

# BÜYÜK VERİ VE SAĞLIK HİZMETLERİNDE BÜYÜK VERİ İŞLEME ARAÇLARI

Arzu KURŞUN \*


## ÖZ

Sağlık hizmetlerinde büyük veri; hastane, doktorlar, hasta ve tıbbi süreçlerin bir bütünüdür. Karmaşıktır ve bu verilerin geleneksel veri analitiği araçları kullanılarak analiz edilmesi oldukça zordur. Sağlık hizmetleri, finansal etüdler, sağlık teknolojileri, sosyal faktörler ve örgütsel süreçler ile ilgili çok disiplinli bir bilimsel araştırma alanı sağlamaktadır. Büyük verinin yardımıyla, büyük miktarda veri depolanabilmekte ve teşhis için etkin bir şekilde kullanılabilir. Ayrıca bu gelişen teknolojilerle hastalıkların uygun tedavisi de izlenebilmektedir. Araştırmacılar sağlık hizmetlerindeki büyük hacimli veriyi inceleyerek yeni model ve eğilimleri bulmaktadırlar. Böylece sağlık hizmetleri maliyetlerini düşürme, sağlık erişimini demokratikleştirme ve insan hayatlarını kurtarma fırsatı sağlamaktadırlar. Son yıllarda, birçok araştırmacı, analitiksel doğruluğu iyileştirmek için sağlık hizmetleri verilerinde büyük veri yaklaşımlarını önermiştir. Sağlık hizmeti verilerinin etkili analizi ve tanımlanması, hastaların durumuna yeni bakış açıları sağlamak ve en uygun tedavi fırsatlarını önermektedir. Bu makalenin amacı, sağlık kurumlarında ve hizmetlerinde büyük veri analitiğine dayalı karar mekanizmalarını araştırmak, sağlık hizmeti liderlerinin kararlarına yardımcı olabilecek temel büyük veri analizlerini belirlemek ve sağlık hizmetleri boyunca verimliliği artırmak için bazı araçlar sunmaktır.

**Anahtar Kelimeler:** Sağlık bilişimi, büyük veri, hadoop, büyük veri araçları, sağlık hizmetleri

## MAKALE HAKKINDA

\* Dr. Öğr. Üyesi, Giresun Üniversitesi, Sağlık Hizmetleri MYO, Tıbbi Hizmetler ve Teknikler Bölümü, a.kurshun@gmail.com

 <https://orcid.org/0000-0001-9169-0577>

Gönderim Tarihi: 09.04.2021

Kabul Tarihi: 27.09.2021

## Atıfta Bulunmak İçin:

Kurşun, A. (2021). Büyük veri ve sağlık hizmetlerinde büyük veri işleme araçları. Hacettepe Sağlık İdaresi Dergisi, 24(4), 921-940

## BIG DATA AND HEALTHCARE BIG DATA PROCESSING TOOLS

Arzu KURŞUN \*

### ABSTRACT

*Healthcare big data is the whole of record of hospital, doctors, patient and medical processes and it is so complex and this data is quite hard to analyze using some traditional data analytics tools. Healthcare services ensures a multi-disciplinary field of scientific investigation in relation to organizational progresses, social factors, health technologies and financial studies. Thanks to big data, the huge volume of data can be processed and stored effectively for diagnosis. Also suitable treatment of diseases can be observed with these emerging technologies. In this way for researchers, there is a chance in healthcare data to find models and tendencies inside of data and ensure a resolution for developing healthcare, thereby decreasing costs, democratizing health access, and rescuing human lives. Recently, many researchers have suggested some big data approaches on healthcare data to develop the accuracy of analytics. Definition of health care data and effective analysis ensures new perspectives of patients status and propose the most proper treatment opportunities. The purpose of this article is to investigate the decision mechanisms based on big data analytics in healthcare institutions and services, to determine essential big data analytics able to help healthcare leaders' decisions and to present some tools to raise efficiency along the healthcare services.*

**Keywords:** Health informatics, big data, hadoop, big data tools, healthcare services

### ARTICLE INFO

\* Assist. Prof., Giresun University, a.kurshun@gmail.com

 <https://orcid.org/0000-0001-9560-7689>

Received: 09.04.2021

Accepted: 27.09.2021

### Cite This Paper:

Kurşun, A. (2021). Büyük veri ve sağlık hizmetlerinde büyük veri işleme araçları. Hacettepe Sağlık İdaresi Dergisi, 24(4), 921-940

## **I. GİRİŞ**

Sağlık biliminin gelişmesi ve yönetim kavramlarının yenilenmesiyle birlikte sağlık hizmetlerinin sağlık hizmetleri kalitesi, doktor-hasta iletişimi, hasta bilgileri gibi alanları gittikçe daha fazla ilgi görmüş ve bununla beraber yönetim bilimi kapsayandan ayrıntıya, gelenekselden yenilikçiliğe doğru bir seyir izlemiştir (Shi vd., 2020). Verilerin hızla sayısallaştırılmasıyla, sağlık sektörü hasta verilerinin elektronik olarak saklandığı, korunduğu ve işlendiği modern forma kaymıştır ve kayıt tutma, sağlık yönetimi, yasal gereklilikler ve hasta bakımı gibi süreçler çok büyük miktarda veri üretilmesine neden olmuştur (Shaikh ve Ali, 2019).

Büyük veri, bilgi olarak tanımlanmaktadır. Ancak bu tanımda büyük boyutlarda bir bilgiden söz edilmektedir. Muazzam büyüklükte ve zamanla katlanarak artan, gelişen bir ürün yelpazesini tasvir etmek için kullanılmaktadır (Vasuki ve Rajiv-Kannan, 2020). Büyük veri, bu bilgiyi incelemek için yeni teknolojileri içeren geleneksel veri tabanını kullanarak işlenmesi zor olan çok sayıda yapılandırılmış ve aynı zamanda yapılandırılmamış veriyi almak, toplamak, yönetmek ve analiz etmek için kullanılan bir tekniktir (Ramya-Devi ve Vijaya-Chamundeeswari, 2020).

Zamanla bu devasa miktarlardaki verilerin görselleştirilmesi gereklilik haline gelmiş; gelişimle birlikte sağlık hizmeti sunumunun kalitesini iyileştirme ve maliyetleri azaltma gibi hedefler ortaya çıkmıştır. Büyük veri, klinik karar desteklerini ve halk sağlığı yönetimini destekleme kapasitesine sahiptir (Shaikh ve Ali, 2019). Veri depolamanın ucuzlaması, bunların gerçek zamanlı analizi, hastanın hayati değerleri ve klinik verilerini toplamanın mümkün hale gelmesi ile birlikte, teşhis ve tedaviler büyük ölçekli bir hale dönüşmüştür. Ayrıca bu veriyi saklamak, analiz etmek ve yorumlamak giderek zorlaşmaktadır (Agrawal ve Prabakaran, 2020).

Tıbbi veriler ise, diğer alanlardaki büyük verilere kıyasla benzersiz özelliklere sahiptir. Görüntüleme, birçok farklı elektronik hasta kayıtları (EHR'ler), klinik kayıtlar, biyobankalar, internet ve hasta öz raporları dahil olmak üzere geniş bir kaynak çeşitliliğine sahiptir (Shilo vd., 2020). Böylece büyük veri mimarisi, birden çok kaynaktan gelen hacimli verileri kolayca yönetebilen karmaşık ve hacimli veriler için ideal bir çerçeve sağlayarak sağlık sistemini iyileştirebilme potansiyeline sahip olmuştur (Naoui vd., 2020).

Çalışmada büyük veri ve büyük verinin sağlık hizmetlerindeki önemi vurgulandıktan sonra literatür taramasıyla mevcut durum açıklanmış, sağlık hizmetlerinde büyük veri kullanım araçları tanımlanmış ve sağlık hizmeti verilerinin çeşitli biçimleri ifade edilerek bunlardan elde edilen verimli ve üretken bilgilerin entegrasyonunu sağlayan bir çerçeve çizilmiştir.

## **II. BÜYÜK VERİNİN TANIMLANMASI**

Bir terim olmasa da bir kavram olarak büyük verinin doğuşu, genellikle 2001 yılında Doug Laney tarafından hazırlanan “3-D data management: controlling data volume, velocity, and variety” adlı META Group raporuyla ilişkilendirilmektedir (Viceconti vd., 2015). Bir tanım olarak ise, odak noktası mevcut teknolojilerle yüksek hızda yakalama, keşif veya analiz sağlayarak çok çeşitli verilerden ekonomik olarak değer elde etmek için tasarlanan bilgidir (Aceto vd., 2020). Boyd ve Crawford (2012)’a göre büyük veri uygun maliyetli, yüksek hacimli, yüksek hızlı veya çok çeşitli bilgi varlığı olarak tanımlanmaktadır (Boyd ve Crawford, 2012).

Uluslararası Veri Kurumu (IDC) 2005 yılında dijital evrenin boyutunu 130 exabytes olarak tahmin etmiştir. 2017 yılında ise dijital evren yaklaşık 16.000 exabytesa genişlemiştir. IDC, dijital evrenin 2020 yılına kadar 40.000 exabytesa genişleyeceğini söylemiştir. Bu her bireyin yaklaşık 5200 gigabayt kullanması anlamına gelmektedir (Dash vd., 2019). Bu tür büyük veriler, politika geliştiriciler tarafından suç ve terörizmi azaltmak, sağlık hizmetlerini iyileştirmek veya şehirleri daha iyi yönetmek için kullanılabilir. Ayrıca araştırmalar, veri odaklı işletmelerin rakiplerinden %5 daha üretken ve %6 daha karlı olduğunu göstermiştir. Bunu fark eden kuruluşlar ve ülkeler büyük veriden yararlanmaya yönelmiştir (Strang ve Sun, 2020). Büyük veriyi oluşturan ve 5V olarak adlandırılan bileşenler vardır.

