



Sosyal medyada otomatik halk sağlığı takibi: Güncel bir derleme

Automatic public health monitoring on social media: A recent survey

Doğan Küçük^{1,*}, Nursal Arıcı², Emine Ela Küçük³

^{1,2} Gazi Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, 06500, Ankara Türkiye

³ Giresun Üniversitesi, Sağlık Bilimleri Fakültesi, Hemşirelik Bölümü, 28340, Giresun, Türkiye

Özet

Yaygın hastalıklar ve salgınlar gibi halk sağlığını durumlarının otomatik olarak belirlenerek takip edilmesi, güncel ve önemli bir araştırma problemidir. Günümüzde, sosyal medya metinleri analiz edilerek halk sağlığı takibi yapılabilmekte, toplumun sağlıkla ilgili eğilimleri ve algıları belirlenebilmektedir. Literatürde bu konularda gerçekleştirilmiş çalışmaların sayısı da hızla artış göstermektedir. Bu çalışmamızda, sosyal medya üzerinde halk sağlığı ile ilgili içerikleri tespit eden ve halk sağlığı takibi yapan çalışmaların güncel bir derlemesi sunulmaktadır. Söz konusu çalışmalar; salgınlar, hastalıklar, tıbbi gelişmeler, aşılar ve tamamlayıcı/alternatif tıp gibi halk sağlığı ile ilgili tüm konuları hedef alabilmektedir. Derlememizde, sosyal medyada otomatik halk sağlığı takibi konusundaki güncel çalışmalar alt konularına göre sınıflandırılarak sunulmuş olup, ilgili dijital kaynakları listelenmiş ve ayrıca ileri çalışma konularına yer verilmiştir. Derlememizin, sağlık bilişimi konusunda hem teorik hem de uygulamaya yönelik önemli bir kaynak olarak ilgili araştırmacı ve uzmanlara hizmet etmesi beklenmektedir.

Anahtar kelimeler: Halk sağlığı takibi, Sağlık bilişimi, Sosyal medya analizi, Makine öğrenmesi, Derin öğrenme

1 Giriş

Günümüzde tıp ve biyoloji alanları, bilgisayar bilimlerinin çeşitli alt alanları bünyesinde sunulan teknolojilerden olabildiğince faydalanmaktadır. Örneğin, görüntü işleme (image processing) ve bilgisayarlı görü (computer vision) teknolojileri kullanılarak tıbbi görüntüler otomatik olarak incelenerek çıkarımlar yapılabilmekte, biyoformatik alanında çok çeşitli veri analizi ve makine öğrenmesi yöntemleri kullanılarak büyük boyutlu veriler analiz edilebilmektedir.

Tıp ve biyoformatik alanları için önem teşkil eden bir diğer yaklaşım da metin biçimindeki verilerin otomatik analiz edilmesidir. Örneğin; tıbbi raporlar ve dokümanlar dijital ortama aktarılabilmişse, bu veriler işlenerek bunların içinden anlamlı veriler otomatik olarak çıkarılabilmektedir.

Halk sağlığı, tıbbin alt alanlarından biridir ve bu alt alanda halk sağlığı takibi (public health monitoring), elektronik sağlık raporları gibi resmi veriler üzerinde otomatik şekilde yapılabildiği gibi [1]; internet ortamındaki siteler, forumlar, arama motorlarının arama geçmişleri ve

Abstract

Automatic detection and monitoring of public health events and phenomena, like common diseases and epidemics, is an important research problem. Today, public health monitoring can be performed automatically on social media and health-related trends and perceptions of the society can be determined by analyzing social media texts. Related studies performed on these topics are increasing. In this study, a recent survey of the studies that detect public health related content on social media and that perform public health monitoring, is presented. Related studies can target at any public health related topics including epidemics, diseases, medical advances, vaccines, and complementary/alternative medicine. In our survey, those studies on automatic public health monitoring on social media are presented after they are categorized by their sub-topics, related digital resources are listed, and additionally, future research topics are included. It is expected that our survey will serve as an important theoretical and application-oriented resource for related researchers and experts.

Keywords: Public health monitoring, Health informatics, Social media analysis, Machine learning, Deep learning

sosyal medya uygulamaları (sosyal ağlar) kullanılarak da yapılabilmektedir. Örneğin, arama motorlarında kullanılan arama terimlerinin sıklığı kullanılarak hastalık takibi yapılabilmektedir. Bu alanda Google Flu Trends (GFT) gibi bazı Web servisleri hatalı sonuçlar vererek başarısız olmuşsa da, hatalarının düzeltilebilir nitelikte olduğu ve bu tip teknolojilerin sonuçlarının geleneksel hastalık takibi yöntemleriyle birleştirildiğinde, tek başına geleneksel yöntemlerden daha başarılı sonuçlar verdiği ilgili çalışmalarda gösterilmiştir [2].

Arama motorlarının tespit ettiği arama desenleri dışında; Twitter, Facebook, Instagram ve Reddit gibi sosyal medya metinlerinin de otomatik incelenerek hastalık, salgın gibi halk sağlığı konularının takibinin yapılması konusunda oldukça fazla sayıda ve güncel akademik yayın yapılmaktadır. Özellikle Twitter, sözü edilen akademik çalışmalarda en yaygın kullanılan sosyal medya sitelerindedir [3]. Hem arama motorları hem de sosyal medya siteleri ticari uygulamalar oldukları için bu

* Sorumlu yazar / Corresponding author, e-posta / e-mail: dogan.kucuk@gazi.edu.tr (D. Küçük)

Geliş / Received: 13.08.2020 Kabul / Accepted: 06.01.2021 Yayınlanma / Published: 27.07.2021

doi: 10.28948/ngumuh.778948

kanallardan toplanan verilerin gerçeği birebir yansıtmayacağı ifade edilmekte, sosyal medyanın ayrıca gerçek olmayan bilgilerin yayılması ve toplumun belirli konulara yönlendirilmesi için kötüye kullanıldığı da bilinmektedir [2]. Bununla birlikte, sosyal medyanın halk sağlığı takibi ve tahmininde oldukça önemli bir veri kaynağı olduğu da belirtilmektedir.

Dünyanın çeşitli bölgelerinde dönem dönem yaygın olarak grip (influenza), Ebola ve Severe Acute Respiratory Syndrome (SARS) gibi salgınlar görülmekte ve bu salgınlar çok ciddi problemlere neden olmakta, çok fazla sayıda insanın hayatını etkilemektedir. Son olarak, 2019 yılı sonunda ortaya çıkan ve 2020 yılının ilk yarısında tüm dünyayı çarpıcı şekilde etkisi altına alan yeni koronavirüs (Covid-19) salgını, halk sağlığı alanının önemini bir kez daha ortaya çıkarmıştır. Salgın tüm dünyayı sosyal ve ekonomik açıdan derinden etkilemiştir, etkileri de halen devam etmektedir. Sosyal medya üzerinde halk sağlığı takibinin özellikle bu gibi salgın durumlarında toplumlar ve devletler tarafından önlemlerin vakitlice alınması açısından önemli bir araç olarak işlev görebilmektedir. Chen vd., Covid-19 konusundaki herkese açık Twitter iletilerinden oluşan veri kümesini oluşturmuş ve tanıtmıştır. Bu çalışmada Türkçe dâhil 9 dilde Covid-19 hakkındaki iletiler derlenmiştir. Ancak salgının etkilerinin uzunca bir süre devam edeceği tahmin edildiğinden Covid-19 takibi konusunda bundan sonra da sosyal medyanın yaygın olarak akademik amaçlarla kullanılabilirliği öngörülmektedir [4]. Sosyal medya dışında, akademik metinler içinde Covid-19 salgınıyla ilgili kavramları otomatik tespit eden sistem önerileri de mevcuttur [5].

Şu noktayı önemle belirtmek gerekir: literatürde otomatik halk sağlığı takibi ile yakın anlama sahip veya halk sağlığı takibi dâhil başka adımları da içeren daha kapsamlı süreçleri ifade eden ilgili birçok terim mevcuttur. Bu terimler arasında halk sağlığı sürveyansı (public health surveillance)¹, sendromik sürveyans (syndromic surveillance), hastalık takibi (disease monitoring), hastalık sürveyansı (disease surveillance), salgınsal zekâ (epidemic intelligence), epidemiyoloji bilişimi (epidemiology informatics), dijital epidemiyoloji (digital epidemiology), sağlık takibi (health monitoring) ve sağlık sürveyansı (health surveillance) sayılabilir.

Bu çalışmamızda, sosyal medya iletileri kullanılarak halk sağlığı takibi konusunda yapılmış güncel çalışmalar derlenmiştir. Söz konusu çalışmalar, 2013-2020 yılları arasında yayınlanmış ilgili yayınlardan oluşmaktadır.

