Matrisi ve Performans Metrikleri*

Görsel 2’de LSTM Modeli Karmaşıklık Matrisi ve Performans Metrikleri verilmiştir. LSTM modelinin performansı, LR modeline kıyasla biraz daha düşük görünmektedir. Karmaşıklık matrisi, LSTM'nin yanlış negatif tahminlerde daha yüksek hata oranına sahip olduğunu göstermektedir. Genel doğruluk oranı % 67 olup, modelin genel performansı tatmin edicidir ancak LR modeline göre biraz daha düşük kalmaktadır. LSTM modeli, verinin bağlamını daha iyi anlayabilme potansiyeline sahip olsa da, elde edilen sonuçlar veri yetersizliği veya modelin optimizasyonu gibi nedenlerle beklenen performansa ulaşamamış olabilir.

**

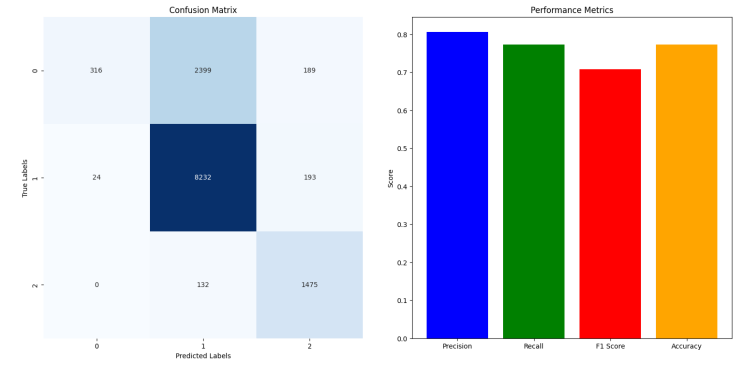
*Görsel 2: LSTM Modeli Karmaşıklık Matrisi ve Performans Metrikleri*

Görsel 3’de Word2Vec & LR Modeli Karmaşıklık Matrisi ve Performans Metrikleri gösterilmektedir. Word2Vec vektörleştirme yöntemi ile LR modelinin performansını göstermektedir. Karmaşıklık matrisi, modelin doğru ve yanlış sınıflandırmalarını detaylı bir şekilde sunmakta olup, modelin genel olarak yüksek Precision, Recall, F1-Score ve Accuracy değerlerine sahip olduğunu göstermektedir. Bu sonuçlar, Word2Vec & LR kombinasyonunun duygu analizi görevinde başarılı olduğunu ve etkili bir performans sergilediğini ortaya koymaktadır.

