



Büyük Veri Teknolojilerinin Sağlık Alanındaki Kullanımı ve Nöroşirürjideki Yansımaları

The Use of Big Data Technologies in the Field of Health and Its Reflections in Neurosurgery

Berk Burak BERKER¹, Barış ALBUZ², Fatih YAKAR¹, Emrah EGEMEN¹, Batuhan BAKIRARAR³

¹Pamukkale Üniversitesi Tıp Fakültesi, Beyin ve Sinir Cerrahisi Anabilim Dalı, Denizli, Türkiye

²Fatih Devlet Hastanesi, Trabzon, Türkiye

³Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara, Türkiye

Yazışma adresi: Fatih YAKAR ✉ yakar@neurosurgery@gmail.com

ÖZ

Klasik veri tabanlarının depolayamayacağı ve yönetemeyeceği büyüklükte verilerin elde edilmesi büyük veri kavramını ortaya çıkarmıştır. Yapılandırılmamış bu verinin işlenmesi ile, veri içerisinde klasik istatistiksel yöntemlerle saptanamayacak modeller ortaya konulmaktadır. Sağlık alanında anlık olarak artan verilerin bir yapay zekâ yöntemi olan makine öğrenmesi ile işlenmesinin tanı koymak, en etkin tedavi yöntemlerini saptamak, prognoza etki eden faktörleri bulmak ve sağkalım hesaplaması yapmak gibi faydaları mevcuttur. Güncel nöroşirürji literatüründe kapsadığı yer günden güne artan büyük veri konusunda her nöroşirürji uzmanının bilgi sahibi olması önemli ve gereklidir. Bu derlemede büyük verinin tanımı yapıldıktan sonra sağlık alanındaki ve nöroşirürjideki kullanım alanları ele alınmış ve getirileri tartışılmıştır.

ANAHTAR SÖZCÜKLER: Büyük veri, Makine öğrenmesi, Yapay zekâ, Nöroşirürji

ABSTRACT

Obtaining data that cannot be stored and managed by classical databases has revealed the concept of big data. Models in the data that classical statistical methods cannot determine are revealed by processing this unstructured data. Processing the instantly increasing data in health with machine learning and artificial intelligence methods, has benefits such as diagnosing, determining the most effective treatment methods, finding the factors affecting the prognosis, and calculating survival. It is crucial and necessary for every neurosurgery specialist to know big data; the area covered by the current neurosurgery literature is increasing day by day. In this review, after the definition of big data, its usage areas in health and neurosurgery are discussed and its benefits are discussed.

KEYWORDS: Big data, Machine learning, Artificial intelligence, Neurosurgery

■ BÜYÜK VERİ NEDİR?

Büyük veri tanımı verinin kullanıldığı alana göre farklı şekillerde yapılabilmektedir. En tipik tanımı verinin terabit veya petabitin yüzlerce katı olmasıdır. Bir başka tanımı

ise organizasyonların farklı dijital veri setlerini bütünleştirerek istatistik ve veri madenciliği teknikleriyle gizli kalmış bilgileri ve sürpriz korelasyonları ortaya çıkarmaktır (12). Çoğunluğu yapılandırılmamış olan ve üstel bir şekilde büyüyen, geleneksel veri tabanlarının (örnek olarak hastanelerde kullanılan hasta kayıt

sistemleri) işleme kapasitesini aşan veri olarak da tanımlanır (22). Sağlık alanında büyük veri analizlerinin kullanılması ile 300 milyon dolarlık tasarruf sağlanacağı tahmin edilmektedir (2).

■ 5V KAVRAMI

Verinin büyüklüğü sadece hacimsel (Volume) olarak ifade edilmemektedir. Verinin ayrıca hızlı hareket ediyor olması (Velocity) veya yapılandırılmamış olması (Variety) da büyük olarak tarif edilmesine neden olur. Ek olarak verinin doğrulanabilir (Verification) ve anlamlı değere sahip olması da (Value) büyük verinin özelliklerindedir (18).

Dünya üzerindeki veri akışının %90'ının son iki yılda gerçekleşmiş olması, dünyada bir dakikada 204 milyon e – posta gönderilmesi, Facebook'a bir dakikada 200 milyon fotoğraf yüklenmesi ve YouTube'a dakikada 100 saatlik veri yüklenmesi verinin hacmini anlamamıza yardımcı olacak bilgilerdir. Dünyada bir günde kaydedilen veriler DVD'lere yüklenip üst üste konulursa aya iki kez ulaşabilmektedir (2).