Bunlar (Demchenko vd., 2014; Viceconti vd., 2015; Molinari ve Nollo, 2020; Sasubilli ve Kumar, 2020):

- Hacim (*Volume*),
- Hız (*Velocity*),
- Çeşitlilik (*Variety*),
- Doğrulama (*Verification*) ve
- Değer (*Value*)

*Hacim:* Birçok şirket büyük miktarlarda arşivlenmiş veriye sahiptir, ancak verileri işleme kapasitesine sahip değildir. Büyük miktarda bilgiyi işleme yeteneğinden elde edilen fayda, büyük veri analitiğinin temel işlevidir. Hacimden kasıt, kaplanan yerdir (Bhubaneswar, 2015). Örneğin, her gün farklı ülkelerden insanlar YouTube'da saatler geçirerek milyarlarca görüntüleme oluşturmaktadır (Pouyanfar vd., 2018).

*Hız:* Büyük veride hız, verinin oluşma ve akış hızının gittikçe artmasını ifade eder. Günümüzde akıllı telefonlar, giyilebilir cihazlar gibi akıllı cihazların yüksek penetrasyon oranı ve kişiselleştirilmiş uygulamalarından ötürü benzeri görülmemiş bir hızda algılanan veriler oluşmakta ve yayınlanmaktadır (Cheng ve Yang, 2017).

*Çeşitlilik:* Veri üzerinde farklı format, içerik ve tiplere yoğunlaşmaktadır (Braun vd., 2016). Bunlara sensörler, giyilebilir cihazlar, akıllı telefonlar, monitörler örnek verilebilir. Veriler karmaşıktır ve verilerin içeriği oldukça çeşitlidir. Ayrıca, görüntü, metin, video, ses, izleme kaydı gibi veri türleri genellikle katmanlar halinde bulunmaktadır (Yaoxue vd., 2017).

*Doğrulama:* Büyük verilerin ve analitiğinin doğruluğu ve güvenilirliğini ifade etmektedir. Sağlık sektöründe, verilerin kalitesi, ilgi düzeyi ve değeri son derece önemlidir. Hayat kurtarıcı kararlar sağlık hizmeti verilerinin kalitesine ve sağlık bilgisinin doğruluğuna bağlı olduğundan, büyük veriler hatasız ve yüksek doğrulukta olmalıdır (Patil vd., 2018).

*Değer:* Oracle'ın tanımına göre, büyük veri genellikle göreceli olarak "düşük değer yoğunluğu" ile karakterize edilmektedir. Başka bir deyişle, orijinal formda alınan veriler genellikle hacmine göre daha düşük bir değere sahiptir. Ancak, bu tür verilerin büyük hacimli analizleri yapılarak yüksek bir değer elde edilebilmektedir (Gandomi ve Haider, 2015).

Büyük verinin sahip olduğu tüm bu bileşenler titizlikle uygulandığında sağlık hizmetlerinde daha doğru tahminler yapılabilecek ve bu tahminler sayesinde maliyetler azalacaktır.

### III. BÜYÜK VERİNİN SAĞLIK HİZMETLERİNDEKİ ROLÜ VE ÖNEMİ

Günümüzde sağlık hizmetleri sistemi, elektronik olarak erişilebilen sağlık kayıtlarını hızla benimsemektedir (Onyemachi ve Nonyelum, 2019). Büyük veri teknolojisi, birçok sağlık kuruluşu tarafından hasta verilerinin toplanması ve analizi için kullanılmaktadır. Böylece hastaların durumu bütüncül bir çerçeveden izlenebilmektedir. Sağlık hizmeti maliyetinin azalması, eksiksiz bilgiye erişim, daha iyi koruyucu bakım, kişiselleştirilmiş teşhis gibi beklentiler, büyük verilerin uygun bir şekilde kullanılmasıyla kolaylaşacaktır (Patil vd., 2018).

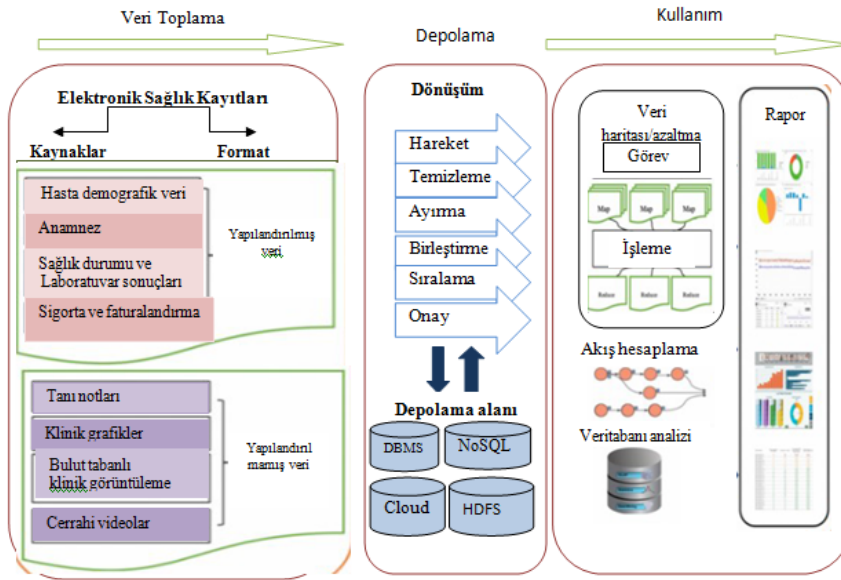
Sağlık sistemi için veri kaynakları incelendiğinde, genel olarak şu şekilde sınıflandırıldığı görülmektedir (Palanisamy ve Thirunavukarasu, 2019):

- Yapılandırılmış veriler: Tanımlanmış veri yapısına uygundur. Sağlık hizmetleri alanında bu tür verilere örnek olarak çeşitli hastalıkların tanı bilgileri, laboratuvar sonuçları, anamnez, ilaç ve fatura bilgileri verilebilir.

- Yarı yapılandırılmış veriler: Tam manası ile organize edilmemiş verilerdir. Bu tür verilere örnek olarak hastanın davranışının etkili bir şekilde izlenmesi için takılan sensör gibi cihazlardan üretilen veriler gösterilebilir.
- Yapılandırılmamış veriler: Tıbbi reçeteleri, klinik mektupları, biyomedikal literatürü, epikriz vb. içerebilen verilerdir.

Şekil 1’de büyük veri analitiği ile sağlık hizmetleri kayıtlarının yönetimi verilmiştir. Elektronik sağlık kayıtları, büyük verinin yönetiminde önemli bir kaynaktır. Veri toplanırken yapılandırılmış (Hasta demografisi, anamnez, sağlık durumu ve laboratuvar sonuçları, faturalandırma) ve yapılandırılmamış (Tanı notları, klinik grafikler, bulut tabanlı klinik görüntüleme, cerrahi videolar) kayıtlardan yararlanılmaktadır. Bu kayıtlar; bulut (Cloud), Hadoop dağıtık dosya sistemi (HDFS), veritabanı yönetim sistemi (DBMS) ve ilişkisel olmayan veritabanlarında (NoSQL) depolanmaktadır. Elde edilen veriler, bazı işlemlerden geçirilerek (Hareket, temizleme, ayırma, birleştirme, sıralama ve onay) işlenmekte ve bu işlenen veriler analiz edilerek raporlanmaktadır.

**Şekil 1: Büyük Veri Analitiği ile Sağlık Hizmetleri Kayıtları\***



**Kaynak:** Sasubilli ve Kumar (2020)

\*(DBMS: Veritabanı yönetim sistemi; NoSQL: İlişkisel olmayan veritabanları; Cloud: Bulut bilişim; HDFS: Hadoop Distributed File System.)

Tıbbi büyük verinin ana veri kaynakları incelendiğinde, üç şekilde açıklandığı görülmektedir. İlki klinik tıbbi verilerdir (Ehrenstein vd., 2017). Bunlar, elektrokardiyografi (EKG), bilgisayarlı tomografi (CT) görüntüsü gibi verileri içermektedir. İkincisi, farmakolojik araştırma ve yaşam bilimi verileridir (Xiang vd., 2020). Üçüncüsü ise kişisel sağlık verileridir. Günümüzde insanlar fiziksel durumlarını yıllık fiziksel muayenelerle anlamaktadır (Ehrenstein vd., 2017). Veriler, birden çok yerde bulunan dâhili ve harici kaynaklardan geldiği için sağlık ağlarının yeniden tanımlanması gerekir. Genel olarak kaynaklar ve veri türleri şu bilgileri içermektedir (Shafqat vd., 2018):

- Sosyal medya platformlarından tıklama akışı gibi veriler, Facebook, Twitter, LinkedIn, bloglar, akıllı uygulamalar ve sağlık planı olan web sitelerinden etkileşimli veriler gibi.
- Makineden makineye veriler: Sensör okumaları, sayaçlar ve diğer hayati sinyal cihazları.
- Büyük işlem verileri: Yarı yapılandırılmış ve yapılandırılmamış formatta bulunan sağlık hizmeti talepleri ve faturaları için kayıtlar.
- Biyometrik veriler: Parmak izleri, genetik, retina taramaları, el yazısı, kan basıncı ve diğer teşhis raporları ve verileri.

- İnsan tarafından oluşturulan veriler: Yarı yapılandırılmış verilerin yanı sıra elektronik tıbbi kayıtlar (EMR'ler), doktor notları, e-posta ve kâğıt belgeler gibi yapılandırılmamış veriler.