Bu çalışmanın geri kalanının yapısı aşağıdaki şekilde düzenlenmiştir: 2. Bölümde genel olarak halk sağlığı ve sosyal medya ilişkisi konusundaki genel kapsamlı ve güncel çalışmalardan bahsedilmiştir. 3. Bölümde ise sosyal medyada otomatik halk sağlığı takibi konusunda yapılmış çalışmalar sınıflandırılarak sunulmuştur. Biz bu

çalışmamızda, çalışmaları konularına göre sınıflandırmayı tercih ettik ve ilgili bölümü bu sınıflandırmaya uygun olarak alt bölümlere ayırdık. Burada, şunu da belirtmek gerekir ki, ilgili çalışmaların çok büyük bir kısmı Twitter üzerinde gerçekleştirilmiştir. Yani, sosyal medya türüne göre bir sınıflandırma yapılacak olursa Twitter üzerinde yapılmış çalışmaların sıklığının diğerlerine göre oldukça fazla olduğu görülecektir. Bu durumun, Twitter'ın diğer sosyal medya sitelerine (Facebook ve Instagram gibi) göre daha çok metin temelli olarak kullanılmasına dayandırılabilirliği düşünülmektedir. 4. Bölümde ilgili yayınlarda kullanıldığı belirtilen alana özgü dijital kaynaklar (sözlükler ve ontolojiler gibi), 5. Bölümde ise ileri çalışma konularından bahsedilmektedir. Son olarak, 6. Bölümde çalışmamızın sonuçları özet şeklinde verilmiştir.

2 Halk sağlığı ve sosyal medya

Halk sağlığı takibi geleneksel olarak resmi kaynaklara dayanılarak yapılmaktadır. Bu süreçte resmi klinik raporlardaki veriler kullanılmaktadır. Ancak sosyal medyanın özellikle son 10 yılda kullanımının çok fazla yaygınlaşmasıyla birçok alanda başarıyla kullanıldığı görülmüştür. Bu nedenle, sosyal medyanın halk sağlığı takibi ve tahmininde de kullanılabilirliği; geleneksel yöntemlere kıyasla sosyal medyadan daha kısa süreler içinde karar alıcıların kullanabileceği çıkarımlar elde edilebileceği öngörülmüştür [6]. Zhou vd. tarafından yayınlanan çalışmada; sosyal medyanın sağlık bilgi yönetiminde kullanımı üzerinde durulmuş, sosyal medyanın hem hasta-doktor ilişkilerini ve iletişimini iyileştirebileceği hem de yeni tıbbi bilgilerin yayılmasını ve keşfini olumlu yönde etkileyeceği belirtilmiştir [7].

Bu konuyla ilgili özellikle vurgulanması gereken bir konu, söz konusu sosyal medya iletilerinin biçimlendirilmiş değil doğal dilde ifadeler olması, dolayısıyla dillere özgü olarak alana özel terimlerin ekler alıp yapılarının değişebileceği, ayrıca sosyal medya iletilerinde dillerin yazım kurallarına riayet edilmeyip resmi olmayan günlük konuşma dilindeki ifadelerin kullanılabilmesidir. Bu durum da, metin-tabanlı diğer diğer sosyal medya analizi problemlerinde olduğu gibi, sistemlerin başarımlarını etkileyen önemli bir faktördür ve birçok çalışmada ifade edilmiştir [8, 9].

Twitter iletileri üzerinde çok çeşitli halk sağlığı konuları (27 tane) dikkate alınarak gerçekleştirilen dilbilimsel analize dayalı bir çalışmada [10], bu tip sosyal medya tabanlı bir yöntemin mevcut geleneksel yöntemleri tamamlayıcı nitelikte olduğu ve bu nedenle mevcut yöntemlerle beraber kullanılabilirliği belirtilmiştir.

Twitter ve halk sağlığı takibinin ilişkili konularından birisi de halk sağlığı konusunda ilgili karar vericilere Twitter yoluyla bilgi sağlanmasıdır. Kapp vd.'nin çalışmasında sağlıkla ilgili bilim insanlarının Twitter yoluyla bilimsel

¹ Halk sağlığı sürveyansı (halk sağlığı gözetimi); halk sağlığı takibini de kapsayan aktif bir süreç olarak tanımlanmaktadır. Halk sağlığı sürveyansında, halk sağlığı takibi dışında, sağlıkla ilgili verilerin analiz edilerek yorumlanması ve bu yolla halk sağlığı problemlerinin çözülmesi adımları da yer almaktadır. Bizim derleme çalışmamızda ele alınan çalışmaların bir kısmı takip

dışında analiz süreçlerini de içerseler de, genel olarak halk sağlığı sürveyansından çok halk sağlığı takibine yakın olduklarından, çalışmamızın halk sağlığı takibi konulu çalışmaların bir derlemesidir demek daha yerindedir. Bu nedenle çalışmamız boyunca, daha uygun bir ifade olan *halk sağlığı takibi* ifadesi kullanılmıştır.

sonuçlarını yayınlamaları ve bunların halk sağlığının karar vericilerine bu yolla ulaştırılması konusuna dikkat çekilmiştir [11].

Phan vd. sağlıkla ilgili sosyal medya sitelerinde insanların davranışlarını tespit edebilmek için derin öğrenme tabanlı bir yaklaşım sunmuşlardır. Söz konusu derin öğrenme yaklaşımına yazarları “Sosyal Kısıtlı Boltzmann Makinesi” adını vermişlerdir [12]. Grover vd. ise sağlıkla ilgili teknolojilerin Twitter iletileri üzerinde ne şekilde tartışıldığı konusunu araştırmıştır. Çalışmada; Twitter’da hangi sağlık teknolojilerinin tartışıldığı, hangi türdeki hastalıkların Twitter’da daha çok ifade edildiği, çeşitli hastalıkların hangi çözümlerle ilişkilendirildiği soruları konu edilmiştir [13].

Halk sağlığı için sosyal medya dâhil dijital kaynakların veri kaynağı olarak kullanılması ve ayrıca da yine halk sağlığı için yapay zekâ algoritmalarının kullanımı konularında yapılmış kapsamlı derleme (survey) çalışmalar da mevcuttur.

İlgili bir çalışmada, sosyal medyanın halk sağlığı takibi için geleneksel yöntemlerin yerini asla alamayacağı ancak tamamlayıcı bilgiler sağlayacakları ifade edilmektedir. Çalışmada ayrıca; sosyal medyanın hastalık ve salgın takibinde, hastalık ve salgınların tahmininde, salgın gibi kriz zamanlarında insanların temiz su ve ilaç gibi ihtiyaçlarını ifade etmede önemli roller oynayacağı belirtilmiştir [14].

Fang vd. tarafından hazırlanan derleme çalışmasında büyük veri işleme problemleri de dikkate alınarak sağlık bilişimi (health informatics) konusundaki makaleler derlenmiştir. Bu çalışmada, sağlık bilişimi, hastaların sağlığını tahmin edebilmek için birçok farklı sağlık alanından bilgi ve iletişim teknolojilerini kullanarak veri toplayan; tıp, biyotıp, hemşirelik ve bilgisayar bilimleri alanlarına dayanan multi-disipliner bir alan olarak tanımlanmıştır [15].

İlgili bir tez çalışmasında, sosyal medyada sağlıkla ilgili bilgi çıkarımı problemlerinde makine öğrenmesi ve derin öğrenme yaklaşımlarının etkinliği araştırılmıştır [16].

Ravi vd. tarafından yayınlanan çalışmada, sağlık bilişimi için derin öğrenme (deep learning) yöntemlerini kullanan akademik çalışmalar derlenmiştir [17]. Derin öğrenme, makine öğrenmesinin bir alt türü olmakla birlikte, klasik makine öğrenmesine göre daha fazla soyutlama katmanı kullanarak verinin üzerinde birden çok soyutlama seviyesinde temsilini sağlayan, ayrıca da genellikle klasik makine öğrenmesi yöntemlerine göre çok daha yüksek başarımla öğrenmeyi sağlayan yöntemlerdir [18-19]. Yine derin öğrenme algoritmalarını resmi sağlık kayıtları halk sağlığı takibi için kullanan çalışmalar üzerine bir derleme de Yao vd. tarafından yapılmıştır [1]. Joshi vd.’nin derleme çalışmasında metin biçimindeki (resmi veya Web üzerindeki) verilerden salgın tespiti konulu çalışmalar dikkate alınmıştır. Söz konusu derlemede, ilgili çalışmalar iki gruba ayrılmıştır: büyük boyutlu verilerden sağlıkla ilgili alt kümeleri tespit etmeyi konu alanlar ve sağlıkla ilgili olduğu belirlenmiş metinlerden salgın tahmini yapanlar [20].

Rabarison vd. Twitter’ın halk sağlığı kurumları ve toplum arasında çift yönlü iletişim için kullanılabilceği konusunu irdelemişlerdir [21].

Guidry vd., sağlıkla ilgili kuruluşların Ebola ve Ebola salgını ile ilgili Twitter iletilerini ve bunlara karşı toplumun yanıtlarını incelemişlerdir. Çalışmada sosyal medya iletileri incelenen kuruluşlar ABD’nin Hastalık Kontrol ve Önleme Merkezleri (Centers for Disease Control and Prevention - CDC), Dünya Sağlık Örgütü (World Health Organization - WHO) ve Sınır Tanımayan Doktorlar’dır (Médecins Sans Frontières - MSF) [22].

