



## Examining the Models Used for Fake News Detection in the Scope of Social Context

Gülsüm KAYABAŞI KORU<sup>1\*</sup> , Çelebi ULUYOL<sup>2</sup> 

<sup>1</sup>Gazi University, Institute of Informatics, Department of Forensics, 06680 Kavaklıdere/ANKARA

<sup>2</sup>Gazi University, Institute of Informatics, Department of Forensics, 06680 Kavaklıdere/ANKARA

### Graphical/Tabular Abstract

In this study, a compilation study was conducted on the social context-based features and models of fake news from a data science perspective..

#### Article Info:

Review article  
Received: 19.07.2022  
Revision: 28.09.2022  
Accepted: 30.11.2022

#### Highlights

- Social Context
- Deep Learning
- Machine Learning

#### Keywords

Misinformation  
Fabricated News  
Fake News Detection  
Social Context

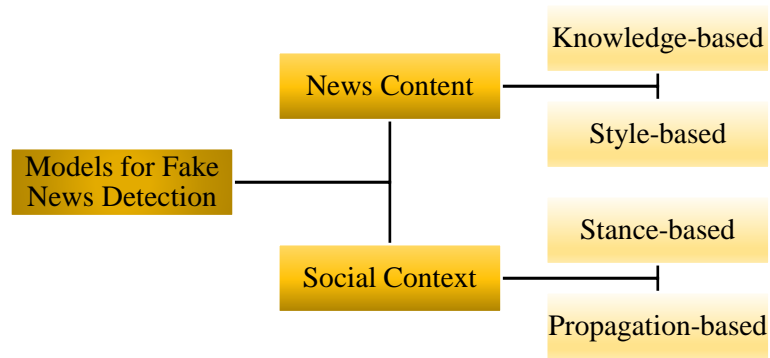


Figure A. Models for fake news detection

**Purpose:** Social context-based models use relationships between users and posts on a social network to capture relevant information to be used for fake news detection. Such models are posture-based and propagation-based models. Our aim here is to give information about these models used in studies that are frequently encountered in the literature and to compare the results of the study.

**Theory and Methods:** Detection of fake news based on social context is the second step in research to prevent the spread of fake news. Accordingly, 3 processes are required for fake news detection. These processes are feature extraction, model building, and dataset generation. The study theoretically explains how and with which methods these processes are performed and is supported by the literature.

**Results:** Studies using the features and models mentioned in the article were examined with both machine learning and deep learning approaches and compared according to performance criteria. Using machine learning models, the study with a rate of 99.2% and the highest performance in comparison of deep learning studies belong to the study with a rate of 95.4%.

**Conclusion:** In this study, the effect of social context on fake news detection is explained with the contribution of literature. The importance of modeling the social context is demonstrated. Information was provided on how social context information is used to detect fake news. The effect of social context has been examined under three main headings: user-based, post-based and network-based. By giving examples and explanations from each point of view, the necessity of social context in detecting fake news is mentioned to the reader. Machine learning and deep learning models used for fake news detection are compared. In addition, 9 different and famous data sets were compared in terms of information and social context features.



## Examining the Models Used for Fake News Detection in the Scope of Social Context

Gülsüm KAYABAŞI KORU<sup>1\*</sup>, Çelebi ULUYOL<sup>2</sup>

<sup>1</sup>Gazi Üniversitesi, Bilişim Enstitüsü, Adli Bilişim Bölümü, 06680 Kavaklıdere/ANKARA

<sup>2</sup>Gazi Üniversitesi, Bilişim Enstitüsü, Adli Bilişim Bölümü, 06680 Kavaklıdere/ANKARA

### Abstract

While in traditional news media, the content of the news is based on fake news detection, social context information in social media can be used to help detect fake news. Considering the social context, the distribution of data on social media and the interaction of online users with each other, it also explains the dissemination of news in the social environment and provides the necessary information to determine whether the news is true or not. Social media supports models based on news content. Developing these models provides some additional resources for researchers. Social context information represents three main topics: user detail, post and network analysis. In this study, a compilation study was conducted on the social context-based features and models of fake news from a data science perspective. Studies using these features and models in the literature have been examined with both machine learning and deep learning approaches and compared according to performance criteria. Using machine learning models, the study with a rate of 99.2% and the highest performance in comparison of deep learning studies belong to the study with a rate of 95.4%. In addition, in the study, 9 known data sets created for feature extraction and fake news detection were analyzed by taking into account the social context features.

### Makale Bilgisi

Derleme makalesi  
Başvuru: 19.07.2022  
Düzeltilme: 28.09.2022  
Kabul: 30.11.2022

### Keywords

Misinformation  
Fabricated News  
Fake News Detection  
Social Context

### Anahtar Kelimeler

Yanlış Bilgilendirme  
Uydurma Haber  
Sahte Haber Tespiti  
Sosyal Bağlam

## Sahte Haber Tespiti için Kullanılan Modellerin Sosyal Bağlam Kapsamında İncelenmesi

### Öz

Geleneksel haber medyasında, sahte haber tespiti için haberin içeriği esas alınırken, sosyal medyada sosyal bağlam bilgileri sahte haberleri tespit etmeye yardımcı olmak için kullanılabilir. Sosyal bağlam, verilerin sosyal medyada dağıtımı ve çevrimiçi kullanıcıların birbirleri ile etkileşimi de göz önünde bulundurularak haberlerin sosyal çevrede yayılımını da açıklayarak haberlerin doğru olup olmadığını tespit etmek amacıyla gerekli bilgileri sağlamaktadır. Sosyal medya, haber içeriğine dayalı modelleri desteklemektedir. Bu modelleri geliştirmek araştırmacılar için ek bazı kaynaklar sunmaktadır. Sosyal bağlam bilgisi kullanıcı detayı, gönderi ve ağ analizi olmak üzere üç ana başlığı temsil etmektedir. Bu çalışmada veri bilimi perspektifinden sahte haberlerin sosyal bağlama dayalı özellikleri ve modelleri konusunda derleme çalışması yapılmıştır. Literatürde bu özellik ve modelleri kullanan çalışmalar hem makine öğrenmesi hem de derin öğrenme yaklaşımıyla incelenmiş ve performans ölçütlerine göre kıyaslanmıştır. Makine öğrenmesi modellerini kullanarak %99.2 oranına sahip çalışmaya ve derin öğrenme çalışmalarının kıyaslanmasında en yüksek performans %95.4 oranına sahip çalışmaya aittir. Ayrıca çalışmada öznitelik çıkarımı ve sahte haber tespitine yönelik oluşturulan 9 adet bilinen veri setinin sosyal bağlam özellikleri göz önünde bulundurularak analizi yapılmıştır.

## 1. GİRİŞ (INTRODUCTION)

Kişilerin bilgi arayışı ve haber elde etme isteği, sosyal medyayı popüler hale getiren etkenler arasında yer almaktadır [1]. Çevrimiçi haber sağlamak ve yaymak, sosyal medya yoluyla daha hızlı ve daha kolay olduğundan sahte haberler, kasıtlı olarak yanlış bilgi dediğimiz dezenformasyon içeren haber makaleleri kapsamında incelenmektedir [2]. Sahte haberlerin yaygınlaşmasıyla bireyler ve toplum ciddi derecede

olumsuz etkilere maruz kalabilmektedir. Yoğun haber tüketimi yapan okuyucular için sahte haberler, haber ekosistemine olan güveni zedelemektedir. Sahte haberler okuyucuları politik veya finansal kazanç maksadıyla kasıtlı olarak önyargılı veya doğru olmayan inançlara ikna etmeye çalışmaktadırlar [3]. Sahte haberler insanların gerçek haberleri yorumlama ve yanıtlama şeklini değiştirerek, doğru olanı olmayandan ayırt etme yeteneklerini engellemektedir. Sahte haberler sadece güveni zedeleme ya da olumsuz duyu geliştirme ile sınırlı kalmayıp, toplumda şiddet olaylarına varacak kadar büyük dalgalanmalara da sebep olabilmektedir. Tarihte bunun örnekleri yaşanmıştır [4]. Bu nedenle, sahte haberlerin nasıl yayıldığını anlamak, etkili ve doğru sahte haber tespiti için veri madenciliği teknikleri geliştirmek ve olumsuz etkileri azaltmak için müdahale etmek büyük önem taşımaktadır [5].

