



LSTM VE BERT MODELLERİ İLE SAHTE HABER TESPİTİ

*Hilal KARTAL ÇOKAL¹, Mevlüt ERSOY²

¹Afyon Kocatepe Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü,
Afyonkarahisar

²Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Isparta

(Geliş/Received: 13.05.2025, Kabul/Accepted: 11.11.2025, Yayınlanma/Published: 26.12.2025)

ÖZ

Dijitalleşen dünyada bilgiye erişimin kolaylaşması ile birlikte bilgi kirliliği büyük bir sorun haline gelmiştir. Sosyal medya platformlarının geniş kitlelere ulaşma potansiyeli ve kullanıcılar tarafından denetimsiz bir şekilde içerik üretilebilmesi, sahte haberlerin yayılımını daha da kolaylaştırmıştır. Özellikle doğruluğu teyit edilmemiş içeriklerin hızla yayılması, bireylerin yanlış bilgilendirilmesine ve toplumsal algıların manipüle edilmesine neden olabilmektedir. Bu tür haberler; bireylerin davranışlarını, kamuoyunu etkileyebilecek bir etki düzeyine sahiptir. Sahte haberlerin yayılması yalnızca bireylerin bilgiye olan güvenini sarsmakla kalmayıp aynı zamanda toplumda kutuplaşma, panik ve yanlış yönlendirme gibi sosyal sorunlara da zemin hazırlamaktadır. Bu nedenle, sahte haberlerin erken tespiti bu sorunların önüne geçilebilmesi için büyük önem taşımaktadır. Gelişen yapay zekâ ve doğal dil işleme teknikleri, bu sorunun tespitine yönelik etkili çözümler sunmaktadır. Bu çalışmada sahte haber tespiti için LSTM ve BERT modelleri kullanılmıştır. Çalışmada gerçek ve sahte haberlerden oluşan iki veri seti dengeli bir şekilde birleştirilerek tek veri setine haline getirilmiştir. Bu veri setinde bulunan verilerin %80'i eğitim ve %20'si test verileri olarak kullanılmıştır. LSTM modeli ile %93 oranında bir başarı elde edilirken BERT modeli ile bu başarı %98 olarak elde edilmiştir.

Anahtar kelimeler: Sahte haber tespiti, LSTM, BERT

FAKE NEWS DETECTION WITH LSTM AND BERT MODELS

ABSTRACT

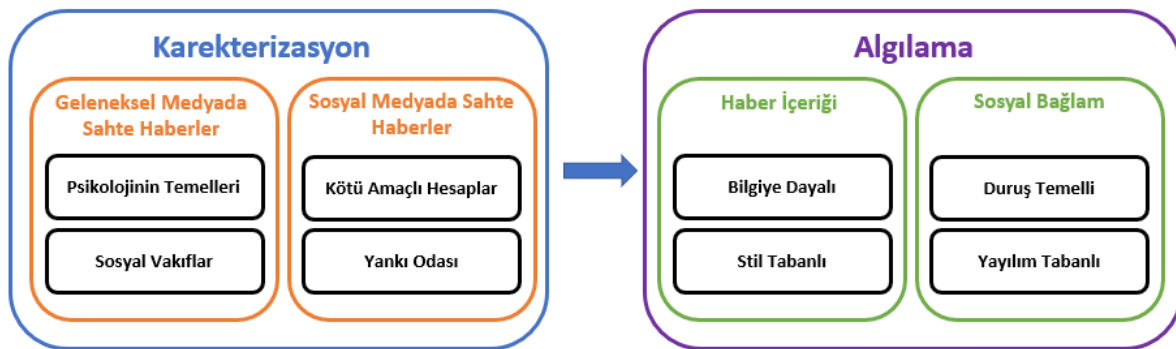
In the digitalised world, information pollution has become a major problem with the ease of access to information. The potential of social media platforms to reach large masses and the ability of users to produce content in an uncontrolled manner have further facilitated the spread of fake news. In particular, the rapid spread of unconfirmed content may cause individuals to be misinformed and social perceptions to be manipulated. Such news has a level of influence that can affect the behaviour of individuals and public opinion. The spread of fake news not only undermines individuals' trust in information, but also paves the way for social problems such as polarisation, panic and misdirection in society. Therefore, early detection of fake news is of great importance to prevent these problems. Developing artificial intelligence and natural language processing techniques offer effective solutions to this problem. In this study, LSTM and BERT models are used for fake news detection. In the study, two data sets consisting of real and fake news were combined in a balanced way and turned into a single data set. In this data set, 80% of the data were used as training data and 20% as test data. While 93% success was achieved with the LSTM model, 98% success was achieved with the BERT model.

Keywords: Fake news detection, LSTM, BERT

1. Giriş (Introduction)

Sahte haber, genellikle kamuoyunu yanıltmak amacıyla kasıtlı olarak oluşturulan ve doğruluğu teyit edilmemiş ya da tamamen yanlış bilgiler içeren içerikler olarak tanımlanmaktadır [1]. Siyasi manipülasyonlardan halk sağlığını tehdit eden dezenformasyonlara kadar geniş bir yelpazede etkili olan sahte haberler, modern çağın en karmaşık problemlerinden biri haline gelmiştir. Sahte haberlerin yaygınlaşması, bireyler ve toplum üzerinde son derece olumsuz etkiler yaratma potansiyeline sahiptir. Bu nedenle, sosyal medyada sahte haber tespiti son zamanlarda büyük ilgi gören bir araştırma haline gelmiştir. Sosyal medyada sahte haber tespiti, geleneksel haber medyasındaki mevcut tespit algoritmalarını etkisiz veya uygulanamaz hale getiren benzersiz özellikler ve zorluklar sunmaktadır. Sahte haberler, bağlamdan koparılmış ifadeler, çarpıtılmış veriler ve kimi zaman görsel manipülasyonlarla desteklenen unsurlar içermektedir. Bu yönüyle sahte haberler, bilgiye erişim kolaylığının olduğu bir dijital çağda, bireylerin eleştirel düşünme becerilerini sıyanan içerikler haline gelmektedir. Bu tür içerikler, çoğu zaman dikkat çekici başlıklar ve duygusal ifadelerle yapılandırılarak, bireyleri hızlıca etkilemeyi ve içeriğin yayılmasını sağlamayı hedeflemektedir [2]. İnternet ve sosyal medya ortamlarında yapılan haberler üzerindeki denetim mekanizmalarının zayıflaması, bilgi ve içeriklerin hızla paylaşılma isteğiyle birleştiğinde, yeni medya alanında içeriklerin yeterince doğrulanmadan ve olgunlaştırılmadan yayımlanması gibi önemli bir etik sorunu gündeme getirmektedir [3]. Sosyal medya platformlarının algoritmik yapısı, sahte haberlerin yayılımını hızlandıran başlıca faktörlerden biridir. Kullanıcıların geçmiş etkileşimlerine göre içerik öneren algoritmalar, benzer düşüncelerin tekrar tekrar karşısına çıkmasına neden olmakta ve bu durum “yankı odası” (echo chamber) ve “bilgi balonu” (filter bubble) olarak adlandırılan bilgi kapanlarını doğurmaktadır [4,5]. Bu tür ortamlarda sahte haberler, bireyler tarafından daha kolay benimsendiği gibi, doğruluğu sorgulanmadan paylaşılma eğilimi de artmaktadır. Ayrıca, otomatik hesaplar (botlar) ve organize trol ağları gibi dijital araçlar sayesinde sahte haberlerin dolaşıma sokulması ve geniş kitlelere ulaşması, artık yalnızca bireysel değil sistematik bir sorun haline gelmiştir. Bu sorunun çözümünde en çok umut vadeden alanlardan biri ise yapay zekâ teknolojileridir. Doğal dil işleme, makine öğrenmesi ve derin öğrenme gibi alt disiplinler, sahte haberlerin tespiti konusunda çeşitli çözümler sunmaktadır.