Jordan vd., Twitter üzerinde 2010-2017 tarihleri arasında yapılan halk sağlığı takibi ve tahmini konularındaki çalışmaları derleyerek sunmuşlardır [23].

Özetlemek gerekirse; sosyal medyanın halk sağlığında birçok farklı kullanım alanı mevcuttur. Bu alanlar arasında; halk sağlığı takibi ve tahmini, halk sağlığı kurumları ve toplum arasında gerekli etkileşimi sağlama (hastalara ve topluma yönelik etkili sağlık iletişimi), bireylerin sağlık riskleri ve daha iyi yaşam tarzı seçenekleri gibi konularda bilgilendirilmesi ile söz konusu sağlık risklerinin azaltılması ve ortadan kaldırılması için gerekli davranış ve motivasyonları sağlama sayılabilir.

Bizim güncel derleme çalışmamız, 2013-2020 tarihleri arasında yayınlanmış sosyal medya üzerinde halk sağlığı takibi konulu çalışmaları hedeflemiştir. Literatürdeki ilgili derleme çalışmalarından şu ölçütlerden biri veya daha fazlası yönüyle farklılık göstermektedir: (i) derlememizde sadece Twitter ile sınırlı kalmamış, genel olarak sosyal medya iletileri üzerinde yapılan güncel çalışmalar derlenmiştir, (ii) çalışmaları sınıflandırma şeklimiz kendimize özgü (Şekil 1) ve farklıdır, (iii) diğer derlemelerle bizim derlememizin dikkate aldıkları çalışmalar farklı zaman aralıklarına aittir, dolayısıyla farklılık göstermektedir.

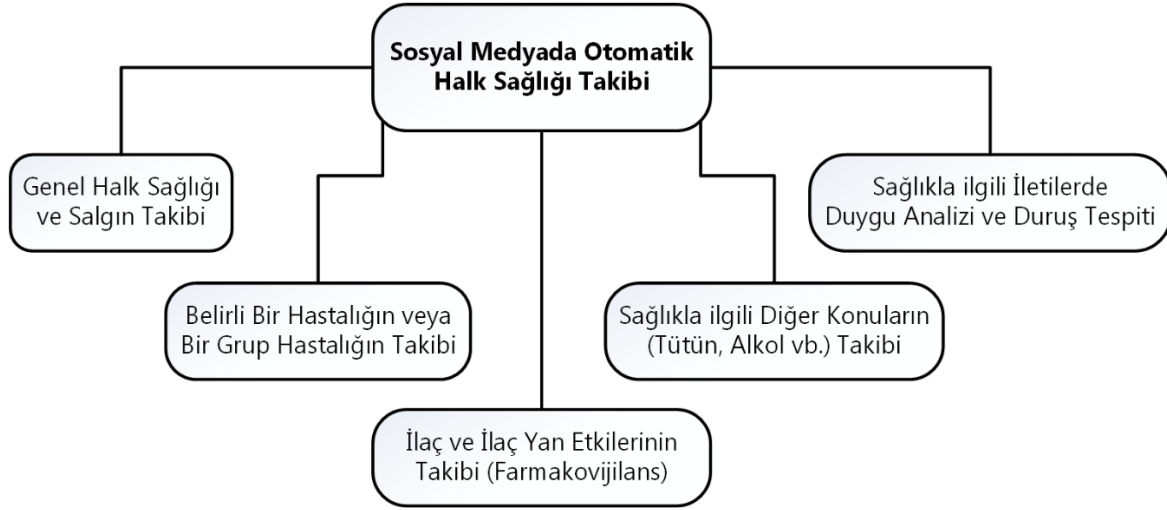
3 Sosyal medyada otomatik halk sağlığı takibi

Derlememizin kapsamına giren ilgili çalışmalar, daha önce belirtildiği gibi 2013-2020 yılları arasında yayınlanmışlardır. Bu çalışmalar, dikkate aldıkları alana özgü konuya göre sınıflandırılmışlardır. Şekil 1’de bu sınıflandırma görsel olarak verilmiş, bu bölümün aşağıdaki alt bölümlerinde de alt sınıflara ait çalışmalar tanıtılmıştır. Tanıtılan bu çalışmalar özet olarak bu bölümün sonunda Tablo 1’de sunulmuştur. Yine bu bölümün sonunda, bölüm içinde tanıtılan bu çalışmaların yayın yıllarına göre dağılımı çubuk grafik olarak Şekil 2’de verilmiştir. Bu çubuk grafik incelendiğinde, çok genel bir ifadeyle, ilgili yayınların son yıllarda artış gösterdiği gözlenebilmektedir. Ayrıca güncel Covid-19 salgını nedeniyle, bu artışın 2020 ve 2021 yıllarında devam edeceği beklenmektedir.

3.1 Genel halk sağlığı ve salgın takibi

Twitter iletileri üzerinde halk sağlığı takibi için Parker vd. mevcut Twitter veri kümelerini ve Twitter iletilerindeki sık kelime gruplarını filtrelemek için Wikipedia’yı kullanarak bir altyapı önermiştir [24]. Yine konuyla ilgili olarak Velardi vd. tarafından anahtar kelime tabanlı bir yaklaşım tanıtılmıştır. Söz konusu çalışmada önce hastaların terminolojisi öğrenilmiş, ardından belirti tabanlı bir yaklaşımla Twitter üzerinde halk sağlığıyla ilgili iletilerin sınıflandırılması gerçekleştirilmiştir [8].

Zhao vd.’nin ilgili bir çalışmasında [25] salgın takibi ve salgının mekânsal ve zamansal olarak gerçek-zamanlı



Şekil 1. Otomatik halk sağlığı takibi konulu güncel çalışmaların konularına göre sınıflandırılması

şekilde modellenmesi için derin öğrenme tabanlı bir yaklaşım sunulmuştur. Simnest adı verilen söz konusu altyapı çok katmanlı algılayıcı (MultiLayer Perceptron - MLP) algoritmasına dayanmaktadır.

Türkçe Twitter iletilerinin halk sağlığı ile ilgili olup olmadıklarını, çalışma kapsamında geliştirilen Türkçe halk sağlığı ontolojisini kullanarak tespit eden bir çalışmayı Küçük vd. sunmuştur [9]. Söz konusu çalışmada kullanılan ve hastalık, hastalık belirtileri ve yan etkileri gibi birçok halk sağlığı terimini içeren ontoloji ayrıca internetten açık olarak erişilebilir hale getirilmiştir.

Choi vd. tarafından gerçekleştirilen çalışmada [26] Kore’de yaşanan 2015 yılının Mayıs ve Temmuz ayları arasında yaşanan ve MERS-CoV adlı virüsün sebep olduğu Middle East Respiratory Syndrome (MERS) adlı hastalığın salgını konu edilmiştir. Bu çalışmada, kitle iletişim araçları ile sosyal medya verilerine dayanılarak halkın bu bulaşıcı hastalık salgınına karşı duygusal tepkileri makine öğrenmesi yöntemleri kullanılarak analiz edilmiştir.

Sarker vd. ise “Sağlık için Sosyal Medya Madenciliği – Social Media Mining for Health (SMM4H) 2017” başlıklı yarışmanın ayrıntılarını tanıtmıştır. Bu yarışmada tıpla ilgili metin sınıflandırma ve kavram normalizasyonu için veri kümeleri ile makine öğrenmesi ve derin öğrenme tabanlı farklı yaklaşımlar sunulmuştur [27].

Tutubalina vd. tarafından yapılan çalışmada, Twitter iletilerinde bahsi geçen tıbbi kavramların normalizasyonu için derin öğrenme (Recurrent Neural Network - RNN) tabanlı bir yaklaşım tanıtılmıştır [28].

Bir diğer çalışmada, bulaşıcı hastalıkların tespiti için istatistiksel bir yaklaşım ile Twitter’in anomali ve ani ortaya çıkış tespit algoritmaları karşılaştırılmıştır [29].

Twitter iletilerinden hastalık isimlerini tespit eden bir diğer çalışmada ontoloji kullanan ve derin öğrenme yaklaşımına dayalı bir yöntem tanıtılmıştır [30].

Jiang vd.’nin ilgili çalışmasında [31] insanların sağlıkla ilgili Twitter iletilerinin tespiti için derin öğrenme (Long Short-Term Memory - LSTM) tabanlı bir yöntem anlatılmıştır. Yine halk sağlığıyla ilgili Twitter iletilerini belirlemek için, metin ayrıştırmaya (parsing) dayalı

sözlüksel-sözdizimsel (lexico-syntactic) desenler gibi doğal dil işleme yöntemlerini kullanan güncel çalışmalar da mevcuttur [32].

Twitter üzerinde salgınların ortaya çıkışının tespit eden, bu amaçla Twitter iletileri üzerinde derin öğrenme tabanlı sınıflandırma yapan SENTINEL adlı güncel bir salgın takip sistemi Şerban vd. tarafından geliştirilmiş ve tanıtılmıştır. Sistemde ayrıca salgınlarla ilgili haber makalelerine de yer verilmektedir [33].