Veriler ne kadar hızlı işlenirse o kadar hızlı bilgi üretilir, anlamlandırılır ve faydalı hâle dönüşür. Örnek olarak marketler akıllı telefonların konum bildirme özelliğini kullanarak otoparklarını kullanan müşterilerinin alışveriş bilgilerini elde ettikten sonra kişiye özel bildirimlerle satışlarını artırabilirler (12). Çeşitlilik verinin heterojen yapısını ifade eder. Elde edilen veriler tabletlerden, akıllı telefonlardan veya sosyal ağlardan farklı formatlarda, dillerde ve kodlarda oluşturulmaktadır. Bu yapılandırılmamış verilerin işlenmesi büyük veri analizleri kullanılmadan olanaksızdır (14). Verilerin değeri ise kurumların hızlı ve doğru karar vermesine katkı sağlaması olarak ifade edilebilir. Sağlık Bakanlığı örneği verilirse, bu kurumun karar alırken anlık olarak hastalık çeşitliğine, ilaç tüketim miktarı ve hasta başına düşen doktor sayısına ulaşabilmesi doğru kararlar vermesine katkıda bulunacaktır (14).

■ BÜYÜK VERİYE GEREK DUYULMA NEDENLERİ

Kurumlarda artan talepler sonucu daha yüksek hacimli ve çeşitli veri saklanması gerçeğini ortaya çıkarmıştır. 2 milyarlık müşteri havuzu olan bir kurumun ödemesini yapmamış bir müşteriyi bulması klasik veri tabanları ile 70 dakika sürerken büyük veri analizi ile bu süre 13 saniyeye inebilmektedir. Bu kadar basit bir sorguda bile zaman ve maliyet tasarrufu sağlaması büyük veri analizlerini ön plana çıkarmıştır (2).

Tıbbi cihazlardan (EKG, USG, mamografi vb.) elde edilen verilerin kaydedilmesi büyük bir trafiğe neden olmaktadır. Bu durum nesnelerin interneti olarak adlandırılmaktadır ve bu verilerin yıldan yıla katlanarak artacağı bir gerçektir (9). Tüm bu verilerin klasik veri tabanları ile depolanıp işlenmesi mümkün değildir. Depolama maliyetlerini düşüren bulut tabanlı sistemlerinin ortaya çıkması ve büyük hacimli verileri birden çok sanal sunucuya dağıtmaya imkân tanıyan NoSQL ve benzeri veri tabanı tasarımlarının ve Hadoop teknolojisinin ortaya çıkması ile büyük veri setleri oluşmaya başlamıştır (2).

■ BÜYÜK VERİDE KULLANILAN TEKNOLOJİLER

Bu alanda kullanılan teknolojiler ile metin, ses ve video gibi klasik yöntemlerle analizi mümkün olmayan verilerin analizine imkân sağlanır (Şekil 1).

Hadoop sistemi çoklu paralel sunucular üzerinden büyük veri işlemesi yapan Java ortamında geliştirilmiş açık kaynak kodlu yazılımdır. Hadoop, Hadoop Distributed File System (HDFS) ve MapReduce bileşenlerinden oluşur. HDFS birden fazla sunucunun disklerini birleştirerek tek bir sanal disk gibi kullanır ve ek bir sunucu üzerinde depolanması imkânsız büyüklükteki verileri saklayabilir. HDFS, NameNode ve DataNode süreçlerinden oluşmaktadır. Veri bloklarının yaratılmasından, dağıtılmasından, gerektiği zaman silinmesinden ve veri erişiminden NameNode sorumlu iken veri bloklarını saklamak DataNode'un görevidir (13).

MapReduce ise Hadoop'un üzerinde çalıştığı mimari altyapıdır. Map fonksiyonu verilerin filtrelenmesini sağlarken, Reduce ise verilerden çıktı elde edilmesini sağlar (13). MapReduce için örnek olarak son yüzyıldaki en yüksek hava sıcaklığı bulunmaya çalışıldığını düşünelim. Sıcakları kaydeden çok fazla sensör bulunduğundan bu verileri tek bir veri tabanında saklamak mümkün değildir. Birden fazla sunucuda depolanan bu verilerin işlenmesi gerektiğinde map ve reduce fonksiyonları kullanılmalıdır (20).

Betik diller ise büyük veri ile uyumlu çalışan programlama dilleridir (Python, Pig, Hive). Makine öğrenmesi bir veri kümesine en iyi uyan modeli hızla bulmak için kullanılan yazılımdır. Analitik sonuçların görsel veya grafik formatta gösterimi görsel analitik olarak ifade edilir. Daha hızlı sonuç almak için büyük verinin bilgisayar belleğinde işlenmesi ise bellek içi analitik olarak tanımlanır.

