

Sahte Haber Tespiti için Derin Bağlamsal Kelime Gömümleri ve Sinirsel Ağların Performans Değerlendirmesi

Cengiz HARK^{1*}

¹Mühendislik Fakültesi, Bilgisayar Mühendisliği, Bilgisayar Bilimleri, Malatya, Türkiye

*¹cengiz.hark@inonu.edu.tr

(Geliş/Received: 06/06/2022;

Kabul/Accepted: 08/09/2022)

Öz: Artan internet ve sosyal medya kullanımı ile sosyal medya ve online haber siteleri bilgi içeriklerini oluşturmada ve yaymada önemli kaynaklar haline gelmişlerdir. Ancak online bilginin miktarı ve üretilme hızından ötürü insan eliyle doğrulanması mümkün olamamaktadır. Dahası rejimler, hükümetler ve etnik kökenler denetimden uzak sahte haberlerin yıkıcı etkisine maruz kalmakta ve bu olumsuz etkilerin minimuma indirgenmesi için yeni teknolojilere ihtiyaç duyulmaktadır. Son yıllarda farklı doğal dil işleme görevlerinde ön-eğitilmiş modeller başarı ile kullanılmaktadır. Derin sinir ağı ve kavramsal kelime gömümlerinin birlikte kullanılması durumunda hangi sınıflandırıcının daha verimli olduğu hususu yeterince net değildir. Bu noktada kapsamlı ve karşılaştırmalı çalışmanın eksikliğinden ötürü Global Vektörler (GloVe) gömülme katmanının sağladığı bağlamsal temsiller ile dört farklı sınıflandırıcı deneysel sürece dahil edilmiştir. GLoVe katmanından sonra Çok Katmanlı Algılayıcı (Multi-Layer Perceptron, MLP), Uzun Ömürlü Kısa Dönem Bellek Ağı (Long-Short Term Memory, LSTM), Yinelemeli Sinir Ağları (Recurrent Neural Network, RNN) ve Evrimsel Sinir ağları kullanılmıştır (Convolutional Neural Network, CNN). Deneysel çalışmalarda açık erişimli COVID-19 isimli sahte haber tespit veri seti kullanılmış, başarımları çeşitli performans metrikleri ile hesaplanmıştır. En yüksek başarımlar %91 ile LSTM tarafından rapor edilmiştir. Ön-eğitilmiş kelime gömümlerinin farklı sinirsel ağlardan bağımsız olarak yüksek bir hassasiyetle sahte haberlerin tespitinde kullanılabileceğini gösteren umut verici sonuçlar sunulmuştur.

Anahtar kelimeler: Doğal Dil İşleme, Kelime Gömümleri, Sahte Haber Tespiti, Kavramsallaştırılmış ön-eğitilmiş kelime Gömümleri, Yapay Zekâ.

Performance Evaluation of Deep Contextual Word Embeddings and Neural Networks for Fake News Detection

Abstract: With the increasing use of the internet and social media, social media and online news sites have become essential sources in creating and disseminating information content. However, due to the amount and production speed of online information, it is impossible to verify it with human hands. Moreover, regimes, governments, and ethnicities are exposed to the devastating effects of uncontrolled fake news, and new technologies are needed to minimize these harmful effects. Recently, pre-trained models have been used successfully in different natural language processing tasks. However, it is not clear enough which classifier is more efficient when deep neural networks and conceptual word embeddings are used together. At this point, due to the lack of comprehensive and comparative study, Global Vectors (GloVe) were included in the experimental process with four different classifiers with the contextual representations provided by the pre-trained word embedding layer. After the GLoVe layer, Multi-Layer Perceptron (MLP), Long-Short Term Memory Network (LSTM), Recurrent Neural Networks (RNN), and Convolutional Neural networks (CNN) are used. In experimental studies, an open access fake news detection dataset named COVID-19 was used, and its performance was calculated with various performance metrics. LSTM Network reported the highest performance of 91%. Promising results are presented, showing that pre-trained word embeddings can be used to detect fake news with a high sensitivity independent of different neural networks.

Key words: Natural Language Processing, Word Embeddings, Fake News Detection, Contextualized Pre-trained, Artificial Intelligence.

1. Giriş

Toplumlar günümüz bilgi çağında, çeşitli kaynaklardan anlık olarak bilgi alıp paylaşabilmektedir. İnternetin yaygınlaşmasıyla birlikte insanların internetten haber almaları diğer geleneksel mecralara göre daha kolay ve elverişli hale gelmiştir [1].

* Sorumlu yazar: cengiz.hark@inonu.edu.tr. Yazarların ORCID Numarası: ¹ 0000-0002-5190-3504