**

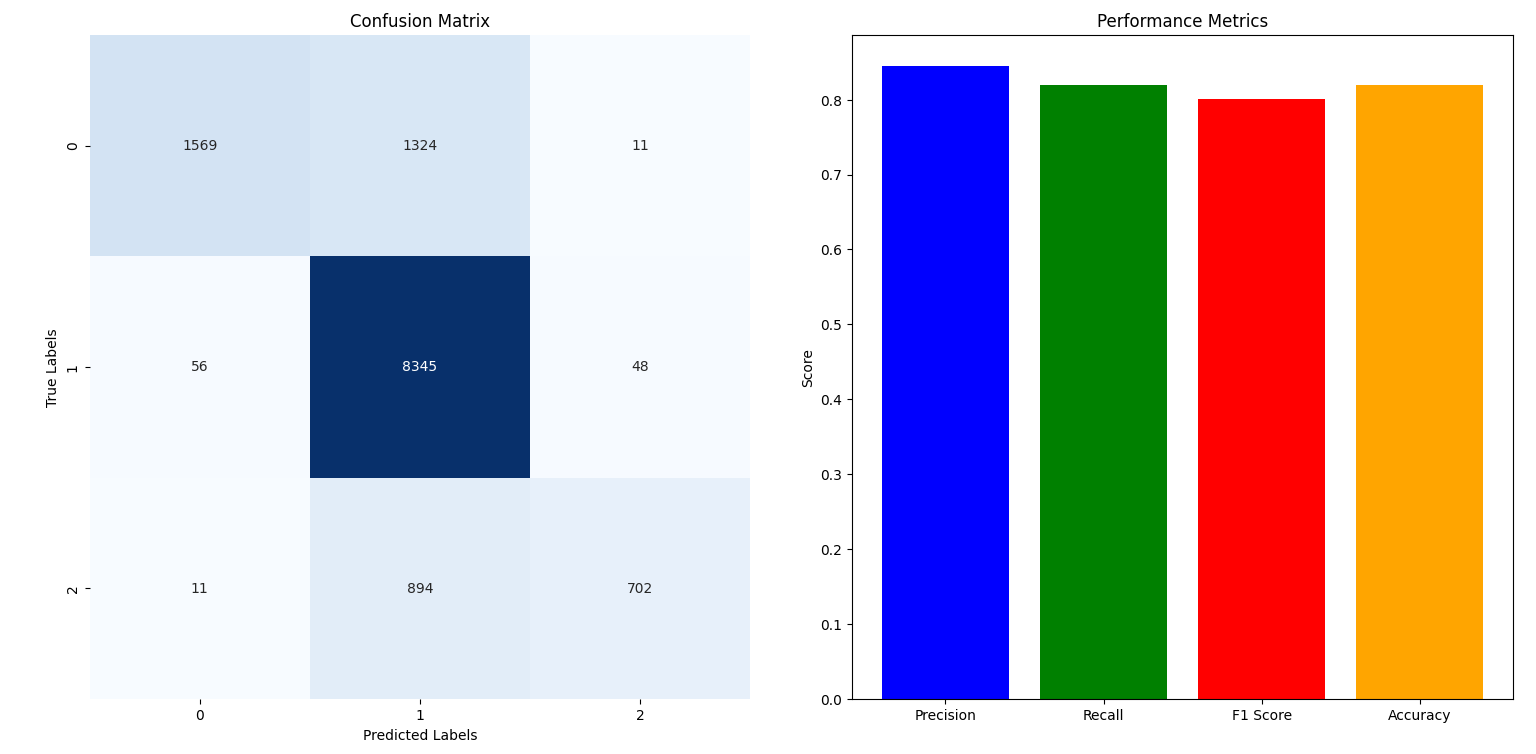
*Görsel 3: Word2Vec & LR Modeli Karmaşıklık Matrisi ve Performans Metrikleri*

Görsel 4’te, Word2Vec vektörleştirme yöntemi ile LSTM modelinin performansını göstermektedir. Precision ve Accuracy oranlarının oldukça yüksek olması, modelin genel doğruluğunun iyi olduğunu, ancak Recall ve F1-Score değerlerinin nispeten daha düşük olması, modelin bazı sınıflarda duyarlılığının yeterince yüksek olmadığına işaret etmektedir. Bu durum, LSTM modelinin daha iyi bir performans gösterse de, genel olarak beklendiği kadar başarılı olmadığını göstermektedir. Sonuç olarak, Word2Vec ve LSTM kombinasyonunun duygu analizi görevinde yeterli performansı gösteremediği söylenebilir.

**

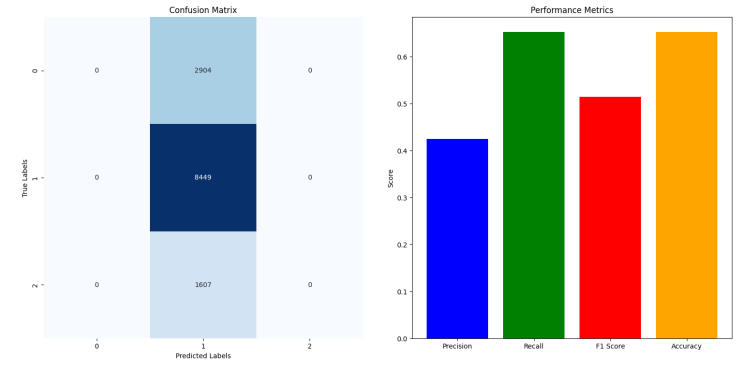
*Görsel 4: Word2Vec & LSTM Modeli Karmaşıklık Matrisi ve Performans Metrikleri*