Khatua vd.’nin çalışmasında, Twitter iletileri ve PubMed makale özetleri üzerinde salgın tespiti yapmak amacıyla kelime vektörleri (word vectors) ve bir ağaç sınıflayıcısı kullanılmıştır. Çalışmada denemeler 2014 Ebola ve 2015 Zika salgınlarına ait veriler kullanılmıştır [34].

Daha güncel çalışmaların genellikle Covid-19 salgını ile ilgili olduğu gözlenmektedir.

İlgili güncel çalışmalardan biri olan Calamusa vd.’nin çalışmasında, Türkçe’ye “bilgi salgını” olarak çevrilebilecek “infodemic” terimi tanımlanmıştır [35]. Özellikle Covid-19 salgını ile yaygın kullanılmaya başlanan bu terim, bilginin istenmeyen ve mantıklı olmayan sonuçlara neden olacak şekilde kontrolsüz bir biçimde yayılması anlamında kullanılmaktadır [35]. Bu çalışmada, İtalya’da özellikle 21 Şubat’tan itibaren Twitter verileriyle kendini gösteren bir bilgi salgınının (infodemic) yayıldığı belirtilmektedir [35].

Guntuku vd.’nin çalışmasında Covid-19 salgını döneminde insanların Twitter üzerindeki paylaşımlarında akıl sağlığı ile ilgili (stres, endişe, gelecek kaygısı gibi) ve Covid-19 belirtileriyle ilgili bilgiler incelenmiştir ve istatistikler çıkarılmıştır [36].

Jahanbin ve Rahmanian’ın çalışmasında, Twitter üzerinde Covid-19 salgını ile ilgili iletilerin otomatik sınıflandırılması için bulanık kural-tabanlı evrimsel (fuzzy rule-based evolutionary) bir yaklaşım kullanılmıştır [37].

Kaur vd.’nin çalışmasında ise Covid-19 salgını hakkındaki Twitter iletileri üzerinde duyguların tespitine yönelik testlerin sonuçların sunulmuştur [38]. Çalışmanın amacının, insanların bu tipte bir salgın durumundaki duygularındaki değişimlerin izlenmesi olduğu ifade edilmiştir [38].

Covid-19 salgını ile ilgili Twitter iletilerinde popüler kelime ve kelime grupları ile yine bu iletiler üzerinde duygu analizi gerçekleştirilen bir çalışma Xue vd. tarafından sunulmuştur [39]. Çalışmada Latent Dirichlet Allocation (LDA) adlı bir makine öğrenmesi algoritması kullanılmıştır [39].

Zhang vd. ise Covid-19 salgını döneminde artış gösteren depresyon konulu Twitter iletilerinden oluşan geniş bir veri kümesi tanıtmış, derin öğrenme yöntemleri kullanılarak Twitter'da depresyon takibi yapılabileceğini ifade etmiştir [40]. Çalışma sonucunda, iletilerde Covid-19 ile ilgili referanslar arttıkça, depresyon ile ilgili ifadelerin de arttığı gözlemlenmiştir [40].

3.2 Belirli bir hastalığın veya bir grup hastalığın takibi

Grip (influenza) ve alt türleri (kuş gribi, domuz gribi) sıklıkla sosyal medya üzerinde takip edilen bir hastalıktır. Lee vd. tarafından sunulan çalışmada, Twitter üzerinde grip ve kanser takibi için mekansal, zamansal ve metin madenciliği gerçekleştiren ve gerçek-zamanlı olarak analiz sonuçlarını raporlayan bir sistem geliştirilmiştir [41].

Twitter üzerinden halk sağlığı takibi konusunda, özellikle de HIV takibi için Twitter'ın kullanılması konusunda Stoové ve Pedrana'nın görüşleri ilgili çalışmalarında sıralanmıştır. Çalışmada, sosyal medyanın diğer geleneksel verilerle birlikte kullanılması düşüncesini savunan çalışmaların varlığından da bahsedilmiştir [42].

Santillana vd.'nin çalışmasında ABD'de grip ve benzeri hastalıkların takibi için sosyal medya iletileri, arama motorlarında yapılan aramalar ve geleneksel (resmi) kaynakları kullanan makine öğrenmesi tabanlı bir yöntem tanıtılmıştır [43]. Byrd vd., Twitter iletilerinin anahtar kelime tabanlı olarak griple ilgili olup olmadıklarının takip edilebileceğini, grip salgınının Twitter üzerinde yapılabileceğini ve bu iletilere dayalı olarak gerçek zamanlı şekilde coğrafi haritalar üzerinde grip vakalarının görselleştirilebileceğini ifade etmişlerdir [44]. Comito vd. İtalya'da 2016-2017 ve 2017-2018 yıllarında mevsimsel grip vakalarını Twitter iletileri kullanarak analiz eden bir çalışma yapmış, elde edilen sonuçlar resmi sonuçlar ile uyumlu bulunmuştur [45]. Wakamiya vd. tarafından sunulan çalışma yine Twitter'da grip tespiti üzerinde durmuştur ve çalışmada 3 yıla yayılmış 7 milyon Japonca tweet üzerinde doğal dil işleme tabanlı bir yaklaşım kullanılmıştır [46]. Wakamiya vd.'nin sunduğu bir diğer çalışmada; belirlenen 8 belirtiyi içeren İngilizce, Japonca ve Çince Twitter iletilerinin gerçekten vaka belirtip belirtmediklerini sınıflandırmaya yönelik MedWeb isimli yarışmanın tanımı, veri kümeleri ve katılımcıların algoritmaları tanıtılmıştır [47].

Gkotsis vd.'nin çalışmasında Reddit adlı sosyal medya platformundaki iletilerin (belirlenen 11 bozukluk durumuna göre) akıl hastalığı durumu belirtip belirtmediklerini sınıflandırma amacına yönelik derin öğrenme tabanlı bir yaklaşım sunulmuştur [48].

Du vd. tarafından gerçekleştirilen derin öğrenme tabanlı çalışmada Twitter üzerinde saman nezlesi (polen alerjisi) ile ilgili iletilerin sınıflandırılması yapılmıştır [49].

3.3 İlaç ve ilaç yan etkilerinin takibi

İlaçların istenmeyen yan etkilerin ve birbirleriyle etkileşimlerinin tespiti, değerlendirilmesi ve önlenmesi çalışmaları olarak tanımlanabilen farmakovijilans da metin madenciliği uygulamalarının yaygın olarak kullanıldığı ve halk sağlığı ile yakından ilişkili bir konudur [50]. Twitter üzerinde antibiyotiklerle ilgili iletilerin sorgulandığı ve makine öğrenmesi yöntemleriyle sınıflandırıldığı bir çalışmayı Kendra vd. sunmuştur [51]. Sarker vd. sosyal medya üzerinde farmakovijilans konulu bir derleme yayınlamıştır [52]. Nikfarjam vd.'nin yine aynı tarihteki çalışmasında, Twitter iletileri üzerinde ilaçların istenmeyen etkilerine dair ifadeleri tespit eden Şartlı Rasgele Alanlar (Conditional Random Fields - CRF) adlı makine öğrenmesi yöntemine dayalı bir yaklaşım anlatılmıştır [53].

Xia vd.'nin tanıttığı bir çalışmada, sağlıkla ilgili sosyal medya forumlarındaki metinlerden ilaç isimleri ile ilaçların istenmeyen etkilerinin otomatik tespit edilmesi amaçlanmış ve derin öğrenme tabanlı bir çözüm sunulmuştur [54].

Twitter iletilerinde ilaçların istenmeyen etkilerinin tespitine yönelik bir diğer çalışmada [55] derin öğrenme (RNN algoritması) tabanlı bir yaklaşım anlatılmıştır.

Zhang vd. hastaların Twitter üzerinde kemoterapi hakkındaki iletilerinden onların bu tedavi hakkındaki algılarını belirlemek için derin öğrenme yaklaşımıyla (LSTM) duygu analizi problemini çözmeye çalışmıştır [56].

Fan vd. tarafından gerçekleştirilen ilgili bir diğer çalışmada, sağlıkla ilgili sosyal medya forumlarındaki metinler üzerinde ilaçların sebep olduğu istenmeyen olayların tespiti ve çıkarımı için derin öğrenme tabanlı bir yaklaşım geliştirilmiştir [57].

3.4 Sağlıkla ilgili diğer konuların (sigara, alkol vb.) takibi

Myślin vd. tarafından tanıtılan tütün konulu bir çalışmada, ilgili Twitter iletilerinden bu iletilerin sahiplerinin sigara içme ve yeni tütün ürünlerine yönelik duyguları belirlenmeye çalışılmış ve genel olarak örnek veri kümesinde tütüne karşı pozitif bir yaklaşım olduğu gözlemlenmiştir. Bu çalışmada birçok farklı makine öğrenmesi yöntemi denenmiştir [58].

