

Büyük Verinin Yapısı ve Eğitimde Kullanılması

DERLEME MAKALESİ

H. Deniz GÜLLEROĞLU¹, Fatma COŞKUN²

1 Doç. Dr., Ankara Üniversitesi, Eğitimde Ölçme ve Değerlendirme, denizgulleroglu@yahoo.com, ORCID: 0000-0001-6995-8223.

2 Dr. Öğr. Üyesi, Kahramanmaraş Sütçü İmam Üniversitesi, Eğitim Fakültesi,, Eğitimde Ölçme ve Değerlendirme, fatmacoskuncf@gmail.com, ORCID: 0000-0002-6388-3504.

Gönderilme Tarihi: 22.12.2022 Kabul Tarihi: 28.07.2023 DOI: 10.37669/milliegitim.1222977

Atf: “Gülleroğlu, H. D., ve Coşkun, F. (2024). Büyük verinin yapısı ve eğitimde kullanılması. *Milli Eğitim*, 53 (242), 1139-1152. DOI: 10.37669/milliegitim.1222977”

Öz

Teknoloji tabanlı sistemlerin yaygınlaşmasıyla birlikte ilkokul çağından, emeklilik yaşına değin tüm bireylerin teknolojik cihazlar ve sanal ortam üzerindeki davranışları veri şeklinde kayıt altına alınmaktadır. Normal yaşam aktivitelerini yerine getiren herhangi bir bireyden ya da gruptan elde edilen bu dev veri yığını “büyük veri” olarak ifade edilmektedir. Günümüzde büyük verinin analizi ile geleceğe yönelik tahminler yapılabilmekte, mevcut duruma ya da geçmişe ait sorular cevaplandırılabilir. Bankacılık, pazarlama, turizm, ulaşım gibi pek çok sektör, gelecek planlarını büyük veri analizine dayandırarak oluşturmaktadır. Benzer şekilde, bilgisayar teknolojilerinin eğitim sistemi üzerindeki etkisinin artışıyla, eğitim alanında da büyük veri patlaması yaşanmıştır. Bu çalışmada, büyük verinin yapısı, önemi ve etkisine ilişkin bilgi verildikten sonra eğitim alanında kullanılması, avantajları, güçlükleri ve mevcut uygulamalar üzerinde durulmuştur. Aynı zamanda, büyük verinin eğitim sistemine nasıl entegre edilebileceği ile ilgili bir takım öneriler sunulmuştur.

Anahtar Kelimeler: büyük veri, eğitimsel veri madenciliği, bireyselleştirilmiş öğrenme, bilgisayar tabanlı eğitim

The Structure of Big Data and It's Use in Education

Abstract

With the spread of technology-based systems, the behaviors of all individuals from primary school age to retirement age on technological devices and virtual environment are recorded in the form of data. This huge pile of data obtained from any individual or group performing normal life activities is expressed as "big data". Today, with the analysis of big data, predictions can be made for the future and questions about the current situation or the past can be answered. Many sectors such as banking, marketing, tourism and transportation base their future plans on big data analysis. Similarly, with the increase in the impact of computer technologies on the education system, there has been a big data explosion in the field of education. In this study, after giving information about the importance and impact of big data in our age, its use in the field of education, its advantages, difficulties and current applications are emphasized. At the same time, some suggestions have been presented on how big data can be integrated into the education system.

Keywords: *big data, educational data mining, individualized learning, computer-based education*