Sahte haberler genellikle kasıtlı olarak ve kesinlikle yanlış olan okuyucuyu yanlış yönlendirebilecek haber makaleleri olarak tanımlanmaktadır [6,7]. Bilginin kötü niyetli olup olmadığını, amacının zarar maksadı içerip içermediğini bilgiyi karakterize ederek bulmak mümkündür. Sahte haberin oluşumunda ve dağıtımında insanların akıllarında veya niyetlerinde belli bir amaç mevcuttur. Bu amaçlara örnek vermek gerekirse:

- insanları, bireyleri, grupları, fikirleri ve gelecekteki eylemleri desteklemeye veya karşı çıkmaya ikna etmek;
- insanların korku, öfke veya neşe gibi duygusal tepkileri üretmesini sağlamak;
- eğitmek;
- utanç verici veya cezai bir fiile olan inanca destek olmak;
- ciddi olmayan konuları abartmak;
- olay ve faaliyetlerin geçmişi konusunda kafa karıştırmak verilebilir [8].

Sahte bir haber makalesinin arkasındaki amaç belirlendikten sonra bu amacın ne kadar başarılı olacağını anlamak daha kolaydır. Karakterize etmeye yardımcı olması için sosyal teorilere dayanan hızlı yayılım ölçümleri kullanılabilir. Sosyal psikoloji, sahte haber tespiti yayılımının itici güçleri için sosyal etkiye (haber makalesinin ne kadar yaygın bir şekilde yayıldığına) ve bir kullanıcının sahip olduğu önceden var olan bilgiye işaret etmektedir [9]. Bir kullanıcının bir haber makalesine güvenmesi, toplumdaki ve kendisinden istemeden yayılma algısı ve davranışını etkilemektedir. Hesaplamalı sosyal ağ analizi, sahte haberlere maruz kalan bireylerin davranışlarına ve/veya inançlarına sosyal bağlamın etkilerini incelemek maksadıyla kullanılmaktadır [10].

Sahte haberleri etkili bir şekilde, açıklanabilir faktörlerle erken bir aşamada tanımlayabilmek tespitinin ilk adımıdır. Sahte haberler, haber içeriğinde yanlış iddialar yaymaya çalıştığından, bunu saptamanın en basit yolu, haber doğruluğuna karar vermek için bir haber makalesindeki ana iddiaların doğruluğunu kontrol etmektir. Haber medyalarında sahte haber tespiti, haber içeriği bilgilerinin ayrıntılı araştırılmasına dayanmaktadır. Haber içeriğinde metin, resim, video gibi birden çok yöntem olabilmektedir. Tek veya birleşik modellerin özelliklerini öğrenmek ve sahte haberleri tespit etmek için makine öğrenmesi modelleri oluşturmak maksadıyla farklı yaklaşımlar bulunmaktadır. Haber makalelerinin içeriğiyle doğrudan ilgili özelliklere ek olarak, sosyal medya platformunda haber tüketiminin kullanıcı güdümlü sosyal katılımlarına ek sosyal bağlam özellikleri elde edilmektedir. Sosyal katılımlar, haber makalelerinin doğruluğunu ortaya çıkarmak için yararlı yardımcı bilgiler sağlayan haber yayma sürecini zamansal olarak temsil etmektedir. Genel olarak, sosyal bağlam üç başlığı temsil etmektedir: kullanıcılar, gönderiler ve ağlar.

İlk olarak, sahte haberlerin sosyal botlar veya insan olmayan cyborgs gibi hesaplar aracılığıyla oluşturulup yayılması olası bir durumdur. Bu sayede, kullanıcı profilleri ve davranışlarını kullanıcı bazlı özellikler kullanarak ele geçirmek sahte haber tespiti için faydalıdır [11]. İkinci olarak, kişilerin sahte haberlerle ilgili duygu ve düşüncelerini ifade şeklinin sosyal medya yayınları aracılığıyla şüpheli ve sansasyonel bir tepki şeklinde olması dikkat çekmektedir. Bu nedenle, gönderi tabanlı özelliklerin çıkarılmasının gerekliliğini doğrulamaktadır. Bu işlem sayesinde halktan gelen tepkiler ışığında olası sahte haberlerin tespiti kolaylaşmaktadır. Son olarak, sahte haberleri tespit etmek için ağ tabanlı özelliklerin çıkarılması gerekmektedir. Bu da, sahte haber yayma süreçlerinde yankı odası döngüsü dediğimiz döngünün

oluşturulması ile olmaktadır. Yani, kullanıcıların ilgi alanları, ilgilendikleri konular ve ilişkiler farklı ağ türlerinin sosyal medya üzerinde oluşmasını gerekli kılmıştır [10].

Üretilen verinin çeşitliliğini, hızlı yayılımını ve veri miktarını artıran en önemli sebep sosyal ağların kullanımının artması ve iletişim teknolojilerindeki gelişmenin hızlanmasıdır [12]. Sahte haberler genellikle metin, resimler, videolar vb. dahil çok yönlü bilgiler içermektedir. Bu nedenle, bilginin çok yönlü kullanılması, algılama performansını iyileştirmek bakımından büyük potansiyele sahiptir. İlk olarak, ikili sınıflandırma için sözcüksel özellikler, sözlük, duyarlılık ve okunabilirlik gibi dilsel özelliklerin çıkarılmasına veya sinir ağır yapıları olan evrişimsel sinir ağları ve tekrarlayan sinir ağları ile sinirsel özelliklerin öğrenilmesine odaklanmaktadır. İkincisi, görsel işaretler temel alınarak görsel istatistiksel özellikler, görsel içerik özellikleri ve sinirsel görsel özelliklerden çıkarılmaktadır [13]. Sahte veya gerçek haberlere eklenen istatistikler görsel istatistiksel özellikleri temsil etmektedir. Görsel içerik özellikleri, görüntülerin içeriğini netlik, tutarlılık, çeşitlilik vb. gibi tanımlayan faktörleri gösterir. Sinirsel görsel özellikler evrişimli sinir ağları gibi sinir ağları kullanılarak öğrenilmektedir. Ayrıca, son gelişmeler görsel içerik grafiğini görsel içerikten büyük ölçüde geliştiren sağduyu bilgisini keşfetmek için görsel sahne grafiğini çıkarmayı amaçlamaktadır [14].

Literatürde, sosyal bağlam özelliklerini kullanarak sahte haber tespit eden başarılı çalışmalar mevcuttur. Sosyal bağlamın etkisi düşünülerek sahte haber tespitinde elde edilen başarının kritik yönü sosyal medyada söylentilerin doğruluk sınıflandırması gibi benzer araştırma alanlarında kullanılan bir dizi ortak özellik sunmasıdır [15].

Sahte haber tespiti ile ilgili mevcut çalışmaların çoğu denetimli yöntemlere dayanmaktadır. Haber içeriği [16], kullanıcı profilleri [17], gönderi yayılımı [18] ve sosyal bağlamlar [19] gibi farklı özellikleri dikkate alarak bir sınıflandırma modeli oluşturmayı amaçlamaktadırlar. Bu denetimli yöntemlere sahip çalışmalar sınıflandırma modelini eğitmek için güvenilir bir veri kümesine ihtiyaç duydıklarından kritik bir sınırlama bulundurmaktadır. Ancak, çok sayıda ek açıklama elde etmenin zaman alıcı olması ve yoğun emek istemesinden dolayı bu süreçte haber içeriğinin yanı sıra başka kanıtların da kontrol edilmesi gerekmektedir. Ek açıklamalar elde etmek için kitle kaynak kullanımını yaklaşımdan yararlanmak uzman kontrolünün yükünü hafifletebilir ancak bu açıklamaların kalitesi düşebilmektedir [20].