Bu tür verilerin ve akıllı cihazların yaygınlaşması ve elektronik iletişimin büyümesi, yeni sağlık hizmeti politikalarının tasarlanmasına fırsat vermiştir. Sağlık idaresi ve politikalarında dört önemli faktör bulunmaktadır. Bunlar; sağlık yönetimi, planlama, kaynak tahsisi, karar vermedir (Pramanik vd., 2020) ve genel olarak, büyük veri, sağlık hizmeti uygulayıcılarının ve araştırmacıların daha iyi karar verebilmek için veri hakkında öngörü kazanmalarına geniş ölçüde olanak tanımaktadır (Chauhan vd., 2020).

Mayer-Schönberger ve Cukier (2013) bu konudaki reformları iki perspektiften ele almaktadır. Bunlardan ilki, gelişme için büyük veri analizi ile birlikte, herkesi kendi hastalıklarının efendisi yapacak kolektif sağlık deneyimini oluşturması; bir diğeri ise, tıbbi verinin sürekli yenilenmesi ile büyük ticari değere sahip endüstriyel etkiler getirmesidir (Mayer-Schönberger ve Cukier, 2013, aktaran Guo ve Chen, 2019). Etkili ve verimli bir sağlık hizmetleri yönetimi sağlamak için (Örneğin, otomatik kaynak planlama, tüm kuruluşlar arasında verimli ve gerçek zamanlı tıbbi kararların doğru şekilde belirlenmesi), sağlık sektöründe büyük veri niteliği taşıyan akıllı araçlar ve teknikler kullanılmaya başlanmıştır (Pramanik vd., 2020).

Dahası bu teknolojiler genel sağlık bakım harcamalarını azaltmaktadır. McKinsey raporlarına göre, bu tür teknolojiler Amerika Birleşik Devletleri sağlık hizmetleri sektöründe yılda 300 milyar dolardan fazla ek maliyetin ortaya çıkmasını önlemiştir (Pramanik vd., 2020). McKinsey için rapor sunan Groves ve arkadaşlarına (2013) göre; büyük verinin sağlık sistemindeki etkisi beş çıkarıma dayanmaktadır (Şekil 1). Bunlar:

- Doğru yaşam (*Right living*)
- Doğru bakım (*Right care*)
- Doğru sağlayıcı (*Right provider*)
- Doğru değer (*Right value*) ve
- Doğru inovasyon (*Right innovation*)

*Doğru Yaşam:* Veriler, hastaların kendi yaşam rutinlerinde aktif rol almalarına yardımcı olmaktadır. Diyet, egzersiz ve ilaç bağımlılığını önlemek gibi (Patil vd., 2018).

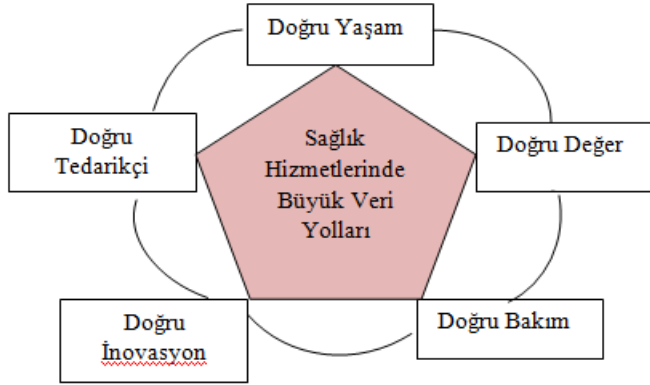
*Doğru Bakım:* Hastalar, zamanında ve uygun tedaviyi almalıdır. Protokollere büyük ölçüde güven duyulmasına rağmen, doğru bakım koordineli bir yaklaşım gerektirmektedir. Tüm bakım sağlayıcılar aynı bilgilere erişmeli ve yetersiz tedavi stratejilerini önlemek için aynı amaca yönelik çalışmalıdır.

*Doğru Sağlayıcı:* Hastaları tedavi eden herhangi bir uzman, güçlü tıbbi kayıtlara sahip olmalı ve en iyi sonuçları elde edebilmelidir. Çalışanlar, iş unvanlarından ziyade beceri ve yeteneklerine göre seçilmelidirler. Örneğin, hemşire veya doktor asistanları, doktor gerektirmeyen birçok görevi yerine getirebilmelidir (Groves vd., 2013).

*Doğru Değer:* Hasta paylarının geri ödenmesinde dolandırıcılığın, israfın ve kötüye kullanımın ortadan kaldırılması için büyük verilerden yararlanılmaktadır. Verilerle kıyaslama yapılarak geri ödemelerde hataların önüne geçilmektedir (Patil vd., 2018).

*Doğru inovasyon:* Paydaşlar, sağlık hizmeti sunumuna yönelik yeni tedaviler ve yaklaşımlar belirlemeye odaklanmalıdır. Ayrıca inovatif girişimleri kendileri de gerçekleştirmeye çalışmalıdırlar. Buna araştırma- geliştirmeye üretkenliğini arttırmak örnek verilmektedir (Groves vd., 2013).

## Şekil 2: Sağlık Hizmetleri İçerisinde Büyük Veri Yolları



**Kaynak:** Patil vd. (2020)

Büyük verinin sağlık sistemindeki bu etkilerinden yüksek düzeyde faydalanmak ve sistemli bir şekilde kullanımını desteklemek için bazı hizmetlerin oluşturulmasına ihtiyaç duyulmuştur. Bunlar (Molinari ve Nollo, 2020):

- Sağlık bilgi sistemleri, büyük veri kümelerine kıyasla daha iyi bir performans sağlayabilmesi için ölçeklenebilir bir mimari üzerine inşa edilmelidir.
- Varlık yönetimini (silme, değiştirme, birleştirme vb.) sağlamak için basit ve kullanılabilir bir web ön yüzü (Web-frontend) kullanılmalıdır.
- Büyük veri için sağlık bilgi sistemlerinde uygulanan teknik çözümler günümüzde çeşitlilik gösterdiğinden, farklı veri kaynaklarına göre farklı veri setleri içeren veri hatları (Data line) bulunmalıdır.
- Ve son olarak, veride belirsizliği gidermek için büyük veri makine öğrenme algoritmalarına dayandırılmalıdır.

Bu kıstaslar ile birlikte büyük veri kullanımı kolaylaşacaktır. Büyük veriyi, insanların sağlıklarına karar vermek için verilere erişebildikleri ve veriyi anlamaları için kullandıkları bireysel bir kaynak olarak düşünmek gerekir (Shah vd., 2019a). Aslında, büyük verinin sağlık hizmetlerinde kullanılmasının, Avrupa Komisyonu'nun Halk Sağlığı, Teletıp ve Sağlık Hizmetlerinde Büyük Veri Araştırması (The Study on Big Data in Public Health, Telemedicine and Healthcare of the European Commission) tarafından bildirildiği üzere farklı düzeylerde katkısı bulunmaktadır. Bu katkılar (Pastorino vd., 2019):

- Erken tanı etkinliğinin artması ile birlikte, erken teşhisin ve hastalığa müdahalenin önü açılmıştır. Tedavilerin kalitesinin iyileştirilmesinde, komplikasyon olasılığının azalmasında ve benzeri durumlarda veriler yol gösterici olmuştur,
- Hastalık için risk faktörlerinin tanımlanması ile hastalıkları oluşmadan önleme olasılığı artmıştır,
- Hastalara doğrudan iletilen bilgilere dayalı olarak daha bilinçli tıbbi kararlar verme yeteneği oluşturulmuş ve hasta güvenliği iyileştirilmiştir ve
- Sonuçlar tahmin edilebilir hale gelmiştir.

## IV. SAĞLIK HİZMETLERİNDE KULLANILAN PLATFORM VE ARAÇLAR

Büyük veri teknolojisi, sağlık hizmetlerinde çeşitli uygulamalara ve kullanımlara sahiptir. Büyük veriyi etkin bir şekilde kullanabilmek için bazı platform ve araçlar bulunmaktadır. Bunlar; hadoop, hbase, hive, zookeeper, pig ve pig latin, jaql, cassandra, oozie, mahout, chukwa, avro, lucenedir. Bütün araçlar verilerin doğru olarak toplanması, korunması, adlandırılması, filtrelenmesi, kümelenmesi, sınıflandırılması ve işlenip bilginin doğru anlamlandırılması için kullanılmaktadır.

#### 4.1. Hadoop

Apache Hadoop, bilgisayar kümeleri arasında paralel büyük veri kümelerini işlemek için MapReduce programlama modelini kullanan dağıtılmış açık kaynaklı bir çerçevedir. Hadoop'un iki temel bileşeni bulunmaktadır. Bunlar (Syed vd.,2019):

HDFS, ağırlıklı olarak büyük veri depolaması için kullanılır ve MapReduce, verileri analiz etmek, işlemek ve almak için kullanılan bir yazılımdır (Syed vd.,2019).

*Map Reduce çerçevesi:* Map ve Reduce olmak üzere iki fonksiyondan oluşur. Bir dizi veri girdi olarak Map fonksiyonunda toplanır ve her birinin bulunduğu başka bir veri kümesine dönüştürülür (Suneetha vd., 2020). Diğer bileşenler Hive, HBase, Pig, Zookeeper, Flume, Sqoop, Mahout, Oozie'dir (Saroja ve Sharma, 2019).

*HDFS:* Büyük miktarda bilgiyi depolayabilmekte, kademeli olarak ölçekleyebilmekte ve veri kaybı olmadan depolama arızalarından kurtulabilmektedir. Hadoop, makine kümeleri oluşturmakta ve bunlar arasındaki işi koordine etmektedir. Biri başarısız olursa, işi kümedeki kalan makinelere kaydırarak veri kaybetmeden veya işi kesintiye uğratmadan kümeyi çalıştırmaya devam etmektedir (Trivedi vd., 2019).