Ne yazık ki, açık internet, etkili bir denetim olmaksızın pek çok sahte haberin yayılmasını kolaylaştırmaktadır. Yalan haber Türk Dil Kurumu sözlüğünde “Gerçek olmayan, uydurma haber” olarak tanımlanmaktadır [2]. Bir başka tanımlamaya göre yalan haber kasıtlı bir biçimde okuyucuları yanlış yönlendiren haberlerdir [3]. [4] ise belirli finansal kazançlar için farklı hizmetler kullanan birden fazla sosyal ağ kullanıcısı aracılığıyla yayılan yanıltıcı bilgiler olarak tanımlanmaktadır. Sahte haberler, kolay erişim ve hızlı yayılma özellikleriyle kamuoyunu kolayca yanıltabilmekte, toplumsal düzene zarar verebilmekte, sosyal medyanın güvenilirliğini azaltmaktadır. Ayrıca tarafların çıkarlarını zedeleyerek güven bunalımlarına neden olabilmektedir. Tüm bu nedenlerden ötürü son yıllarda sahte ve yalan haberlerin yayılması dünya çapında bir sorun haline gelmiştir. Yalan haberlerin tespitine yönelik yöntemlerin geliştirilmesi oldukça önemlidir. Sahte haber tespit problemini en basit haliyle bir metin sınıflandırma problemine indirgemek mümkündür [5]. Sahte haberlerin tespiti için birçok çalışmada araştırmacılar yazım stillerine odaklanarak metinleri gerçek veya yalan haberler olarak sınıflandırmayı denemiş olsalar da [6] bu yaklaşımlar ile incelenen haberlerde tam bir netlik sağlayamamışlardır. Buna karşılık bilgi bazlı sahte haber tespit sistemleri stil ve biçimden ziyade içeriğe odaklandıkları için daha umut verici neticelere ulaşmışlardır [7]. Son zamanlarda donanım, veri setleri ve algoritmalarındaki iyileştirmeler gibi üç teknik noktada büyük ilerleme kaydedilmiştir. Bu nedenlerden ötürü, araştırmacılar farklı doğal dil işleme (DDİ) görevlerinde metin içeriklerinden değerli bilgileri çıkarmak için derin sinir ağlarına yönelebilmişlerdir [8]. Sahte haber tespitinde araştırmacılar, oldukça popülerleşen kavramsallaştırılmış ön-eğitilmiş kelime gömülme yöntemlerini kullanarak, denetimli veya denetimsiz öğrenme yöntemlerini uygulamaktadırlar. 2017 yılında önerilen Transformer mimarisinin [9] temellerini oluşturduğu ön-eğitilmiş kelime gömülmeleri dikkatleri üzerine toplamıştır. Ek olarak bu yükseliş dizi kodlama ve kod çözme mimarisinin oluşmasında da önemli bir rol oynamıştır. Ön eğitilmiş modellerin farklı özelliklere sahip olan sinir ağları ile beraber kullanılarak yalan haber tespiti gibi güncel ve zorlu bir görevde ortaya konulan performansların ayrı ayrı değerlendirilmesi yazarları motive etmiştir. Çalışma kapsamında yalan haber tespiti amacı ile farklı sinirsel sınıflandırıcıların kullanılmasına yönelik kapsamlı ve karşılaştırmalı çalışmanın eksikliği yazarların bir başka motivasyonu olmuştur.

Sahte haber tespiti görevi için kelime gömülme modelleri ve derin sinirsel ağların beraber kullanılması hususunda; önceden eğitilmiş modellerin hangi sinir ağı türleri ile daha verimli olduğunun araştırılması oldukça önemlidir.

Bu çalışma kapsamında MLP, LSTM, RNN ve CNN gibi dört farklı sinir ağı mimarisi etkili bir kavramsallaştırılmış metin temsil modeli olan GloVe ile beraber kullanılmaktadır. Kavramsallaştırılmış metin temsiline ve derin sinir ağlarına dayalı sahte haber tespiti için bir dizi deneysel süreç gerçekleştirilmiştir. Ön eğitilmiş kelime gömülmelerinin ve sinir ağlarının farklı kombinasyonlarını detaylı bir şekilde rapor eden bir dizi deneysel süreç gerçekleştirilmiştir.

Çalışmanın sonraki bölümleri şu şekilde organize edilmiştir: ikinci bölüm sahte haber tespiti ile ilgili literatürdeki mevcut çalışmaların bir incelemesini sunulduğu, üçüncü bölüm ise kullanılan materyal ve metodların açıklandığı bölümdür, deneysel süreçlere dair detaylı bulgular dördüncü bölümde yer alırken beşinci bölümde çalışmaların sonuçları yorumlanmıştır.

2. İlişkili Çalışmalar

Sahte haberlerin tespit edilebilmesi oldukça yeni bir çalışma alanı olmasına rağmen araştırmacıların oldukça yoğun ilgisi söz konusudur. Sahte haber biçimindeki kötü niyetli içeriğin tespiti, araştırmacılar için oldukça zorlayıcı bir süreçtir. Çeşitli veri türlerinde sahte haberleri tespit etmek için çeşitli yaklaşımlar önerilmiştir. Bu bölümde sahte haber tespiti ile ilgili literatürdeki mevcut çalışmaların bir incelemesi sunulmuştur.

Araştırmacılar genelde, sahte haberlerdeki kalıpları belirlemek için denetimli makine öğrenimi (MÖ) algoritmalarını kullanmaya odaklanarak yalan haberleri yayılmadan önce tespit etmeye çalışırlar. Bu yaklaşım ile doğru ve yalan haberlerden oluşan etiketli veri setleri belirli algoritmalar ile kullanılmaktadır. Sonraki aşamada sunulan model daha önce karşılaşılmayan haber metinlerinin doğru veya yanlış olduğuna yönelik olarak çıkarımlarda bulunabilmektedir [10]. Dünya çapında araştırmacılar tarafından sahte haberleri sınıflandırmak için Denetimli öğrenme kapsamında CNN, destek vektör makineleri, karar ağaçları gibi algoritmalar kullanılmaktadırlar [4][11]. Bu bağlamda gerçekleştirilen çalışmalara özellik seçme ve özellik çıkarma gibi birçok DDİ teknikleri eşlik etmektedir. Bahsedilen teknikler Terim frekansı, kelime çantası, Word2Vec vb.. metin işleme teknikleridir. Ancak bu sayede metinler algoritmalar tarafından kullanılabilen sayısal temsillere dönüştürülerek metinlere ait özellik alt kümeleri elde edilebilmektedir [4]. Denetimli öğrenme mimarileri arasında CNN mimarileri en yaygın kullanılan yöntem olarak kabul görmektedir. Sahte haberlerin sınıflandırılmasında da hayati bir rol oynamaktadır. Kalıyar ve ark. [5] FNDNet adını verdikleri modellerinde kullanıcı tarafından belirlenen özelliklere bağlı kalmadan parametre optimizasyonu ile derin öğrenmeyi beraber kullanmışlardır. Ahmed ve ark.

[12] Farklı MÖ algoritmalarını karşılaştırdıkları çalışmalarında karar ağaçlarını kullanmışlardır. Tanıttıkları modeli iyi bilinen üç farklı veri seti ile test etmişlerdir.