Sahte haberlerin sınıflandırılması, içerik yapılarının ve yayılma amaçlarının anlaşılmasını gerektirmektedir. Bazı içerikler yanlış bilginin istemsiz aktarımına dayanırken, bazıları bilinçli dezenformasyon stratejilerinin bir parçasıdır. Bu ayırım, sahte haber tespitinde yapay zekâ tabanlı yöntemlerin etkinliğini doğrudan etkileyen bir faktördür. Zira içerik temelli analiz yöntemleri dil ve yapı üzerinden çalışırken, kasıtlı dezenformasyonun arkasındaki bağlamsal ve niyet temelli katmanları yakalamakta zorlanabilmektedir. Bu nedenle sahte haberlerin yapısal özellikleri kadar, bu yapıların arkasındaki motivasyonların da anlaşılması önemlidir [6].



Şekil 1. Sosyal medyada sahte haberler (Fake news on social media) [6]

Şekil 1, sosyal medyada sahte haber olgusunun hem karakterizasyon hem de tespit aşamalarını bütüncül bir çerçevede sunmaktadır. Karakterizasyon aşamasında, sahte haberlerin geleneksel medyada psikolojik ve sosyal temellere dayandığı; sosyal medyada ise kötü niyetli hesaplar ve yankı odası etkisiyle şekillendiği görülmektedir. Tespit süreci ise iki ana ekseninde ilerlemektedir: haber içeriği ve sosyal bağlam. Haber içeriği, bilgiye dayalı doğrulama ve dil-üslup analizi yöntemleriyle incelenirken; sosyal bağlam, paylaşılan haberin kullanıcı tutumları ve yayılma biçimi üzerinden analiz edilmektedir. Bu yapı, sahte haberle mücadelede multidisipliner bir yaklaşımın gerekliliğine işaret etmektedir. Bu durumda sahte haber niteliği taşımayan haberler arasında aşağıdaki haber türleri dahil edilebilmektedir:

- Eğlence Odaklı Haberler: Kullanıcıyı yanıltma amacı taşımayan, gerçek dışı olduğu kolayca belirlenebilen, eğlence amaçlı haberlerdir.
- Söylentiler: Haber niteliği taşımayan, gerçekliği teyit edilememiş bilgiler.
- Komplo Teorileri: Gerçek veya yanlış olduğu kesin olarak kanıtlanamayan spekülasyona dayalı bilgiler.
- İstem Dışı: Kasti olmayan yanlış bilgiler.
- Şaka/Dolandırıcılık İçerikli Haberler: Eğlence veya dolandırıcılık amacıyla üretilen gerçek dışı haberler [7]

Sahte haberlerin yayılım hızındaki artış, geleneksel doğrulama mekanizmalarının yetersiz kalmasına yol açmaktadır. Bu nedenle, sahte haberlerin otomatik olarak tespit edilmesine yönelik yapay zekâ tabanlı yöntemlere olan ilgi son yıllarda önemli ölçüde artmıştır. Yapay zekâ sistemleri, özellikle makine öğrenmesi ve doğal dil işleme teknikleri sayesinde haber metinlerini analiz edebilmekte ve içeriğin doğruluğu hakkında çıkarımlarda bulunabilmektedir [8]. Metin tabanlı tespit yöntemleri, genellikle haber başlığı ve içeriğindeki dilsel özelliklerin analizine dayanmaktadır. Bu yaklaşımlar, kelime frekansı, sözdizimi, duygusal tonlama ve anlamsal tutarlılık gibi özellikleri kullanarak sahte ve gerçek haberleri ayırt etmeye çalışır [9]. Doğal dil işleme tabanlı yöntemler arasında özellikle "word embeddings" (Word2Vec, GloVe) ve derin öğrenmeye dayalı modeller (LSTM, GRU, BERT) öne çıkmaktadır. Transformer tabanlı modellerin kullanımı, özellikle bağlamı anlamada sağladığı üstün performans ile sahte haber tespitinde yeni bir çığır açmıştır [10].

Sahte haber tespitinde yalnızca metin değil, görsel ve çoklu ortam (multimodal) içerikler de analiz edilmektedir. Özellikle sosyal medyada yer alan sahte haberler, sıkça manipüle edilmiş görseller ya da uydurma videolar ile desteklenmektedir. Bu bağlamda, görüntü işleme ve video analizine dayalı derin öğrenme yöntemleri de kullanılmaktadır [11]. Örneğin, Convolutional Neural Networks (CNNs) görsel anormallikleri yakalamada etkili olurken, haber metni ile görsel arasındaki tutarsızlıkları analiz eden çok modlu modeller sahte içerik tespitinde önemli katkı sağlamaktadır [12]. Bununla birlikte, bazı yaklaşımlar yalnızca içeriğe değil, haberin sosyal bağlamına da odaklanmaktadır. Bu tür yöntemler, bir haberin kim tarafından, ne zaman, nasıl ve kimler tarafından yayıldığını analiz ederek, sosyal etkileşim ağları üzerinden sahte içerik sinyalleri üretmektedir [13]. Bu yaklaşım, özellikle sosyal medya platformlarında yayılan dezenformasyonun tespiti açısından büyük önem taşımaktadır.

Yapay zekâ tabanlı sahte haber tespitinde kullanılan algoritmalar zaman içerisinde önemli bir evrim geçirmiştir. Başlangıçta Naive Bayes, Support Vector Machines (SVM) ve Random Forest gibi klasik makine öğrenmesi yöntemleri, özellikle TF-IDF ve n-gram temsilleriyle birlikte, sahte haberlerin dilsel özelliklerine dayalı olarak sınıflandırma amacıyla yaygın biçimde kullanılmıştır [14]. Bu yöntemler, ISOT gibi temel veri setleri üzerinde önemli ölçüde başarı sağlamış olsa da, bağlamsal ilişkileri yakalama konusunda sınırlı kalmışlardır. Bu eksiklikler, daha sonra Long Short-Term Memory (LSTM) ve Gated Recurrent Unit (GRU) gibi derin öğrenme yaklaşımları ile giderilmeye çalışılmıştır. Ruchansky [15] tarafından geliştirilen CSI modeli, yalnızca haber içeriğini değil, aynı zamanda sosyal etkileşim verilerini ve yazar güvenilirliğini de dikkate alan bir LSTM mimarisine öne çıkmıştır. Son yıllarda ise dil modelleme alanında devrim yaratan Transformer tabanlı yaklaşımlar, özellikle BERT ve RoBERTa gibi modeller aracılığıyla sahte haber tespitinde anlamlı ilerlemeler sağlamıştır [10,16]. Kaliyar [17] tarafından geliştirilen FakeBERT modeli, hem LIAR hem de ISOT veri setlerinde klasik yöntemleri geride bırakarak dikkat çekici başarılar elde etmiştir. Bununla birlikte, görsel ve metin içeriğini birlikte analiz eden çok modlu modeller ve haberin sosyal medya üzerindeki yayılım dinamiklerini grafik yapılar aracılığıyla işleyen Graph Convolutional Networks gibi yöntemler [18] de sahte haberin çok boyutlu doğasını yansıtmaya potansiyeli açısından önem taşımaktadır. Tablo 1.'de literatürde incelenen son 3 yılda sahte haber tespitine yönelik yapılan çalışmalar, kullanılan yöntemler ve veri setleri verilmektedir.