#### 4.2. Hbase

HBase (Apache HBase) güvenilir, yüksek performanslı, ölçeklenebilir, anahtar-değer olgusuna sahip olan bir açık kaynak veritabanıdır. Özellikle, Google Bigtable için açık kaynaklı bir uygulamadır. HBase, bir depolama sistemi olarak Hadoop tarafından sağlanan dağıtılmış dosya sistemi HDFS'yi kullanmaktadır. HBase, esas olarak bir ana bilgisayar ve bir dizi bölge sunucusundan oluşmaktadır (Liu vd., 2019).

#### 4.3. Hive

Hive, yapılandırılmış verileri Hadoop'ta işleyen bir tür veri ambarı altyapı aracıdır. Sorgulama ve analiz etmeyi kolaylaştırmakta ve Hadoop MapReduce çerçevesi, sorguların yürütülmesini hızlandırmaktadır. Hive, Hadoop çerçevesi üzerinde problemleri çözmek için depo sistemi olarak önerilmiştir ve Hive, SQL'den MapReduce'a dili çevirebilmektedir (Garg, 2015).

#### 4.4. ZooKeeper

Apache ZooKeeper, yapılandırma bilgilerinin korunması, adlandırma ve dağıtılmış senkronizasyon ve grup hizmetleri sağlamak için kurulmuş bir hizmettir. Birden fazla istemci aynı anda ZooKeeper'a erişebilmektedir (Artho vd., 2019).

#### 4.5. PIG and PIG Latin

Apache Pig, veri analiz programlarını ifade etmek için yüksek seviyeli bir dilden oluşmuş ve büyük veri setlerini analiz etmek için kullanılmakta olan bir platformdur. Pig programlarının göze çarpan özelliği, yapıları önemli ölçüde paralelleştirmeye uygundur ve bu da çok büyük veri setlerini işlemelerine olanak tanımaktadır (Moussa, 2012).

#### 4.6. Jaql

Jaql, yarı yapılandırılmış büyük veri kümelerini Hadoop'u kullanarak analiz eden bildirim temelli bir MapReduce çerçevesidir. Jaql günümüzde IBM'in InfoSphere BigInsights ve Cognos Consumer Insight ürünlerinde kullanılmaktadır. Jaql'in tasarım özellikleri şunlardır: Esnek bir veri modeli, yeniden kullanılabilirlik, farklı düzeylerde soyutlama ve ölçeklenebilirlik (Beyer vd., 2011).



#### **4.7. Cassandra**

Apache Cassandra önde gelen dağıtılmış veritabanlarından biridir. Kesintisiz çalışma, doğrusal ölçeklenebilirlik ve sorunsuz çoklu veri merkezini devreye alma gibi özellikleri ile büyük veri yönetimi söz konusu olduğunda çokça tercih edilmektedir. Dünya'nın en büyük veri kümelerinden bazılarını yönettiği bilinmektedir (Chebotko vd., 2015).

#### **4.8. Oozie**

Oozie, dağıtılmış bir ortamda Hadoop işlerini denetleyen bir zamanlayıcı çerçevesidir ve sıralı bir şekilde çalıştırılması gereken çeşitli karmaşık işleri kolaylıkla halledebilmektedir. Oozie'de en az iki iş herhangi bir belirsizlik olmadan birbirine paralel olarak çalışabilmektedir (Bandi ve Anitha, 2018).

#### **4.9. Mahout**

Mahout, hızlı ve ölçeklenebilir bir performans ortamı oluşturmak için kullanılan makine öğrenim uygulamasıdır (Xhafa vd., 2017). Mahout, işbirliğine dayalı filtreleme, kümeleme ve sınıflandırma alanlarında ölçeklenebilir ve dağıtılmış makine öğrenimi algoritmalarının ücretsiz uygulamalarını sağlayan açık kaynaklı bir projedir. Mahout, Java ile yazılmıştır. Hem dağıtılmamış hem de dağıtılmış (Map-Reduce) algoritmaları sağlamaktadır (Bagchi, 2015).

#### **4.10. Chukwa**

Chukwa dağıtılmış günlük izlem ve analiz için tasarlanmış bir hadoop alt projesidir. Chukwa, günlüklerin izlenmesi ve analizi için kullanılan Apache Lisansı kapsamındaki bir Hadoop tasarımıdır (Jose ve Binu, 2014). Chukwa, ölçeklenebilirlik ve sağlamlık özellikleri olan açık kaynaklı bir veri koleksiyonudur. Sonuçları görüntülemek, izlemek ve analiz etmek, toplanan verileri daha iyi kullanmak için esnek ve güçlü bir araç haline dönüşmektedir (Jia vd., 2010).

#### **4.11. Avro**

Avro, satır tabanlı bir depolama formatıdır ve aynı zamanda Java Serialization (Serileştirme)'a benzer veri serileştirme sistemidir. Avro'da veri dosyalarını okumak veya yazmak için kod üretimi gerekli değildir. Avro, optimize edilmiş çalışma performansı sağlayabilmektedir (Plase vd., 2016).

#### **4.12. Lucene**

Lucene, bilgi aramayı destekleyen genel bir veritabanı sistemi olarak kabul edilmektedir. Sistem yapısı, zengin ve nesneye yönelik tasarım fikirlerini benimsemekte ve çekirdek veriler, bir dosya dizini formatı tanımlanarak sınıflar halinde özetlenmektedir. Daha sonra eksiksiz bir bilgi arama motoru oluşturmak için süreçler programlanmaktadır (Ji, 2020).

Lucene, Apache Software Foundation tarafından oluşturulmuş açık kaynak kodlu bir araçtır. Tam metin tabanlı bir arama motoru olmamakla birlikte kullanıcıların sorgu ve metin indeksleme motorundan yararlanmasını sağlayan bir çerçeve sunmaktadır. Kullanımı oldukça basittir (Lakhara ve Mishra, 2017). Sonuç olarak, Lucene veri arama verimliliğini artırmak için oluşturulmuş bilgi sorgulama kütüphanesi olarak adlandırılmaktadır (Ji, 2020).

Büyük veri teknolojisinin ve bahsi geçen tüm araçların, biyomedikal ve sağlık alanlarında hızla uygulanmaya başlandığı bu "Büyük veri" çağında; insanların sağlık hizmetlerine bakış açısı tamamen değişmiştir.

Sağlık profesyonelleri, hastanın tıbbi geçmişi (tanı ve reçetelerle ilgili veriler) ve klinik verilerinden (görüntüleme ve laboratuvar muayenelerinden elde edilen veriler) sorumludur. Önceden, bu tür tıbbi kayıtları saklamak için kullanılan yaygın yöntem, ya elle yazmak ya da daktiloya geçirmektir. Ancak teknolojinin gelişmesi, kayıt sistemini de değiştirmiş; dijital olarak kayıt altına

alınan veriler daha değerli hale gelmiştir (Dash vd., 2019) ve veriler büyüdükçe kontrol edebilme ihtiyacı doğmuştur. Çünkü artık sağlık hizmeti sağlayıcılarının güçlü karar destek araçlarını mobil bilgi işlem cihazlarına taşıyabilecekleri, klinisyenlere birden fazla kaynaktan veri sentezi yapabilecekleri ve karar verme konusunda sağlık hizmetlerine yardımcı olabilecek bir teknolojiye söz edilmektedir (Tawalbeh vd., 2016).

Uygun şekilde kullanılırsa büyük verilerin halk sağlığını garanti altına almak, hastalar için uygun tedavi yollarını belirlemek ve uygulamak, klinik iyileştirmeyi desteklemek, sağlık sistemlerinin güvenliğini izlemek, yönetsel kontrolü sağlamak ve sağlık sisteminin halka karşı hesap verebilirliğini teşvik etmek gibi bir takım misyonu bulunmaktadır (Je ve Kim, 2013).

## V. SAĞLIK HİZMETLERİNDE BÜYÜK VERİ UYGULAMALARI

Tıbbi büyük veri, diğer alanların büyük verileri ile karşılaştırıldığında farklı özelliklere sahiptir. Tıbbi büyük veriye ulaşmak zordur ve kullanımının yasal komplikasyonları vardır (Elizabeth vd., 2018). Temel olarak, bu tıbbi büyük veriler; klinik bilişim, halk sağlığı informatiği, sağlık hizmetleri risk tahmini, biyoinformatik uygulamalar ve genom dizilimi, halk sağlığı, mobil sensörler ve giyilebilir cihazlar, hastalıkların teşhisi ve tedavisi, bulaşıcı hastalıkların tahmini, tıbbi görüntüleme ve ruh sağlığı durumunun değerlendirilmesi gibi alanları kapsamaktadır.

### 5.1. Klinik Bilişim

Sağlık hizmeti sağlayıcıları hastalarını değerlendirmek, tedavi etmek ve izlemek için karmaşık biyomedikal cihazlara güvenmektedir (Kobayashi vd., 2019). Klinik bilişimciler bakım kalitesini artırmak, sağlık sonuçlarını iyileştirmek ve klinisyen-hasta ilişkilerini desteklemek için bilgi sistemlerini değerlendirir, analiz eder, tasarlar, uygularlar (Desai vd., 2020). Acil servisler, yoğun bakım üniteleri ve ameliyathaneler gibi akut bakım ortamları hasta durumunu, stabiliteyi, müdahaleye yanıtı, zaman içindeki değişimi ve beklenen seyri ölçmek için çok çeşitli gelişmiş hasta izleme sistemlerine bağlı olarak çalışmaktadır (Kobayashi vd., 2019).