Sosyal bağlama dayalı yöntemler, kullanıcı profilleri, gönderi içerikleri ve sosyal ağların özelliklerini kullanarak tespit yapmaktadır. Kullanıcı profilleri, kullanıcıların özelliklerini ve kullanıcı güvenilirliğine dayanarak ölçmek için kullanılmaktadır [17]. Her ikisi de etiketli haber içeriği ve sosyal bağlam bilgisi içeren Buzzfeed ve Politifact platformlarında doğruluk kontrolü yapılarak toplanan FakeNewsNet [15,21] isimli kapsamlı sahte haber tespit veri seti mevcuttur. Gönderi içeriği, haberin değişim için üretilen nesnel özelliklerini (ör. gövde metni), sosyal bağlam ise, haber öğelerinin ilgili kullanıcının sosyal etkileşimlerini (ör. Twitter’da kullanıcı gönderme/haber paylaşma) içermektedir. Doğrulanmış web sitelerindeki sahte ve gerçek haberlerle ilgili kullanıcı etkileşimleri, Twitter’ın gelişmiş arama Uygulama Programlama Arayüzü (Application Programming Interface - API) üzerinden 19 adet özellik kullanılarak toplanmıştır. Kullanıcı etkileşimlerini toplamak amacıyla yapılan arama sorguları, başlıklardan meydana gelmiştir. Arama sorgusundan veri ön işleme amacıyla özel karakterler kaldırılmıştır. Haber parçalarını doğrudan yayan sosyal medya gönderileri elde edildikten sonra, bu gönderilere verilen yanıtlar, beğeniler ve yeniden paylaşımlar gibi kullanıcı yanıtları alınmıştır. Haber yayan kullanıcılar için, kullanıcı profilleri, kullanıcı gönderileri ve sosyal ağ bilgileri gibi pek çok meta veri toplanmıştır [21].

Sahte haberleri haber içeriğindeki yanlış iddiaları yaymaya çalıştığından, bunu tespit etmenin en basit yolu, haber makalesindeki ipuçlarını bulmaktır. [15]’de ilk olarak, sahte ve gerçek haber makalelerinin konu dağılımı analiz edilmiştir. Politifact veri setinde sahte ve gerçek haberlerin daha çok siyasi kampanya ile ilgili olduğu görülmüştür. GossipCop veri setinde ise, sahte ve gerçek haberlerin daha çok ünlüler arasındaki ilişki hakkındaki dedikoduları içerdiği gözlemlenmiştir. Ayrıca, sahte ve gerçek haber konularının birbirlerinden farklı olduğu görülmüştür. Bazı haberler için, haber içeriklerinin konularına

bakarak tespit yapmak sahte haber tespiti için yeterli olmamakla birlikte zordur. Bu da sosyal bağlam bilgilerinin kullanılması gerekliliğini doğurmaktadır [15].

Metin madenciliği yöntemleri ve denetimli yapay zeka algoritmalarının birleştirilmesiyle sosyal medyadaki sahte haberlerin tespiti için önerilen modelde metin madenciliği analizini ve denetimli yapay zeka algoritmaları ayrı ayrı kullanılmıştır. Bu hibrit model, üç farklı gerçek dünya veri seti üzerinde test edilmiş, doğruluk, kesinlik, duyarlılık ve f-ölçütü değerlerine göre değerlendirme yapılmıştır. Değerlendirme sonuçlarına göre ortalama performansları hesaplanmıştır. Elde edilen sonuçlara bakıldığında, doğruluk, kesinlik ve f-ölçütü açısından en iyi ortalama değerler Karar Ağacı algoritmasından sağlanmıştır. Duyarlılık metriği açısından ise en iyi algoritmaların ZeroR, Çapraz Doğrulama Parametre Seçimi (Cross Validation Parameter Selection-CVPS) ve Ağırlıklı Örnek İşleyici Sarmalayıcı (Weighted Instances Handler Wrapper -WIHW) algoritmaları olduğu görülmüştür [22].

Sahte haber tespit problemini ilk kez bir optimizasyon problemi olarak ele alan, Gri Kurt Optimizasyon (Grey Wolf Optimization-GWO) ve Salp Sürü Optimizasyon (Salp Swarm Optimization-SSO) adlı iki metasezgisel algoritma kullanan yaklaşımda üç aşama yer almaktadır. Veri ön işleme aşamasından sonra yeni bir sahte haber tespit modeli oluşturulması için GWO ve SSO'ya uyarlanmış ve son olarak test için önerilen modellerini kullanmışlardır. Çalışmaları yedi denetimli yapay zeka algoritması ile karşılaştırılmıştır. Tüm veri setlerinde en iyi doğruluk GWO'dan elde edilmiştir. GWO ayrıca üç veri setinden ikisinde en iyi kesinlik ve f-ölçütü değerlerini vermiştir. SSO, üç veri kümesinden ikisinde hassasiyet açısından tüm algoritmalarından daha iyi performans göstermiştir. GWO algoritmasının SSO algoritması ve diğer yapay zeka algoritmalarından performans olarak daha iyi olduğu görülmüştür. Birçok farklı hedefi aynı anda ve verimli bir şekilde ele alan temsili ve esnek uygunluk fonksiyonu nedeniyle, önerdikleri iki algoritmadan elde edilen sonuçların umut verici olduğu görülmektedir [23].

SSO ve GWO algoritmalarının uyarlandığı sosyal medyadaki sahte haberleri tespit etmek için metin madenciliği ve optimizasyon yöntemlerini birleştiren model oluşturulmuştur. Bu modelde, popüler ve yeni bir çevrimiçi sosyal medya problemi bir optimizasyon problemi olarak ele alınmıştır. Ayrıca önerilen algoritmaların performansını doğrulamak için literatürdeki başka bir çalışmada dört farklı gerçek dünya sahte haber veri seti üzerinde test yapılmıştır. Bu yöntemler arasındaki farklılıkları ayırt etmek amacıyla Friedman testi uygulanmıştır. Elde edilen sonuçların da gerçek dünya veri setlerinde önemli sonuçlar elde ettiği görülmüştür [24].

Sahte haberlerin hükümet ve toplum üzerindeki etkisini belirlemek, anlamak ve sahte haberleri orijinal olanlardan ayırt etmek amacıyla derin öğrenme tabanlı yaklaşımları benimsemişlerdir. Toplumun sahte haberlerin yayılmasına karşı yardımcı olmayı, insanların sahte haber yayıcılarından ve günümüz dünyasındaki yayımlarından haberdar etmeyi amaçlayan çalışmalarında GloVe tekniği kullanmışlardır. Bu teknik sayesinde, kelime yerleştirmeler için her bir kelime vektör biçiminde temsil edilmiştir. Vektör biçimine getirilen veri mimarilerindeki çeşitli seviyelerde Uzun-Kısa Vadeli Bellek (Long Short Term Memory-LSTM) sinir ağı kullanılmıştır. Değerlendirme ölçütü olarak doğruluk seçilmiş ve önerdikleri model ile %99,88 doğruluk elde edilmiştir [25].

Makalenin ikinci bölümünde araştırma sürecinden ayrıntılı olarak bahsedilmiş, üçüncü bölümünde sosyal bağlam kapsamında sahte haber tespiti yapan çalışmaların kullandığı modellerden ve bu modellerin makine öğrenmesi ve derin öğrenme yaklaşımları bakımından incelenmesine yer verilmiş ve dördüncü bölümünde literatürde bilinen 9 veri setinin analizi yapılmıştır.