Tablo 1. Son 3 yılda sahte haber tespitine yönelik yapılan çalışmalar, kullanılan yöntemler ve veri setleri (Studies, employed methods and datasets devoted to fake news detection of last three years)

Yıl	Yöntem	Teknikler & Veri Setleri	Performans
2023 [19]	MMFN	Çok modlu füzyon ağı (Multi-grained Multi-modal Fusion Network) önerilmiş ve Weibo, Twitter, ve Gossipcop veri setleri kullanılmıştır.	MMFN modeli ile Weibo veri setinde 0.92, Twitter veri setinde 0.93 ve Gossipcop veri setinde 0.89 doğruluk elde edilmiştir.
2023 [20]	COOLANT	Görüntü-metin hizalaması elde etmeyi amaçlayan, çok modlu sahte haber tespiti için çapraz modal kontrastlı bir öğrenme çerçevesi olan COOLANT'ı önerilmiş ve Twitter, Weibo veri setleri kullanılmıştır.	COOLANT modelinde Twitter veri setinde 0.9 ve Weibo veri setinde 0.92 doğruluk elde edilmiştir.
2023 [21]	DeBERTa, Swinv2, Factify 2	Veri seti Twitter, Doğruluk Kontrolü web siteleri ve Satirik haber web sitelerinden elde edilen verilerin birleştirilmesiyle oluşturulmuştur.	Metin için önceden eğitilmiş DeBERTa ve görüntü katıştırılmaları için Swinv2 içeren model ile 81.82% doğruluk elde edilmiştir.
2023 [22]	MisRoBÆRTa	9,4 milyondan fazla veri içeren FAKENEWSCORPUS veri seti kullanılmıştır.	Genel olarak en iyi performans gösteren model %92,50 doğruluk oranıyla yeni MisRoBÆRTa olurken, genel olarak en iyi performans gösteren transformatör modeli %91,94 doğruluk oranıyla BART large olmuştur.
2023 [23]	CNN, RNN, Hibrit CNN+RNN	44898 veriden oluşan ISOT veri seti ve 804 başlıktan oluşan FA-KES veri seti kullanılmıştır.	FA-KES veri seti ile CNN 0.5, RNN 0.5 Hibrid CNN+RNN 0.6 doğruluk elde edilmiş, ISOT veri setinde ise CNN 0.99, RNN 0.98 ve Hibrid CNN+RNN ile 0.99 doğruluk elde edilmiştir.
2023 [24]	BERT, ALBERT, RoBERTa	BERT, ALBERT, RoBERTa modelleri ile sahte haber tespiti yapılmıştır.	ALBERT modeli ile 87.6%, RoBERTa ile 83.3%, IndoBERT ile 86.6%, BERT-Multilingual ile 78% doğruluk elde edilmiştir.
2024 [25]	AraBERT, MARBERT, AraELECTRA, ARBERT, AraGPT2	Arabic Multisource Fake News Detection (AMFND) veri seti kullanılarak modeller test edilmiştir.	AraBERT v2 modeli ile 0.85, MARBERT ile 0.89, AraELECTRA ile 0.93, AraGPT2 ile 0.90 ve ARBERT ile 0.77 doğruluk elde edilmiştir.
2024 [26]	BERT, RoBERTa, Distil-BERT	GPT-4 ile etiketlenmiş ve güvenilirliği sağlamak için insan uzmanlar tarafından doğrulanmış haber makalelerinden oluşan bir veri seti oluşturulmuş ve BERT modelleri ile test edilmiştir.	Batch-size 16 için BERT ile 88.16%, RoBERTa ile 92.12% ve Distil-BERT ile 85.16% doğruluk elde edilmiştir.
2024 [27]	BERT, BiLSTM	BERT embedding + dikkat mekanizması ile ISOT, LIAR, TI-CNN veri setlerinde test edilmiştir.	ISOT, LIAR ve TI-CNN doğrulukları %99,47, %83,38 ve %87,90 olarak elde edilmiştir.
2025 [28]	UNITE-FND framework	Uni-Fakeddit-55k veri seti genişletilmiş ve UNITE-FND ile görüntü-metin dönüşümlü sahte haber tespiti gerçekleştirilmiştir.	Uni-Fakeddit-55k veri setinde, UNITE-FND ile 92.52% doğruluk elde edilmiştir.
2025 [29]	SVM, Lojistik Regresyon, Rastgele Orman, k-NN, Karar Ağacı ve Naive Bayes	TRT ve teyit.org dan elde edilen 5325 kayıttan oluşan veri seti kullanılmıştır.	Hiperparametre ayarı ve test doğruluğu ile TF-IDF yönteminin SVM modelinde %93.12'lik en yüksek doğruluk oranını verdiğini ve TF-IDF'nin daha etkili olduğu sonucu elde edilmiştir.
2025 [30]	TFIDF+RF, BiLSTM, XLNet, BERT-Base, RoBERTa	62308 veriden oluşan WELFake news veri seti kullanılmıştır.	TFIDF+RF'nin F 1 puanı 0,891, BiLSTM,modelinde ise 0.905, XLNet ile 0.927, BERT-Base ile 0.922 ve RoBERTa ile 0.930 elde edilmiştir.

Yapay zekâ tabanlı sahte haber tespiti sistemlerinin başarısı, büyük ölçüde kullanılan veri setlerinin kalitesi ve çeşitliliğine bağlıdır. Bu alanda yaygın olarak kullanılan veri setleri arasında LIAR, FakeNewsNet, BuzzFeedNews, PolitiFact ve ISOT Fake News Dataset gibi açık kaynaklı koleksiyonlar yer almaktadır [13,31]. Bu veri setleri genellikle haber metinleri, başlıklar, etiket bilgileri (doğru/yanlış) ve bazı durumlarda sosyal medya etkileşimleri içermektedir. Tablo 1. 'de literatürde genellikle kullanılan sahte haber veri setleri ve veri sayıları ile sınıfları verilmektedir.