Bir kısım araştırmacı ise denetimsiz öğrenme algoritmalarını kullanarak, etiketlenmemiş girdi verilerindeki anlamlı ve faydalı kümeleri tespit etmek ve ardından yalan haberleri tespit etmek yoluna gitmişlerdir. Sahte haberlerin tespiti için mevcut en modern denetimsiz öğrenme algoritmalarından birisi Zhang ve ark. [13] tarafından tanıtılmıştır. İlgili çalışmada araştırmacılar, metodolojiler ve algoritmalar üzerinde çalışarak sahte haberler, sahte haberlerin yazarları ve sahte haber konularını belirlemeye çalışmışlardır. Yine aynı çalışma kapsamında haberlerin, haber yazarlarının ve haber konularının arasındaki çeşitli ilişkilere ve sahte haberlerin bilinmeyen karakteristiklerine odaklanmışlardır.

Denetimli ve denetimsiz öğrenmeye ilaveten bir grup araştırmacı pekiştirmeli öğrenme metotlarını kullanarak sahte haberlerin tespitine yönelik araştırmalar yapmışlardır. Pekiştirmeli öğrenmenin her adımda uygulandığı sistemden geri bildirim olarak çıkış veriyor olması sahte haber tespiti noktasında verimli olabilmesine olanak tanımaktadır. [14]'de araştırmacılar blockchain ve Bi-LSTM tabanlı bir sınıflandırıcı model sunmuşlar ve sahte haber sınıflandırma problemine katkı sunmuşlardır.

Araştırmacıların bir grubu yarı denetimli öğrenme çerçevesinde sahte haber tespiti görevine odaklanmışlardır. Mansouri ve ark. [15] hem etiketli hem de etiketsiz verileri kullandıkları bir CNN modeli önermişlerdir. Modellerinde resim ve metin verilerinin özelliklerini çıkarmışlardır. Sonraki aşamada ise etiketsiz verilerin sınıf bilgilerini tahmin etmek için olasılık tabanlı bir konu modelleme yöntemi kullanmışlardır.

Ayrıca çok sayıda deneysel çalışma topluluk yaklaşımlarının tek bir modele göre daha başarılı olabileceğini ortaya koymuştur. Belirtilen görev kapsamında topluluk öğrenme yöntemleri ile başarılı sonuçların rapor edildiği çalışmalar gerçekleştirilmiştir. Böyle bir topluluk yaklaşımı Mahabub ve ark. [16] tarafından tanıtılmıştır. Yaygın olarak kullanılan on bir MÖ yaklaşımını çalışmalarına dahil ederek en iyi üç MÖ yöntemini çalışmalarında kullanmışlardır. Doğruluk açısından yüksek değerler elde edebilmişlerdir.

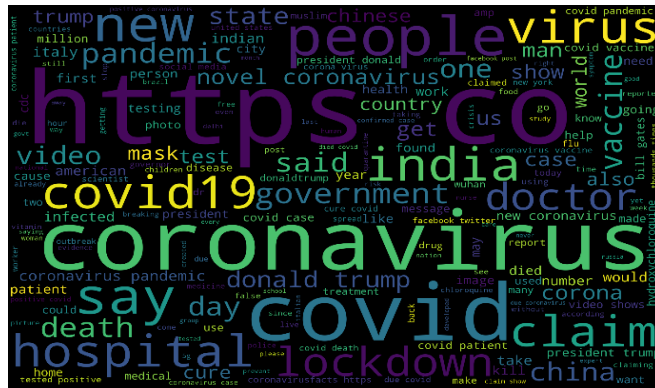
Bu çalışmada sahte haber tespiti gibi karmaşık ve zorlayıcı bir amaç için, sinirsel sınıflandırıcılar ve ön-eğitilmiş kelime gömülme yöntemlerinin hangi kombinasyonlarının daha iyi performans rapor ettiği araştırılmıştır. Çalışmada COVID-19 adlı sahte haber tespit veri seti [17] ve GloVe [18] ön eğitilmiş kelime gömülme katmanının sağladığı bağlamsal temsiller kullanılmıştır. Karşılaştırmalı performans değerlerinin elde edildiği bu kapsamlı çalışmada MLP, LSTM, RNN ve CNN kullanılarak haber metinleri sınıflandırılmıştır.

3. Materyal ve Metotlar

Bu çalışmada sahte haber tespiti için ön-eğitilmiş kelime gömülme ve derin sinirsel sınıflandırıcılara dayalı bir model önerilmiştir.

3.1. Veri Seti

COVID-19 pandemisi tüm dünyayı etkisi altına aldığına teyide muhtaç haber ve bilgiler sosyal medya kullanıcıları tarafından büyük bir hızla yayılmaya başlamıştır. Bu zararlı ve yıkıcı etkinin motive ettiği bir araştırmacı grubu [17] twitter dan sahte ve gerçek haber sınıflarını içeren oldukça kapsamlı bir veri seti hazırlamıştır. Araştırmacıların COVID-19 adını verdikleri veri kümesi 10700 gönderiden oluşmaktadır. Sahte haberler, haber doğrulama internet sayfalarından elde edilirken gerçek haberler ise 14 resmi sosyal medya hesabı kullanılarak doğrulanmış ve bu bilgilere dayanarak iki sınıflı olacak biçimde etiketlenmişlerdir. Şekil 1, COVID-19 veri setinde yer alan kelimelerin sıklıklarını vurgulamaktadır.



Şekil 1. COVID-19 veri setine ait kelime bulutu

Real sınıfı COVID-19 hakkında resmi hesaplar tarafından doğrulanabilen Twitter gönderilerini içermektedir. Fake sınıfı ise doğrulanamayan dolayısı ile spekülasyon ve iddia olduğu düşünülen gönderilerden oluşmaktadır. Fake sınıfı kayıtları içinde Facebook ve Twitter’ında bulunduğu farklı sosyal medya platformlarından ve çeşitli internet sayfalarından toplanmışlardır. Kullanılan veri seti sınıf sayıları bakımından oldukça dengeli ve tutarlıdır. Tablo 1 veri seti hakkında bazı istatistikî bilgiler içermektedir.