Giriş

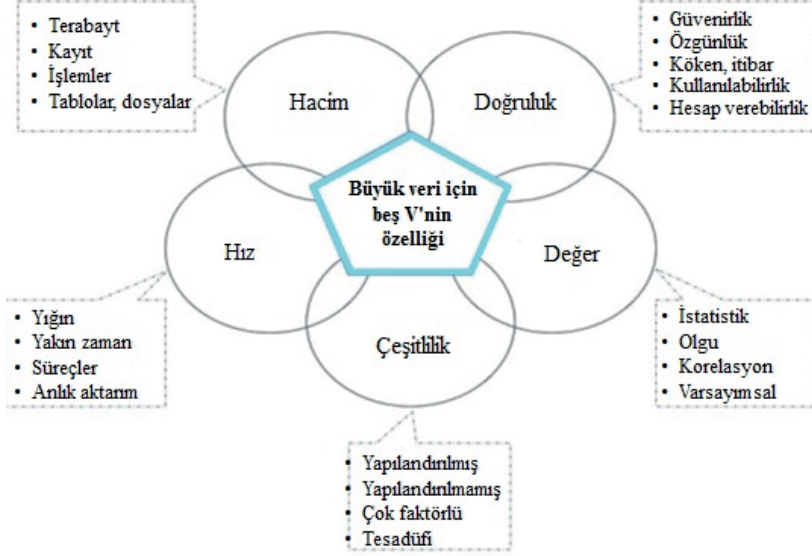
İçinde bulunduğumuz yüzyılın başından itibaren bilgisayar teknolojileri ve sanal dünyada yaşanan hızlı gelişmelerle beraber, herhangi bir bireyin gün içerisinde neredeyse yaptığı tüm aktiviteler dijital ortamda kayıt altına alınmaktadır. Bir bireyin sabah uyandığı saat, evden çıkarken giydiği kıyafet, kırmızı ışıkta beklerken yüz ifadesinin aldığı hal, sosyal medyada geçirdiği süre, hatta gün içerisinde kullandığı kelimeler, ses tonu gibi pek çok davranışı kişisel teknolojik cihazları veya çevresinde bulunan sensörler, kameralar, mikrofonlar tarafından kaydedilmektedir. Bu denli geniş bir zaman diliminden elde edilen, çok fazla bilgiyi ve çeşitliliği içeren muazzam boyuttaki veriler, büyük veri (big data) olarak ifade edilmektedir. Doğan ve Arslantekin'in (2016) belirttiği gibi büyük veri, bir anlamda yeni bir çağın başlangıcına işaret etmektedir. Bunun farkına varan pek çok mevcut kurum ve kuruluş, gelecek planlarını büyük veri üzerine inşa etmektedir. Bu çalışmada, günümüzde gücü yadsınamaz bir gerçek olan büyük veri konusu, genel bir bakış açısı ile ele alınarak, avantaj ve güçlükleri, kullanım alanları ve eğitim alanındaki uygulamalarından örnekler sunulurken, büyük verinin eğitimde kullanımına ilişkin bir takım öneriler sunulmuştur.

Büyük Veri Nedir?

Büyük veri terimi ilk defa, 2005 yılında O'Reilly Medyadan Roger Magoulas tarafından kullanılmıştır. Roger Magoulas, “büyük veri” (big data) terimini karmaşık yapısı, ulaştığı büyük boyutlar ve yönetilmesinin güçlüğü nedeniyle geleneksel veriden farklı olan verileri tanımlamak için kullanmıştır (Hadi, Shnain, Hadishaheed, ve Ahmad, 2015). Sin ve Muthu (2015) ise büyük veriyi, geleneksel analiz teknikleriyle işlenemeyecek kadar büyük veya karmaşık veri kümeleri olarak tanımlamışlardır. Büyük veri teriminin kullanımı, tarihsel olarak çok yakın bir geçmişe sahip olsa da, veri kaynaklarının sürekli çeşitlenmesi ve gelişmesiyle beraber büyük veri tanımı ve modellenmesinde çok kısa süre içerisinde güncellenmeler yaşandığı görülmektedir. Zikopoulos, Eaton, deRoos, Deutch ve Lapis (2012), “Büyük Veriyi Anlamak (Understanding Big Data)” adlı kitapta büyük veri terimini, “geleneksel veriden hacim (volume), çeşitlilik (variety) ve hız (velocity) açısından farklı olan veri” biçiminde tanımlamışlar ve bu durumu 3V (ya da V3) şeklinde modellemişlerdir. Hadi ve diğerleri (2015), büyük veriyi tanımlarken önceki tanımlamayı kabul etmekle birlikte doğruluk (veracity) ve değer (value) özelliklerini de ekleyerek büyük veriyi 5V modeli ile açıklamışlardır. Bozkurt (2016), büyük veriyi tanımlarken doğrulama (verification) özelliğini de kullanarak 6V modeli ile açıklamıştır. Sonuç olarak büyük veri, sadece yeni bir terim değil, aynı zamanda, yeni ve geniş bir çalışma alanının da temel konusu olduğundan sürekli olarak gelişmekte ve güncellenmektedir. Bu çalışmada, Hadi ve diğerlerinin (2015) ortaya koyduğu 5V modeli üzerinden büyük veri ile ilgili detaylı tanımlamalar yapılmıştır. Hadi ve diğerlerinin (2015) oluşturduğu 5V modeli Şekil 1’de sunulmuştur.