## 2. ARAŞTIRMA SÜRECİ (SURVEY PROCESS)

Araştırma süreci, [26,27] çalışmalarında da kullanılan Google Scholar, Science Direct, IEEE Explore, Wiley, Web of Science, Scopus ve ACM Digital Library gibi dijital kütüphanelerden alınan yaptığımız çalışmanın konusuna uygun olan yayınların bir araya getirilmesiyle oluşmaktadır. İngilizce dilinde yazılmış orijinal araştırma ve tarama/derleme makaleleri incelenmiştir. Makaleler, sahte haberlerle ilgili anahtar kelimeler ve “sahte haber”, “aldatmaca”, “yanlış bilgilendirme”, “uydurma haber”, “dezenformasyon”,

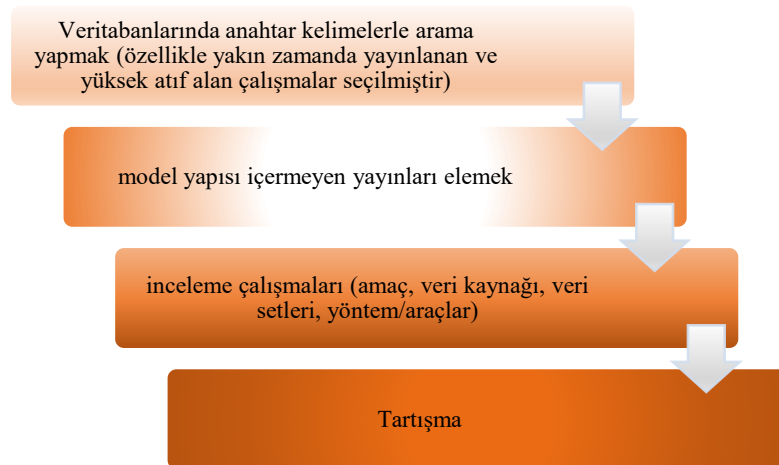
“sosyal bağlam”, “sahte haber tespiti” gibi ilişkili terimler kullanılarak yukarıda bahsi geçen dijital kütüphanelerden elde edilmiştir. Tablo 1’de tüm araştırma ögesi ve zaman aralıklarına karşılık kullanılan anahtar sözcükler sunulmuştur. Burada zaman aralığı olarak 2010 ile 2022 yılları arası seçilmiş ancak son 5 yıla ait yayınlar güncelliğini halen koruyabildiği düşünülerek tercih sebebi olmuştur. Kullanılan anahtar sözcükler birlikte kullanılma (aralarında ve bağlacı bulunan) ve tamamı kullanılan (aralarında veya bağlacı bulunan) şeklinde Tablo 1’de açıklanmıştır. İncelenen yayınlar İngilizce dilinde yazıldığından Türkçe dilindeki karşılıklarıyla birlikte tabloda yerini almıştır

**Tablo 1.** Arama dizesi içerisinde kullanılan anahtar sözcükler

Arama Ögesi ve Zaman Aralığı	Kullanılan Anahtar Sözcükler
2010 ile 2022 yılları arasında tüm arama soruları için	<p>‘social context(sosyal bağlam)’ VE (‘social context(sosyal bağlam)’ VEYA ‘fake news detection(sahte haber tespiti)’</p> <p>‘disinformation (dezenformasyon)’ VE (‘social context(sosyal bağlam)’ VEYA ‘fake news detection(sahte haber tespiti)’</p> <p>‘misinformation (yanlış bilgilendirme)’ VE (‘social context(sosyal bağlam)’ VEYA ‘fake news detection(sahte haber tespiti)’</p> <p>‘fabricated news (uydurma haber)’ VE (‘social context(sosyal bağlam)’ VEYA ‘fake news detection(sahte haber tespiti)’</p> <p>‘hoax (aldatmaca)’ VE (‘social context(sosyal bağlam)’ VEYA ‘fake news detection(sahte haber tespiti)’</p>

Araştırma sürecinde 2010 yılından 2022 yılları arasında yayınlanan ve mükerrer olmayan 1303 makaleye ulaşılmıştır. Makaleler seçilirken yakın zamanda yayınlanana öncelik verilmiştir. Araştırma sürecinin akışı Şekil-1’de gösterilmiştir. Belirtilen kriterlere göre makaleler filtrelendikten sonra 261 makale tam metin incelemesi için değerlendirilmiştir. İçeriğinde model yapısı içermeyen yayınlar elenmiştir. Yayınların sosyal bağlam ile ilgili olmasına dayalı olarak, inceleme çalışmalarında amaç, veri kaynağı, veri setleri ve yöntem/araçlar metodolojik ve sistematik bir organizasyon şeklinde anlatılmıştır.

İncelenen çalışmaların özeti Tablo 2 ve Tablo 3’de model inşa edilmesi başlığının altında, Tablo 4’te ise söz konusu çalışmalarda kullanılan 9 bilinen veri seti sosyal bağlam kapsamında incelenerek literatürdeki veri kümeleri başlığı altında sunulmuştur. Sunulan çalışmanın son olarak sonuç bölümü çalışmamızda yerini almaktadır.



**Şekil 1.** Araştırma sürecinin akışı

### 3. SOSYAL BAĞLAMA DAYALI SAHTE HABER TESPİTİ (FAKE NEWS DETECTION BASED ON SOCIAL CONTEXT)

Sahte haber tespit yöntemleri genellikle haber içeriklerini ve sosyal bağlam bilgilerini kullanarak tespit işlemini gerçekleştirmeye odaklıdır [15]. Haber içerikleri, sahte haber ile gerçek haberi ayırt etmeye yönelik ipuçları içermektedir ancak tek başına yeterli değildir. Haber içeriği kadar, haberlerle ilgili sosyal bağlam da, sahte haberlerin tespit edilmesine yardımcı olacak zengin bilgiler içermektedir. Sosyal bağlam özelliklerini kullanan yaklaşımlar; kullanıcı tabanlı, gönderi tabanlı ve ağ tabanlı yaklaşımlardır.

Kullanıcı tabanlı yaklaşımlar, kullanıcı özellikleri ve güvenilirliğini ölçmek maksadıyla kullanıcı profillerinden çıkarılmaktadır [29,17,28]. Kullanıcı özellikleri, Twitter'ın sunmuş olduğu API sayesinde elde edilerek analiz edilebilmektedir [30,31]. Örneğin, Shu vd., sahte haberleri gerçek haberlerden ayırt etmek maksadıyla kullanıcı profillerini çeşitli yönlerden anlamayı önermektedir [29]. Yang vd., kullanıcıların sosyal medyadaki görüşlerini kullanarak ve güvenilirliklerini tahmin ederek denetimsiz bir sahte haber tespit algoritması önermiştir [28]. Gönderi tabanlı özellikler, bakış açısı (duruş) [32], konular [19] veya güvenilirlik [17, 18] bakımından kullanıcıların sosyal tepkisini temsil etmektedir. Ağ tabanlı özellikler [10], yayılma ağı [33] vb. gibi belirli ağlar oluşturularak çıkarılmaktadır. Sosyal bağlamı benimseyen yaklaşımlardan gönderi tabanlı yaklaşım temel olarak duruş tabanlı, duygu tabanlı ve yayılma tabanlı modelleri içermektedir. Duruşa dayalı modeller, haberin doğruluğunu anlamak için kullanıcıların haberlere yönelik görüşlerini kullanmaktadır. Yayılma temelli modeller, haberlerin güvenilirliğinin, çeşitli yayma yöntemlerinin uygulanabileceği ilgili sosyal medya gönderilerinin güvenilirliği ile yüksek oranda ilişkili olduğunu varsaymaktadır [32]. Literatürdeki son çalışmalarda, haberlerin temsil eden zamansal ve dilsel özellikleri öğrenmek maksadıyla derin öğrenme modelleri uygulanmaktadır [32,34]. Shu vd., tıklama tuzağı tespitini iyileştirmeyi sağlamak için eğitim verilerini artırmak maksadıyla yapay (sentetik) veriler oluşturmayı önermiştir. Yayılım temelli yaklaşımları doğrudan karşılaştırmak sadece kullanıcı eylemlerine (haber göndermek, yayınlamak vb.) sahip olan bir sistem için mümkün değildir. Bunun sebebi metinden çıkarılan yayılma sinyallerinin aynı olması ve bu nedenle de etkisiz hale gelmiş olmasıdır [35].

Sosyal bağlama dayalı sahte haberlerin tespiti, sahte haberlerin yayılmasını engellemeye yönelik araştırmaların ikinci adımındır. Buna göre sahte haber tespiti için 3 işlem gereklidir. Bu işlemler: özellik çıkarımı, modelin inşa edilmesi ve veriseti oluşturmadır.