Tablo 1. COVID-19 veri setine ait istatistikî bilgiler

Veri Kümeleri	Etiketler		Toplam
	Real	Fake	
Eğitim	3360	3060	6420
Doğrulama	1120	1020	2140
Test	1120	1020	2140

3.2. Metin Gömülmeleri

Sinir ağları ham metin verilerini girdi olarak alamazlar. Sadece sayısal tensörleri girdi olarak alabilirler. Metinleri vektörlerle temsil etmek için sayısal tensörlere dönüştürmek gerekmektedir. Sözü edilen dönüştürme işlemi kelime, karakter veya n-gram bazlı olarak yapılmaktadır. Nihai amaç metinleri tokenlarına (andiçlerine) ayırmaktır. Daha sonra her bir andiç sayısal tensörlerle ilişkilendirilir. Böylelikle metinlerden andiçlere, andiçlerden de vektörlere temsil gerçekleştirilir. Bu aşamada bir-eleman-bir kodlama sık kullanılan yöntemlerden birisidir. Ancak belirtilmelidir ki bir-eleman-bir kodlama ikili ifade şeklinden ötürü seyrek bilgi içerir ve kelime sayısı kadar boyuta sahiptir. Derin sinir ağları bu vektörlerle beslenebilmektedir. Metinleri vektörlerle temsil etme noktasında bir başka tercih ise oldukça güncel bir temsil şekli olan kelime gömülmeleridir. Bir-eleman-bir kodlamadan farklı olarak vektörler küçük boyutlu ondalık değerlerden oluşmaktadır. Çok boyutlu bir temsil söz konusu değildir [19].

Kelime gömülmeleri ile çok büyük miktarlara sahip veri havuzları sıkıştırılarak daha kabul edilebilir boyutlara indirgenebilmektedir. Kelime veya andiç sayısından farklı olarak 256-512 gibi sabit uzunlukta vektörler ile çalışılabilir. Kelime gömülmeleri ile temsil deki en büyük yenilik vektörlerin metinlerden öğrenilmesidir. Belirtilen tüm bu yenilik ve avantajlarından ötürü bu çalışma kapsamında sabit uzunluklu, sıkıştırılmış kelime vektörleri kullanılmaktadır. Kelime gömülmeleri ile üç boyutlu uzayda geometrik yakınlığın anlamsal yakınlık ile uyumlu olduğu bir ilişki kümesi elde edilmektedir (zıt anlamlı kelimelerin kelime vektör uzayında da uzak noktalarda konumlandırılması beklenmektedir) [4],[18].

Kelime gömülmeleri elde edilirken rastgele değerlerin atandığı vektörler ile çalışılarak ağız bu başlangıç değerlerini süreç içinde öğrenmesi ile vektörler elde edilebilmektedir. Tasarlanan model öğrenme sürecinde alacağı geri besleme ile vektörleri güncelleyecek, modele ve/veya eğitim verisine bağlı olarak kelime gömülme uzayını oluşturacaktır. Bu yaklaşım biçimi ile görev ve eğitim verisine en uygun kelime uzayı elde edebilmek mümkün olmaktadır (yeterli eğitim verisi mevcut olması koşulu ile). Bu durum göreve özel gömülmelerin elde edilmesini sağlamaktadır. Ağız ne denli iyi tasarlandığından bağımsız olarak yeterli eğitim verisi bulunmadığında ise yüksek temsil gücüne sahip başarılı kelime gömülmeleri elde edilememektedir. Yeterli eğitim verisinin mevcut olmaması durumunda ön eğitilmiş metin gömülmelerini kullanmak yerinde bir yaklaşımdır. Göreve özel olarak önceden oluşturulmuş gömülmeler dilin ve görevin genel yapısına ait nitelik ve özellikler taşıyabilmektedir. Farklı ön eğitilmiş kelime gömülmelerini farklı görevlerde kullanabilmek mümkündür. Kelime gömülmelerinin araştırma alanlarında kullanılarak yaygınlaşması ise Mikolov ve ark. [20] nın geliştirdiği word2vec algoritmasından sonra mümkün olmuştur. Akabinde araştırmacılar ve endüstri tarafından farklı kelime gömülme çalışmaları sunulmuştur ve yaygınlaşmıştır.

Çalışma kapsamında sahte haber tespiti için GloVe kelime gömülme tekniği kullanılmıştır. GloVe ön eğitilmiş modeli paralel çalışma prensibinden ötürü kolaylıkla eğitilebilmektedir. Wikipedia verileri kullanılarak oluşturulan GloVe ile anlamsal ilişkilerin tespiti mümkün olabilmektedir. Çalışma kapsamında 400000 kelimelik bir kütüphane üzerinde hazırlanan ve 100 boyutlu gömülmeler içeren GloVe kelime gömülme modeli kullanılmıştır.

GloVe modeline ait amaç fonksiyonu X_{ij} matrisi için;

$$J = \sum_{i,j=1}^n f(X_{ij})(w_i^T w_j + b_i + b_j - \log x_{ij})^2 \quad (1)$$

Denklem 1'deki f ağırlık fonksiyonu, w_i ve w_j GloVe modelinin minimize ederek öğrendiği iki vektördür. b_i ve b_j ağırlık değerleridir.

3.3. Sınıflandırıcılar

3.3.1 Çok Katmanlı Algılayıcı Sinir Ağı

Farklı uygulama alanlarındaki sınıflandırma ve regresyon problemlerini çözmek için kullanılan en yaygın mimarilerden birisi MLP sinir Ağlarıdır. Bir girdi katmanından, bir veya daha fazla katmanlardan oluşan ara katmanlardan ve son olarak bir çıktı katmanından oluşmaktadır. Gizli katmandaki nöronlar, önceki ve sonraki katmanların nöronlarına tamamen bağlıdır. Kabaca girdi katmanına yakın olan katmanlar alt katmanlar, çıktı katmanına yakın olan katmanlar ise üst katmanlar olarak adlandırılmaktadır. MLP sinir ağları ileri beslemeli sinir ağlarına bir örnektir [21]. Deneysel süreçler boyunca 128 birimden oluşan 1 gizli katman kullanılmıştır. Tam bağlı gizli katmanda Relu aktivasyon fonksiyonu kullanılırken çıkış katmanında ise sigmoid aktivasyon fonksiyonuna yer verilmiştir.