Tablo 2. Literatürde kullanılan sahte haber veri setleri (Datasets of using fake news on literature)

Veri Seti	Veri Sayısı	Sınıf Sayısı
LIAR [31]	12.8K	6
FEVER [32]	185,445	3
COVID-19 [33]	10,700	2
PHEME [34]	6425	2
CRED BANK [35]	4856	2
Twitter [36]	1111	2
Weibo [36]	816	2
PolitiFact [37]	488	2
FakeNewsNet [38]	çeşitli	2
ISOT [14]	44898	2
George McIntire (GM) [39]	14151	2
Fakeddit [40]	1063106	2,3,6
MediaEval15 [41]	15821	2

Sahte haber tespiti çalışmalarında genellikle, dilsel çeşitlilik ve kültürel farklar gibi önemli zorluklarla karşılaşmaktadır. Dil yapıları, kelime seçimleri ve deyimsele ifadeler farklı dillerde farklılık gösterdiği için, bir dilde doğru şekilde eğitilmiş bir modelin başka bir dilde etkili olamayabilir [7]. Kültürel farklılıklar da bu sorunda önemli bir rol oynamaktadır. Bir toplumda sahte olarak algılanan bir haber, diğer kültürlerde doğru kabul edilebilir, bu da sınıflandırma algoritmalarının doğruluğunu etkileyebilmektedir [42]. Gerçek zamanlı sahte haber tespiti, özellikle haberlerin hızla yayıldığı sosyal medya platformlarında önemli bir zorluk yaratmaktadır. Veri akışının sürekli ve anlık olması, modelin doğru sonuçlar verebilmesi için büyük bir problem teşkil etmektedir [15]. Ayrıca, adversarial saldırılar [43] sahte haberlerin algoritmalar tarafından doğru şekilde tespit edilmemesine yol açabilmektedir. Algoritmik yanlışlık da bu alandaki önemli sınırlamalardan biridir; veri setlerinin önyargılı olması, tespit sistemlerinin yanıltıcı sonuçlar üretmesine neden olabilmektedir.

2. Materyal ve Metot (Material and Method)

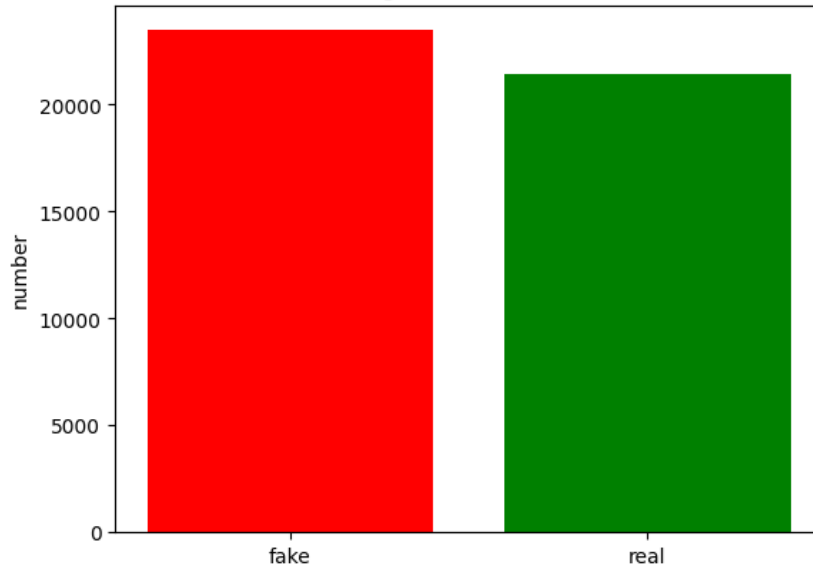
2.1. Veri seti (Dataset)

Bu çalışma kapsamında gerçek ve sahte haberlerden oluşan iki farklı veri seti birleştirilerek kullanılmıştır [44]. İlgili veri setleri herhangi bir lisans kısıtlaması içermemektedir. Birleştirilen veri seti, içerik türü bakımından iki farklı sınıfa ayrılmış makalelerden oluşmaktadır: gerçek haberler ve sahte haberler. Gerçek haber içerikleri, Reuters.com adlı haber sitesinden derlenmiş makalelerden oluşmaktadır. Öte yandan, sahte haber içerikleri ise Politifact (Amerika Birleşik Devletleri merkezli bir doğruluk kontrol kuruluşu) ve Wikipedia tarafından güvenilir olmadığı belirtilen çeşitli internet sitelerinden elde edilmiştir. Veri seti; politika, dünya gündemi gibi çeşitli konulara odaklanan farklı türde haber verileri içermektedir. Şekil 2.'de veri setinde bulunan haber verileri örneği verilmektedir. Her bir haber verisi; başlık (title), içerik metni (text), tür (subject) ve yayın tarihi(date) olmak üzere dört temel bilgi alanı içermektedir.

	title	text	subject	date
0	Trump Won't Be At The Fox News Debate, Undero...	Well, it looks like Republican presidential fr...	fake	January 26, 2016
1	Trump Is Talking To FOX News About Appearing O...	We re thinking Trump shouldn t give the rating...	fake	Aug 10, 2015
2	Obama to deliver statement on the economy: Whi...	WASHINGTON (Reuters) - President Barack Obama ...	real	May 6, 2016
3	WATCH: NY TOWN REJECTS Pledge Of Allegiance..."T...	Remember when it wasn t acceptable to be an an...	fake	Mar 21, 2016
4	UK's Johnson makes fuller apology for remarks ...	LONDON (Reuters) - British Foreign Secretary B...	real	November 13, 2017

Şekil 2. Veri setinde bulunan haber verileri örneği (Example of news datas based on datasets)

Modelin eğitimi ve test süreci birleştirilmiş veri seti kullanılarak gerçekleştirilmiştir. Şekil 2.'de Gerçek ve Sahte haberlerin veri sayısı grafiği verilmektedir. Gerçek haberlerden oluşan veri setinde 21417 veri, sahte haberlerde ise 23481 veri bulunmaktadır.



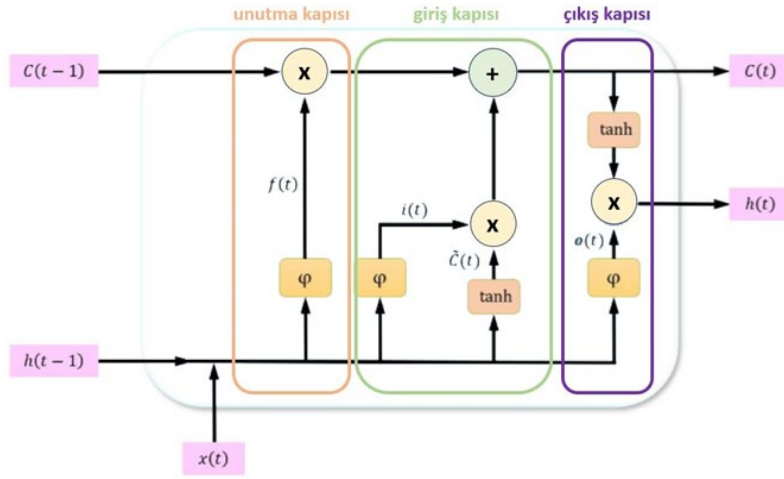
Şekil 3. Gerçek ve Sahte haberlerin veri sayısı grafiği (Graph of data numbers of real and fake news)

Veri seti Google Colab ortamına yüklenmiş ve pandas kütüphanesiyle okunmuştur. Veri kümesindeki "subject" sütunundaki etiketler "fake ve real" olmak üzere yeniden tanımlanmıştır. Veri ön işleme aşamalarında;

- Tüm metinler küçük harfe dönüştürülerek sözcük biçimleri standartlaştırılmıştır.
- Noktalama işaretleri, sayılar ve özel karakterler kaldırılmıştır.
- Anlamsız kelimeler çıkartılmıştır.
- Metinler, modelin girdi olarak alabileceği biçimde sayısal dizilere dönüştürülmüştür.