Genel bilişim müfredatı iç hastalıkları, patoloji, pediatri, psikiyatri ve radyoloji alanlarındaki örnekler ile birlikte 1990'larda ortaya çıkmış (Turer vd., 2020) ve bazı konulara odaklanmıştır. Bunlar (Silverman vd., 2019):

- Araştırma, kanıta dayalı tıp, halk sağlığı ve nüfus sağlığını desteklemek için elektronik sağlık kayıtlarındaki verileri kullanmaya daha fazla odaklanma,
- Yenilikçi veri analitik metodolojilerinin geliştirilmesini ve uygulanmasını teşvik eden bilimsel gelişmeleri destekleme,
- Sağlık bilgi teknolojisinin klinik süreçlere entegre edilmesi sağlanarak klinisyen üretkenliğini ve hasta memnuniyetini artırma,
- Kullanıcılar arasında artan beklentilerin karşılanması (hem klinisyenler hem de hastalar için)

Büyük veri analitiği, büyük verileri işleme ve verileri gerçek zamanlı veya gerçek zamanlıya yakın olarak ayrıştırma- görselleştirme konusundaki benzersiz yeteneği nedeniyle geleneksel klinik karar destek sistemlerinden giderek farklılaşmıştır (Wang vd., 2018).

### 5.2. Halk Sağlığı İformatiği

Geleneksel halk sağlığı hizmetleri sağlık kayıtları yönetimi, çocuk sağlığı, anne sağlığı, yaşlı sağlığı, kronik hastalığı olan hastaların sağlığı, ağır psikiyatri sağlığı dâhil olmak üzere toplum sağlığı merkezleri tarafından sağlanmaktadır. Fakat gelişen teknoloji ile birlikte kâğıt kayıtların kitlelere verimli bir şekilde hizmet etmesi zorlaşmıştır (Zhao, 2020). Günümüzde ise elektronik sağlık kayıtlarının benimsenmesi sağlık hizmetleri sektöründe giderek artmakta ve hastanelerin % 80,5'i temel bir elektronik sağlık kaydı kullanmaktadır (McFarlane vd., 2019).

Her seviyedeki halk sağlığı pratisyenleri araştırma, müdahale ve sağlığı iyileştirmek için bol miktarda veri ve bilgiye sahiptir (Shah ve Rogers, 2019b). Halk sağlığı bilişiminin genel kabul gören tanımı şöyledir: “Bilginin, bilgisayar biliminin ve teknolojinin halk sağlığı uygulamalarına, araştırmalarına ve öğrenmeye sistematik olarak uygulanmasıdır.” Halk sağlığı bilişimi, tanımlanmış popülasyonların sağlık yönetiminden sorumlu grupların ve kuruluşların bilgi teknolojisi ve analitik ihtiyaçlarını ele almaktadır (Gamache vd., 2017).

### 5.3. Sağlık Hizmetleri Risk Tahmini

Caruana ve arkadaşları (2015) yaptığı bir çalışmada otuz gün süreyle hastaneye tekrar gelişleri inceleyerek makine öğrenimi yardımı ile pnömöni riskini tahmin etmeye çalışmışlardır (Caruana vd., 2015). Manogaran ve arkadaşları (2017) ise bulut ve büyük veri teknolojileri aracılığıyla klinik testlerden ve sensör verilerinden elde edilen Elektronik Sağlık Kaydı'nı (EHR) oluşturan bir model önermişlerdir (Manogaran vd., 2017). Mun ve arkadaşları da (2018) radyasyon terapisi ile tedaviden sonra prostat kanseri hastalarının sağkalım oranlarını tahmin etmek için bir sınıflandırma geliştirmişlerdir. Hastaları yaş, klinik evresi, prostrata özgü antijen, takip süresi gibi birçok faktöre göre risk gruplarına ayırmış ve Cox regresyon analizi ile incelemişlerdir (Mun vd., 2018). Son yıllarda yapılan bu ve benzeri birçok araştırma, sağlık hizmetlerinin risk tahmininde büyük veri araçlarının kullanılabileceğini göstermiştir.

### 5.4. Biyoinformatik Uygulamalar ve Genom Dizilimi

Biyoinformatik uygulamalarında büyük veri tekniklerinin rolü biyolojik bilgileri toplamaları ve analiz etmeleri için araştırmacılara etkili veri işleme araçları; bilgisayarlarla yürütülebilen altyapı ve veri havuzları sağlamaktır. Hadoop ve MapReduce, şu anda biyomedikal alanda yaygın olarak kullanılmaktadır (Luo vd., 2016). İnsan genomunu sekanslama maliyeti (30.000 ila 35.000 geni kapsayan), yüksek verimli sıralama teknolojisinin gelişmesiyle azalmıştır. Genom teknolojisinde uzun yıllardır kullanılan ve P4 olarak anılan (tahmin edebilen, önleyici, katılımcı ve kişiselleştirilmiş sağlık) faaliyetler yürütülmektedir. Bu faaliyetlerin bazı amaçları bulunmaktadır. Bunlar (Belle vd., 2015):

- Hastalık durumlarını belirlemek için genom ölçeğinde veri kümelerini analiz etmek,
- Bir deneğin sürekli izlenmesi için kan temelli teşhis araçlarına geçmek,
- İlaç keşfine yönelik yeni yaklaşımları incelemek, geliştirmek için bir sistem yaklaşımı kullanmak ve her bir birey için verileri yakalama, doğrulama, depolama, bütünleştirme işlemi yapmak ve son olarak
- Modelleme yaparak büyük veride karşılaşılan zorlukları kolaylaştırmak için araçlar geliştirmektir.

Kanser araştırmalarında, araştırmacılar tarafından üretilen veriler kanser araştırmalarını desteklemek için analiz edilebilmektedir. Evrimsel bağları belirlemek ve moleküler yapıları tahmin etmek için çoklu protein dizileri analiz edilmekte (Pastur-Romay vd., 2016) ve biyolojik araştırmalar, yeni nesil yüksek verimli DNA dizilemesi ile üretilen genomların sıralanması ve sınıflandırılmasında oldukça gelişme göstermektedir.

Biyoinformatik, biyolojik büyük verinin heterojenliğinde, hacminde ve coğrafi dağılımında ivme kazandıran bir büyümeye yol açmıştır. Üretilen verinin yıllık büyümesi göz önüne alındığında 2020'de 44 zettabayta ulaştığı tahmin edilmektedir (Ocaña vd., 2020). Genomik tıpta, dizilenen ilk insan genomu yaklaşık 3 milyar dolara mal olmuştur. Günümüzde ise 1000 dolardan daha az bir fiyata yapılabilmektedir. Özetle, kanser araştırmalarında, araştırmacılar tarafından üretilen veriler bu araştırmaları desteklemek için analiz edilmektedir (Pastur-Romay vd., 2016).

### 5.5. Halk Sağlığı

Halk sağlığı sürveyans sistemleri, kaynakları tahsis etmek amacıyla hastalık insidansı, sağlık davranışları ve çevresel koşullardaki eğilimleri izlemektedir. Sürveyans için veri kategorileri

yetkililerin halk sađlığı hakkında bilgilendirilmesine katkı sađlamaktadır (Mooney ve Pejaver, 2018). Epidemiyolojide büyük verinin muazzam potansiyeli, Deiner tarafından Google Arama Motoru ve Twitter gibi sosyal medya araçlarını kullanarak hastalık belirtileriyle ilgili çevrimiçi soruları takip eden ve salgınların yayılmasının erken tespit edilebileceğini gösteren öncü bir çalışma ile gösterilmiştir (Benke ve Benke, 2018).

Özellikle Covid-19 pandemisinin etkisiyle, halk sađlığını izleme araçları deđişmeye başlamıştır. Örneğin, Unacast ve Google, farklı bölgelerdeki bireylerin sosyal mesafe hükümlerine nasıl uyduğunu izlemek için herkese açık siteler oluşturmuş ve bunun için akıllı telefonların konum verilerini kullanmıştır. Bu verileri işlemek için de büyük veri araçlarından faydalanılmıştır (Martinez-Martin, 2020).

Özetle, sađlık hizmeti sađlayıcıları büyük veri analizini kullanarak tedaviden sonra hastanın tekrar hastaneye yatma olasılığını hesaplayabilmekte; doktorlar, hastaların daha kısa hastanede kalış süresinden fayda sađlayıp sađlamayacağını tahmin edebilmektedir. Ayrıca, büyük veri hastalıklarının yeniden ortaya çıkmasından sorumlu faktörlerin araştırılmasına da yardımcı olmaktadır (Roy, 2016).

### 5.6. Mobil Sensörler ve Giyilebilir Cihazlar

Giyilebilir cihazlar, hastalarla fiziksel etkileşim engellerini azaltarak sađlık hizmetleri sektöründe devrim yaratma potansiyeline sahip olmuş ve bu cihazlar, tıbbi verilerin toplanma şeklini deđiştirmiştir. Ayrıca sađlık hizmeti sonuçlarını iyileştirmiş; verileri uzaktan okuyarak tüm hastane süreçlerini hızlandırmıştır (Mezghani vd., 2015).

Mobil sensörler ise "Akıllı" devrimin çarklarıdır. Hemen hemen tüm yeni akıllı telefonlar, hali hazırda yerleşik olan gömülü bir sensör seti ile birlikte gelmektedir (Han vd., 2015). Akıllı telefonlar, veri yakalamak için giderek daha fazla sayıda gömülü sensör içermekte ve çeşitli dijital biyobelirteçleri toplamak için diđer cihazlarla arayüz oluşturabilmektedir. Akıllı telefon tabanlı sensörler arasında kameralar, ivme ölçerler, yakınlık sensörleri, ışık sensörleri, barometreler, termometreler ve adım ölçerler bulunmaktadır (Purswani vd., 2019).