3.3.2 Yinelemeli Sinir Ağları

RNN'ler daha önce öğrendiği bilgiyi korurken aynı anda yeni bilgiler geldikçe arttırımlı olarak öğrenme sürecini sürdürme prensibinden yola çıkılarak tasarlanmış sinir ağlarıdır. Öğrenme sürecinde o ana kadar öğrenilen bilgiyi durum bilgisi olarak saklarlar. Özetle yapısında döngü bulunduran bir sinir ağıdır. CNN'lerden farklı olarak veriyi bir defa da değil sıralı olarak iç döngüsü içinde işlerler [22]. Girdi boyutu ile uyumlu olacak biçimde kurulan model 32 birimli RNN'de Relu aktivasyon fonksiyonu kullanılmıştır. Çıkış katmanında ise sigmoid aktivasyon fonksiyonu tercih edilmiştir.

3.3.3 Uzun Ömürlü Kısa Dönem Bellek

RNN mimarilerinde her bir dönüşüm anında sisteme yeni ihmaller katılmakta ve verinin bir kısmı kaybedilmektedir. Bu durum uzun veri dizilerinde başarımın düşmesine neden olmaktadır. Kısa dönemli bellek problemlerinin aşılması amacı ile uzun ömürlü bellekler içeren farklı hücrelerden oluşan yeni sinir ağı mimarileri tanıtılmıştır. Bu geliştirim zaman içinde aşamalar halinde gerçekleştirilmiştir. Bu tarz uzun ömürlü bellek hücrelerinden birisi LSTM hücreleridir. Çalışma prensibi bakımından RNN'lerin gelişmiş bir versiyonudur. Geçmiş bilginin unutulmaması ve sonraki adımlara aktarılabilmesi durumu mümkün olmaktadır [23]. Çalışma kapsamında LSTM katmanlarından oluşan bir ağ oluşturulmuş ve 128 birimli hücre kullanılmıştır. RNN ile uyumlu olacak şekilde aktivasyon fonksiyonları seçilmiştir.

3.3.4 Evrimsel Sinir Ağları

CNN mimarileri bilgisayarlı görü uygulamalarında özellikle resim sınıflandırma uygulamalarında oldukça verimli bir biçimde kullanılmaktadır. Ayrıca CNN'ler dizi işleme problemlerinde de kullanılabilir. Bu durumda tek boyutlu dizilere ilaveten zamanı ikinci boyut olarak düşünmek gerekmektedir. Güncel çalışmalar incelendiğinde metin sınıflandırma çalışmalarında CNN mimarileri RNN mimarilerini geride bırakabilmektedir. CNN'ler resimlerde olduğu gibi bir boyutlu diziler üzerinde de bölgesel dizileri öğrenebilmektedir. Metinler içerisinde özel bir kısımda öğrenilen örüntü ve ilişkiler metnin farklı noktalarında da tekrar tanımlanabilir. Böylelikle CNN modelleri metinlerin morfolojisini öğrenebilmektedir [24]–[28].

Çalışma kapsamında metinleri oluşturan andıçların ilişki ve örüntülerini keşfetmek için modellenen evrimsel sinir ağımızın giriş katmanına ön-eğitilmiş GloVe modelinin çıktısı bağlanmıştır. Modellenen CNN, boyutları 3, 4 ve 5 olan üç adet paralel evrim katmanından oluşturulmuştur. Oluşturulan paralel katmanların her biri ayrı ayrı max-pooling katmanında birleştirilerek ayrı ayrı çıkarılan özelliklerin oluşturduğu bir öz nitelik

vektör elde edilir. Haberlerin sahte veya doğru şeklinde iki etiket ile sınıflandırılması görevi için binary cross-entropi kayıp fonksiyonu kullanılmıştır.

Tablo 2. Kullanılan Sinir ağları ve bazı parametreleri

Sinir Ağı Mimarisi	Optimizer	Kayıp Fonksiyonu	Çıkış Katmanı Aktivasyon Fonksiyonu
MLP	adam	binary_crossentropy	sigmoid
RNN	adam	binary_crossentropy	sigmoid
LSTM	adam	binary_crossentropy	sigmoid
CNN	adam	binary_crossentropy	sigmoid

Tablo 2 kullanılan algoritmaları ve parametrelerini içermektedir.

3.4. Performans Metrikleri

Bu çalışmada sahte haber tespiti için derin bağlamsal kelime gömülmeleri ve derin sinir ağlarının performanslarını değerlendirmek için hata matrisi ve alıcı-işletim karakteristiği (AİK) eğrilerinden yararlanılarak sonuçlar elde edilmiştir. Temel olarak dört değerden oluşturulan hata matrisi sınıflandırma sonucunda var olan durumlar için sınıflandırmanın doğru ve yanlış tahmin sayılarının verildiği tablodur. Bu dört değer doğru pozitif (DP), doğru negatif (DN), yanlış pozitif (YP) ve yanlış negatif (YN) değerleridir. Karışıklık matrisi değerlerine göre sınıflandırma işlemi performans değerlendirme için tanımlanmış bazı önemli metrikler Tablo 3’de özetlenmiştir. Hata matrisinden türetilen doğruluk, duyarlılık ve özgünlük metrikleri ilgili tabloda yer almaktadır.