Şekil 1

5V'nin Büyük Veri Özellikleri (Kaynak: Hadi ve diğ., 2015)



Hadi ve diğerlerine (2015) göre, büyük veri; hacim, çeşitlilik, hız, değer ve doğruluk açısından geleneksel veriden ayrılmaktadır. Büyük verinin ayırt edici özelliklerinin tanımı şu şekildedir:

Hacim: büyük verinin hacim özelliği ile kastedilen, geleneksel veriden farklı olarak basit teknik ve sistemler ile depolanamayacak kadar geniş bir boyut oluşturmaktadır. 2000 yılına kadar bütün dünyada üretilen veri hacmi 800.000 perapayt (PB) olarak kaydedilmiştir. Bugün ise Twitter tek başına günlük 7 terabayt (TB), Facebook ise 10 TB veri üretmektedir. En basit şekliyle hacim, verilerin TB'den PB'ye ve en sonunda kaçınılmaz olarak zettebayt (ZB) ile ifade edilebilecek boyutlara ulaşmasıdır (Zikopoulos ve diğ., 2012).

Çeşitlilik: büyük verinin en önemli özelliklerinden biri, geleneksel veride görmeye alışkın olduğumuz ilişkili yapılandırılmış veriler yerine; web sayfaları, arama dizinleri, sosyal medya platformları, e-posta, belge, aktif ve pasif sensor girdileri gibi çok çeşitli kaynaklardan gelen veriyi içermesidir. Bu açıdan, özetlenecek olursa, büyük verinin yapılandırılmış geleneksel veriden farklı olarak ham, yarı yapılandırılmış ve yapılandırılmamış veriden oluşması çeşitlilik özelliği olarak ifade edilmiştir (Zikopoulos ve diğ., 2012).

Hız: geleneksel verideki hız anlayışı, tipik olarak, verilere ne kadar sürede ulaşıldığı, depolandığı ve geri bildirim sağlandığı ile ilgilidir. Ancak büyük verideki hız özelliği, sürekli bilgi akışından dolayı her saniyede, hatta her salisede değişen verinin akış hızından etkilenmeden analiz edilebilmesini belirtir. Büyük verinin hacim ve çeşitlilik özellikleri ile başa çıkılabilmesi için verinin durgunken değil, akış halindeyken analiz edilebilmesi gerekir. Bu da hız özelliğidir (Zikopoulos ve diğ., 2012).

Değer: büyük verinin problem durumu karşısındaki istatistiksel anlamını, kurulan hipotezler ve modeller açısından önemini belirtir (Hadi ve diğ., 2015).

Doğruluk: büyük verinin güvenilirlik, geçerlik, kullanılabilirlik, gerçeği yansıtma, hesap verilebilirlik gibi özelliklerini temsil eder (Hadi ve diğ., 2015).

Büyük verinin toplanması, depolanması ve analizi pek çok açıdan geleneksel veriden farklılık göstermektedir. Büyük verinin gelişimi, veri madenciliği biliminin ortaya çıkmasına neden olmuştur. Veri madenciliği, büyük verinin gizlediği önemli bilgilerin ortaya çıkarılması amacıyla, kullanılacak analiz tekniklerinin belirlenmesi, matematiksel modellerin oluşturulması, destek sağlayacak yazılım ve teknolojilerin geliştirilmesini konu alan bilim dalıdır (Baykal, 2006). Büyük verinin ortaya çıkması ve öneminin artmasıyla birlikte, veri tabanlarında depolanan ham veriyi düzenlemek, analiz etmek ve yorumlamak amacıyla matematik, istatistik, bilgisayar bilimleri gibi pek çok bilim alanından destek alınarak veri madenciliği biliminin ortaya çıkması bu sürecin doğal bir sonucudur. Büyük verinin gelişim hızının, giderek artan bir ivme kazanmasıyla, zaman içerisinde veri madenciliği bilimi de farklı bilim alanları açısından özelleşmiştir. Eğitim alanında ise eğitimin planlanması, stratejilerin belirlenmesi, uygulamaya dönüştürülmesi gibi birçok aşamada büyük verinin kullanılması, depolanması ve analizi gibi işlemler için eğitsel veri madenciliği bilimi ortaya çıkmıştır (Hotaman, 2020).