Üniversite öğrencilerinin Facebook ve Twitter iletilerinde yaptıkları alkol referansları Moreno vd. tarafından yapılan çalışmada incelenmiştir. Çalışmada öğrencilerin Facebook üzerinde Twitter'a kıyasla daha fazla alkol referansı kullandıkları ve bu bulgunun izleme ve önleme süreçlerinde kullanılabileceği ifade edilmiştir [59].

Unger vd.'nin ilgili çalışmasında adolesan ve genç yetişkinlerin Twitter üzerinde tütün hakkında yayınladıkları iletiler incelenmiş ve Twitter'da bu şekilde ileti yayınlamakla tütün kullanımı arasında ilişki saptanmıştır [60].

3.5 Sağlıkla ilgili iletilerde duygu analizi ve duruş tespiti

Literatürdeki bazı çalışmalarda; halk sağlığı konularıyla ilgili toplumun sosyal medya iletileri yoluyla belirttikleri algılarının otomatik olarak tespit edilmesi amaçlanmaktadır. Bu konuyla ilgili olarak iki önemli probleme değinilmiştir: duygu analizi (sentiment analysis) ve duruş tespiti (stance detection).

Tablo 1. Sosyal medyada otomatik halk sağlığı takibi konulu çalışmalar

ÇalışmalarYayın	Konu	Yaklaşım
Parker vd. (2013)	Genel Halk Sağlığı	Wikipedia tabanlı kendine özgü yaklaşım
Velardi vd. (2013)	Genel Halk Sağlığı	Anahtar kelime tabanlı kendine özgü yaklaşım
Zhao vd. (2015)	Genel Halk Sağlığı	Derin öğrenme (MLP)
Küçük vd. (2017)	Genel Halk Sağlığı	Ontoloji tabanlı kendine özgü yaklaşım
Choi vd. (2017)	MERS Salgını	Makine öğrenmesi
Sarker vd. (2018)	Genel Halk Sağlığı	Makine öğrenmesi ve derin öğrenme
Tutubalina vd. (2018)	Genel Halk Sağlığı	Derin öğrenme (RNN)
Wiemken vd. (2018)	Genel Halk Sağlığı	İstatistiksel bir yaklaşım
Magumba vd. (2018)	Genel Halk Sağlığı	Derin öğrenme
Jiang vd. (2018)	Genel Halk Sağlığı	Derin öğrenme (LSTM)
Doan vd. (2019)	Genel Halk Sağlığı	Doğal dil işlemeye dayalı bir yaklaşım
Şerban vd. (2019)	Genel Halk Sağlığı	Derin öğrenme
Khatua vd. (2019)	Ebola ve Zika Salgınları	Ağaç sınıflandırıcı tabanlı yaklaşım
Calamusa vd. (2020)	Covid-19 Salgını	İstatistiksel bir yaklaşım
Guntuku vd. (2020)	Covid-19 Salgını ve Akıl Sağlığı	İstatistiksel bir yaklaşım
Jahanbin ve Rahmanian (2020)	Covid-19 Salgını	Bulanık kural-tabanlı evrimsel bir algoritma
Kaur vd. (2020)	Covid-19 Salgını	Doğal dil işlemeye dayalı bir yaklaşım
Xue vd. (2020)	Covid-19 Salgını	Latent Dirichlet Allocation (LDA)
Zhang vd. (2020)	Covid-19 Salgını ve Depresyon	Derin öğrenme
Lee vd. (2013)	Grip ve Kanser	Metin madenciliği
Stoové ve Pedrana (2014)	HIV	Sosyal medya ile geleneksel verilerin kullanımı
Santillana vd. (2015)	Grip ve Benzeri Hastalıklar	Makine öğrenmesi
Byrd vd. (2016)	Grip	Anahtar kelime tabanlı kendine özgü yaklaşım
Gkotsis vd. (2017)	Akıl Sağlığı	Derin öğrenme
Comito vd. (2018)	Grip	İstatistiksel bir yöntem ve duygu analizi yöntemi
Wakamiya vd. (2018)	Grip	Doğal dil işlemeye dayalı bir yaklaşım
Wakamiya vd. (2019)	8 Hastalık Belirtisi	Çeşitli algoritmalar
Du vd. (2019)	Saman Nezlesi	Derin öğrenme
Abacha (2015)	Farmakovijilans	Makine öğrenmesi
Kendra vd. (2015)	Antibiyotikler	Makine öğrenmesi
Sarker vd. (2015)	Farmakovijilans	Çeşitli algoritmalar (derleme)
Nikfarjam vd. (2015)	İlaçların İstenmeyen Yan Etkileri	Makine öğrenmesi (CRF)
Xia vd. (2017)	İlaçlar ve İstenmeyen Yan Etkiler	Derin öğrenme
Cocos vd. (2017)	İlaçların İstenmeyen Yan Etkileri	Derin öğrenme (RNN)
Zhang vd. (2018)	Kemoterapi	Derin öğrenme (LSTM)
Fan vd. (2020)	İlaçların İstenmeyen Yan Etkileri	Derin öğrenme
Myslín vd. (2013)	Sigara ve Tütün Ürünleri	Makine öğrenmesi
Moreno vd. (2016)	Alkol	İstatistiksel bir yaklaşım
Unger vd. (2018)	Tütün	İstatistiksel bir yaklaşım
Seltzer vd. (2017)	Zika Virüs Hakkındaki Duygular	İstatistiksel bir yaklaşım
Zhang vd. (2017)	Önleyici ve Tamamlayıcı Tedaviler	Derin öğrenme
Du vd. (2019)	HPV Aşısı Hakkındaki İnançlar	Derin öğrenme (RNN)

Duygu analizi genel olarak kutuplaşma tespiti (polarity detection) olarak tanımlanır ve verilen bir metnin pozitif, negatif veya nötral şeklinde sınıflandırılmasına karşılık gelir [61].

Duruş tespitinde ise belirli bir hedefe (kavram, olay vb.) karşı verilen metnin duruşu belirlenmeye çalışılır ve işlem sonuçta lehinde, aleyhinde veya hiçbiri şeklinde bir sınıflandırma sonucu elde edilir [62]. Halk sağlığı konusuyla ilgili bir çalışmada; sosyal medya uygulamalarından

Instagram üzerinde Zika virüs konulu resimler indirilerek konunun uzmanları tarafından incelenmiş, bu resimler hakkında çok çeşitli istatistiksel bilgiler çıkarılmış ve resimlerin hangi duyguları ifade ettikleri belirlenmiştir [63].

Daha önce bahsi geçen tütün konulu çalışmada da [58] yine bir halk sağlığı konusunda duygu analizi çalışması gerçekleştirilmiştir.

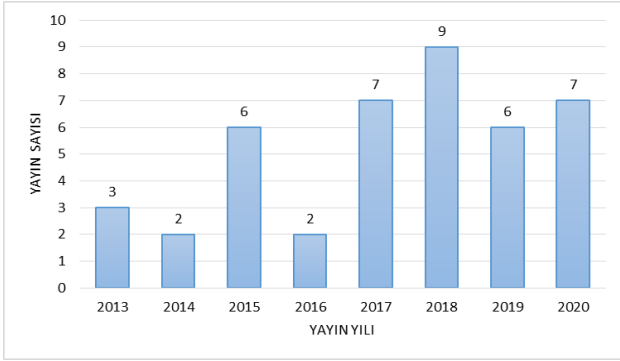
Zhang vd. tarafından gerçekleştirilen çalışmada, meme kanseri hakkındaki bir internet forumunda hastaların tartışmalı tamamlayıcı ve alternatif tıp tedavileri hakkındaki ifadeleri incelenmiştir. Bu tip tartışmaların ve hastaların bu tartışmalardaki duruşlarının tespiti için derin öğrenme tabanlı sınıflandırıcılar geliştirilmiş, her iki problem için de başarılı sonuçlar elde edilmiştir [64].

Daha önce bahsi geçen ve Comito vd. tarafından gerçekleştirilen grip salgını takibi konulu çalışmada ayrıca gribe karşı kullanılan tedaviye yönelik duygular da analiz edilmiş, genel olarak bu duyguların negatif oldukları gözlemlenmiştir [45].

Yine daha önce bahsedilen ve Zhang vd.'nin yayınladığı çalışmada, hastaların Twitter üzerinde kemoterapi hakkındaki duyguları bir derin öğrenme algoritması ile analiz edilmeye çalışılmıştır [56].

İlgili sayılabilecek güncel bir diğer çalışmada [65] Twitter iletilerinden Human Papillomavirus (HPV) aşısına karşı insanların inançlarını belirlemeye yönelik derin öğrenme (RNN) tabanlı bir yaklaşım sunulmuştur.

Daha önce Bölüm 3.1'de Covid-19 salgını ile ilgili bahsedilen çalışmalardan Kaur vd.'nin çalışmasında Covid-19 ile ilgili Twitter iletilerinde duyguların tespit edilmesine [38], Xue vd.'nin çalışmasında ise duygu analizine yönelik deneyler yapılmış ve sonuçları sunulmuştur [39].