2.2. LSTM modeli (LSTM model)

Tekrarlayan Sinir Ağlarının (RNN) uzun vadeli bağımlılık problemini çözmek için tasarlanan LSTM [45] bir bellek hücresi, bir giriş kapısı, bir çıkış kapısı ve bir unutma kapısından oluşmaktadır. Şekil 3.'te LSTM hücresi modeli verilmektedir. Hücre, rastgele zaman aralıklarında değerleri hatırlar ve üç kapı, hücreye giren ve çıkan bilgi akışını düzenlemektedir [46]. Bu farklı kapılar sayesinde LSTM uzun vadeli bağımlılıkları daha etkin bir şekilde ele alabilir. Bu özellik dil işleme, metin sınıflandırma ve zaman serisi analizi gibi problemlerde daha iyi performans sağlar. LSTM, RNN'e göre daha derin ve karmaşık bir yapıya sahiptir ancak bu ek bileşenler sayesinde daha güçlü bir hafıza ve daha verimli bilgi akışı sağlamaktadır.

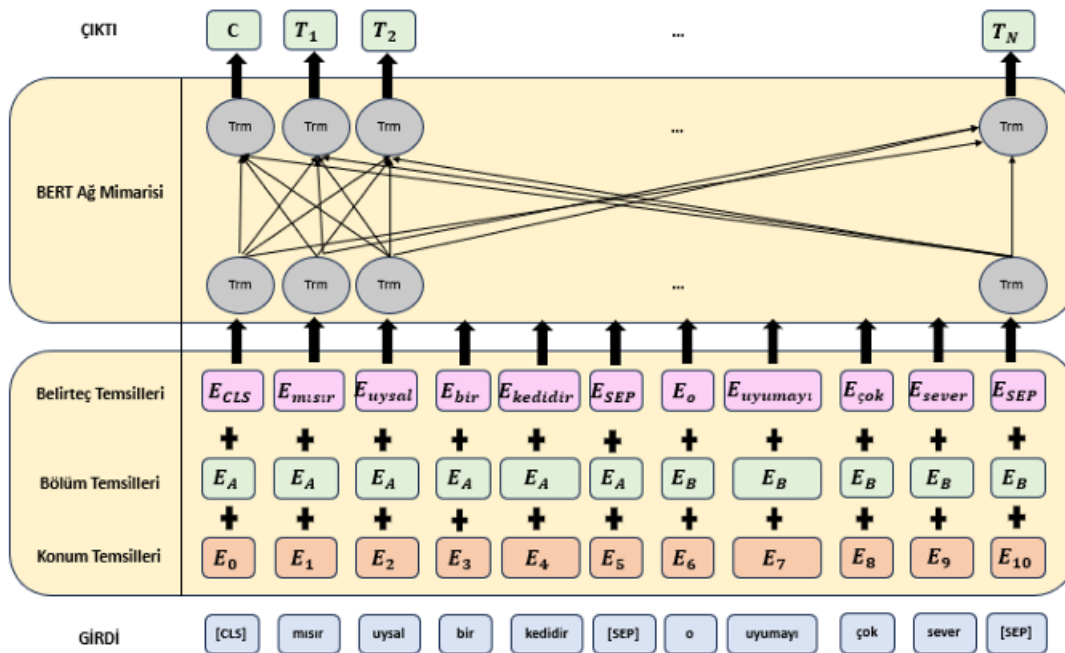


Şekil 4. LSTM hücresi (Cell of LSTM)[47]

LSTM hücresine girdileri $x(t)$ ve önceki durumu $h(t-1)$ uzun süreli bellek hücre durumu olarak da adlandırılır ve uzun süreli durumların hangi kısımlarının unutulması gerektiğini tanımlayan unutma kapısı $f(t)$ tarafından kontrol edilmektedir. Giriş kapısı, $i(t)$, uzun süreli durumlara hangi parçaların eklenmesi gerektiğini kontrol ederken, çıkış kapısı, $o(t)$, mevcut zaman durumlarında hangi parçaların üretilmesi gerektiğini kontrol etmektedir. $x(t)$ giriş sinyaline karşılık gelmektedir, $C(t)$ bellek hücresinin mevcut durumunu ve $C(t-1)$ bellek hücresinin önceki anını, $h(t)$ mevcut gizli durumunu ve $h(t-1)$ önceki gizli durumunu temsil etmektedir [48].

2.3. BERT modeli (BERT model)

BERT (Bidirectional Encoder Representations from Transformers), çok çeşitli doğal dil işleme görevlerini yerine getirebilen bir sinir ağıdır. BERT modeli, derin çift yönlü ön eğitim elde etmek için RNN'lere yeni ve güçlü bir alternatif olan Transformer mimarisini kullanmaktadır [49]. Transformer, birçok doğal dil işleme görevi için yeni bir mimari sağlamaktadır. CNN'ler ve RNN'ler eğitim seti için büyük ve etiketli veri setlerine ihtiyaç duyduğundan, Transformer matematiksel bir yaklaşımla örüntüleri tespit ederek bu gereksinimi ortadan kaldırmaktadır [50].



Şekil 5. BERT model mimarisi (Architect of BERT model) [51]

Şekil 3.'te BERT model mimarisi verilmektedir. Ağ mimarisi katmanında dönüştürücüler ifade edilmektedir. Bu katmanda bulunan dönüştürücüler, modelin metin verilerini işleyip anlamlandırmasını

sağlayan karmaşık bir sinir ağı yapısıdır. BERT modelinin çift yönlü olması, her kelimenin hem solundaki hem de sağındaki kelimelerle etkileşimini göz önünde bulundurmasına olanak tanımaktadır.

2.4. Değerlendirme metrikleri (Assessment metrics)

Sınıflandırma ve tespit görevlerinde model performanslarını değerlendirme metriklerinde çoğunlukla pozitif ve negatif olmak üzere iki sınıf tanımlanmaktadır. Buna göre, tahmin edilen sonuçlar pozitif sınıfta yer almakta ve doğru biçimde pozitif olarak sınıflandırılmakta ise doğru pozitif (true positive, TP) olarak nitelendirilmektedir. Gerçekte pozitif sınıfa ait olup model tarafından reddedilen örnekler yanlış negatif (false negative, FN) olarak değerlendirilmektedir. Benzer şekilde, negatif sınıfta yer almakta olup pozitif olarak sınıflandırılan örnekler yanlış pozitif (false positive, FP) olarak tanımlanmakta, diğer örnekler ise doğru negatif (true negative, TN) olarak sınıflandırılmaktadır [52].

Her iki modelin de performansını değerlendirebilmek için sınıflandırma görevleri için yaygın olarak kullanılan [53] doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skor (F1-score) değerlendirme metrikleri kullanılmıştır. Kesinlik ve duyarlılık gibi metrikler tek başlarına modelin performansını tüm hatlarıyla değerlendirmede eksik kalacağı için [54] bu metrikler bir arada değerlendirilmiştir. Tablo 3. 'te modellerin performansını değerlendirmek amacı ile kullanılan değerlendirme metrikleri ve formülleri [52,55,56] verilmektedir.