Tablo 3. COVID-19 veri setine ait istatistiki bilgiler

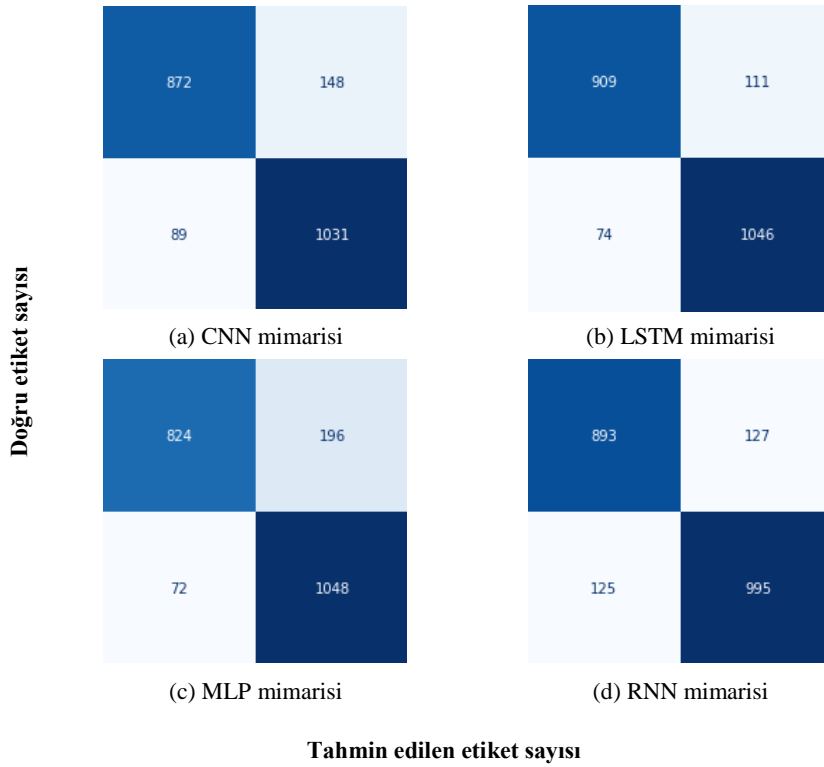
Değerlendirme Metriği	Formül	Tanım
Doğruluk	$\frac{DP + DN}{DP + DN + YP + YN}$	Tüm örnekler bazında doğru gerçekleşen tahmin oranıdır.
Duyarlılık	$\frac{DP}{DP + YN}$	Pozitif örnekler içinde doğru pozitif tahmin oranıdır.
Özgünlük	$\frac{DN}{DN + YP}$	Negatif örnekler içinde doğru negatif tahmin oranıdır.

AİK eğrileri ikili sınıflandırıcılar için yaygın olarak kullanılan araçlardan biridir. Gerçek pozitif oranı yanlış pozitif oranına göre elde edilen bu eğri deneysel çalışma boyunca başvuru bir başka araçtır. Bu araç ile yanlış pozitifleri azaltma doğru pozitiflerin tespit edilme kabiliyeti değerlendirilir. AİK değerinin artması sistemin tespit performansının iyileştiğini göstermektedir.

4. Bulgular ve Tartışma

Çalışmada kapsamında ön-eğitilmiş kelime gömülmelerinin sahte haber tespiti gibi zorlayıcı bir görev üzerindeki başarımı farklı derin öğrenme mimarileri üzerinde incelenmiştir. Bu kapsamda önerilen tüm modeller üst düzey yapay sinir ağları oluşturmak için imkan sağlayan keras kütüphanesi kullanılarak gerçekleştirilmiştir. Ek olarak sklearn, numpy gibi programlama kütüphaneleri de kullanılmıştır. Kullanılan ön-eğitilmiş kelime modeli GloVe, 6 milyon andıç, 400000 kelime ve 100 boyutlu gömülme vektörü olacak biçimde seçilmiştir.

Sahte haber tespiti için ikili sınıflandırma görevi biçiminde modellenen çalışmayı değerlendirmek için, sınıf sayıları bakımından oldukça dengeli veriler içeren COVID-19 [17] veri seti kullanılmıştır. Şekil 2 ve Tablo 4 GLoVe kelime gömülmeleri kullanılarak farklı sinir ağı mimarileri üzerinde sunulan modelin sınıf düzeyinde performanslarını sunmaktadır. Gerçekleştirilen deneysel süreçler neticesinde Şekil 2’de sunulan hata matrisleri elde edilmiştir. Tablo 4’de ise Şekil 2’deki hata matrislerinden elde edilen f-puan metrikleri rapor edilmiştir. Hata matrislerini oluşturan satır ve sütunlar sırasıyla doğru ve yanlış sınıflarını temsil etmektedirler. Tablo 4, COVID-19 veri seti üzerinde GloVe kelime gömülmeleri ile farklı modellerin sınıflandırma düzeyinde elde ettikleri başarımı göstermektedir.

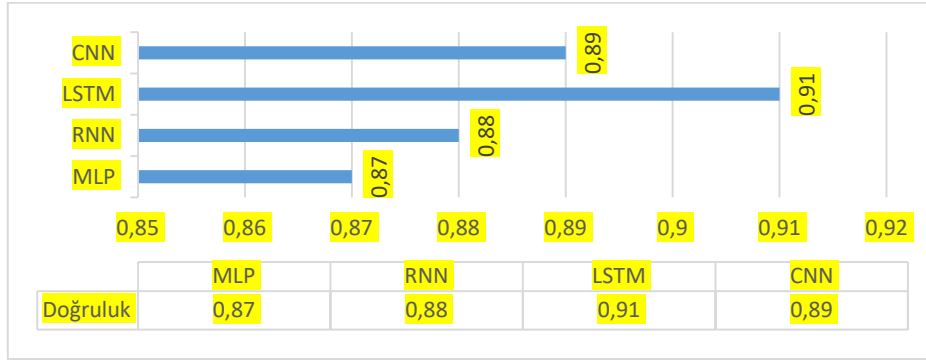


Şekil 2. Modellere ait hata matrisleri

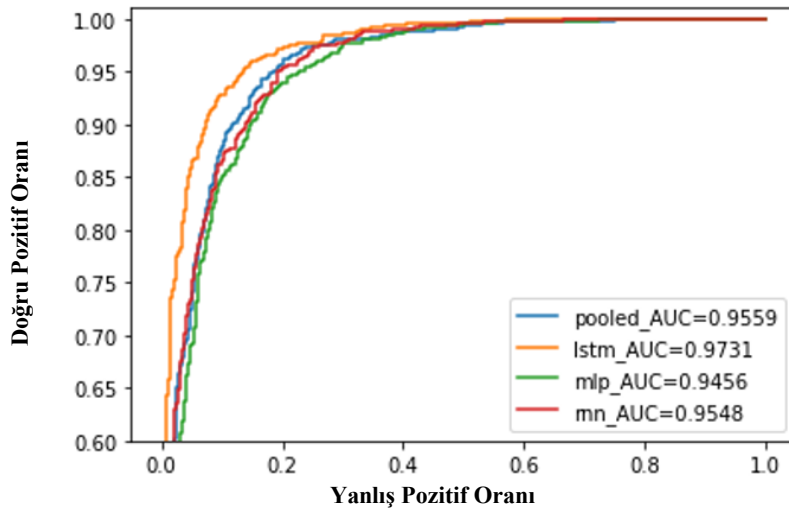
Tablo 4. Farklı mimarilerin, COVID-19 veri seti'nin test kümesinde ki her sınıf için hesaplanan f1-puanı