#### 3.1. ÖZELLİK ÇIKARIMI (FEATURE EXTRACTION)

Sosyal bağlama dayalı sahte haberlerin tespitine yönelik ilk adım, haber makalesinden özelliklerin çıkarılmasıdır. Bahse konu çalışmaların özet hali Tablo 2 ve Tablo 3'de sunulmuştur. Bir haber makalesi, sosyal bağlamla ilgili aşağıdaki üç özellikten oluşmaktadır:

##### 3.1.1. Kullanıcı Detayları

Yaş, cinsiyet, üyelik, konum, takipçi sayısı, tweet sayısı vb. gibi kullanıcı ayrıntılarından sahte haberlerin tespit edilmesine yardımcı olabilecek özellikler çıkarılabilmektedir [11]. Yang ve diğ. İçerik tabanlı, yayılım tabanlı, uygulama programı tabanlı ve konum tabanlı özelliklerin eşlik ettiği kullanıcı tabanlı özelliklerin bir kombinasyonunu kullanan Sina Weibo adında bir uygulama geliştirmişlerdir [36].

Ayrıca, bu özelliklere dayalı olarak, bir tweet-retweet grafiği oluşturulabilmekte ve belirli bir haber makalesiyle ilgili kullanıcı grubunun özelliklerinin daha iyi anlaşılması için farklı kullanıcılar bir grup altında kümelenebilmektedir [37].

##### 3.1.2. Gönderi Analizi

Kullanıcılar sosyal medyada bir haber okuduktan sonra tepkilerini, duygularını veya düşüncelerini genellikle yorum ya da paylaşım şeklinde ifade etmektedirler. Bu nedenle, gönderilerden özelliklerin çıkarılması, sosyal bağlama dayalı sahte haber tespit sürecinde önemli bir faktör olmaktadır. Jin ve diğ.

duruşun belirli bir haber makalesinin kabul veya reddinin bir göstergesi olduğunu gözlemlemişlerdir [38]. Rubin ve diğ. absürtlük, noktalama işaretleri ve anlambilim gibi özelliklerin bir arada kullanılmasının hiciv içeriğini tahmin etmeye yardımcı olabileceği sonucuna varmıştır [39]. Castillo ve diğ. konu düzeyindeki özellikleri analiz etmiştir [37]. Ma ve diğ. Latent Dirichlet Allocation (LDA)'yı kullanırken, Verma ve diğ. güvenilirliği değerlendirmek amacıyla görevlerden TF-IDf ve CV gibi özellikleri çıkarmıştır [19,40].

Bununla birlikte, bu tür yaklaşımlar, yayılma sırasında mesaj özelliklerinin değişimini yakalayamamıştır. Bu sorunu en aza indirebilmek amacıyla zaman serisine dayalı yaklaşım formüle edilmiş ve yayılma sırasında söylenti olmayanlara kıyasla söylentilerde bulunan belirli özelliklerin diğerlerine göre farklı olduğunu göstermişlerdir [19].

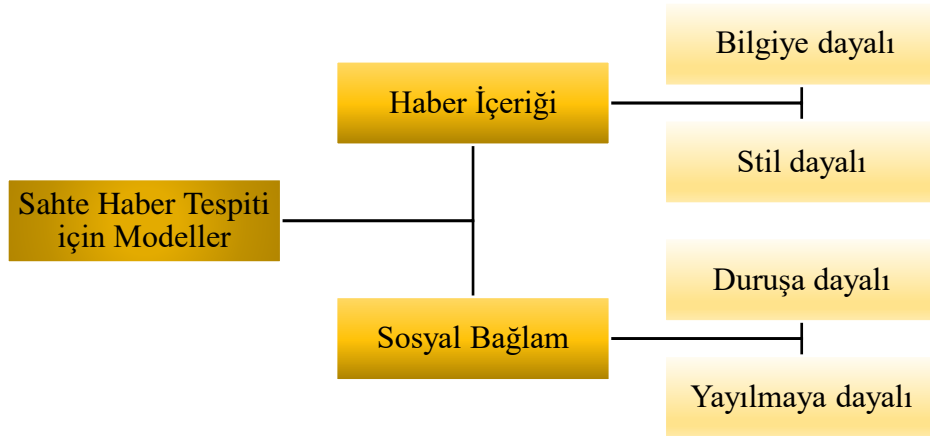
### 3.1.3. Ağ Analizi

Ağ analizine dayalı yöntemlerde amaç, haber akışını haritalamak ve ilgili kullanıcı kümesini belirlemektir. Düğümlerin bir haberle ilgili gönderileri ve kenarların duruş benzerliğine bağlı olarak ağırlıkları temsil edebileceği durum duruş temellidir. Kwon ve diğ. ağı, haber yayma yörüngelerine dayalı olarak bir ağı kurduğu bir yayılma temelli eşlik eden takipçi-takip edilen ilişkileri üzerine inşa edildiği, arkadaşlık temelli bir yaklaşım tasarlamışlardır [33]. Örneğin, [38] ve [41], ağı güvenilirliğini belirlemek için kitle kaynaklı bir yaklaşım benimsemektedir. Ruchansky ve diğ., gönderileri benzer makalelerdeyse, kullanıcıların ağ üzerinden bağlandığı eşlik tabanlı bir teknik benimsemiştir [42].

Bu çalışmalardan, içerik tabanlı özelliklerin, kullanıcı ayrıntılarının ve bilginin ağ üzerinden yayılmasının analizinin bir araya getirilmesinin, sosyal bağlama dayalı sahte haber tespiti için çok yönlü ve sağlam bir çözüm sağlayabileceği çıkarılabilmektedir.

## 3.2. MODEL İNŞA EDİLMESİ (MODEL CONSTRUCTION)

Sahte haber tespiti için inşa edilen modeller iki kategoriye ayrılmaktadır. Bu kategoriler haber içeriği ve sosyal bağlamdır. Biz bu çalışmada sosyal bağlam kategorisine değineceğiz. Sahte haber tespiti için model inşası bakımından sunulan kategoriler Şekil 2'de gösterilmiştir.



Şekil 2. Sahte haberleri tespiti için modeller

Sosyal bağlama dayalı modeller, sahte haberlerin tespiti için kullanılacak ilgili bilgileri yakalamak için bir sosyal ağdaki kullanıcılar ve gönderiler arasındaki ilişkileri kullanır. Bu tür modeller aşağıdaki tiplerden olabilir:

### 3.2.1 Duruşa Dayalı

Duruşlar (bakış açıları), kullanıcıların haberlere yönelik destekleme, karşı çıkma vb. fikirlerini göstermektedir. Tipik olarak, sahte haberler sosyal medya kullanıcıları arasında tartışmalı görüşlere neden



olabilmektedir; burada sorgulama duruşlarını reddetmenin, iddiaların sahte olduğunu belirtmede çok önemli bir rol oynadığı görülmüştür. Kullanıcılara ait gönderilerin duruşu belirgin ve üstü kapalı olabilmektedir. Belirgin bakış açıları Facebook'un "beğeni" eylemlerine benzer şekilde doğrudan duygu veya fikir ifadesidir. Üstü kapalı bakış açıları ise, sosyal medya paylaşımları incelenerek otomatik olarak çıkarılabilmektedir.

Bu tür modeller, sahte veya gerçek olduğunu doğrulamak maksadıyla bir haber makalesiyle ilgili farklı kullanıcıların görüşlerini almaktadır. Görüşler, belirli bir gönderinin aldığı beğeni sayısından, bir makaleyle ilgili olarak sağlanan yorumlardan oluşabilmektedir. Görüş tespiti için, kullanıcıların açık veya örtük duruşlarını bulmak amacıyla daha fazla madencilik yapılmaktadır. Örneğin, Jin ve diğ., çelişen duruşları analiz ederek haber içeriğini doğrulamak için bir yöntem geliştirmiştir [38]. Başka bir çalışmada Mohammad ve diğ. n-gram ve kelime yerleştirme gibi giriş özellikleri bulunan SVM'ye dayalı daha işlenmiş bir duruş algılama tekniği önermiştir [43].