Görsel 5, Doc2Vec vektörleştirme yöntemi ile LR modelinin performansını göstermektedir. Karmaşıklık matrisi ve performans metrikleri, modelin genel olarak dengeli bir performans sergilediğini ve duygu analizi görevinde etkili olduğunu göstermektedir. Precision, Recall ve Accuracy oranlarının yüksek ve birbirine yakın olması, modelin genel olarak dengeli bir performans sergilediğini işaret etmektedir. Özellikle, F1-Score’un yüksek olması, modelin hem duyarlılık hem de kesinlik açısından başarılı olduğunu göstermektedir.



*Görsel 5: Doc2Vec & LR Modeli Karmaşıklık Matrisi ve Performans Metrikleri*

Görsel 6, Doc2Vec & LSTM modelinin karmaşıklık matrisini ve performans metriklerini göstermektedir. Karmaşıklık matrisinde, doğru ve yanlış sınıflandırılan örneklerin dağılımı net bir şekilde görülmektedir. Performans metrikleri grafiğinde ise modelin doğruluk oranı %65’dir, kesinlik, geri çağırma ve F1 skoru gibi önemli performans göstergeleri yer almaktadır. Doc2Vec & LSTM kombinasyonunun özellikle doğruluk oranında beklenen performansı yakalayamadığı, ancak geri çağırma oranının diğer metriklere kıyasla daha yüksek olduğu dikkat çekmektedir.

**

*Görsel 6: Doc2Vec & LSTM Modeli Karmaşıklık Matrisi ve Performans Metrikleri*

Tablo 2: Yapılan çalışmalardaki doğruluk oranları

|  |  |  |
| --- | --- | --- |
| Model | Vektörleme Yöntemi | Doğruluk Oranı |
| LR | Klasik NLP işlemleri | %70 |
| LSTM | Klasik NLP işlemleri | %67 |
| LR | Word2Vec | %81 |
| LSTM | Word2Vec | %77 |
| LR | Doc2Vec | %83 |
| LSTM | Doc2Vec | %65 |

Tablo 2’ten anlaşıldığı gibi LR modeli, sinir ağı modeline kıyasla daha yüksek doğruluk oranlarına ulaşmıştır. LR modeli, klasik NLP işlemleriyle %70 doğruluk oranı elde ederken, sinir ağı modeli olan LSTM ise %67 doğruluk oranına ulaşmıştır. Bu durum, LR modelinin sinir ağı modeline göre daha başarılı olduğunu göstermektedir.

Word2Vec ve Doc2Vec yöntemlerinin kullanımıyla yapılan deneylerde ise, LR modelinin sırasıyla %81 ve %83 doğruluk oranlarına ulaştığı görülmüştür. LSTM modeli ise Word2Vec için %77, Doc2Vec için ise %65 doğruluk oranı sağlamıştır. Bu sonuçlar, LR modelinin Word2Vec ve Doc2Vec vektörleme yöntemleriyle daha tutarlı bir performans sergilediğini ortaya koymaktadır. Özellikle, Doc2Vec'in duygusal içeriği daha etkili bir şekilde temsil ettiği ve LR modeliyle birlikte daha yüksek doğruluk oranları elde edildiği gözlemlenmiştir. LSTM modelinin daha düşük doğruluk oranlarına sahip olması, kelime bağlamını eksik anlama yeteneğinden kaynaklanabilir ve derin öğrenme modellerinin karmaşık ilişkileri öğrenme konusundaki zorlukları bu sonuçları etkileyebilir.

LR modeli, sosyal medya gönderileri üzerinde daha etkili bir duygu analizi sağlamıştır. Sinir ağı modelinin daha düşük doğruluk oranlarına sahip olması, kelimeler arasındaki bağlamı eksik anlama yeteneğinden kaynaklanabilir. Ayrıca, derin öğrenme modellerinin daha karmaşık ilişkileri öğrenme konusundaki zorlukları, Doc2Vec vektörleme yöntemiyle oluşturulan sinir ağı modelinin istenilen performans seviyesine ulaşmasını engellemiştir.