Mimari	F-puanı	
	Doğru Haber	Yanlış haber
MLP	0,89	0,86
RNN	0,89	0,88
LSTM	0,92	0,91
CNN	0,90	0,88

Tablo 4 incelendiğinde her iki sınıf için (doğru ve yanlış sınıfları) LSTM mimarisinin diğer yöntemleri geride bıraktığı gözlemlenmektedir. Ayrıca CNN'in ikinci en yüksek f-puan değerlerini elde ettiği görülmektedir. MLP ve RNN ağları birbirlerine oldukça yakın, 0.89 f-puanı elde edebilmişlerdir. Sınıflar bazında rapor edilen f-puanlarının birbirlerine oldukça yakın olduğu gözlemlenmektedir. Farklı sinirsel sınıflandırıcıların kavramsallaştırılmış ön-egitimli kelime gömümleri ile birlikte kullanılması durumunda elde edilen performans değerlerine Şekil 3'de yer verilmiştir. Sunulan görsel ve raporlanan sonuçlar incelendiğinde en yüksek performansı 0.91 sınıflandırma doğruluğu ile LSTM modelinin elde ettiği ve en düşük sonuçların ise 0.87 sınıflandırma doğruluğu ile MLP tarafından üretildiği görülmüştür. Raporlanan doğruluk performans değeri açısından yaklaşımlar arasında çok büyük farklılıklar söz konusu değildir. Her ne kadar veri setini oluşturan sınıfların kayıt sayıları son derece dengeli olsa da bulguları netleştirmek adına modellerin AİK eğrilerine Şekil 4'de yer verilmiştir. Buna göre, göre en başarılı sınıflandırma işlemi LSTM modeli tarafından gerçekleştirilmiştir. AİK eğrisinin altında kalan alan değeri 0.9731 olarak elde edilmiştir. AİK değerleri bakımından RNN ve CNN'ler yakın değerler elde etmişlerdir. AİK eğrisinin altında kalan alan bakımından en düşük başarıyı MLP sinir ağı elde etmiştir.



Şekil 3. Modellere ait performans sonuçları



Şekil 4. Modellere ait AİK eğrileri

5. Sonuç

Sahte haber tespit problemi sosyal medya ve iletişim araçlarının yaygınlaşması ile oldukça önemli ve zorlayıcı bir görev haline gelmiştir. COVID-19 salgını sırasında sahte haberlerin yıkıcı ve olumsuz etkileri ile mücadele etmek çok daha önemli hale gelmiştir. Katlanarak artan bilgi akış trafığı, sahte haberlerin tespiti için makine öğrenmesi metodlarının kullanılmasını kaçınılmaz kılmıştır. Yapay zeka temelli modellerin sahte haber tespiti görevinde yüksek başarıyı yakalaması modellerin kapsamlı veriler ile eğitilmesini gerektirir. Bu pratikte oldukça zorlayıcı bir iştir. Yeterli eğitim verisinin mevcut olmaması durumunda önceden oluşturulmuş ön eğitilmiş metin gömülmelerini kullanmak oldukça avantajlıdır. Son yıllarda, farklı DDİ görevlerinde Göreve özel olarak oluşturulan dilin ve görevin genel yapısını bilen ön-eğitilmiş modeller kullanılmaktadır. Ancak Derin sinir ağı ve kavramsal kelime gömülmelerinin birlikte kullanılması durumunda hangi sınıflandırıcı ile daha yüksek başarımlar elde edildiği hususu yeterince açık değildir. Çalışma kapsamında GloVe ön eğitilmiş kelime gömülme katmanının sağladığı bağlamsal temsiller kullanılmış ve farklı sınıflandırıcılar deneysel süreçte dahil edilmiştir. Deneysel çalışma ile sunulan modellerin hangi kombinasyonlarının daha verimli olduğunu ortaya koymak için çeşitli performans değerlendirme metrikleri kullanılmıştır. Ayrıca deneysel çalışmalarda açık erişimli ve oldukça yaygın olarak kullanılan COVID-19 sahte haber tespit veri seti kullanılmıştır. Veri seti ikili sınıflar içermektedir. Sınıflar bazında karşılaştırmaların rapor edildiği ilgili tabloda f-puanlarının birbirlerine oldukça yakın olduğu gözlemlenmektedir. Bu durumun veri setinin dengeli yapısı ile ilişkilendirmek mümkündür. İlgili veri setinde bulunan sınıflar Tablo 1’de görüldüğü üzere eşit sayıda ve eşit dağılımdadır. Çalışma kapsamında ön-eğitilmiş modellerin farklı özelliklere ve çalışma prensiplerine sahip olan sinir ağlarının hangileri ile daha verimli çalışarak yalan haber tespiti gibi güncel ve zorlu bir görevde ortaya konulan performansların ayrı ayrı değerlendirilmesi mümkün olmuştur. Sonuçlara göre kavramsallaştırılmış kelime gömülme dil modeli kısa haber metinlerinin başarılı bir şekilde sınıflandırılmasında oldukça büyük paya sahip olduğu değerlendirilmektedir. Haberlerin

içeriklerini ve bazı kavramları keşfetmek için kısa metinlerden özelliklerin çıkarılması büyük bir zordur ancak sonuçlara göre içerik bazlı dil modeli performans üzerinde olumlu etkiler oluşturmaktadır. Ayrıca ilgili tablo ve grafiklerden de görüleceği üzere LSTM mimarileri en yüksek sınıflandırma başarımına sahip olmuşlardır. Bu bağlamda bir dizi deneysel süreç yürütülerek metin gömülmeleri ve farklı sinirsel sınıflandırıcıların başarımlarına yönelik kapsamlı ve karşılaştırmalı bir çalışma ile alandaki eksiklik giderilmeye çalışılmıştır.