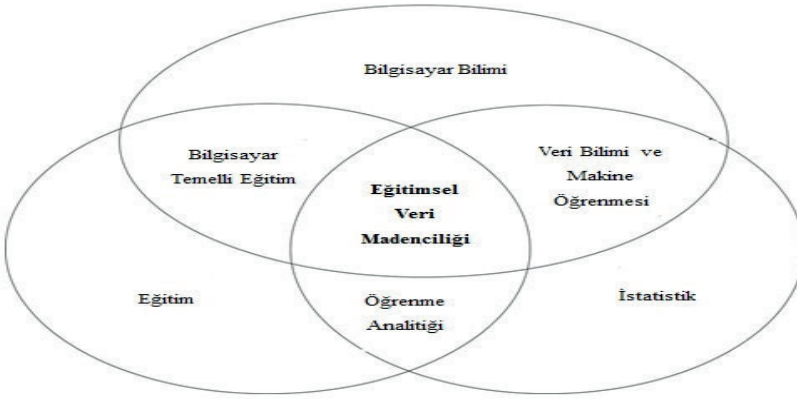
Eğitimsel Veri Madenciliği

Eğitim araştırmacıları için büyük veri, bazı açılardan geleneksel veriden oldukça farklıdır ve bu durum bir dizi metodolojik sorunu beraberinde getirmektedir (Cope ve Kalantzis, 2016). Bu nedenle, eğitim bilimleri alanında veri bilimi yöntemlerinin uygulanması ve teori geliştirmek gerekmektedir. Bu durum, iki nedenle çok önemlidir. İlki, sanal ortam üzerinden performans değerlendirmenin psikometrik olarak bir takım zorluklar barındırması, ikincisi ise, eğitim araştırması hazırlama ve uygulamada yeni ölçme araçlarının, karmaşık sistemlerdeki yapılarının ortaya çıkartılması gerekliliğidir (Gibson, Webb ve Ifenthaler, 2015). Bu tür soru ve sorunlara yanıt üretebilmek amacıyla, eğitimsel veri madenciliği alanı ortaya çıkmıştır. Eğitimsel veri madenciliğinin ortaya çıkışı çok yeni olmasına rağmen hızla gelişen bir alandır. Eğitimsel veri

madenciliği, eğitim soru ve sorunlarına cevap verebilmek amacıyla veri madenciliği yöntem ve tekniklerinin eğitimden elde edilen ham verilere uygulanması yolu ile verilerdeki gizli bilgilerin ve örüntü şemalarının ortaya çıkarılması olarak tanımlanabilir (Bousbia ve Belamiri, 2014). Eğitimsel veri madenciliği alanı, eğitim alanında elde edilen büyük verinin analizi ve yorumlanması için yöntem ve tekniklerin geliştirilmesini konu alan disiplinler arası bir alandır. Romero ve Ventura (2013), eğitsel veri madenciliği alanını ve ilişkili olduğu bilim dallarını Şekil 2’deki gibi modellemiştir.

Şekil 2

Eğitimsel Veri Madenciliği ile İlişkili Alanlar (Kaynak: Romero ve Ventura, 2013)



Şekil 2’de görüldüğü üzere, eğitimsel veri madenciliği, bilgisayar bilimi, eğitim ve istatistik ana alanları olmak üzere bu üç alanın birleşiminden oluşan bilgisayar temelli eğitim, öğrenme analitiği ve veri bilimi/ makine öğrenmesi alanları ile doğrudan ilişkili yeni bir alandır (Romero ve Ventura, 2013). Cope ve Kalantzis (2016), eğitimde kullanılan büyük veri kaynaklarını; makine değerlendirmeleri, yapılandırılmış gömülü veri ve yapılandırılmamış tesadüfi veri şeklinde üç tema altında açıklamıştır. Buna göre, eğitimde kullanılan büyük veri kaynakları şunlardır:

1-Makine değerlendirmeleri: bu tema ile kastedilen bilgisayar ortamında bireye uyarlanmış test (Computer Adaptive Testing) uygulamalarıdır. Örneğin; verilen yanıtı göre değerlendirme yapılması, sınavlar (örn: okuduğunu anlama, kelime bilgisi, gramer bilgisi) bu kategori altında yer almaktadır.

2-Yapılandırılmış, gömülü veriler: doğal dil işleme, prosedür tanımlı süreçler, argüman tanımlı süreçler ve makine öğrenmesi süreçleridir. Örneğin; otomatik makale puanlama, dil özellikleri hakkında geri bildirim veren oyunlar, yapay zeka (öğretme)

sistemleri, rubrik tabanlı akran değerlendirme sistemleri, semantik etkileme ve açıklama, metin görselleştirme, kabul edilen metinsel değişiklik önerileri gibi işlemler bu kategori altında bulunur.

3- Yapılandırılmamış/ tesadüfi veriler: tesadüfi “veri çıkışı” ve yapılandırılmamış veri elde etmek için özel cihazlardan elde edilen verilerdir. Örneğin: tuş vuruşu desenleri, kayıtların düzenlenmesi, tıklama akışı, navigasyon kayıtları, sosyal etkileşim kalıpları, video yakalama, göz takip cihazları, hareket detektörleri vb. veri kaynaklarından elde edilen büyük veri bu sınıfta yer alır.