■ BÜYÜK VERİNİN GETİRDİĞİ YENİLİKLER

En önemli katkısı maliyetlerin düşürülmesi ve karmaşık / yapılandırılmamış verilerin işlenerek anlamlı modeller oluşturabilmesidir. Geleneksel veri tabanları verileri basit sütun satır formatında kaydederken büyük veri farklı formatlarda kayıt yapabilmektedir. Büyük verileri tek bir sunucu üzerinden işlemek imkânsız olduğundan Hadoop kullanılarak veriler birden çok sunucuya dağıtılıp eş zamanlı işlenir ve sonuçlar elde edilir. Verilerin birden çok sunucuya dağıtılması veri işleme süresini yüz kat veya daha fazla düşürmektedir. Binlerce sunucusu olan ucuz fiyatlı ticari sunucuların yükselişi ile de büyük veri daha da çok kullanılmaya başlanmıştır (11).

■ SAĞLIK ALANINDA BÜYÜK VERİ KULLANIMI

Sağlıkta büyük veri yöntem bilimi; kavramsal çalışma, öneri geliştirme, metodoloji ve test / değerlendirme olmak üzere dört basamaktan oluşmaktadır (8).

Büyük veri kullanımı ile elde edilecek faydalar:

1. Erken evrede tanı koyma ve tedavi etme fırsatı sunması
2. Kişilerin sağlık durumlarının kontrolünün sağlanabilmesi

■ YAPAY ZEKÂ NEDİR?

İnsana özgü olan anlama, öğrenme, algılama, hatırlama, kavramlar arasında etkili bağlantı kurabilme, sorun ve problemleri başarıyla çözme, diğer birimler ile iletişim kurabilme gibi üst düzey bilişsel (kognitif) fonksiyonları insan seviyesinde sergileyebilen işletim sistemi yapay zekâ olarak isimlendirilir (15). Amaç insan zekâsını bilgisayarlarla taklit ederek makinelere öğrenme becerisi kazandırabilmektedir (5).

Yapay zekânın insan zekâsına üstünlükleri: asla unutmaması, statik yapısından dolayı tutarlı olması, kopyalanabilir ve paylaşılabilir olması, doğal zekâdan çok daha az maliyetli olması, öğrendiği bilgilerin ve tecrübelerin kayıt altına alınabiliyor olmasıdır (3,6).

İnsan zekâsının üstün yanları ise: edindiği bilgi ve becerileri karşılaşılan bir durum karşısında hızlıca ve belirli bir alana hapsolmeden sınırsızca kullanabilmesi, yaratıcı olması, hissiyatlı ve güçlü önsezilere sahip olması, eleştirel düşünebilmesi, önsezi, bilgelik, basiret, etik ve inanış gibi özellikleridir (3,15).

Farklı kullanımlar için yapay zekâ yöntemlerinden bazıları: Makine öğrenmesi, Uzman Sistemler, Yapay Sinir Ağları, Görüntü İşleme, Doğal Dil İşleme ve Genetik Algoritmalarıdır. Makine öğrenmesi ile yapay zekâ, her bir karar için ayrı ayrı yazılımlara ihtiyaç duymak yerine önceden belirlenmiş algoritmalar ve kümeler yardımıyla bir eğitim kümesi oluşturur ve sorunlara bu küme üzerinden kendisi çözümler üretir. Tanımlanan eğitim kümeleri arttıkça yapay zekâ daha iyi çıkarımlar yaparak kendi mantık alanını genişletebilir.

Günlük yaşamımızda en basit olarak gelen e-postaların filtrelenmesi, arama motorları, navigasyon sistemleri ve takvim ayarı için kullanılan sanal organizasyonlar yapay zekâya örnektir.

■ YAPAY ZEKÂ KULLANMANIN SAĞLADIĞI AVANTAJLAR

Yapay zekâ sistemlerinin kullanılmasıyla hızlı teşhis, doğru tedavi ve etkin kararlar alınarak hata payı eskiye oranla ciddi şekilde azaltılabilir. Yapay zekânın sağlık ekonomisine etkisinin incelendiği çalışmalardan birisi olan Centerstone Araştırma Enstitüsü çalışmasına göre yapay zekâ kullanarak tanı koymak, geleneksel tanı koymaya göre daha ucuza mal olmaktadır. Rastgele örneklem metodu kullanılarak 500 hastanın tanı ve tedavilerine ait veriler yapay zekânın dizilime göre karar verme modelleriyle kıyaslanarak analiz yapılmıştır. Sonuçta yapay zekâ bir hastanın tüm işlemlerini 189 dolara yapabilirken, klasik yöntemler ile bu işlemler 497 dolara mal olmuştur (19).