Kaynaklar

- [1] Hark C, Karıcı A. Karıcı summarization: A simple and effective approach for automatic text summarization using Karıcı entropy. *Information processing & management* 2020; vol. 57, no. 3: 102187.
- [2] Türk Dil Kurumu. Türk Dil Kurumu Sözlükleri., *Tdk*, 2021. <https://sozluk.gov.tr/> (accessed Nov. 01, 2021).
- [3] Pan JZ, Pavlova S, Li C, Li N, Li Y, Liu J. Content Based Fake News Detection Using Knowledge Graphs., 2018 669–683.
- [4] Gravanis G, Vakali A, Diamantaras K, Karadais P. Behind the cues: A benchmarking study for fake news detection. *Expert Syst. Appl.* 2019; vol. 128 201–213.
- [5] Kaliyar RK, Goswami A, Narang P, Sinha S. FNDNet – A deep convolutional neural network for fake news detection. *Cogn. Syst. Res.* 2020; vol. 61 32–44.
- [6] ShuKai, SlivaAmy, WangSuhang, TangJiliang, LiuHuan. Fake News Detection on Social Media. *ACM SIGKDD Explor. Newsl.* 2017; vol. 19, no. 1: 22–36.
- [7] Gilda S. Notice of Violation of IEEE Publication Principles: Evaluating machine learning algorithms for fake news detection.,in *2017 IEEE 15th Student Conference on Research and Development (SCORED)*,Dec. 2017,110–115.
- [8] Chollet *FDeep Learning with Python*. Buzdağı Yayınevi,2021.
- [9] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez AN. Attention is all you need. *Adv. neural Inf. Process. Syst.* 2017; 5998–6008..
- [10] Kaur S, Kumar P, Kumaraguru P. Automating fake news detection system using multi-level voting model. *Soft Comput.* 2020; vol. 24, no. 12: 9049–9069.
- [11] Verma C, Stoffova V, Illes Z, Tanwar S, Kumar N. Machine Learning-Based Student’s Native Place Identification for Real-Time. *IEEE Access* 2020; vol. 8 130840–130854.
- [12] Ahmed H, Traore I, Saad S. Detecting opinion spams and fake news using text classification. *Secur. Priv.* 2018; vol. 1, no. 1: e9.
- [13] Zhang J, Dong B, Yu PS. FakeDetector: Effective Fake News Detection with Deep Diffusive Neural Network.,in *2020 IEEE 36th International Conference on Data Engineering (ICDE)*,Apr. 2020,1826–1829.
- [14] Bhattacharya P, Patel SB, Gupta R, Tanwar S, Rodrigues JJPC. SaTYa: Trusted Bi-LSTM-Based Fake News Classification Scheme for Smart Community. *IEEE Trans. Comput. Soc. Syst.* 2021; 1–10.
- [15] Mansouri R, Naderan-Tahan M, Rashti MJ. A Semi-supervised Learning Method for Fake News Detection in Social Media.,in *2020 28th Iranian Conference on Electrical Engineering (ICEE)*,Aug. 2020,1–5.
- [16] Mahabub A. A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers. *SN Appl. Sci.* 2020; vol. 2, no. 4: 525.
- [17] Patwa P *et al.* Fighting an Infodemic: COVID-19 Fake News Dataset., 2021 21–29.
- [18] Fu X, Liu W, Xu Y, Cui L. Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis. *Neurocomputing* 2017; vol. 241 18–27.
- [19] Zhou X, Zafarani R. A Survey of Fake News. *ACM Comput. Surv.* 2020; vol. 53, no. 5:
- [20] Mikolov T, Chen K, Corrado G, Dean J. Efficient Estimation of Word Representations in Vector Space. 2013; [Online]. Available: <http://arxiv.org/abs/1301.3781>
- [21] Samadi M, Mousavian M, Momtazi S. Deep contextualized text representation and learning for fake news detection. *Information processing & management* 2021; vol. 58, no. 6: 102723.
- [22] Aydoğan M, Kocaman V. TRSAv1: A new benchmark dataset for classifying user reviews on Turkish e-commerce websites. *J. Inf. Sci.* 2022; 016555152210743.
- [23] KARACA YE, ASLAN S. Sentiment Analysis of Covid-19 Tweets by using LSTM Learning Model. *Comput. Sci.* 2021; no. Special: 366–374.
- [24] Seyyarer E, Uckan T, Hark C, Ayata F, Inan T, Karci A. Applications and Comparisons of Optimization Algorithms Used in Convolutional Neural Networks.,in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*,Sep. 2019,1–6.
- [25] KARACA YE, ASLAN S, HARK C. SEPARATION OF DOMESTIC WASTE WITH DEEP LEARNING TECHNIQUES. *Comput. Sci.* 2022; vol. 7, no. 1: 10–19.
- [26] SEYYARER E, KARCI A, ATEŞ A. Stokastik ve deterministik hareketlerin optimizasyon süreçlerindeki etkileri. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Derg.* 2021; vol. 37, no. 2: 949–966.
- [27] SÖYLEMEZ ÖF, ERGEN B. Farklı Evrimsel Sinir Ağı Mimarilerinin Yüz İfade Analizi Alanındaki Başarımlarının

- İncelenmesi. *DMF Mhendislik Derg.* 2020; vol. 11, no. 1: 123–133.
- [28] Tumen V, Soylemez OF, Ergen B. Facial emotion recognition on a dataset using convolutional neural network.,in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*,Sep. 2017,1–5.