Bunun yanı sıra öğrencilerin sosyal medya hesapları üzerinden elde edilen veri de, eğitim sistemlerinin geliştirilmesi için büyük veri kaynaklarına dahil edilebilir (Özen, Kartal ve Emre, 2017). Özen, Kartal ve Emre (2017) eğitimde büyük verinin kullanımını üç adımda özetlemişlerdir. Bu adımlar;

1-Büyük verinin elde edilmesi ve depolanması için ilgili yazılım ve donanım sistemlerinin geliştirilmesi,

2-Elde edilen büyük verinin analiz edilmesi,

3-Analiz sonucunda elde edilen bilgilerden hareketle bireyselleştirilmiş öğrenme sistemlerinin geliştirilmesidir.

IBM, “Başarı için Analitik (Analytics for achievement)” adı altında gerçekleştirdiği bir incelemede, büyük verinin eğitim uygulamalarındaki kullanımını sekiz madde altında özetlemiştir (Akt. Picciano, 2012). Bunlar sırasıyla şöyledir:

1-Bireysel öğrenci performansının incelenmesi,

2-Öğrenci performansının ana dal, eğitim yılı, eğitim düzeyi, etnik köken gibi değişkenlere göre ayrıştırılması,

3-Erken müdahale için uç değerlerin tespiti,

4-Tüm öğrencilerin maksimum başarı gösterebilmesi amacıyla sahip oldukları potansiyelin tahmin edilmesi,

5-Bir dersten veya programdan başarısız olma ihtimalinin önlenmesi,

6-Etkili öğretim tekniklerinin belirlenmesi ve geliştirilmesi,

7-Standart değerlendirme araçları ve tekniklerinin test edilmesi ve

8-Öğretim müfredatının test edilmesi ve değerlendirilmesidir (Akt. Picciano, 2012).

Büyük Verinin Eğitimde Kullanımının Avantajları

Eğitimde büyük veri kullanımının yaygınlaşması, özellikle çevrim-içi eğitim uygulamalarının artışıyla gündeme gelmiştir. 2000’li yılların başından itibaren eğitimde bilgisayar teknolojilerinin kullanımı, öğrenme süreçleri ile ilgili pek çok verinin kaydedilmesini sağlamış ve eğitim alanında büyük verinin oluşumunu hızlandırmıştır (Bozkurt, 2016). Eğitimde yaşanan güçlüklerin aşılabilmesi ve mevcut sorunların çözümü için büyük verinin toplanması, işlenmesi ve analiz edilmesi büyük bir fırsat olarak nitelendirilmektedir (Har Carmel, 2016).

Uluslararası eğitim karşılaştırmaları için günümüz ülkelerinin büyük bir önem attığı TIMSS, PISA, PIRLS gibi uygulamalar, ciddi bir büyük veri kaynağıdır. Nitekim son dönemde büyük ölçekli sınav verilerinin analizinde, veri madenciliği tekniklerinin kullanılması bu görüşü desteklemektedir. Aksu ve Güzeller (2016), PISA 2012 uygulamasına ait verileri analiz etmek amacıyla veri madenciliği yöntemlerinden olan J48 karar ağacı tekniğini kullanmışlardır. Çetintav, Altun-Tot ve Yılmaz (2022), TIMSS 2019 uygulaması verilerinden hareketle teknoloji kullanımının, uluslararası düzeydeki öğrenci başarısına etkisini yine eğitsel veri madenciliği yöntemlerinden biri olan karar ağacı tekniği ile incelemişlerdir. Martínez-Abad, Gamazo ve Rodríguez-Conde (2020), PISA 2015 verisinden hareketle İspanya’daki yüksek ve düşük etkiye sahip okulları veri madenciliği yöntemlerinden karar ağacı tekniğini kullanarak belirlemiştir. Ayrıca araştırmacılar özellikle büyük ölçekli uygulamalardan elde edilen verilerin analizinde karar ağaçları tekniğinin %90 gibi büyük bir olasılıkla daha doğru sonuçlar sağladığını vurgulamıştır (Martínez-Abad, Gamazo ve Rodríguez-Conde 2020). Büyük ölçekli sınavlardan elde edilen verilerin, araştırmacıların erişimine açık olarak sunulması, araştırmacılar tarafından veri madenciliği yöntem ve teknikleri kullanılarak eğitim değişkenlerinin farklı açılardan değerlendirilmesine imkân sağlamaktadır. Bu nedenle PISA, TIMSS, PIRLS gibi uluslararası uygulamaları eğitim sistemlerindeki sorun, eksiklik veya olumlu yönlerin ortaya çıkartılması açısından önemli bir büyük veri kaynağıdır.