Kullanıcı meta profili özelliklerini açık ve örtük olmak üzere iki ana açıdan toplamakta ve analiz etmektedir [44]. Açık özellikler, doğrudan sosyal medya sitesi APİsi sorgulanarak döndürülen meta verilerden elde edilmektedir. Örtük özellikler doğrudan mevcut değildir, ancak kullanıcı meta bilgilerinden veya geçmiş tweetler gibi çevrimiçi davranışlardan çıkarılmaktadır. Seçilen özellik setleri, hiçbir şekilde tüm olası özelliklerin kapsamlı bir listesi değildir. Ancak, kolayca erişilebilen ve hemen hemen tüm genel kullanıcılar için mevcut olan açık özelliklere ve kullanıcı özelliklerini daha iyi anlamak için literatürde yaygın olarak kullanılan örtük özelliklere sahiptir. İlk olarak, FakeNewsNet verilerine [21] dayanarak sahte ve gerçek haberleri paylaşma olasılığı daha yüksek olan iki kullanıcı alt kümesi seçilmiş ve bu iki küme üzerinden birbirlerinin toplu istatistikleri karşılaştırılmıştır.

#### *Açık Özellikler*

Temsili açık profil özneliklerinin listesi aşağıdakileri içerir:

- **Profille İlgili.** Temel kullanıcı açıklama alanları:
  - Doğrulanmış: kullanıcının doğrulanmış bir kullanıcı olup olmadığı;
  - KayıtZamanı: hesabın kaydından bu zamana kadar geçen gün sayısı;
- **İçerikle İlgili.** Kullanıcı aktivitelerinin özellikleri:
  - DurumSayısı: gönderi sayısı;
  - FavoriSayısı: beğeni sayısı;
- **Ağ ile İlgili.** Sosyal medya nitelikleri:
  - TakipçiSayısı: kullanıcıyı takip eden sayısı;
  - TakipEdilenSayısı: kullanıcının takip ettiği kullanıcı sayısı.

#### *Örtük Özellikler*

Doğrudan kullanıcı meta verileri aracılığıyla sağlanmayan, ancak kullanıcı demografisini açıklamak ve anlamak için yaygın olarak kullanılan birkaç örtük profil özelliği de kullanılmaktadır [45]. Bu örtük özellikleri denetimsiz bir şekilde tahmin etmek için yaygın olarak kullanılan araçları benimsendiği unutulmamalıdır. Bazı temsili özellikler aşağıdaki gibidir.

- **Yaş:** Araştırmalar, yaşın insanların psikolojisi ve bilişselliği üzerinde büyük etkileri olduğunu göstermektedir.
- **Kişilik:** Kişilik, bir bireyi diğerlerinden farklı kılan özellikleri ifade etmektedir.
- **Siyasi önyargı:** Kullanıcıların profillerini şekillendirmede ve sosyal medyadaki haber tüketim tercihlerini etkilemede siyasi önyargı önemli bir rol oynamaktadır.

DeneySEL olarak yapılan karşılaştırma analizinden, açık ve örtük profil özelliklerinin çoğunun farklı özellik dağılımlarını ortaya çıkardığını ve bu da onları sahtekarlığı tespit etmek için kullanma potansiyelini ortaya koyduğunu gözlemlemekteyiz.

### 3.2.2. Yayılmaya Dayalı

Bu tür modeller, aralarında bir dereceye kadar benzerlik içeren sosyal medya gönderileri arasındaki ilişkileri ortaya çıkarmaktadır. Haber makalelerinin çeşitli ağlar arasındaki akışını izler ve sahte veya gerçek olup olmadığını doğrulamak için bu ağlarda bulunan kullanıcıların güvenilirliğini kullanır [32]. Yang ve diğ. yayılmaya dayalı bilgiyi inanılabilirliği belirleyen bir özellik olarak değerlendirmiştir [36]. Jin ve diğ. olay, alt olay ve mesaj tabanlı yayılma analizinden oluşan üç katmanlı bir hiyerarşik model önermiştir [46].

Mekansal yayılımı analiz etmenin yanı sıra, birkaç çalışma zamansal yayılmayı da hayati bir güvenilirlik göstergesi olarak değerlendirmiştir. Örneğin, yayılım sırasında zamansal bilgileri yakalamak için zaman serisine dayalı bir yaklaşım tasarlamıştır [19]. Bunun dışında, bazı çalışmalar bir ağda sahte haberlerin yayılımını belirlemek için istatistiksel önlemler uygulamıştır. Örneğin, bir ağda mesaj yayılımının önemini belirlemek için bir parametre olarak temel taklit sayısını elde etmiştir [47].

Makine öğrenimi modelleri kullanılarak sahte haberlerin tespiti için oluşturulan ilgili çalışmaların performans bakımından kıyaslanması Tablo 2’de sunulmuştur. Derin öğrenme modelleri kullanılarak sahte haberlerin tespiti için oluşturulan ilgili çalışmaların performans kıyaslanması Tablo 3’te verilmiştir.

**Tablo 2.** Makine öğrenmesi yaklaşımlarının performans kıyaslaması

Model	Metot	Performans
Mukherjee ve diğ. (2013)[48]	SVM	Doğruluk=0.861 F1-Skor=0.857
Kwon ve diğ. (2013) [33]	SVM, RF ve DT	Doğruluk=0.897 F1-Skor=0.878 (RF)
Castillo ve diğ. (2013) [37]	LR ve RF	F1-Skor=0.824 ROC Alan=0.816
Ma ve diğ. (2015) [19]	SVM	Doğruluk=0.871 F1-Skor=0.87
Rubin ve diğ. (2016) [39]	SVM	Doğruluk=0.87
Granik ve diğ. (2017) [49]	NB	Doğruluk=0.74
Tacchini ve diğ. (2017) [41]	LR ve HBLC	Doğruluk=0.992 (HBLC)
Gravanis ve diğ. (2019) [50]	SVM ve AdaBoost	Doğruluk=0.787
Ozbay ve diğ. (2019) [23]	GWO	Doğruluk=0.926
Ozbay ve diğ. (2020) [22]	DT	Doğruluk=0.968 F1-Skor=0.968
Taşkın ve diğ. (2021) [51]	RF	F1-Skor=0.86

Tablo 2’de görüldüğü üzere, en iyi performans değeri 0.991 doğruluk değeri ile Tacchini ve diğerlerinin yaptığı çalışmaya aittir. Bu çalışmada kullanılan HBLC (Harmonic boolean label crowdsourcing) sayesinde bu değer elde edilmiştir [41]. Ayrıca en iyi performanslardan biri olan Özbay ve arkadaşlarının çalışmaları ISOT veri seti kullanılarak karar ağacı algoritmasıyla elde edilmiştir [22]. Çalışmada söz konusu veri seti kullanılarak 23 adet algoritma denenmiştir. Karar ağacı algoritması bu algoritmalar içerisinde en iyisi sonucu vermiştir.

**Tablo 3.** Derin öğrenme yaklaşımlarının performans kıyaslanması

Model	Metot	Performans
Long ve diğ. (2017) [52]	LSTM	Doğruluk=0.415
Wang ve diğ. (2017) [16]	CNN ve Bi-LSTM	Doğruluk=0.274
Ruchansky ve diğ. (2017) [42]	RNN	Doğruluk=0.922 F1-Skor=0.924
Roy ve diğ. (2018) [53]	CNN ve Bi-LSTM	Doğruluk=0.45 F1-Skor=0.43
Jwa ve diğ. (2019) [54]	BERT	F1-Skor=0.746
Polignano ve diğ. (2019) [55]	CNN	Doğruluk=0.918
Baruah ve diğ. (2020) [56]	BERT	Doğruluk=0.69
Song ve diğ. (2021) [57]	CNN	Doğruluk=0.922
Paka ve diğ. (2021) [58]	BERT, RoBERT	Doğruluk=0.954 F1-Skor=0.953
Nasir ve diğ. (2021) [59]	CNN, RNN	Doğruluk=0.795 F1-Skor=0.79
Taşkın ve diğ. (2022) [60]	GRU	F1-Skor=0.83

Tablo 3 incelendiğinde en düşük performans CNN ve Bi-LSTM metotları kullanılarak Wang ve diğerleri elde etmiştir. Buradaki performans değerinin düşük çıkma sebebi veri setinde aşırı uydurma (overfitting) durumudur [16]. Roy ve arkadaşlarına ait diğer düşük performansın sebebi ise kullandıkları veri setinde anlam bakımından ilişkili sınıfların tahmin sırasını örtüşme yaşamasıdır [53]. Doğruluk performansı 0.954 olan Paka ve arkadaşlarına ait çalışmada BERT kullanılmıştır. BERT, veri setinde bulunan metni çift taraflı incelemekte ve metin içerisindeki herhangi bir kelimenin sağ ve solundaki kelimelerle olan ilişkilerini çok iyi kavrayabilecek bir yapı sunmaktadır [58].