Şekil 2. Bölüm 3'te sunulan çalışmaların yayın yıllarına göre dağılımı

4 Halk sağlığı takibi için dijital kaynaklar

Çalışmamızda bahsi geçen çalışmaların bir bölümü, açıkladıkları yaklaşımları kapsamında bir takım dijital kaynaklar kullanmışlardır ve bu durum diğer derleme çalışmalarında da ifade edilmiştir [20].

Söz konusu kaynaklardan en önemlileri, sağlık alanına özgü ontolojilerdir. Alan ontolojileri, ilgili alandaki kavram ve bunların birbiriyle ilişkilerini içeren anlamsal (semantik) kaynaklardır. Bu alandaki öncül ontolojilerden biri Birleştirilmiş Tıbbi Dil Sistemi'dir (Unified Medical Language System - UMLS) [66]. Benzer bir diğer dijital

kaynak türü de sözlüksel kaynaklar ve kelime listeleridir. Tıp alanıyla ilgili yaygın kullanımı olan bir kelime listesi Tıbbi Konu Başlıkları'dır (Medical Subject Headings - MeSH) [67].

Derlememizde değindiğimiz çalışmalardan Küçük vd., Tutubalina vd. ve Magumba vd. kullandıkları yöntemlerinde ontolojilerden yararlandıklarını ifade etmişlerdir [9,28,30].

Ontolojiler, sözlüksel kaynaklar ve kelime listeleri bilgi-tabanlı birçok sistemde olduğu gibi halk sağlığı takibine yönelik sistemler için de faydalı olabileceklerinden, ilgili araştırmacı ve geliştiricilerin bu tip kaynakları araştırmaları ve geliştirecekleri uygulamalarda kullanmalarının başarımlarını arttıracığı düşünülmektedir.

5 İleri çalışma konuları

Çalışmamızda derlenen araştırmalardan çıkarılacak önemli sonuçlardan bazıları şunlardır:

- Sosyal medyada halk sağlığı takibi ve tahmini, geleneksel olarak yapılan resmi kaynaklara dayalı takip ve tahmine kıyasla umut vaat edici sonuçlara ulaşılmasını sağlamaktadır ancak halen sosyal medya üzerinde takibin doğruluğunun artırılmasına gereksinim vardır. Bazı çalışmalarda, sosyal medyanın resmi raporlarla beraber kullanılmasının daha uygun olduğu ve daha başarılı sonuçlara neden olduğu vurgulanmaktadır.
- Sosyal medya üzerinde halk sağlığı, salgın veya hastalık tespiti ve tahmini için; metin ayrıştırma gibi doğal dil işlemeye dayalı yöntemlere başvurulduğu gibi, makine öğrenmesi ve son dönemde derin öğrenme yöntemlerine de başvurulabilmektedir.

Aşağıda sosyal medya üzerinde halk sağlığı takibi konusundaki üç önemli ileri araştırma konusuna yer verilmiştir.

5.1 Az çalışılan diller üzerinde çalışmalar yapılması

Çalışmamızda derlenen ilgili literatür incelendiğinde Türkçe gibi dillerdeki sosyal medya iletileri üzerinde daha az sayıda çalışma yapıldığı gözlenmektedir. Bu nedenle, salgınların ve diğer halk sağlığı konularının global ölçekte değerlendirilip izlenebilmesi için Türkçe ve benzeri az kaynaklı (low-resource) ve az çalışılmış diller üzerinde de geniş-ölçekli çalışmalar yapılması da önemli bir ileri araştırma konusudur.

Türkçe içerik üzerinde yapılmış çalışmalardan bir tanesi, daha önceki bölümlerde bahsi geçen ontoloji tabanlı sistemdir [9]. Türkçe ilgili literatürdeki kayda değer bazı çalışmalarda ise, T.C. Sağlık Bakanlığı gibi sağlıkla ilgili belirli Twitter hesaplarının paylaşımları üzerinde içerik analizi yapılarak bu analizlerin sonuçlarını sunulmuştur [68,69].

5.2 Değişken terimler ve normalizasyon

Sosyal medya, kullanıcıların istedikleri şekilde yazabildikleri ortamlar olduğundan; kullanıcılar sıklıkla yazım hataları, kısaltmalar, takma adlar kullanabilmektedir. Sağlıkla ilgili sosyal medya iletilerinde de tıbbi terimlere karşılık değişken terimler/isimler (variant terms/names) veya kısaltmalar kullanılabilmekte, bunlar yazılırken de

yazım hataları yapılabilmektedir. Bu nedenlerle, alana özgü terimlerin tespitinde sözü edilen durumlar dikkate alınmalı ve bozulmuş terimler düzgün formlarına dönüştürülmelidir, yani normalizasyon işlemi gerçekleştirilmelidir.

Sosyal medyada sağlık takibinin yüksek başarımla yapılabilmesi için yukarıda sayılan durumlar göz önünde bulundurularak tıbbi terimleri belirleme işlemi dikkatle yapılmalı ve terimler normalize edilmelidir. Örneğin, “grip” kelimesi “grip” olarak düzeltilmelidir.

5.3 Gerçek bilgiyle gerçek olmayan bilginin ayrıştırılması

Sosyal medyada yaygın olarak hastalık isimleri veya sağlıkla ilgili ifadeler abartma amacıyla kullanılabilir. Benzer şekilde yine hastalık belirten ifadeler deyimler içerisinde yer alabilmektedir. Bazen de sosyal medya kullanıcıları gerçek bir vaka olmadan belli bir hastalıktan korkularını iletilerinde ifade edebilmektedir.

Sosyal medya üzerinde sağlık takibinin yüksek başarımla yapılabilmesi için bu tip durumlar tespit edilerek filtrelenmelidir. Örneğin; “bu takım insana kalp krizi geçirir” gibi bir iletide gerçek “kalp krizi” vakasından bahsedilmeyip muhtemelen bir futbol takımının başarımlarından memnuniyetsizlik ifade edilmek istenmiştir.

6 Sonuçlar

Bu çalışmada, sosyal medya üzerinde halk sağlığı (veya sağlık) takibi konusunda yapılmış güncel çalışmaların bir derlemesi yer almaktadır. Kapsama alınan çalışmalar 2013-2020 yılları arasında yayınlanmış ilgili çalışmalardır. Bu güncel çalışmalar konularına göre kategorize edilerek beş farklı başlık altında sunulmuştur. Söz konusu başlıklar şunlardır: (1) *Genel Halk Sağlığı ve Salgın Takibi*, (2) *Belirli Bir Hastalığın veya Bir Grup Hastalığın Takibi*, (3) *İlaç ve İlaç Yan Etkilerinin Takibi*, (4) *Sağlıkla ilgili Diğer Konuların (Sigara, Alkol vb.) Takibi*, (5) *Sağlıkla ilgili İletilerde Duygu Analizi ve Duruş Tespiti*.

Taranan çalışmaların yıllara göre dağılımı incelendiğinde genel olarak ilgili yayınlarda bir artış olduğu gözlemlenmektedir. Özellikle 2019 yılı sonunda ortaya çıkan ve halen sürmekte olan Covid-19 salgını nedeniyle 2020 ve 2021 yıllarında bu salgının sosyal medya (özellikle Twitter) üzerinden takibi konusunda daha fazla çalışmanın yayınlanması beklenmektedir.

Son dönem çalışmalarında sadece halk sağlığı takibi değil, ilgili konular hakkında çoğunlukla duygu analizi ve duygu tespiti gibi analizler yapılmakta ve başarılı sonuçlar elde edilmektedir. Böylelikle, insanların halk sağlığını ilgilendiren konulardaki tutum, inanç ve duyguları belirlenmeye çalışılmaktadır.

Derin öğrenme yöntemlerinin birçok farklı alanda (geleneksel makine öğrenmesi yöntemlerine kıyasla) oldukça başarılı sonuçlara ulaşılmasını sağladığı bilinmektedir. Derlememizde yer alan halk sağlığı takibi alanındaki çalışmalar incelendiğinde de, özellikle yakın dönemde yayınlanan çalışmalarda derin öğrenme yöntemlerinin sıklıkla ve başarılı şekilde kullanıldıkları gözlemlenmektedir.

Derlememizde ayrıca, sosyal medyada halk sağlığı takibi alanındaki, ontolojiler ve kelime listeleri gibi, önemli dijital kaynaklardan bahsedilmiştir. Son olarak da, birçok dildeki

sosyal medya üzerinde çalışmalar yapılması, normalizasyon konusu ve gerçek bilginin gerçek olmayandan ayrımı gibi ileri çalışma konularına yer verilmiştir.

Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmektedir.