2020 yılının mart ayında bütün dünyayı etkisi altına almaya başlayan koronavirüs salgını, insanlığın eğitim tarihi boyunca baş etmesi gereken en büyük problemlerden birini, beraberinde getirmiştir. Bu dönemde dünya nüfusunun önemli bir kısmını oluşturan 1.6 milyar öğrenci yüz yüze eğitimden mahrum kaldığından birçok ülke zorunlu olarak acil uzaktan eğitim sistemine geçiş yapmıştır (Özer ve Suna, 2020). Koronavirüs salgını, 1990’lardan itibaren eğitimde etkisini göstermeye başlayan dijitalleşmenin hızına ivme kazandırmıştır. Bu dönemde uzaktan eğitim sistemlerinde kullanılan bilgi iletişim teknolojileri aracılığıyla eğitim alanında kayıt altına alınan verinin hacminde ve çeşitliliğinde muazzam bir artış yaşanmıştır. Bu durum eğitim

alanında ciddi oranda büyük verinin oluşmasını sağlamıştır (Tosunoğlu, Yılmaz, Özeren ve Sağlam, 2021). Bu süreç, yaşamın her alanında olduğu gibi eğitimde de büyük veri ve buna dayalı sistemlerin kullanımına dair gerçekleşen dönüşümün kaçınılmaz olduğunu göstermiştir (Akçapar, 2020).

Eğitimde büyük veri kullanımının, öğrencilerin akademik başarılarının artırılmasından, devamsızlık ya da okuldan ayrılma gibi problemlerin çözümüne kadar birçok konuda fayda sağlayacağı düşünülmektedir (Özen, Kartal ve Emre, 2017). Ayrıca, eğitim alanındaki büyük verinin analizi ile elde edilen bilgiler, program geliştiriciler ve eğitim sistemi tasarlayıcıları için altın madeni değerindedir. Eğitim alanındaki büyük veri, eğitimin niteliğinin artırılması ve öğrenci performansının yükseltilmesi için mükemmel bir fırsat sunmaktadır (Bousbia ve Belamiri, 2014). Eğitimde büyük verinin kullanımı, öğrencilerin mevcut durumunun yanı sıra, gelecekteki davranışlarının ve performanslarının tahmin edilmesini de mümkün kılmaktadır. Bu durum, gelecekte başarısız olması veya olumsuz davranışlar edinmesi muhtemel olan öğrenciler için önceden düzeltici önlemlerin alınmasına olanak sağlamaktadır. Öte yandan, eğitimde büyük veri kullanımının sağladığı en önemli katkılardan birisi de bireyselleştirilmiş eğitim sistemleridir. Bireyselleştirilmiş eğitim sistemleri, doğru öğrenci performansının yükselmesine katkı sunar (Özen, Kartal ve Emre, 2017).

Eğitimde büyük veri kullanımının doğrudan katkı sağladığı bir diğer alan, ölçme ve değerlendirme alanıdır. Yapılandırmacı eğitim felsefesinin eğitim alanında etkili olmasının ardından, eğitimde kullanılan ölçme ve değerlendirme anlayışında da önemli değişimler yaşanmıştır. Sonuç odaklı geleneksel değerlendirme yaklaşımından, süreç odaklı değerlendirme yaklaşımına geçiş, tamamlayıcı (alternatif) ölçme ve değerlendirme araçlarının eğitimde kullanımını yaygınlaştırmıştır. Bu durum, ölçme ve değerlendirme işlemlerini sürece yaymakla kalmamış, aynı zamanda öğrenciler ile ilgili büyük verinin oluşmasına da yol açmıştır (Hotaman, 2020). Söz gelimi, geleneksel eğitim anlayışında kullanılan kağıt-kalem testlerinin yanı sıra, günümüzde kullanılan elektronik portfolyo gibi uygulamalar öğrencilere ait akran değerlendirme, öz değerlendirme, ürünler gibi uzun bir periyoda dayalı pek çok bilgiyi içermektedir (Kutlu, Doğan ve Karakaya, 2014). Bu yönüyle, tamamlayıcı ölçme ve değerlendirme araçları büyük veri kaynağı olarak görülebilir. Öte yandan, bu türden hacim ve çeşitlilik özelliklerine sahip verilerin analizi için eğitsel veri madenciliği tekniklerinin kullanılması gereklidir (Hotaman, 2020). Eğitim alanındaki büyük verinin diğer bir önemli özelliği, bilinçli olarak elde edilen verilerin dışında, tesadüfi ya da plansız olarak da elde edilen verilerden oluşmasıdır. Örneğin, bilgisayar teknolojilerinin eğitimde kullanılmasıyla birlikte kamera, mikrofon ve benzeri sensörler aracılığıyla elde edilen verilerin büyük bir kısmı, toplanması amaçlanan veri dışındaki tesadüfi veriyi

de içermektedir (Toptaş ve Şen, 2021). Bu açıdan, eğitim alanındaki büyük verinin, ölçme ve değerlendirme süreçlerinde elde edilen bilgiden daha gerçek bir durumu yansıttığı savunulabilir. Nitekim planlanmış ölçme ve değerlendirme süreçlerinde, sınav kaygısı, değerlendirilme kaygısı gibi tesadüfi hata kaynaklarının ölçme sonuçlarına karışma ihtimalinin yüksek olduğu bilinen bir gerçektir.