Ayrıca elektronik sađlık kayıtlarının sistematik olarak izlenmesi ve güncellenmesi, internet özellikli nesnelerin yardımı ile gerçekleştirilmektedir. Gartner, bađlı nesnelerin akıllı halini "Nesnelerin İnterneti (IoT)" olarak tanımlamaktadır. Gartner'a göre IoT, "Dâhili durumlarıyla veya dış ortamla iletişim kurmak, bunları algılamak veya etkileşim kurmak için gömülü teknoloji içeren fiziksel nesnelere ađıdır." (Vuppalapati vd., 2016).

Bununla birlikte ek biyometrik veriler, çok işlevli sensör platformları kullanılarak yakalanabilmektedir. Bunlar arasında nabız, kalp atış hızı, solunum hızı, kan basıncı, vücut sıcaklığı, ağırlık, çevresel deđişkenler, uyku davranışı, iletişim kalıpları, yürüyüş deđişiklikleri ve düşmeler gibi hareketlilikle ilgili olaylar yer almaktadır (Purswani vd., 2019). Bu mobil sensör ve cihazlar, verinin toplanmasında ve işlenmesinde büyük öneme sahiptir.

### 5.7. Hastalıkların Teşhisi ve Tedavisi

Elektronik tıbbi kayıtlara geçiş hasta verilerinin kullanılabilirliği, dođru teşhis ve tedavi için artan beklentiler, hasta bilgilerinin hacmi ve karmaşıklığındaki artış ile ilişkilendirilmiştir. Ne yazık ki, sađlık hizmeti sađlayıcılarına yönelik bu artan talepler, teşhis ve tedavi hataları için büyük bir risk oluşturmaktadır (Dilsizian ve Siegel, 2014). Çünkü sađlık hizmetlerinde farklı hastanelerde farklı standartlar kullanılmaktadır ve bir hastanenin standartları diğeri için kabul edilebilir düzeyde olmayabilir. Bu nedenle farklı hastanelerden toplanan tıbbi veri setleri için bir standardizasyon süreci gerekmektedir (Chen vd., 2018).

Örneğin; mevcut tanı teknolojileri (manyetik rezonans görüntüleme, elektroensefalogram) nörolojik hastalıkların tespiti, izlenmesi ve tedavisi için çok büyük miktarda veri üretir. Genel olarak,

bu tıbbi büyük verilerin analizi, anormallikleri tanımlamak ve anlamak için uzmanlar tarafından manuel olarak gerçekleştirilmektedir (Siuly ve Zhang, 2016).

Yakın gelecekte ise yapay zeka/makine öğrenimi hekimlere farklı hastalık teşhisi, tedavi seçenekleri önerileri ve tıbbi görüntüleme görüntü yorumlama ipuçları ile yardımcı olacaktır (Dilsizian ve Siegel, 2014). Bu durum sağlık hizmetlerine bir standardizasyon getirecektir. Çünkü bir kişinin bu kadar büyük hacimli verileri görsel inceleme ile biriktirmesi, yönetmesi, analiz etmesi ve asimile etmesi zor bir iştir. Genel olarak, tıbbi görüntü verileri bile tek bir çalışma için birkaç megabayttan çalışma başına yüzlerce megabayta kadar değişmektedir (Siuly ve Zhang, 2016) ve bu veriyi işlemek için büyük veri uygulamaları kullanılmaktadır.

### **5.8. Bulaşıcı Hastalıkların Tahmini**

Bulaşıcı hastalık, bir kişinin başka bir kişiden veya bir hayvandan gelen patojenler tarafından enfekte edildiğinde ortaya çıkan klinik tablodur. Sadece bireylere zarar vermekle kalmaz; makro ölçekte de zararlar verir ve bu nedenle bulaşıcı hastalıklar sosyal bir sorun olarak kabul edilmektedir (Chae vd., 2018).

Salgın durumlarında insanlar sıklıkla internette tıbbi bilgi aramaktadırlar ve bu toplu arama verileri halka açık olarak kullanılmaktadır. Böylece bu arama motoru verilerinin halk sağlığı araştırmalarında potansiyel olarak yararlı olduğu ortaya çıkmıştır (Young ve Zhang, 2018). Veriler salgın sırasında hastalık dinamiklerini ölçmek, kaynakları etkili bir şekilde tahsis etmek ve hızlı bir şekilde müdahale etmek için oldukça önemlidir.

Halk sağlığı açısından bir diğer konu, bir hastalığın etkisinin küçümsenmesi yetersiz bir tepkiye yol açabilirken; onu fazla tahmin etmek, sınırlı kaynaklarda yanlış tahsise yol açabilmesidir (Chen vd., 2016). Yaşanan Covid-19 salgını sürecinde Dünya Sağlık Örgütü'nün yanlış direktifleri ve veriyi etkin kullanamaması sebebiyle ülkeler hala Covid-19 ile mücadele etmektedir. Covid-19 süreci, halk sağlığı verilerini doğru kullanmanın ne kadar önemli olduğunu bizlere göstermiştir.

### **5.9. Tıbbi Görüntüleme**

Bir görüntü veritabanı görüntü raporlarını, patolojik, genetik ve diğer klinik bilgileri içermektedir. Genellikle görüntüler, görüntü içermeyen bilgilerden ayrılmakta ve görüntüleme bilgileri "meta veriler" olarak anılmaktadır. Örneğin, kardiyak görüntünün yorumlamasında, görsel verilere büyük veri araçları ile yaklaşılmaktadır (Morris vd., 2018). Hastalıklı dokunun büyük ölçekli radyografik görüntülemesi, bilimsel ve tıbbi keşif için inanılmaz derecede zengin bir veri kaynağı sunmaktadır. Görüntüleme verileri rutin klinik uygulama sırasında toplandığından, büyük veri kümeleri potansiyel olarak tıbbi araştırmalar için hazır halde bulunmaktadır (Parmar vd., 2018) ve tıbbi görüntüleme de bu veri kümelerinin işlenmesi ve yorumlanması söz konusudur.

### **5.10. Ruh Sağlığı Durumunun Değerlendirilmesi**

Dünya çapında sosyal medya kullanıcılarının sayısındaki hızlı artış, 2019'da küresel nüfusun neredeyse yarısını oluşturmaktadır. Sosyal medya platformları, kullanıcılardan büyük miktarlarda pasif veri oluşturarak insan davranışını incelemekte ve ruh sağlığı hastalıklarındaki risk faktörlerine odaklanmaktadır (Naslund vd., 2019). Ruh sağlığı hastalıkları, "ruh halini, düşünmeyi, davranışı ve başkalarıyla ilişkiyi etkileyen ve kişisel iyilik yerine yıkıcı tehditler oluşturan çok çeşitli akıl sağlığı koşullarını" ifade etmektedir (Liang vd., 2019).

Twitter, Facebook, Reddit veya Instagram'dan alınan sosyal medya gönderileri stres ya da iyilik halini saptayarak, öfke, kafa karışıklığı, tiksinti, korku, mutluluk, üzüntü, sürpriz gibi duyguları belirlemekte ve akut stres ifadelerini tanımlayabilmektedir. Bu yöntemler aynı zamanda depresyon gibi zihinsel bozukluklara ilişkin ön görümler de sağlayabilmektedir (Naslund vd., 2019). Bu tarz verilerin işlenmesinde büyük veri araçları kullanılmaktadır.

## VI. SONUÇ VE ÖNERİLER

Sağlık hizmetlerinde verinin işlenmesi zorluk olarak görülmemelidir. Buradaki asıl zorluk, yargıları, zamanlamayı ve taktikleri desteklemek için kullanılan veri gereksinimidir. Büyük verinin operasyonel uygulamasının önündeki en büyük engel, sağlık hizmeti verilerinin temel karakteridir. Kullanıcılar, tedarikçiler ve diğer bileşenlerin tümü kendi veri depolarına sahiptir. Gizlilik, verilerin bileşik ve ayrık doğası, etik nedenler ve her depoda meta verilerin olmayışı nedeniyle bu verilerin işlenmesi son derece sorunlu bir hale gelmiştir.

Bu nedenle büyük veri, verileri farklı şekillerde yönetmek için tasarlanan son dönemlerin en önemli mekanizmalarından biridir. Başlangıçta sağlık hizmetleri verileri farklı depolardan toplamakta ve depolama mekanizmaları verileri izlemektedir. Daha sonra dosya formatına göre depolanmakta ve kaydedilen veriler ihtiyaca göre kullanılmaktadır. Farklı kaynaklardan toplanan tüm veriler, veri madenciliği mekanizmaları kullanılarak önceden işlenecek ve bu veri özellikleri uygulamaya ve sorunlara göre diğer kanallara iletilecektir. Bu veriler tahmin modellerine tabi tutuldukları için orijinal olmalı ve içinde veri kaybı olmamalıdır.

Araştırmacılar için büyük veri büyük hacimli veriyi inceleme, verilerin içindeki modelleri ve eğilimleri bulma ve sağlık hizmetlerini geliştirmek için bir çözüm sağlama, böylece maliyetleri düşürme, sağlık erişimini demokratikleştirme ve insan hayatlarını kurtarma fırsatı sağlamaktadır.

Son yıllarda, birçok araştırmacı, analitiksel doğruluğu iyileştirmek için sağlık hizmetleri verilerinde bazı büyük veri yaklaşımları önermiştir. Sağlık hizmeti verilerinin etkili analizi ve tanımlanması, hastaların durumuna yeni bakış açıları sağlamak ve en uygun tedavi fırsatlarını önermektedir. Bu makalede, farklı büyük veri platformlarının tıbbi uygulaması tartışılmıştır.