Sonuç olarak, LR modeli sosyal medya verileri üzerinde daha etkili bir duygu analizi sağlamıştır. Bu çalışma, makine öğrenmesi ve derin öğrenme algoritmalarının sosyal medya verileri üzerindeki performansını değerlendirerek, LR karmaşık veri setlerinde bile yüksek doğruluk oranlarına sahip olduğunu göstermiştir. Bu bulgular, pazarlama ve müşteri hizmetleri gibi alanlarda, sosyal medya verilerinin analizinde LR modelinin tercih edilebileceğini ortaya koymaktadır.

# SONUÇ

Bu çalışma, LR ve LSMT algoritmalarının, Word2Vec ve Doc2Vec vektörleme yöntemleriyle, duygu analizi modellerindeki performansını karşılaştırılmıştır. LR algoritması, basit yapısı ve hesaplama verimliliği sayesinde tatmin edici doğruluk oranları sağlarken, LSTM sinir ağı modelinin bu testler özelinde LR algoritması kadar başarılı olmadığı gözlemlenmiştir. Doc2Vec vektörleme ile oluşturulan sinir ağı modeli, duygu analizi görevinde beklenen performansı sağlayamamıştır. Modelin duygusal içeriği etkili bir şekilde temsil edememesi ve LR modeline kıyasla daha düşük doğruluk oranlarına sahip olması, kelimeler arasındaki bağlamı eksik anlama yeteneğinden kaynakladığı şeklinde düşünülmektedir. Word2Vec ve Doc2Vec vektörleme yöntemlerinin karşılaştırılması, kelimelerin bağlamını ve duygusal içeriği temsil etme yetenekleri açısından değerli bulgular sunmuştur. Duygu analizi modeli için yapılan testler sonucunda, Doc2Vec yönteminin, Word2Vec yönteminden daha etkili bir vektörleşme yöntemi olduğu görülmektedir. Doc2Vec'in, duygu analizi görevlerinde kelime bağlamını daha iyi anlayarak daha etkili sonuçlar elde etme eğiliminde olması, bu yöntemin potansiyel avantajlarını vurgulamaktadır. Vektörleşme yöntemlerinin entegrasyonu, duygu analizi doğruluk oranlarını artırmaktadır ve vektörleşme yöntemleri, LR ile daha uyumlu çalışmaktadır. Çalışma sonucunda, Doc2Vec ve LR’nin birlikte kullanımı ile %83’lik bir başarım oranı elde edilmiştir. Bu çalışmanın gelecekteki araştırmalar için temel teşkil edebileceği ve duygu analizi modellerinin ve vektörleme yöntemlerinin daha büyük veri setleri üzerinden ve farklı algoritmalarla incelenerek daha kapsamlı karşılaştırma sonuçları elde edilebileceği düşünülmektedir.