Eğitim alanında büyük verinin kullanılmaya birlikte, bireyselleştirilmiş eğitim sistemleri ve öğrenme analitiği gibi alanların ortaya çıkışı, eğitim sistemine önemli katkılar getirmiştir (Bozkurt, 2016). Hor Camel (2016), eğitimde büyük veri kullanımının avantajlarını; bireyselleştirilmiş öğrenme, anında ve etkili geri bildirim, uygulanabilir öğrenme, geçerli değerlendirme ve performans tahmini şeklinde beş maddede açıklamıştır. Büyük veriye dayalı eğitim uygulamaları geliştiren yazılım şirketi INOXOFT ise eğitimde büyük veriye dayalı olarak yapılabilecek çalışmaları; eğitimi kişiselleştirmek, harmanlanmış öğrenmeyi geliştirmek, sürece yayılmış değerlendirme sistemleri, yaşam boyu öğrenmeyi teşvik etmek şeklinde kategorileştirmiştir (INOXOFT, 2022).

Eğitimde Büyük Veri Kullanımında Karşılaşılabilecek Güçlükler

Önceki bölümlerde belirtildiği gibi büyük verinin kullanımı, pek çok açıdan önemli avantajlar sağlamaktadır. Ancak büyük verinin kullanımı sanıldığı kadar kolay değildir. Büyük veri kullanımı sırasında karşılaşılabilecek bir takım önemli güçlükler ve problemler mevcuttur. Gibson, Webb ve Ifenthaler (2015), eğitim değerlendirmeleri özelinde, büyük veri kullanımı sırasında yaşanabilecek sorunları zaman duyarlılığı, dijital performans mekân ilişkisi, toplama katmanları-dönüştürme ve dinamiklerin temsilleri olmak üzere dört tema üzerinden açıklamıştır. Bu dört temanın tanımı kısaca şöyledir; zaman duyarlılığı, analizlerin eş zamanlı analiz edilmesi gerekmektedir. Çünkü büyük verinin önemli bir özelliği de 3V de hız olarak belirtilen sürekli değişimdir. Büyük veri her an her saniye değişmektedir. Dijital performans mekân ilişkisi, sanal ortamdaki verinin standart ölçme ortamlarından bağımsız olmasıdır. Toplama katmanları ve dönüştürme, toplanan verinin yapısı, düzenlenmesi ve ayıklanmasıdır. Dinamiklerin temsilleri, kullanılan matematiksel modellerin geçerliğini ifade etmektedir (Gibson, Webb ve Ifenthaler, 2015).

Büyük veri, yapısı gereği boyut ve çeşitlilik açısından çok geniş veri setlerinden oluşur. Dolayısıyla büyük veride saklı bilgi de, geleneksel verideki sade ve yorumlanabilir düzeyin çok üstündedir. Bu açıdan günümüz koşullarında, büyük verinin tam anlamıyla analiz edilip yorumlanabilmesi imkânsız olarak görülmektedir. Başka bir deyişle, günümüz koşullarında yapılan büyük veri analizi ve yorumlanması kısmi düzeyde kalmaktadır (Işıklı, 2014).

Teknik konuların yanı sıra, eğitim alanında büyük verinin kullanılması, depolanması ve analizi ile ilgili etik bir takım güçlükler de mevcuttur. Büyük veri kullanımı ile ilgili sayısız farklı yasal gizlilik kısıtlaması söz konusudur. Bu durum, farklı ülkelerde büyük veri kullanımının en ciddi zorluklarından birine dönüşmektedir (Buhl, Röglinger, Moser ve Heidemann, 2013). Alanyazında büyük verinin etik açıdan kullanılabilirliğine yönelik iki farklı görüş mevcuttur. Bir grup, büyük verinin pek çok problemin yaşanmadan ön görülüp çözülmesine olanak sağlayan büyük bir keşif olarak nitelendirirken; diğer grup büyük verinin kişisel gizliliği ve hakları ihlal eden tehlikeli bir güç olarak değerlendirmektedir (Işıklı, 2014; Özen, Kartal ve Emre, 2017). Büyük veri, kişisel ve özel hayatı olumsuz etkileme potansiyeline sahip bir olgudur (Cope ve Kalantzis, 2016). Bu durum, günlük yaşam aktivitelerini sürdüren bir yetişkin için de, okul görevlerini yerine getiren bir öğrenci için de geçelidir. Büyük veri, bireylerde her an gözetim altında olduğu, her davranışının değerlendirildiği ve bir sonuç çıkarılmaya çalışıldığı düşüncesini doğurabilir, bu durum da kişilerin psikolojilerini olumsuz yönde etkileyebilir. Öte yandan büyük veri sayesinde riskli gruptaki öğrencilerin tespit edilerek erken müdahale ile yaşanabilecek problemlerin ortadan kaldırılması da mümkündür.