■ GÜNCEL NÖROŞİRÜRJİ LİTERATÜRÜNDE BÜYÜK VERİ VE MAKİNE ÖĞRENMESİ

Mevcut Surveillance, Epidemiology, and End Results (SEER) veri tabanları üzerinden güncel nöroşirürjikal hastalıkların prognostik faktörlerini belirlemek mümkündür. 65 yaş altındaki gliyoblastom hastalarının kemoterapiden fayda görüp görmeyeceğini hesaplayan bir çalışmada 25 binin üzerinde hasta değerlendirilmiştir (24). Hasta gruplarında prognozu tahmin

etmek için de bu veri tabanları kullanılmaktadır (17). Erken mortalitenin sebepleri de binlerce hastanın bulunduğu bu sistemlerle saptanabilmektedir (25). Ayrıca görüntüleme yöntemlerinden elde edilen veriler bu sistemlere öğretildiğinde tanı koyma becerilerinden faydalanmak mümkün olacaktır. Atıcı ve ark.nın yaptığı çalışmada T2 ağırlıklı MR görüntülerinin öğretildiği makinelerin yüksek dereceli gliyal tümörleri saptama oranı %81 – 98 olarak belirlenmiştir (4). Halihazırda makine öğrenmesi sistemleri hastane kayıt sistemlerine entegre edildiğinde, yapılan görüntülemelerde (BT, MRG vb.) kanama, tümör ya da inme tanısı konularak ilgili hekimlerin anında bilgilendirilmesi de sağlanabilmektedir.

Kraniyosinostoz varlığını saptamak için hastaların yüzlerinin fotoğraflarının öğretildiği sistemler ile kraniyosinostoz tanısı konulabilmiştir (21). Bir Ommaya rezervuarının kalıcı şanta dönüşme olasılığı da benzer sistemler yardımıyla hesaplanabilmiştir (1). Sellar bölge tümörlerinin transsfenoidal rezeksiyonu sonrası görsel sonuçları tahmin etmek için de makine öğrenmesi kullanılabilir (23). Güncel bir sistematik derlemede ise (10) makine öğrenmesinin cerrahi yararlanım oranını hesaplamada, derin beyin uyarımı elektrotlarının yerlerini doğru saptamada, endovasküler tedavide vazospazm oranını saptamada, epilepsi cerrahisine uygun adayları belirlemede ve gliyom hastalarında sağkalım hesaplama gibi birçok alanda kullanıldığı saptanmıştır. Sözgelimi bu veri tabanlarından elde edilen hesaplama sistemlerine poliklinikte değerlendirdiğimiz bir hastanın verilerini girdiğimizde hastanın ne kadar süre hayatta kalacağını, mortalite için risk faktörlerini ve cerrahi yapılırsa ne kadar fayda göreceğini saptayabileceğiz. Yakın gelecekte bu sistemler günlük rutimize eklendiğinde hastaların tedavisinde standart yaklaşımların oluşturulabileceği açıktır.

■ KAYNAKLAR

1. Alegre MS, Martín RM, Flores AP, Benito FC, Alegre JS, Negre GS, Bauer R, Feijoo PG, Gutiérrez JS: Development of machine learning-based predictor algorithm for conversion of an Ommaya reservoir to a permanent cerebrospinal fluid shunt in preterm posthemorrhagic hydrocephalus. World Neurosurg, 2022 (In press)
2. Altunışık R: Büyük veri: Fırsatlar kaynağı mı yoksa yeni sorunlar yumağı mı? Yıldız Social Science Review 1:45-76, 2015
3. Armağan Y: Mobilya Tasarımında Yapay Zekâ: Tasarım ve Ar-Ge Merkezleri Üzerinden Bir Değerlendirme (Yüksek Lisans Tezi), Ankara: Hacettepe Üniversitesi Güzel Sanatlar Enstitüsü, 2019
4. Atıcı MA, Sağıroğlu S, Çeltikçi P, Ucar M, Börcek AO, Emmez H, Çeltikçi E: A novel deep learning algorithm for the automatic detection of high-grade gliomas on t2-weighted magnetic resonance images: A preliminary machine learning study. Turk Neurosurg 30(2):199-205, 2020
5. Aydın A: Acil Servis Triajında Yapay Zekâ Yöntemlerinin Güvenilirliği (Uzmanlık Tezi), Antalya: Akdeniz Üniversitesi Tıp Fakültesi Acil Tıp Anabilim Dalı, 2011
6. Aydın ŞE: İşletme ve Teknoloji Yönetimi Yüksek Lisans Programı (Yüksek Lisans Dönem Projesi), Adana: Çukurova Üniversitesi Sosyal Bilimler Enstitüsü, 2017