## KAYNAKÇA

- Aceto, G., Persico, V. & Pescapè, A. (2020). Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0. *Journal of Industrial Information Integration*, 18, 1-13.
- Agrawal, R. & Prabhakaran, S. (2020). Big data in digital healthcare: lessons learnt and recommendations for general practice. *Heredity*, 124(4), 525-534.
- Artho, C., Banzai, K., Gros, Q., Rousset, G., Ma, L., Kitamura, T., Hagiya, M., Tanabe, Y. & Yamamoto, M. (2019). Model-based testing of Apache ZooKeeper: Fundamental API usage and watchers. *Software Testing Verification & Reliability*, 30(7-8) 1-29.
- Bagchi, S. (2015). Performance and quality assessment of similarity measures in collaborative filtering using Mahout. *Procedia Computer Science*, 50, 229-234.
- Bandi, R. & Anitha, G. (2018). *Machine learning based oozie workflow for hive query schedule mechanism*. International Conference on Smart Systems and Inventive Technology (ICSSIT), Tamil Nadu, India.
- Belle, A., Thiagarajan, R., Soroushmehr, S.M.R., Navidi, F., Beard, D.A. & Najarian, K. (2015). Big data analytics in healthcare. *BioMed Research International*, 1-16.
- Benke, K. & Benke, G. (2018). Artificial intelligence and big data in public health. *International journal of environmental research and public health*, 15(12), 2796.
- Beyer, K.S., Ercegovic, V., Gemulla, R., Balmin, A., Eltabakh, M.Y., Kanne, C.C., Özcan, F. & Shekita, E.J. (2011). Jaql: A scripting language for large scale semistructured data analysis. *PVLDB*, 4(12), 1272-1283.

- Bhubaneswar, O. (2015). A brief introduction on big data 5Vs characteristics and Hadoop technology. *Procedia Computer Science*, 48, 319-324.
- Boyd, D. & Crawford, K. (2012). Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679.
- Braun, P., Cuzzocrea, A., Leung, C.K., Pazdor, A.G.M. & Tran, K. (2016). Knowledge discovery from social graph data. *Procedia Computer Science*, 96, 682-691.
- Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015, August). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1721-1730).
- Chae, S., Kwon, S. & Lee, D. (2018). Predicting infectious disease using deep learning and big data. *International journal of environmental research and public health*, 15(8), 1596-1616.
- Chauhan, R., Kaur, H. & Chang, V. (2020). An optimized integrated framework of big data analytics managing security and privacy in healthcare data. *Wireless Personal Communications*, 117(1), 1-22.
- Chebotko, A., Kashlev, A., & Lu, S. (2015, June). A big data modeling methodology for Apache Cassandra. In *2015 IEEE International Congress on Big Data* (pp.238-245). IEEE. <https://doi.org/10.1109/BigDataCongress.2015.41>.
- Chen, Y., Crespi, N., Ortiz, A.M. & Shu, L. (2016). Reality mining: A prediction algorithm for disease dynamics based on mobile big data, *Information Sciences*, 379, 82-93.
- Chen, J., Li, K., Rong, H., Bilal, K., Yang, N. & Li, K. (2018). A disease diagnosis and treatment recommendation system based on big data Mining and Cloud Computing. *Information Sciences*, 435, 124-149.
- Cheng, X. & Yang, L. (2017). Mobile big data: The fuel for data-driven wireless. *IEEE Internet of Things Journal*, 4(5), 1489- 1516.
- Dash, S., Shakyawar, S.K., Sharma, M. & Kausnik, S. (2019). Big data in healthcare: Management, analysis and future prospects. *Journal of Big Data*, 6(1), 1-25.
- Demchenko, Y., De Laat, C., & Membrey, P. (2014, May). Defining architecture components of the Big Data Ecosystem. In *2014 International conference on collaboration technologies and systems (CTS)* (pp. 104-112). IEEE. <https://doi.org/10.1109/CTS.2014.6867550>.
- Desai, S., Mostaghimi, A., & Nambudiri, V. E. (2020). Clinical informatics subspecialists: characterizing a novel evolving workforce. *Journal of the American Medical Informatics Association*, 27(11), 1711-1715.
- Dilsizian, S.E. & Siegel, E.L. (2014). Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current cardiology reports*, 16(1), 441, 1-8.
- Ehrenstein, V., Nielsen, H., Pedersen, A.B., Johnsen, S.P. & Pedersen, L. (2017). Clinical epidemiology in the era of big data: New opportunities, familiar challenges. *Clinical Epidemiology*, 9, 245-250.

- Elezabeth, L., Mishra, V. P., & Dsouza, J. (2018, August). The Role of Big Data Mining in Healthcare Applications. In *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 256-260). IEEE.
- Gamache, R., Kharrazi, H. & Weiner, J. P. (2018). Public and population health informatics: The bridging of big data to benefit communities. *Yearbook of medical informatics*, 27(1), 199–206.
- Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Garg, V. (2015, December). Optimization of multiple queries for big data with apache Hadoop/Hive. In *2015 International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 938-941). IEEE. <http://dx.doi.org/10.1109/CICN.2015.184>.
- Groves, P., Kayyali, B., Knott, D. & Van Kuiken, S. (2013). *The 'big data' revolution in health care: Accelerating value and innovation*. McKinsey & Company, New York.
- Guo, C. & Chen, J. (2019). Big data analytics in healthcare: Data-driven methods for typical treatment pattern mining. *Journal of Systems Science and Systems Engineering*, 28(6), 694–714.
- Han, Q., Liang, S. & Zhang, H. (2015). Mobile cloud sensing, big data, and 5G networks make an intelligent and smart world, *IEEE Network*, 29(2), 40-45.
- Je, K. & Kim, G.H. (2013). Potentiality of big data in the medical sector: Focus on how to reshape the healthcare system. *Healthcare Informatics Reseach*, 19(2), 79-85.
- Ji, W. (2020, June). Research and Application of Information Data Retrieval System in Station Based on Lucene Technology. In *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* (pp. 687-690). IEEE.
- Jia, B., Wlodarczyk, T. W., & Rong, C. (2010, November). Performance considerations of data acquisition in hadoop system. In *2010 IEEE Second International Conference on Cloud Computing Technology and Science* (pp. 545-549). IEEE.
- Jose, A. S., & Binu, A. (2014, August). Automatic detection and rectification of dns reflection amplification attacks with hadoop mapreduce and chukwa. In *2014 Fourth International Conference on Advances in Computing and Communications* (pp. 195-198). IEEE.
- Kobayashi, L., Oyalowo, A., Agrawal, U., Chen, S. L., Asaad, W., Hu, X., ... & Merck, D. L. (2018). Development and deployment of an open, modular, near-real-time patient monitor datastream conduit toolkit to enable healthcare multimodal data fusion in a live emergency department setting for experimental bedside clinical informatics research. *IEEE Sensors Letters*, 3(1), 1-4.
- Lakhara, S., & Mishra, N. (2017). Desktop full-text searching based on Lucene: A review. In *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)* (pp. 2434-2438). IEEE.
- Liang, Y., Zheng, X. & Zeng, D.D. (2019). A survey on big data-driven digital phenotyping of mental health. *Information Fusion*, 52, 290–307.
- Liu, W., Zhang, T., Shen, Y. & Wang, P. (2019). Fast correlation coefficient estimation algorithm for HBase-based massive time series data. *Frontiers of Computer Science*, 13(4), 864–878.
- Luo J., Wu, M., Gopukumar, D. & Zhao, Y. (2016). Big data application in biomedical research and health care: A literature review. *Biomedical Informatics Insights*, 8, 1–10