Benzerlik oranı (iThenticate): %4

Kaynaklar

- [1] Z. J. Yao, J. Bi, and Y. X. Chen, Applying deep learning to individual and community health monitoring data: a survey, *International Journal of Automation and Computing*, 15 (6), 643-55, 2018. <https://doi.org/10.1007/s11633-018-1136-9>
- [2] D. Lazer, R. Kennedy, G. King, and A. Vespignani, The parable of Google Flu: Traps in big data analysis, *Science*, 343 (6176), 1203-1205, 2014. <https://doi.org/10.1126/science.1248506>
- [3] Y. Pershad, P. T. Hangge, H. Albadawi, and R. Oklu, Social medicine: Twitter in healthcare, *Journal of Clinical Medicine*, 7 (6), 121, 2018. <https://doi.org/10.3390/jcm7060121>
- [4] E. Chen, K. Lerman, and E. Ferrara, #Covid-19: The first public coronavirus Twitter dataset, arXiv preprint arXiv:2003.07372, 2020.
- [5] M. Aydoğan, and A. Sener, An Artificial Intelligence Application in Health Developed on Covid-19 Documents, *Journal of Health, Medicine and Nursing*, 75, 58-66, 2020. <https://doi.org/10.7176/JHMN/75-08>
- [6] R. Thiébaud, and F. Thiessard, Artificial Intelligence in Public Health and Epidemiology, *Yearbook of Medical Informatics*, 27 (01), 207-10, 2018. <https://doi.org/10.1055/s-0038-1667082>
- [7] L. Zhou, D. Zhang, C. C. Yang, and Y. Wang, Harnessing social media for health information management, *Electronic Commerce Research and Applications*, 27, 139-51, 2018. <https://doi.org/10.1016/j.elerap.2017.12.003>
- [8] P. Velardi, G. Stilo, A. E. Tozzi, and F. Gesualdo, Twitter mining for fine-grained syndromic surveillance, *Artificial Intelligence in Medicine*, 61 (3), 153-63, 2014. <https://doi.org/10.1016/j.artmed.2014.01.002>
- [9] E. E. Küçük, K. Yapar, D. Küçük, and D. Küçük, Ontology-based automatic identification of public health-related Turkish tweets, *Computers in Biology and Medicine*, 83, 1-9, 2017. <https://doi.org/10.1016/j.combiomed.2017.02.001>
- [10] A. Culotta, Estimating county health statistics with Twitter, *SIGCHI Conference on Human Factors in Computing Systems*, pp. 1335-1344, 2014. <https://doi.org/10.1145/2556288.2557139>
- [11] J. M. Kapp, B. Hensel, and K. T. Schnoring, Is Twitter a forum for disseminating research to health policy makers?, *Annals of Epidemiology*, 25 (12), 883-7, 2015. <https://doi.org/10.1016/j.annepidem.2015.09.002>
- [12] N. Phan, D. Dou, B. Piniewski, and D. Kil, Social restricted Boltzmann machine: Human behavior

- prediction in health social networks, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 424-431, 2015. <https://doi.org/10.1145/2808797.2809307>
- [13] P. Grover, A. K. Kar, and G. Davies, Technology enabled Health–Insights from Twitter analytics with a socio-technical perspective, *International Journal of Information Management*, 43, 85-97, 2018. <https://doi.org/10.1016/j.ijinfomgt.2018.07.003>
- [14] I. C. H. Fung, Z. T. H. Tse, and K. W. Fu, The use of social media in public health surveillance, *Western Pacific Surveillance and Response Journal: WPSAR*, 6 (2), 3, 2015. <https://doi.org/10.5365/WPSAR.2015.6.1.019>
- [15] R. Fang, S. Pouyanfar, Y. Yang, S. C. Chen, and S. S. Iyengar, Computational health informatics in the big data age: a survey, *ACM Computing Surveys (CSUR)*, 49 (1), 1-36, 2016. <https://doi.org/10.1145/2932707>
- [16] A. Nikfarjam, Health information extraction from social media, Ph. D. thesis, Arizona State University, Tempe, AZ, 2016.
- [17] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G. Z. Yang, Deep learning for health informatics, *IEEE Journal of Biomedical and Health Informatics*, 21 (1), 4-21, 2017. <https://doi.org/10.1109/JBHI.2016.2636665>
- [18] I. Goodfellow, Y. Bengio and A. Courville. *Deep Learning*. Cambridge: MIT Press, 2016.
- [19] D. Küçük, and N. Arıcı, Doğal dil işleme derin öğrenme uygulamaları üzerine bir literatür çalışması, *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 2 (2), 76-86, 2018.
- [20] A. Joshi, S. Karimi, R. Sparks, C. Paris, and C. R. Macintyre, Survey of text-based epidemic intelligence: a computational linguistics perspective, *ACM Computing Surveys (CSUR)*, 52 (6), 1-19, 2019. <https://doi.org/10.1145/3361141>
- [21] K. M. Rabarison, M. A. Croston, N. K. Englar, C. L. Bish, S. M. Flynn, and C. C. Johnson, Measuring audience engagement for public health Twitter chats: insights from # LiveFitNOLA, *JMIR Public Health and Surveillance*, 3 (2), 34, 2017.
- [22] J. P. Guidry, Y. Jin, C. A. Orr, M. Messner, and S. Meganck, Ebola on Instagram and Twitter: How health organizations address the health crisis in their social media engagement, *Public Relations Review*, 43 (3), 477-486, 2017. <https://doi.org/10.1016/j.pubrev.2017.04.009>
- [23] S. E. Jordan, S. E. Hovet, I. C. H. Fung, H. Liang, K. W. Fu, and Z. T. H. Tse, Using Twitter for public health surveillance from monitoring and prediction to public response, *Data*, 4 (1), 6, 2019. <https://doi.org/10.3390/data4010006>
- [24] J. Parker, Y. Wei, A. Yates, O. Frieder, and N. Goharian, A framework for detecting public health trends with Twitter, *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 556-563, 2013 <https://doi.org/10.1145/2492517.2492544>
- [25] L. Zhao, J. Chen, F. Chen, W. Wang, C. T. Lu, and N. Ramakrishnan, Simnest: Social media nested epidemic simulation via online semi-supervised deep learning, *IEEE International Conference on Data Mining*, pp. 639-648, 2015. <https://doi.org/10.1109/ICDM.2015.39>
- [26] S. Choi, J. Lee, M. G. Kang, H. Min, Y. S. Chang, and S. Yoon, Large-scale machine learning of media outlets for understanding public reactions to nation-wide viral infection outbreaks, *Methods*, 129, 50-59, 2017. <https://doi.org/10.1016/j.ymeth.2017.07.027>
- [27] A. Sarker et al., Data and systems for medication-related text classification and concept normalization from Twitter: Insights from the Social Media Mining for Health (SMM4H)-2017 shared task, *Journal of the American Medical Informatics Association*, 25 (10), pp. 1274-1283, 2018. <https://doi.org/10.1093/jamia/ocy114>
- [28] E. Tutubalina, Z. Miftahutdinov, S. Nikolenko, and V. Malykh, Medical concept normalization in social media posts with recurrent neural networks, *Journal of Biomedical Informatics*, 84, 93-102, 2018. <https://doi.org/10.1016/j.jbi.2018.06.006>
- [29] T. L. Wiemken et al., Methods for computational disease surveillance in infection prevention and control: statistical process control versus Twitter's anomaly and breakout detection algorithms, *American Journal of Infection Control*, 46 (2), 124-132, 2018. <https://doi.org/10.1016/j.ajic.2017.08.005>
- [30] M. A. Magumba, P. Nabende, and E. Mwebaze, Ontology boosted deep learning for disease name extraction from Twitter messages, *Journal of Big Data*, 5 (1), 31, 2018. <https://doi.org/10.1186/s40537-018-0139-2>
- [31] K. Jiang, S. Feng, Q. Song, R. A. Calix, M. Gupta, and G. R. Bernard, Identifying tweets of personal health experience through word embedding and LSTM neural network, *BMC Bioinformatics*, 19 (8), 210, 2018. <https://doi.org/10.1186/s12859-018-2198-y>
- [32] S. Doan, E. W. Yang, S. S. Tilak, P. W. Li, D. S. Zisook, and M. Torii, Extracting health-related causality from Twitter messages using natural language processing, *BMC Medical Informatics and Decision Making*, 19 (3), 79, 2019. <https://doi.org/10.1186/s12911-019-0785-0>
- [33] O. Şerban, N. Thapen, B. Maginnis, C. Hankin, and V. Foot, Real-time processing of social media with SENTINEL: a syndromic surveillance system incorporating deep learning for health classification, *Information Processing & Management*, 56 (3), 1166-1184, 2019. <https://doi.org/10.1016/j.ipm.2018.04.011>
- [34] A. Khatua, A. Khatua, and E. Cambria, A tale of two epidemics: Contextual Word2Vec for classifying Twitter streams during outbreaks, *Information Processing & Management*, 56 (1), 247-257, 2019. <https://doi.org/10.1016/j.ipm.2018.10.010>
- [35] A. Calamusa, S. Tardelli, M. Avvenuti, S. Cresci, I. Federigi, M. Tesconi, M. Verani, and A. Carducci, Twitter monitoring evidence of Covid-19 infodemic in