Tablo 3. Değerlendirme metrikleri ve formülleri (Assessment metrics and formulas)

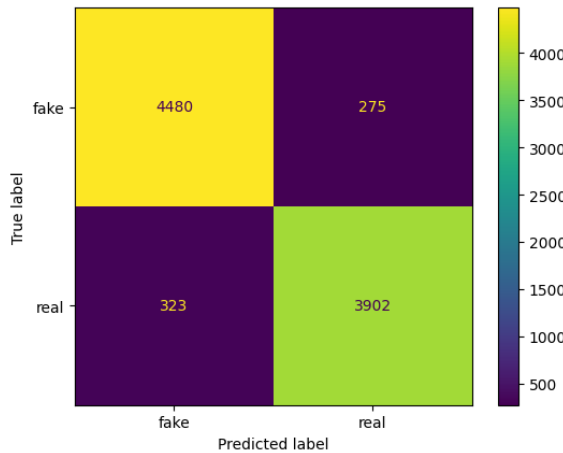
Değerlendirme Metriği	Metriklerin Hesaplama Denklemleri	Açıklama
Kesinlik	$kesinlik = \frac{TP}{TP + FP}$	Doğru olarak pozitif sınıflandırılan örneklerin, tüm pozitif sınıflandırılan örneklere oranıdır.
Duyarlılık	$duyarlılık = \frac{TP}{TP + FN}$	Doğru olarak pozitif sınıflandırılan örneklerin tüm gerçeklere oranıdır.
F1-Skor	$F1 = 2 \times \left(\frac{kesinlik \times duyarlılık}{kesinlik + duyarlılık} \right)$	F1 puanını kesinlik ve duyarlılık değerleri ile hesaplanmaktadır. F1 puanı hassasiyet ve duyarlılığın ortalamasıdır.
Doğruluk	$doğruluk = \frac{TP + TN}{TP + TN + FP + FN}$	Doğru tahmin edilen etiketler ile veri kümesindeki tüm örnekler arasındaki orandır.

3. Araştırma Bulguları (Research Findings)

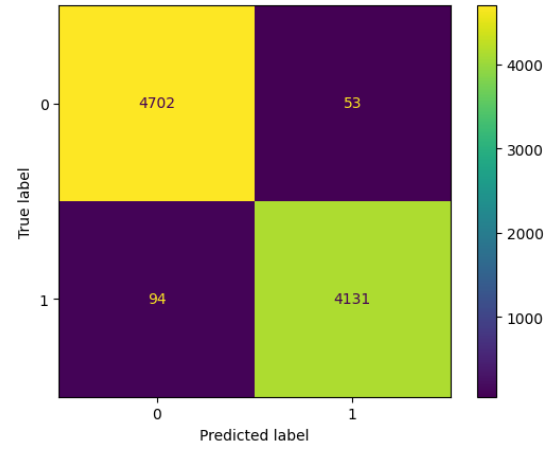
Makine öğrenimi algoritmaları eğitim için sayısal verileri kullandığından, metin önce bir etiket kodlayıcı kullanılarak sayısal forma dönüştürülmüştür. Daha sonra veri seti %80 eğitim seti ve %20 test setine bölünmüş ve ardından LSTM ve BERT algoritmaları uygulanmıştır. LSTM ile eğitim sürecinin gerçekleştirilmesinde 128 bellek hücresi kullanılmış olup, çıkış katmanında çıktı olarak sigmoid değer ile 0-1 olarak sahte-gerçek şeklinde ikili sınıflandırma gerçekleştirilmiştir. BERT ile eğitim sürecinin gerçekleştirilmesinde BERT-Based-uncased modeli kullanılmıştır. BERT-Based-uncased dizini olarak ön bellek, optimizasyon algoritması olarak ise AdamW kullanılmıştır. Her iki model de eğitim döngüsü (Epoch) sayısı 5, her iterasyonda işlenen örnek sayısı (Batch size) 32, öğrenme oranı olarak Adam optimizasyon algoritmasının varsayılan değeri olan 0.001 hiperparametreleri ile eğitilmiştir.

LSTM ve BERT modellerinin eğitimi sonucunda elde edilen sonuçlar karışıklık matrisi aracılığı ile her iki model için de incelenmiştir. Grafik 1.'de LSTM modeline ait karışıklık matrisi verilmiştir. Modelin sınıflandırma başarımı, sunulan karışıklık matrisi aracılığıyla değerlendirildiğinde, genel olarak yüksek doğrulukla çalıştığı gözlemlenmektedir. Gerçek etiketleri "fake" (sahte) olan 4755 örnekten 4480 tanesi doğru şekilde sınıflandırılmış, buna karşılık 275 tanesi "real" (gerçek) olarak etiketlenmiştir. Diğer taraftan, gerçek etiketleri "real" olan 4225 örneğin 3902'si doğru tahmin edilmiş, 323 örnek ise "fake" olarak hatalı sınıflandırılmıştır. Bu dağılım, modelin hem sahte hem de gerçek haberleri yüksek başarı oranıyla ayırt edebildiğini göstermektedir. Ancak, modelin sahte haberleri gerçek haber olarak tanıma eğilimi (false negative = 275) ile gerçek haberleri sahte olarak etiketleme eğilimi (false positive = 323) arasındaki fark, kritik uygulama senaryoları açısından önemlidir. Grafik 2. 'de BERT modeline ait

karışıklık matrisi verilmiştir. Gerçek etiketleri "fake" (sahte) olan 4755 örnekten 4702 tanesi doğru şekilde sınıflandırılmış, buna karşılık 53 tanesi "real" (gerçek) olarak etiketlenmiştir. Diğer taraftan, gerçek etiketleri "real" olan 4225 örneğin 4131'i doğru tahmin edilmiş, 94 örnek ise "fake" olarak hatalı sınıflandırılmıştır. Bu dağılım, modelin hem LSTM modeline kıyasla daha yüksek başarı oranıyla tespit yapabildiğini göstermektedir.



Grafik 1. LSTM karışıklık matrisi (LSTM confusion matrix)



Grafik 2. BERT karışıklık matrisi (BERT confusion matrix)

Tablo 2. 'de LSTM ve BERT modelleri ile sahte haber tespiti sınıflandırma raporu verilmiştir. Sahte haber tespiti için kullanılan iki farklı modelin sınıflandırma performanslarını detaylı olarak incelediğimizde modellerin doğruluk, kesinlik, duyarlılık ve F1-skoru açısından önemli farklılıklar gösterdiği görülmektedir.