7. Bakırarar B, Yüksel C, Yavuz Y: Prediction of new prescription requirements for diabetes patients using big data Technologies. *J Health Res* 36(2):334-344, 2022
8. Beyan T: Sağlıkta Dönüşüm ve Büyük Veri Uygulamaları. Available at: <https://saglikvebilisim.wordpress.com/2016/02/18/saglikta-donusum-ve-buyuk-veri-uygulamaları>. Accessed June 10, 2016.
9. Court D: Getting Big Impact from Big Data. Available at: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/getting-big-impact-from-big-data>. Accessed January 1, 2015
10. Çeltikci E: A systematic review on machine learning in neurosurgery: The future of decision-making in patient care. *Türk Neurosurg* 28(2):167-173, 2018
11. Davenport T: Big Data @ Work. Available at: <https://store.hbr.org/product/big-data-at-work-dispelling-the-myths-uncovering-the-opportunities/16574>. Accessed February 25, 2014.
12. Demirtaş B, Argan M: Büyük veri ve pazarlamadaki dönüşüm: Kuramsal bir yaklaşım. *Pazarlama ve Pazarlama Araştırmaları Dergisi* 8(15):1-22, 2015
13. ENACORE: Yapısal Olmayan Verinin Analizi ve Big Data. Available at: <http://www.enacore.com.tr/tr/dokumantasyon/19-yapisal-olmayan-verinin-analizi-ve-big-data>. Accessed June 13, 2016
14. Göksu C: Datawarehouse Türkiye. Available at: <http://datawarehouse.gen.tr/big-data-nedir-geleneksel-veri-yonetimine-etkisi-ne-olur/>. Accessed August 19, 2016
15. Gönenç SA, Arslan S: Yapay Zekâ Araştırmaları ve Biomimesis Kavramlarının Günümüzde Mimarlık Alanındaki Uygulamaları: Akıllı Mekanlar, Yapı ve Kente Bilişim Kongresi, 2004
16. Kelion L: Q & A: NSA's Prism internet surveillance scheme. Available at: <http://www.bbc.com/news/technology-23051248>. Accessed June 6, 2016
17. Li H, He Y, Huang L, Luo H, Zhu X: The nomogram model predicting overall survival and guiding clinical decision in patients with glioblastoma based on the SEER database. *Front Oncol* 10:1051, 2020
18. Mam G: Big data. The next big thing in innovation. *Research-Technology Management* 56(1):64-66, 2013
19. Mesko B: Yapay zekayla tıbbi karar almak. İçinde: Mesko B (ed), *Tıbbın Geleceğine Yolculuk*. İstanbul: Optimist Yayın Grubu, 2017:174-183
20. Özkan E: Büyük Veri (Big Data) ve Map Reduce. Available at: <http://www.erencanozkan.com/2013/08/buyuk-veri-big-data-ve-map-reduce.html>. Accessed June 5, 2016
21. Paro M, Lambert WA, LeClair NK, Romano R, Stoltz P, Martin JE, Hersh DS, Bookland MJ: Machine learning-driven clinical image analysis to identify craniosynostosis: A pilot study of telemedicine and clinic patients. *Neurosurgery*, 2022 (Epub ahead of print)
22. Partners N: Big Data Executive Survey: Creating a Big Data Environment to Accelerate Business Value. NewVantage Partners, 2012
23. Qiao N, Ma Y, Chen X, Ye Z, Ye H, Zhang Z, Wang Y, Lu Z, Wang Z, Xiao Y, Zhao Y: Machine learning prediction of visual outcome after surgical decompression of sellar region tumors. *J Pers Med* 12(2):152, 2022
24. Snoad L: Is data safe in brands' hands? Available at: <http://www.marketingweek.co.uk/analysis/data-strategy/is-data-safe-in-brands-hands/3032564.article>. Accessed June 5, 2016.
25. Wen J, Chen W, Zhu Y, Zhang P: Clinical features associated with the efficacy of chemotherapy in patients with glioblastoma (GBM): A surveillance, epidemiology, and end results (SEER) analysis. *BMC Cancer* 21(1):81, 2021
26. Zhou X, Niu X, Sun K, Li J, Mao Q, Liu Y: Pediatric glioma outcomes: Predictors of early mortality. *World Neurosurg* 139: e700-e707, 2020