### 3.3. LİTERATÜRDEKİ VERİ KÜMELERİ

Bu bölümde, sahte haber tespiti çalışmalarında kullanılan mevcut veri setlerinin bir listesi sunulmuştur. Ayrıca bu veri setlerinin açıklamaları ve sosyal bağlama dayalı özelliklerinin karşılaştırılması Tablo-4 de verilmiştir.

**Tablo 4.** Sahte haber tespiti için oluşturulan veri setlerinin sosyal bağlama dayalı olarak karşılaştırılması

Veri Seti	Sosyal Bağlam Özelliği			Açıklama
	Kullanıcı Detayı	Gönderi Analizi	Ağ Analizi	
BuzzFeedNews	-	-	-	BuzzFeed'den gerçekliği doğrulanmış 1627 adet haber makalesinden oluşmaktadır.
LIAR	-	-	-	PolitiFact'ten bir API kullanılarak çeşitli alanlardan 12836 adet makaleden oluşmaktadır[16].
BS Detector	-	-	-	Belirli bir web sayfasındaki tüm bağlantıları tarayarak ve önceden sağlanan bir alan listesiyle karşılaştırılarak güvenilir olmayan kaynakları tespit eden araç setinin kendi kendine açıklama özelliği ile oluşturulmuştur.

ISOT	-	-	-	21417 adet reuters.com'dan alınmış gerçek haber, 23481 adet farklı haber kaynaklarından alınan sahte haberden oluşmaktadır [61,62].
Fake News Challenge (FNC)	-	Evet	-	Siyaset, sağlık, çevre, yaşam tarzı gibi çeşitli konuları içeren eğitim seti 49972, test seti 25413 adet kayıttan oluşmaktadır.
CREDBANK	Evet	Evet	-	Ekim 2015'e ait 96 günlük bir süre boyunca çekilen 60 milyon tweetten oluşmaktadır.
BuzzFace	-	Evet	-	1,6 milyon yorum ve 2263 makaleden oluşan BuzzFeed'in bir uzantısıdır [63]
FacebookHoax	Evet	Evet	-	1 milyonun üzerinde beğeniye sahip, bilimsel ve komplo teorileriyle ilgili 15500 gönderiden oluşmaktadır [41]
FakeNewsNet	Evet	Evet	Evet	PolitiFact ve GossipCop'tan elde edilen haber-bağlam, sosyal-bağlam ve zamansal bilgiler gibi çeşitli kategorilerdeki bilgileri içeren çok boyutlu bir veri setidir.

#### 4. SONUÇ (CONCLUSION)

Sahte haberlerin toplumu olumsuz şekilde etkilediği endişesiyle bu sorunu durdurmak maksadıyla önlemler alınması gerekmektedir. Bu makale, kapsamlı bir yayın olmakla birlikte sahte haberlerin üretilmesini ve yayılmasını anlamaya ve azaltılmasına yönelik yapılan araştırmaların ne şekilde ilerlediğine dair bir fikir sunmaktadır. Spesifik olarak, çalışmamızın teoride ve pratikte çeşitli sonuçları bulunmaktadır.

Teoride, bu makale aktarılan bilginin amacına ve türüne dayalı olarak sahte haberlerle ilişkili tüm terimler hakkında bir fikir vermektedir. Sahte haberlerin yayılmasında psikolojik zafiyetlere, toplumsal faktörlere ve sosyal medyanın hayatımızdaki rolüne ışık tutmaktadır. Ayrıca, sahte haberlerin ayırt edilebileceği haber içeriği ve sosyal bağlam gibi çeşitli yönleri araştırmaktadır. Sosyal bağlama dayalı olarak sahte haberlerin tespiti için önceki çalışmalar hakkında kapsamlı bir sunum yapmaktadır. Bu araştırma aracılığıyla, sahte haberleri azaltmaya yönelik tekniklerin çoğunun, denetimli bir yaklaşım izlediği gözlemlenmiştir. Bu makalede, okuyucular arasında sahte haberlere ilişkin sağlam bir bilgi birikimi oluşturabilmek amacıyla kavramlar, uygun alıntılarla desteklenerek mevcut çalışmaların ayrıntılı bir araştırması yapılmıştır. Pratikte, incelenen veri setleri, sınıflandırıcıları geliştirmek ve bu çalışmayı ileriye taşımak isteyen okuyucular için yol gösterici olacaktır.

Bu çalışmada, sosyal bağlamın sahte haber tespiti üzerindeki etkisi literatür katkısıyla anlatılmıştır. Sosyal bağlamı modellemenin önemi gösterilmiştir. Sosyal bağlam bilgilerinin sahte haber tespiti yapmakta nasıl kullanıldığı konusunda bilgilendirme yapılmıştır. Sosyal bağlamın etkisi kullanıcı tabanlı, gönderi tabanlı ve ağ tabanlı olmak üzere üç ana başlıkta incelenmiştir. Her bakış açısından örnekler ve açıklamalar verilerek okuyucuya sahte haber tespitinde sosyal bağlamın gerekliliğinden bahsedilmiştir. Sahte haber tespiti için kullanılan makine öğrenmesi ve derin öğrenme modelleri kıyaslanmıştır. Ayrıca 9 birbirinden farklı ve ünlü veri seti ile ilgili bilgilendirme ve sosyal bağlam özellikleri bakımından kıyaslama yapılmıştır.

#### KAYNAKLAR (REFERENCES)

- [1] Savaş, S. (2020). @Sosyal Medya. Kutlu Yayınevi.
- [2] Çömlekçi, M. F. (2019). Sosyal medyada dezenformasyon ve haber doğrulama platformlarının pratikleri. Gümüşhane Üniversitesi İletişim Fakültesi Elektronik Dergisi, 7(3), 1549-1563.
- [3] Kiraz, E. (2020). Sosyal Medyada Sahte Haberin Yayılmasında Kullanıcı Faktörü. İnönü Üniversitesi İletişim Fakültesi Elektronik Dergisi (İNİF E-Dergi), 5(1), 9-24.

- [4] Savaş, S., & Topaloğlu, N. (2015). Sosyal Medya Verileri Üzerinden Siber İstihbarat Faaliyetleri. VIII. Uluslararası Bilgi Güvenliği Ve Kriptoloji Konferansı, 1-7.
- [5] Shu, K., Wang, S., Lee, D., & Liu, H. (2020). Mining disinformation and fake news: Concepts, methods, and recent advancements. In *Disinformation, Misinformation, and Fake News in Social Media* (pp. 1-19). Springer, Cham.
- [6] Edson, C. T. Jr., Lim, Z.W., & Ling, R. (2018). Defining fake news a typology of scholarly definitions. *Digital journalism*, 6(2):137{153.
- [7] Shu, K., & Liu, H. (2019). Detecting fake news on social media. *Synthesis Lectures on Data Mining and Knowledge Discovery*.
- [8] Hidey, C., T., & Kathleen, M. (2018). Persuasive inuence detection: The role of argument sequencing. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- [9] Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019) Fake news: Fundamental theories, detection strategies and challenges. In *WSDM*.
- [10] Shu, K., Bernard, H. R. and Liu, H. (2018). Studying fake news via network analysis: Detection and mitigation. *CoRR*, abs/1804.10233.
- [11] Shu, K., Wang, S., & Liu, H. (2018). Understanding User Profiles on Social Media for Fake News Detection. In *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE.
- [12] Savaş, S., & Topaloğlu, N. (2016). Siber Güvenlikte Yeni Bir Boyut: Sosyal Medya İstihbaratı. presented at the XVIII. Akademik Bilişim Konferansı, Aydın, 8.
- [13] Cao, J., Guo, J., Li, X., Jin, Z., Guo, H., & Li, J. (2018). Automatic rumor detection on microblogs: A survey. *arXiv preprint arXiv:1807.03505*.
- [14] Bosselut, A., Rashkin, H., Sap, M., Malaviya, C., Celikyilmaz, A., & Choi, Y. (2019). Comet: Commonsense transformers for automatic knowledge graph construction. *arXiv preprint arXiv:1906.05317*
- [15] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1),22-36.
- [16] Wang, W. Y. (2017). "liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*.
- [17] Castillo, C.; Mendoza, M.; and Poblete, B. (2011). Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*, 675–684. ACM.
- [18] Wu, L., and Liu, H. (2018). Tracing fake-news footprints: Characterizing social media messages by how they propagate. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM)*, 637–645. ACM.
- [19] Ma, J., Gao, W., Wei, Z.; Lu, Y.; and Wong, K.-F. (2015). Detect rumors using time series of social context information on microblogging websites. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, 1751–1754. ACM.
- [20] Kim, J.; Tabibian, B.; Oh, A.; Schölkopf, B.; and Gomez-Rodriguez, M. (2018). Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM)*, 324–332. ACM.