Tablo 4. LSTM ve BERT modelleri ile sahte haber tespiti sınıflandırma raporu (Fake news detection classification report by models of LSTM and BERT)

		Kesinlik	Duyarlılık	F1-Score	Veri Sayısı
LSTM Modeli	0 (Gerçek)	0.93	0.94	0.94	4755
	1 (Sahte)	0.93	0.92	0.93	4225
	Doğruluk			0.9334	8980
BERT Modeli	0 (Gerçek)	0.98	0.99	0.98	4755
	1 (Sahte)	0.99	0.98	0.98	4225
	Doğruluk			0.98	8980

LSTM modelinin performans değerlendirmesinde genel doğruluk oranının %93,34, BERT modelinin doğruluk oranının %98 elde edilmesi BERT modelinin LSTM modeline kıyasla sahte haber tespiti görevinde oldukça başarılı bir performans sergilediği görülmektedir. LSTM modeli, gerçek haberleri temsil eden "0" sınıfı için %93 kesinlikte, %94 duyarlılıkta ve %94 F1-skorda sonuçlar elde etmiştir. Sahte haberleri temsil eden "1" sınıfında ise %93 doğruluk, %92 duyarlılık ve %93 F1-skora ulaşılmıştır. Bu değerler, modelin her iki sınıfı da dengeli biçimde öğrenebildiğini ve önyargısız sınıflandırma yaptığını göstermektedir. Doğruluk metriğinin de %93 seviyelerinde olması, modelin genel başarımının tutarlı olduğunu desteklemektedir. Ancak BERT modelinde elde edilen performans metriklerinin sonucuna baktığımızda gerçek haberleri temsil eden "0" sınıfı için %98 kesinlikte, %99 duyarlılıkta ve %98 F1-skorda sonuçlar elde etmiştir. Sahte haberleri temsil eden "1" sınıfında ise %98 doğruluk, %98 duyarlılık ve %98 F1-skor elde edilerek LSTM modeline göre daha iyi bir performans elde edilmiştir.

4. Tartışma ve Sonuç (Results and Discussion)

Bu çalışmada LSTM ve BERT modelleri ile sahte haber tespiti gerçekleştirilmiştir. Modellerin eğitimi için sahte ve gerçek haberlerden oluşan toplam 44898 veri kullanılmıştır. Veri ön işleme sürecinde metinler küçük harfe dönüştürülmüş, noktalama işaretleri, sayılar, özel karakterler ve anlamsız kelimeler çıkarılmış; ardından modelin girdi olarak kullanabileceği sayısal dizilere dönüştürülmüştür. Böylece

veriler modeller için uygun hale getirilmiştir. Her iki model de yüksek doğruluk değerlerine ulaşmış olması ile birlikte BERT modeli %98 ile sahte haber tespiti için önemli bir başarı sağlamıştır. Sonuç olarak, modelin yüksek doğruluk ve dengeli sınıf performansı, sahte haber tespiti gibi kritik bir görevde kullanılabilirliğini güçlü biçimde desteklemektedir.

Sahte haber tespitinde, dilsel çeşitlilik ve kültürel farklar dil yapıları, kelime seçimleri ve deyimsele ifadeler farklı dillerde farklılık göstermesi, bir toplumda sahte olarak algılanan bir haberin diğer kültürlerde doğru kabul edilebilmesi gibi zorluklarla karşılaşmaktadır. Sahte haber tespiti, özellikle haberlerin hızla yayıldığı sosyal medya platformlarında önemli bir zorluk yaratmaktadır. Dolayısıyla, sahte haberlerin özellikle sosyal medya platformlarında ne kadar hızlı dolaştığı göz önüne alındığında, hızlı tanımlama ihtiyacı açıktır. Ayrıca, artık yapay zekâ teknolojilerini kullanan sahte haber yaratıcıları yenilikçi ve uyumludur, bu da başka bir zorluk olarak nitelendirilebilmektedir. Sahte haber tespit yöntemlerindeki en büyük zorluklardan bir diğeri ise içerik değişkenliğidir. Sahte haberler metin, fotoğraf, video ve ses dahil olmak üzere çok çeşitli içerik formatlarında olabilmektedir. Dolayısıyla bu çeşitlilik esnek tespit yöntemlerini gerektirmektedir. Gelecekte, gelişen yapay zekâ yöntemleri ve insan-yapay zekâ iş birliği ile bu engellerin aşılması beklenmektedir.

5. Kaynaklar (References)

- [1] M. Öztunç, O. Kartav, Sosyal medyada yalan haber sorunu ve doğrulama platformları
- [2] H. Allcott, M. Gentzkow, Social media and fake news in the 2016 election, *Journal of Economic Perspectives* 31 (2017) 211–236.
- [3] Ünal, R., Taylan, A., Sağlık iletişimde yalan haber–yanlış enformasyon sorunu ve doğrulama platformları, *Atatürk İletişim Dergisi* 14 (2017) 81–100.
- [4] M.G. Samuels, Review: The filter bubble: What the internet is hiding from you by Eli Pariser, *InterActions: UCLA Journal of Education and Information Studies* 8 (2012).
- [5] D. Spohr, Fake news and ideological polarization: Filter bubbles and selective exposure on social media, *Business Information Review* 34 (2017) 150–160.
- [6] K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media: A data mining perspective, *SIGKDD Explorations* 19 (2017) 22–36.
- [7] K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media, *ACM SIGKDD Explorations Newsletter* 19 (2017) 22–36.
- [8] X. Zhou, R. Zafarani, A survey of fake news: Fundamental theories, detection methods, and opportunities, *ACM Computing Surveys* 53 (2020) 1–40.
- [9] H. Rashkin, E. Choi, J.Y. Jang, S. Volkova, Y. Choi, Truth of varying shades: Analyzing language in fake news and political fact-checking, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (2017) 2931–2937.
- [10] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, *arXiv:1810.04805* (2018).
- [11] J. Li, H. Shi, S. Tang, F. Wu, Y. Zhuang, Informative visual storytelling with cross-modal rules, in: *Proceedings of the 27th ACM International Conference on Multimedia*, ACM, 2019, pp. 2314–2322.
- [12] Z. Jin, J. Cao, H. Guo, Y. Zhang, J. Luo, Multimodal fusion with recurrent neural networks for rumor detection on microblogs, in: *Proceedings of the ACM Multimedia Conference*, ACM, 2017, pp. 795–816.