##### KAYNAKÇA

1. Go, A., Huang, L., and Bhayani, R., Twitter sentiment analysis, Entropy, 17, (2009).
2. Liu, B. and Lei, Z., Mining Text Data: A survey of opinion mining and sentiment analysis, Mining Text Data, Springer, USA, pp. 415-463. ISBN: 978-1-4614-3223-4.Prabowo, R. and Thelwall, M., "Sentiment analysis: A combined approach", Journal of Informetrics, 3(2), 143-157, (2009).
3. Akgül, E.S., Ertano, C. ve Diri, B., Twitter verileri ile duygu analizi,Pamukkale University Journal of Engineering Sciences, 22(2), 106-110,(2016)
4. Kumaş, E. (2021). Türkçe twitter verilerinden duygu analizi yapılırken sınıflandırıcıların karşılaştırılması. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi*, *2*(2), 1-5.
5. B. Pang, L. Lee, “Opinion Mining and Sentiment Analysis,” Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1–135, 2008.
6. M. Meral and B. Diri, “Sentiment Analysis on Twitter”, in 22nd Signal Processing and Communications Applications Conference (SIU), Trabzon, 2014.
7. Ö. Çoban, B. Özyer, and G. T. Özyer, “Sentiment Analysis For Turkish Twitter Feeds,” in 23nd Signal Processing and Communications Applications Conference (SIU), Malatya, 2015.
8. S. E. Şeker, "Duygu Analizi (Sentimental Analysis)," YBS Ansiklopedi, vol. 3, no. 3, pp. 21-36, 2016.
9. A. Yeşilyurt and S. E. Şeker, "Metin Madenciliği Yöntemleri ile Twitter Duygu Analizi," YBS Ansiklopedi, vol. 4, no. 2, 2017.
10. R. S. Gound, P. V. Tikone, S. S. Suryawanshi and D. Nagpal, "Twitter Data Sentiment Analysis and Visualization," International Journal of Computer Applications, vol. 180, no. 20, pp. 14-16, 2018.
11. T. T. Astono Putri, M. D. Mendoza and M. F. Alie, "Sentiment Analysis On Twitter Using The Target-Dependent Approach And The Support Vector Machine (SVM) Method," Jurnal Mantik, vol. 4, no. 1, pp. 20-26, 2020.
12. R. Xia, C. Zong and S. Li, "Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification," Information Sciences, vol. 181, no. 6, pp. 1138-1152, 2011.
13. M. Anjaria and R. M. R. Guddeti, "Influence factor based opinion mining of Twitter data using supervised learning," in 2014 Sixth International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, 2014.
14. R. Sharma, S. Nigam and R. Jain, "Opinion Mining of Movie Reviews at Document Level," International Journal on Information Theory (IJIT), vol. 3, no. 3, 2014.
15. P.-W. Liang and B.-R. Dai, "Opinion Mining on Social Media Data," in 2013 IEEE 14th International Conference on Mobile Data Management, Milan, Italy, 2013.
16. <https://github.com/enesyurt/SentimentyBot/tree/main>
17. Mikolov T., Chen K., Corrado G., Dean J., Efficient estimation of word representations in vector space, International Conference on Learning Representations, Arizona, 1-12, 2013.
18. Mikolov T., Sutskever I., Chen K., Corrado G, Dean J., Distributed representations of words and phrases and their compositionality, Neural Information Processing Systems Conference, Lake Tahoe, 3111–3119, 2013.
19. Go, A., Huang, L., and Bhayani, R., Twitter sentiment analysis, Entropy, 17, (2009).
20. Szomszor, M.N., Kostkova, P. and Quincey, E.D., # Swineflu: Twitter predicts swine flu outbreak in 2009, 3rd International ICST Conference on Electronic Healthcare for the 21st Century. 2012.
21. Bian J, Topaloglu, U, Yu, F., Towards large-scale Twitter mining for drugrelated adverse events, International Workshop on Smart Health and Wellbeing (SHB’12), Maui, Hawaii, USA, 29 October-2 November 2012.
22. Witten, I.H., et al., Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2016.
23. Indra, S. T., Wikarsa, L., & Turang, R. 2016. Using Logistic Regression Method to Classify Tweets Into the Selected Topics. In 2016 international conference on advanced computer science and information systems, 15-16 Ekim, Malang, Indonesia, 385-390.
24. Prabhat, A., & Khullar, V., Sentiment classification on big data using Naïve Bayes and logistic regression, In 2017 International Conference on Computer Communication and Informatics (ICCCI), IEEE ,1-5, Coimbatore- India, January, 2017
25. Salazar, D. A., Vélez, J. I., & Salazar, J. C., Comparison between SVM and logistic regression: Which one is better to discriminate?, Revista Colombiana de Estadística, 35 (SPE2), 223-237, 2012.
26. Brownlee J., “Develop Sequence Prediction Models With Deep Learning”. Machine Learning Mastery, 2017.
27. Yılmaz, M. C., & Orman, Z. (2021). LSTM derin öğrenme yaklaşımı ile Covid-19 pandemi sürecinde twitter verilerinden duygu analizi. *Acta Infologica*, *5*(2), 359-372.
28. Bilgin, M. (2017). Gerçek veri setlerinde klasik makine öğrenmesi yöntemlerinin performans analizi. *Breast*, *2*(9), 683.