- [21] Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2018). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. arXiv preprint arXiv:1809.01286.
- [22] Ozbay, F. A., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*, 540, 123174.
- [23] Ozbay, F. A., & Alatas, B. (2019) "A novel approach for detection of fake news on social media using metaheuristic optimization algorithms." *Elektronika ir Elektrotechnika* 25.4: 62-67.
- [24] Ozbay, F. A., & Alatas, B. (2021). Adaptive Salp swarm optimization algorithms with inertia weights for novel fake news detection model in online social media. *Multimedia Tools and Applications*, 80(26), 34333-34357.
- [25] Chauhan, T., & Palivela, H. (2021). Optimization and improvement of fake news detection using deep learning approaches for societal benefit. *International Journal of Information Management Data Insights*, 1(2), 100051.
- [26] Ansar, W., & Goswami, S. (2021). Combating the menace: A survey on characterization and detection of fake news from a data science perspective. *International Journal of Information Management Data Insights*, 1(2), 100052.
- [27] Kabudi, T., Pappas, I., & Olsen, D. H. (2021). Ai-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, 2, 100017.
- [28] Yang, S., Shu, K., Wang, S., Gu, R., Wu, F., and Liu, H. (2019). Unsupervised Fake News Detection on Social Media: A Generative Approach. In *AAAI'19*.
- [29] Shu, K., Mahudeswaran, D., and Liu, H. (2018). FakeNewsTracker: a tool for fake news collection, detection, and visualization. *Computational and Mathematical Organization Theory* (2018), 1–12.
- [30] Savaş, S. (2021, October). Big Data Analysis on Twitter for 2017 Turkey Referendum: TRefendum. In *2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (pp. 524-528). IEEE.
- [31] Savaş, S. (2022). Metin madenciliği yöntemiyle dijital katılım analizi ve 2017 Türkiye sosyal medya referandumunu. *Acta Infologica*.
- [32] Karimi, H., Roy, P., Saba-Sadiya, S., and Tang, J. Multi-Source Multi-Class Fake News Detection. In *COLING'18*.
- [33] Kwon, S., Cha, M., Jung, K., Chen, W., and Wang, Y. 2013. Prominent features of rumor propagation in online social media. In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*. IEEE, 1103–1108.
- [34] Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., Su, L., and Gao, J. EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection. In *KDD'18*.
- [35] Shu, K., Zhou, X., Wang, S., Zafarani, R., & Liu, H. (2019, August). The role of user profiles for fake news detection. In *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 436-439).
- [36] Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic detection of rumor on Sina Weibo. In *Proceedings of the ACM SIGKDD workshop on mining data semantics* (pp. 1–7).

- [37] Castillo, C., Mendoza, M., & Poblete, B. (2013). Predicting information credibility in time-sensitive social media. *Internet Research*.
- [38] Jin, Z., Cao, J., Zhang, Y., & Luo, J. (2016, March). News verification by exploiting conflicting social viewpoints in microblogs. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 30, No. 1).
- [39] Rubin, V. L., Conroy, N., Chen, Y., & Cornwell, S. (2016). Fake news or truth? Using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection* (pp. 7–17).
- [40] Verma, P. K., Agrawal, P., Amorim, I., & Prodan, R. (2021). Welfake: Word embedding over linguistic features for fake news detection. *IEEE Transactions on Computational Social Systems*.
- [41] Tacchini, E., Ballarin, G., Della Vedova, M. L., Moret, S., & de Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. *arXiv preprint arXiv:1704.07506*
- [42] Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on conference on information and knowledge management* (pp. 797–806).
- [43] Mohammad, S. M., Sobhani, P., & Kiritchenko, S. (2017). Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT)*, 17(3), 1–23.
- [44] Shu, K., Wang, S., & Liu, H. (2019, January). Beyond news contents: The role of social context for fake news detection. In *Proceedings of the twelfth ACM international conference on web search and data mining* (pp. 312-320).
- [45] Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9), e73791.
- [46] Jin, Z., Cao, J., Jiang, Y. & Zhang, Y. “News credibility evaluation on microblog with a hierarchical propagation model,” In *Data Mining (ICDM), 2014 IEEE International Conference, 2014*, 230–239.
- [47] Shrivastava, G., Kumar, P., Ojha, R. P., Srivastava, P. K., Mohan, S., & Srivastava, G. (2020). Defensive modeling of fake news through online social networks. *IEEE Transactions on Computational Social Systems*, 7(5), 1159–1167.
- [48] Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013, June). What yelp fake review filter might be doing? In *Proceedings of the international AAAI conference on web and social media* (Vol. 7, No. 1).
- [49] Granik, M., & Mesyura, V. (2017, May). Fake news detection using naive Bayes classifier. In *2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON)* (pp. 900-903). IEEE.
- [50] Gravanis, G., Vakali, A., Diamantaras, K., & Karadais, P. (2019). Behind the cues: A benchmarking study for fake news detection. *Expert Systems with Applications*, 128, 201-213.
- [51] Taşkın, S. G., Küçüksille, E. U., & Topal, K. Twitter üzerinde Türkçe sahte haber tespiti. *Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 23(1), 151-172.
- [52] Long, Y., Lu, Q., Xiang, R., Li, M., & Huang, C. R. (2017, November). Fake news detection through multi-perspective speaker profiles. In *Proceedings of the eighth international joint conference on natural language processing (volume 2: Short papers)* (pp. 252-256).

- [53] Roy, A., Basak, K., Ekbal, A., & Bhattacharyya, P. (2018). A deep ensemble framework for fake news detection and classification. arXiv preprint arXiv:1811.04670.
- [54] Jwa, H., Oh, D., Park, K., Kang, J. M., & Lim, H. (2019). exbake: Automatic fake news detection model based on bidirectional encoder representations from transformers (bert). *Applied Sciences*, 9(19), 4062.
- [55] Polignano, M., de Pinto, M. G., Lops, P., & Semeraro, G. (2019, September). Identification Of Bot Accounts In Twitter Using 2D CNNs On User-generated Contents. In *Clef (working notes)*.
- [56] Baruah, A., Das, K. A., Barbhuiya, F. A., & Dey, K. (2020, September). Automatic Detection of Fake News Spreaders Using BERT. In *CLEF (Working Notes)*.
- [57] Song, C., Ning, N., Zhang, Y., & Wu, B. (2021). A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Information Processing & Management*, 58(1), 102437.
- [58] Paka, W. S., Bansal, R., Kaushik, A., Sengupta, S., & Chakraborty, T. (2021). Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection. *Applied Soft Computing*, 107, 107393.
- [59] Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, 1(1), 100007.
- [60] Taskin, S. G., Kucuksille, E. U., & Topal, K. (2022). Detection of Turkish Fake News in Twitter with Machine Learning Algorithms. *Arabian Journal for Science and Engineering*, 47(2), 2359-2379.
- [61] Ahmed, H., Traore, I., & Saad, S. (2017, October). Detection of online fake news using n-gram analysis and machine learning techniques. In *International conference on intelligent, secure, and dependable systems in distributed and cloud environments* (pp. 127-138). Springer, Cham.
- [62] Ahmed, H., Traore, I., & Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Security and Privacy*, 1(1), e9.
- [63] Santia, G., & Williams, J. (2018). BuzzFace: A news veracity dataset with facebook user commentary and egos. In *Proceedings of the international AAAI conference on web and social media: vol. 12*.