- Manogaran, G., Lopez, D., Thot, C., Abbas, K.M., Pyne, S. & Sundarasekar, R. (2017). *Big Data Analytics in Healthcare Internet of Things, in Innovative Healthcare Systems for the 21st Century*, Eds. Qudrat-Ullah, H. and Tsasis, P., New York: Springer International Publishing.
- Martinez-Martin, N. (2020). Big data, corporate surveillance and public health, *The American Journal of Bioethics*, 20(10), 79-81.
- Mezghani, E., Exposito, E., Drira, K., Da Silveira, M., & Pruski, C. (2015). A semantic big data platform for integrating heterogeneous wearable data in healthcare. *Journal of medical systems*, 39(12), 1-8.
- Mooney, S.J. & Pejaver, V. (2018). Big data in public health: Terminology, machine learning, and privacy. *Annual Review of Public Health*, 39, 95-112.
- McFarlane, T.D., Dixon, B.E., Grannis, S.J. & Gibson, P.J. (2019). Public health informatics in local and state health agencies: An update from the public health workforce interests and needs survey. *Journal of Public Health Management and Practice*, 25(2), s67-s77.
- Molinari, A., & Nollo, G. (2020, June). The quality concerns in health care Big Data. In *2020 IEEE 20th Mediterranean Electrotechnical Conference* (pp. 302-305). IEEE. <https://ieeexplore.ieee.org/document/9140534>
- Morris, M.A., Babak, S., Brian, B., Jackson, G. & Elliot L., S. (2018). Reinventing radiology: Big Data and the future of medical imaging. *Journal of Thoracic Imaging*, 33(1), 4-16.
- Moussa, R. (2012, June). Tpc-h benchmarking of pig latin on a hadoop cluster. In *2012 International Conference on Communications and Information Technology* (pp. 85-90). IEEE.
- Mun, S., Park, J., Dritschilo, A., Collins, S., Suy, S., Choi, I. & Rho, M. (2018). The prostate clinical outlook (PCO) classifier application for predicting biochemical recurrences in patients treated by stereotactic body radiation therapy (SBRT). *Applied Sciences*, 8(9), 1620. MDPI AG. <http://dx.doi.org/10.3390/app8091620>.
- Naoui, M.A., Lejdel, B., Ayad, M., Belkeiri, R. & Khaouazm, A.S. (2020). Integrating deep learning, social networks, and big data for healthcare system. *Bio-Algorithms and Med-Systems*, 16(1), 1-14.
- Naslund, J. A., Gonsalves, P. P., Gruebner, O., Pendse, S. R., Smith, S. L., Sharma, A., & Raviola, G. (2019). Digital innovations for global mental health: opportunities for data science, task sharing, and early intervention. *Current treatment options in psychiatry*, 6(4), 337-351.
- Ocaña, K., Galheigo, M., Osthoff, C., Gadelha Jr., L.M.R., Porto, F., Gomes, A.T.A., Oliveira, D. & Vasconcelos, A.T. (2020). BioinfoPortal: A scientific gateway for integrating bioinformatics applications on the Brazilian national high-performance computing network. *Future Generation Computer Systems*, 107, 192–214.
- Onyemachi, N. C., & Nonyelum, O. F. (2019). Big Data Analytics in Healthcare: A Review. In *2019 15th International Conference on Electronics, Computer and Computation* (pp. 1-5). IEEE. <https://doi.org/10.1109/ICECCO48375.2019.9043183>.
- Palanisamy, V., & Thirunavukarasu, R. (2019). Implications of big data analytics in developing healthcare frameworks—A review. *Journal of King Saud University-Computer and Information Sciences*, 31(4), 415-425.
- Parmar, C., Barry, J. D., Hosny, A., Quackenbush, J., & Aerts, H. J. (2018). Data analysis strategies in medical imaging. *Clinical cancer research*, 24(15), 3492-3499.

- Pastorino, R., De Vito, C., Migliara, G., Glocker, K., Binenbaum, I., Ricciardi, W., & Boccia, S. (2019). Benefits and challenges of Big Data in healthcare: an overview of the European initiatives. *European journal of public health, 29*(3), 23-27.
- Pastur-Romay, L. A., Cedrón, F., Pazos, A., & Porto-Pazos, A. B. (2016). Deep artificial neural networks and neuromorphic chips for big data analysis: pharmaceutical and bioinformatics applications. *International journal of molecular sciences, 17*(8), 1313.
- Patil, S., Patil, K. R., Patil, C. R., & Patil, S. S. (2020). Performance overview of an artificial intelligence in biomedics: a systematic approach. *International Journal of Information Technology, 12*(3), 963-973.
- Plase, D., Niedrite, L., & Taranovs, R. (2016). Accelerating data queries on Hadoop framework by using compact data formats. In *2016 IEEE 4th Workshop on Advances in Information, Electronic and Electrical Engineering* (pp. 1-7). IEEE. <https://doi.org/10.1109/AIEEE.2016.7821807>.
- Pramanik, M. I., Lau, R. Y., Azad, M. A. K., Hossain, M. S., Chowdhury, M. K. H., & Karmaker, B. K. (2020). Healthcare informatics and analytics in big data. *Expert Systems with Applications, 152*, 113388.
- Pouyanfar, S., Yang, Y., Chen, S. C., Shyu, M. L., & Iyengar, S. S. (2018). Multimedia big data analytics: A survey. *ACM computing surveys, 51*(1), 1-34.
- Purswani, J. M., Dicker, A. P., Champ, C. E., Cantor, M., & Ohri, N. (2019). Big data from small devices: the future of smartphones in oncology. In *Seminars in radiation oncology* (pp. 338-347). WB Saunders. <https://doi.org/10.1016/j.semradonc.2019.05.008>.
- Devi, R. R., & Chamundeeswari, V. V. (2020). Triple DES: privacy preserving in big data healthcare. *International Journal of Parallel Programming, 48*(3), 515-533.
- Roy, A.K. (2016). Impact of big data analytics on healthcare and society. *Journal of Biometrics & Biostatistics, 7*(3), 300-307.
- Saroha, M., & Sharma, A. (2019). Big Data and Hadoop Ecosystem: A Review. In *2019 International Conference on Smart Systems and Inventive Technology* (pp. 1-5). IEEE. <https://doi.org/10.1109/ICSSIT46314.2019.8987848>.
- Sasubilli, G., & Kumar, A. (2020). Machine Learning and Big Data Implementation on Health Care data. In *2020 4th International Conference on Intelligent Computing and Control Systems* (pp. 859-864). IEEE. <https://doi.org/10.1109/ICICCS48265.2020.9120906>.
- Shah, G., Shah, A., & Shah, M. (2019a). Panacea of challenges in real-world application of big data analytics in healthcare sector. *Journal of Data, Information and Management, 1*(3), 107-116.
- Shah, G. H., & Rogers, V. N. (2019b). Publishing on topics in public health informatics for a quarter century. *Journal of Public Health Management and Practice, 25*(1), 27-29.
- Shaikh, T. A., & Ali, R. (2019). Big data for better Indian healthcare. *International Journal of Information Technology, 11*(4), 735-741.
- Shafqat, S., Kishwer, S., Rasool, R. U., Qadir, J., Amjad, T., & Ahmad, H. F. (2020). Big data analytics enhanced healthcare systems: a review. *The Journal of Supercomputing, 76*(3), 1754-1799.

- Shi, M., Jiang, R., Hu, X. & Shang, J. (2020). A privacy protection method for health care big data management based on risk access control. *Health Care Management Science*, 23(3), 427–442.
- Shilo, S., Rossman, H. & Segal, E. (2020). Axes of a revolution: Challenges and promises of big data in healthcare. *Nature medicine*, 26(1), 29–38.
- Silverman, H.D., Steen, E.B., Carpenito, J.N., Ondrula, C.J., Williamson, J.J. & Fridsma, D.B. (2019). Domains, tasks, and knowledge for clinical informatics subspecialty practice: results of a practice analysis. *Journal of the American Medical Informatics Association*, 26(7), 586–593.
- Siuly, S. & Zhang, Y. (2016). Medical big data: Neurological diseases diagnosis through medical data analysis. *Data Science and Engineering*, 1(2), 54–64.
- Strang, K.D. & Sun, Z. (2020). Hidden big data analytics issues in the healthcare industry. *Health Informatics Journal*, 26(2), 981-998.
- Suneetha, V., Suresh, S., & Jhananie, V. (2020). A novel framework using apache spark for privacy preservation of healthcare big data. In *2020 2nd International Conference on Innovative Mechanisms for Industry Applications* (pp. 743-749). IEEE. <https://doi.org/10.1109/ICIMIA48430.2020.9074867>.
- Syed, L., Jabeen, S., Manimala, S. & Alsaeedi, A. (2019). Smart healthcare framework for ambient assisted living using IoMT and big data analytics techniques. *Future Generation Computer Systems*, 101, 136-151.
- Tawalbeh, T.A., Mehmood, R., Benkhelifa, E. & Song, H. (2016). Mobile cloud computing model and big data analysis for healthcare applications. *IEEE Access*, 4, 6171-6180.
- Trivedi, K., Kumari, S. & Jain, S. (2019). Big data processing with hadoop. *International Journal of Business & Engineering Research*, 10, 1-5.
- Turer, R.W., Arribas, M., Balgord, S.M., Brooks, S., Hopson, L.R., Bassin, B.S. & Medlin, R. (2020). Clinical informatics training during emergency medicine residency: The university of Michigan experience. *AEM Education and Training*, 1-8. <https://doi.org/10.1002/aet2.10518>
- Vasuki, N. & Rajiv-Kannan, A. (2020). Big healthcare data for trivial client having Novel Smart Attire (NSA). *Soft Computing*, 24, 18367-18378.
- Viceconti, M., Hunter, P. & Hose, R. (2015). Big data, big knowledge: Big data for personalized healthcare. *IEEE Journal of Biomedical and Health Informatics*, 19(4), 1209-1215.
- Vuppalapati, C., Ilapakurti, A., & Kedari, S. (2016). The role of big data in creating sense ehr, an integrated approach to create next generation mobile sensor and wearable data driven electronic health record (ehr). In *2016 IEEE second international conference on big data computing service and applications (BigDataService)* (pp. 293-296). IEEE. <https://doi.org/10.1109/BigDataService.2016.18>.
- Young, S.D. & Zhang, Q. (2018) Using search engine big data for predicting new HIV diagnoses. *PLoS ONE*, 13(7), 1-8.
- Yaoxue, Z., Ju, R., Jiagang, L., Chugui, X., Hui, G. & Yaping, L. (2017). A survey on emerging computing paradigms for big data. *Chinese Journal of Electronics*, 26(1), 1-12.
- Xhafa, F., Bogza, A. & Caballe, S. (2017). *Performance evaluation of mahout clustering algorithms using a twitter streaming dataset*. IEEE 31st International Conference on Advanced Information Networking and Applications, Taipei, Taiwan.

- Xiang, Z., Jinghua, C. & Tao, W. (2020). *Review of machine learning algorithms for health-care management medical big data system*. International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India. <https://doi.org/10.1109/ICICT48043.2020.9112458>.
- Zhao, Y., Liu, L., Qi, Y., Lou, F., Zhang, J. & Ma, W. (2020). Evaluation and design of public health information managementsystem for primary health care units based on medical and health information. *Journal of Infection and Public Health*, 13(4), 491-496.
- Wang, Y., Kung, L.A., Chung Wang, W.Y. & Cegielski, C.G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, 55(1), 64-79.