- [13] K. Shu, S. Wang, H. Liu, Beyond news contents: The role of social context for fake news detection, in: Proceedings of the 12th ACM International Conference on Web Search and Data Mining, ACM, 2019, pp. 312–320.
- [14] H. Ahmed, I. Traore, S. Saad, Detecting opinion spams and fake news using text classification, Security and Privacy 1 (2018).
- [15] N. Ruchansky, S. Seo, Y. Liu, CSI: A hybrid deep model for fake news detection, in: Proceedings of the ACM Conference on Information and Knowledge Management, ACM, 2017, pp. 797–806.
- [16] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, RoBERTa: A robustly optimized BERT pretraining approach, arXiv:1907.11692 (2019).
- [17] R.K. Kaliyar, A. Goswami, P. Narang, FakeBERT: Fake news detection in social media with a BERT-based deep learning approach, Multimedia Tools and Applications 80 (2021) 11765–11788.
- [18] F. Monti, F. Frasca, D. Eynard, D. Mannion, M.M. Bronstein, Fake news detection on social media using geometric deep learning, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (2019).
- [19] A. Yadav, S. Gaba, H. Khan, I. Budhiraja, A. Singh, K.K. Singh, ETMA: Efficient transformer-based multilevel attention framework for multimodal fake news detection, yayımlanma bilgisi eklenmelidir.
- [20] L. Wang, C. Zhang, H. Xu, Y. Xu, X. Xu, S. Wang, Cross-modal contrastive learning for multimodal fake news detection, in: Proceedings of the 31st ACM International Conference on Multimedia, ACM, 2023, pp. 5696–5704.
- [21] S. Suryavardan, S. Mishra, M. Chakraborty, P. Patwa, A. Rani, A. Chadha, A. Reganti, A. Das, A. Sheth, M. Chinnakotla, A. Ekbal, S. Kumar, Findings of Factify 2: Multimodal fake news detection, CLEF Working Notes (2023).
- [22] C.-O. Truică, E.-S. Apostol, MisRoBERTa: Transformers versus misinformation, Mathematics 10 (2023).
- [23] J.A. Nasir, O.S. Khan, I. Varlamis, Fake news detection: A hybrid CNN–RNN based deep learning approach, International Journal of Information Management Data Insights 1 (2021).
- [24] S. Fitria, N. Azizah, H.D. Cahyono, S.W. Sihwi, W. Widiarto, Performance analysis of transformer-based models (BERT, ALBERT and RoBERTa) in fake news detection, yayımlanma bilgisi eklenmelidir.
- [25] L. Al-Zahrani, M. Al-Yahya, Pre-trained language model ensemble for Arabic fake news detection, Mathematics 12 (2024).
- [26] S. Raza, D. Paulen-Patterson, C. Ding, Fake news detection: Comparative evaluation of BERT-like models and large language models with generative AI-annotated data, arXiv:2412.14276 (2024).
- [27] I.Q. Abduljaleel, I.H. Ali, Detecting fake news using BERT word embedding, attention mechanism, partition and overlapping text techniques, TEM Journal 14 (2025) 1152–1165.
- [28] A. Mukherjee, S. Ghosh, UNITE-FND: Reframing multimodal fake news detection through unimodal scene translation, arXiv:2502.11132 (2025).
- [29] İ. Kulaksız, A. Coşkunçay, Fake news detection on mainstream media using natural language processing, Black Sea Journal of Engineering and Science 8 (2025) 214–224.

- [30] X. Xu, P. Yu, Z. Xu, J. Wang, A hybrid attention framework for fake news detection with large language models, arXiv:2501.11967 (2025).
- [31] W.Y. Wang, “Liar, liar pants on fire”: A new benchmark dataset for fake news detection, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL, 2017, pp. 422–426.
- [32] J. Thorne, A. Vlachos, C. Christodoulopoulos, A. Mittal, FEVER: A large-scale dataset for fact extraction and verification, in: Proceedings of NAACL-HLT, 2018.
- [33] P. Patwa, S. Sharma, S. Pykl, V. Guptha, G. Kumari, S. Akhtar, A. Ekbal, A. Das, T.Chakraborty, Fighting an infodemic: COVID-19 fake news dataset, Springer Lecture Notes in Computer Science (2021).
- [34] L. Derczynski, K. Bontcheva, M. Lukasik, T. Declerck, A. Scharl, G. Georgiev, P. Osenova, P. Lobo, A. Kolliakou, R. Stewart, S.-J. Terp, G. Wong, C. Burger, A. Zubiaga, R. Procter, PHEME: Computing veracity in social media, ACM Transactions on Social Computing (2017).
- [35] T. Mitra, E. Gilbert, CREDBANK: A large-scale social media corpus with associated credibility annotations, Proceedings of the 24th International World Wide Web Conference (2015).
- [36] J. Ma, W. Gao, P. Mitra, S. Kwon, B.J. Jansen, K.-F. Wong, M. Cha, Detecting rumors from microblogs with recurrent neural networks, Proceedings of IJCAI (2016).
- [37] B.D. Horne, S. Adalı, This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, Proceedings of ICWSM (2017).
- [38] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, H. Liu, FakeNewsNet: A data repository with news content, social context and spatiotemporal information for studying fake news on social media, arXiv:1809.01286 (2019).
- [39] F. Farhangian, R.M.O. Cruz, G.D.C. Cavalcanti, Fake news detection: Taxonomy and comparative study, Information Fusion 103 (2024).
- [40] K. Nakamura, S. Levy, W.Y. Wang, r/Fakeddit: A new multimodal benchmark dataset for fine-grained fake news detection, in: Proceedings of LREC (2020).
- [41] C. Boididou, K. Andreadou, S. Papadopoulos, D.-T. Dang-Nguyen, G. Boato, M. Riegler, Y. Kompatsiaris, Verifying multimedia use at MediaEval 2015, CEUR Workshop Proceedings (2015).
- [42] T. Bolukbasi, K.-W. Chang, J. Zou, V. Saligrama, A. Kalai, Man is to computer programmer as woman is to homemaker? Debiasing word embeddings, Advances in Neural Information Processing Systems (2016).
- [43] Z. Zhou, H. Guan, M.M. Bhat, J. Hsu, Fake news detection via natural language processing is vulnerable to adversarial attacks, arXiv (2020).
- [44] Kaggle, Fake news detection datasets, <https://www.kaggle.com/datasets/emineyettm/fake-news-detection-datasets> (accessed 10 March 2025).
- [45] U. Ergün, S. Orcin, S. Barın, Extraction of clinical entities from chest radiology reports using NLP methods, Journal of Materials and Mechatronics 6 (2025) 1–14.
- [46] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9 (1997) 1735–1780.

- [47] X.H. Le, H.V. Ho, G. Lee, S. Jung, Application of long short-term memory neural network for flood forecasting, *Water* 11 (2019).
- [48] A.P.M. Diniz, P.M. Ciarelli, E.O.T. Salles, K.F. Coco, Use of deep neural networks for clogging detection in the submerged entry nozzle of the continuous casting, *Expert Systems with Applications* 238 (2024).
- [49] M. Khadhraoui, H. Bellaaj, M. Ben Ammar, H. Hamam, M. Jmaiel, Survey of BERT-base models for scientific text classification: COVID-19 case study, *Applied Sciences* 12 (2022).
- [50] A. Wang, Y. Pruksachatkun, N. Nangia, A. Singh, J. Michael, F. Hill, O. Levy, S.R. Bowman, SuperGLUE: A stickier benchmark for general-purpose language understanding systems, *Advances in Neural Information Processing Systems* (2019).
- [51] X. Zhang, J. Fan, M. Hei, Compressing BERT for binary text classification via adaptive truncation before fine-tuning, *Applied Sciences* 12 (2022).
- [52] S. Orozco-Arias, G. Isaza, R. Guyot, R. Tabares-Soto, A systematic review of the application of machine learning in the detection and classification of transposable elements, *PeerJ* 7 (2019).
- [53] S. Orozco-Arias, J.S. Piña, R. Tabares-Soto, L.F. Castillo-Ossa, R. Guyot, G. Isaza, Measuring performance metrics of machine learning algorithms for detecting and classifying transposable elements, *Processes* 8 (2020).
- [54] M.A. Karabıyık, Kontrollü dengesizlik senaryolarında topluluk öğrenme modellerinin sistematik karşılaştırması, *Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi* 9 (2025) 41–50.
- [55] R. Izhar, S.N. Bhatti, S.A. Alharthi, Bridging precision and complexity: A novel machine learning approach for ambiguity detection in software requirements, *IEEE Access* 13 (2025) 12014–12031.
- [56] S.P. Taş, S. Barin, G.E. Güraksin, Detection of retinal diseases from ophthalmological images based on convolutional neural network architecture, *Acta Scientiarum Technology* 44 (2022).