- Italy, *European Journal of Public Health*, 30 (5), ckaa165-066, 2020.
- [36] S. C. Guntuku, G. Sherman, D. C. Stokes, A. K. Agarwal, E. Seltzer, R. M. Merchant, L. H. Ungar. Tracking mental health and symptom mentions on Twitter during COVID-19, *Journal of General Internal Medicine*, 35 (9), 2798-800, 2020. <https://doi.org/10.1007/s11606-020-05988-8>
- [37] K. Jahanbin and V. Rahmadian. Using Twitter and web news mining to predict COVID-19 outbreak. *Asian Pacific Journal of Tropical Medicine*, 13 (8), 378-380, 2020. <https://doi.org/10.4103/1995-7645.279651>
- [38] S. Kaur, P. Kaul, and P. M. Zadeh. Monitoring the Dynamics of Emotions during COVID-19 Using Twitter Data, *Procedia Computer Science*, 177, 423-430, 2020. <https://doi.org/10.1016/j.procs.2020.10.056>
- [39] J. Xue, J. Chen, R. Hu, C. Chen, C. Zheng, Y. Su and T. Zhu. Twitter Discussions and Emotions about the COVID-19 Pandemic: Machine Learning Approach, *Journal of Medical Internet Research*, 22 (11), e20550, 2020. <https://doi.org/10.2196/20550>
- [40] Y. Zhang, H. Lyu, Y. Liu, X. Zhang, Y. Wang, and J. Luo. Monitoring Depression Trend on Twitter during the COVID-19 Pandemic, *arXiv preprint arXiv:2007.00228*, 2020.
- [41] K. Lee, A. Agrawal, and A. Choudhary, Real-time disease surveillance using Twitter data: demonstration on flu and cancer, *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1474-1477, 2013. <https://doi.org/10.1145/2487575.2487709>
- [42] M. A. StooVé, and A. E. Pedrana, Making the most of a brave new world: Opportunities and considerations for using Twitter as a public health monitoring tool, *Preventive Medicine*, 63, 109-111, 2014. <https://doi.org/10.1016/j.ypmed.2014.03.008>
- [43] M. Santillana, A. T. Nguyen, M. Dredze, M. J. Paul, E. O. Nsoesie, and J. S. Brownstein, Combining search, social media, and traditional data sources to improve influenza surveillance, *PLoS Computational Biology*, 11 (10), 2015. <https://doi.org/10.1371/journal.pcbi.1004513>
- [44] K. Byrd, A. Mansurov, and O. Baysal, Mining Twitter data for influenza detection and surveillance, *International Workshop on Software Engineering in Healthcare Systems*, pp. 43-49, 2016. <https://doi.org/10.1145/2897683.2897693>
- [45] C. Comito, A. Forestiero, and C. Pizzuti, Twitter-based influenza surveillance: An analysis of the 2016-2017 and 2017-2018 seasons in Italy, *International Database Engineering & Applications Symposium*, pp. 175-182, 2018. <https://doi.org/10.1145/3216122.3216128>
- [46] S. Wakamiya, Y. Kawai, and E. Aramaki, Twitter-based influenza detection after flu peak via tweets with indirect information: text mining study, *JMIR Public Health and Surveillance*, 4 (3), e65, 2018.
- [47] S. Wakamiya, M. Morita, Y. Kano, T. Ohkuma, and E. Aramaki, Tweet classification toward Twitter-based disease surveillance: New data, methods, and evaluations, *Journal of Medical Internet Research*, 21 (2), e12783, 2019. <https://doi.org/10.2196/12783>
- [48] G. Gkotsis et al., Characterisation of mental health conditions in social media using informed deep learning, *Scientific Reports*, 7, 45141, 2017. <https://doi.org/10.1038/srep45141>
- [49] J. Du, S. Michalska, S. Subramani, H. Wang, and Y. Zhang, Neural attention with character embeddings for hay fever detection from Twitter, *Health Information Science and Systems*, 7 (1), 21, 2019. <https://doi.org/10.1007/s13755-019-0084-2>
- [50] A. B. Abacha, M. F. M. Chowdhury, A. Karanasiou, Y. Mrabet, A. Lavelli, and P. Zweigenbaum, Text mining for pharmacovigilance: Using machine learning for drug name recognition and drug-drug interaction extraction and classification, *Journal of Biomedical Informatics*, 58, 122-132, 2015. <https://doi.org/10.1016/j.jbi.2015.09.015>
- [51] R. L. Kendra, S. Karki, J. L. Eickholt, and L. Gandy, Characterizing the discussion of antibiotics in the Twittersphere: What is the bigger picture?, *Journal of Medical Internet Research*, 17 (6), e154, 2015. <https://doi.org/10.2196/jmir.4220>
- [52] A. Sarker et al., Utilizing social media data for pharmacovigilance: a review, *Journal of Biomedical Informatics*, 54, 202-212, 2015. <https://doi.org/10.1016/j.jbi.2015.02.004>
- [53] A. Nikfarjam, A. Sarker, K. O'Connor, R. Ginn, and G. Gonzalez, Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features, *Journal of the American Medical Informatics Association*, 22 (3), 671-681, 2015. <https://doi.org/10.1093/jamia/ocu041>
- [54] L. Xia, G. A. Wang, and W. Fan, A deep learning based named entity recognition approach for adverse drug events identification and extraction in health social media, *International Conference on Smart Health*, pp. 237-248, 2017. https://doi.org/10.1007/978-3-319-67964-8_23
- [55] A. Cocos, A. G. Fiks, and A. J. Masino, Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts, *Journal of the American Medical Informatics Association*, 24 (4), 813-821, 2017. <https://doi.org/10.1093/jamia/ocw180>
- [56] L. Zhang, M. Hall, and D. Bastola, Utilizing Twitter data for analysis of chemotherapy, *International Journal of Medical Informatics*, 120, 92-100, 2018. <https://doi.org/10.1016/j.ijmedinf.2018.10.002>
- [57] B. Fan, W. Fan, and C. Smith, Adverse drug event detection and extraction from open data: A deep learning approach, *Information Processing & Management*, 57 (1), 102131, 2020. <https://doi.org/10.1016/j.ipm.2019.102131>
- [58] M. Myslín, S. H. Zhu, W. Chapman, and M. Conway, Using Twitter to examine smoking behavior and perceptions of emerging tobacco products, *Journal of*

- Medical Internet Research, 15 (8), e174, 2013. <https://doi.org/10.2196/jmir.2534>
- [59] M. A. Moreno, A. Arseniev-Koehler, D. Litt, and D. Christakis, Evaluating college students' displayed alcohol references on Facebook and Twitter, *Journal of Adolescent Health*, 58 (5), 527-532, 2016. <https://doi.org/10.1016/j.jadohealth.2016.01.005>
- [60] J. B. Unger et al., Talking about tobacco on Twitter is associated with tobacco product use, *Preventive Medicine*, 114, 54-56, 2018. <https://doi.org/10.1016/j.ypmed.2018.06.006>
- [61] B. Pang, and L. Lee, Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval*, 2 (1-2), 1-135, 2008.
- [62] S. Mohammad, S. Kiritchenko, P. Sobhani, X. Zhu, and C. Cherry, SemEval-2016 task 6: Detecting stance in tweets, *International Workshop on Semantic Evaluation (SemEval-2016)*, pp. 31-41, 2016.
- [63] E. K. Seltzer, E. Horst-Martz, M. Lu, and R. M. Merchant, Public sentiment and discourse about Zika virus on Instagram, *Public Health*, 150, 170-175, 2017. <https://doi.org/10.1016/j.puhe.2017.07.015>
- [64] S. Zhang, L. Qiu, F. Chen, W. Zhang, Y. Yu, and N. Elhadad, We make choices we think are going to save us: Debate and stance identification for online breast cancer CAM discussions, *International Conference on World Wide Web Companion*, pp. 1073-1081, 2017. <https://doi.org/10.1145/3041021.3055134>
- [65] J. Du et al., Leveraging deep learning to understand health beliefs about the Human Papillomavirus Vaccine from social media, *NPJ Digital Medicine*, 2 (1), 27, 2019. <https://doi.org/10.1038/s41746-019-0102-4>
- [66] D. A. Lindberg, B. L. Humphreys, and A. T. McCray, The unified medical language system, *Yearbook of Medical Informatics*, 2 (01), 41-51, 1993. <https://doi.org/10.1055/s-0038-1634945>
- [67] C. E. Lipscomb, Medical subject headings (MeSH), *Bulletin of the Medical Library Association*, 88 (3), 265, 2000.
- [68] A. A. Çobaner, S. Köksoy. Sağlık alanında sosyal medyanın kullanımı: Twitter'da sağlık mesajları, *Akademik Konferans Bildirileri*, 899-906, 2014.
- [69] A. Bilgiç, S. S. Akyüz. Türkiye'de Covid-19 pandemisi döneminde Sağlık Bakanı Fahrettin Koca'nın sosyal medya kullanımı: Twitter paylaşımları içerik analizi, *Gaziantep Üniversitesi Sosyal Bilimler Dergisi*, 19(Covid-19 Özel Sayı), 230-243, 2020.