Işıklı (2014), özellikle büyük verinin kullanımı ile ilgili etik ilkelerin yetersizliğine vurgu yaparak, bu konuda ivedilikle büyük veriden kaynaklı oluşabilecek etkilerin tartışıldığı yeni bir felsefi yaklaşımın geliştirilmesi gerektiğini belirtmiştir. Alanyazındaki değerlendirmelerden de anlaşıldığı üzere, büyük verinin kullanımı çok hızlı yaygınlaşan bir gelişim izlemesine rağmen, büyük verinin kullanımı konusunda yaşanan teknik, teknolojik, felsefi ve etik gibi bir takım güçlükler henüz tam anlamıyla aşılabilmiş değildir. Ancak bu durum, büyük verinin kullanımı için bir engel değildir. Picciano'nun (2012) belirttiği gibi günümüzde büyük verinin kullanımı henüz emekleme dönemini yaşamaktadır. Dolayısıyla bu tür güçlüklerin yaşanması oldukça doğaldır. Bu açıdan bu konuda yapılacak her çalışma ve atılacak her adım oldukça önemli görülmektedir.

Eğitimde Büyük Veri Kullanımı İçin Mevcut Örnekler

Günümüzde büyük verinin eğitimde kullanımına dair pek çok örnek görülmektedir. Bu başlık altında ele alınan birkaç örnek şu şekildedir:

- Büyük verinin eğitim sistemlerine en önemli katkılarından biri de yapay zeka sistemlerine dayalı eğitim sistemleridir. Amerika Birleşik Devletleri'nde bulunan Silikon Vadisi'nde, büyük veriye dayanan akıllı öğrenme yazılım sistemlerini test etmek için Başlangıç Okulları (Startup Schools) kurulmuştur (Williamson, 2015).

- Clutch tarafından yılın en iyi Web geliştiricileri ödülünü alan INOXOFT şirketi büyük veri ve makine öğrenmesi alanlarından yararlanarak yapay zeka destekli eğitim içerikleri hazırlamaktadır. Şirket eğitimde büyük veri kullanımını ile ilgili uygulamalarını şu şekilde sıralamıştır: öğrencilerin başarılarının ölçülmesi, davranış algılama ve tahmine dayalı modelleme, uygulamada öğrencilerin davranış tespiti, iyileştirmeleri artırmada öğretmenlerin yardımı, kariyer yolu eşleştirme ve işe alım uygulamaları, kişiselleştirilmiş öğrenmenin geliştirilmesi, bilginin dar alanlara dönüştürülmesi (INOXOFT, 2022).
- Amerika Birleşik Devletleri'ndeki Purdue Üniversitesi, öğrencilerinin akademik başarısını arttırmak amacıyla büyük veriye dayalı "Kurs İşaret Sistemi (Course Signals System)" isimli bir uygulama geliştirmiştir. Bu uygulama sayesinde belli bir akademik düzeye erişilmeden önce, akademik düzey olarak riskli durumdaki öğrenciler tespit edilerek sistem üzerinden erken geribildirim uyarıları ile müdahale edilmektedir. Bu uygulamanın üç önemli özelliği; gelecek zamana yönelik geribildirim, erken müdahale ve sürekli geri bildirim şeklinde belirtilmiştir (Picciano, 2012).

Büyük verinin eğitimde kullanılmasına ilişkin sonuç ve öneriler

Büyük veri, eğitimde karşılaşılabilecek bütün problemlerin çözümü için sihirli bir güç değildir; ancak eğitimin yönetsel ve eğitsel süreçlerine entegre edilerek çözümün önemli bir parçası haline getirilebilir (Picciano, 2012). Büyük verinin eğitimde kullanılmasına bir anda geçilmesi, bilimsel ve pratik açıdan çok mümkün görünmemektedir. Bu süreç, ancak adım adım planlanarak başarıya ulaşılabilir. Bu çalışma kapsamında yapılan alanyazın okumalarına dayanarak, önerilen adımlar şu şekilde sıralanabilir:

1- Gerekli teknolojik alt yapının sağlanması: okullarda büyük verinin toplanması ve depolanması için kullanılacak uygun kamera, mikrofon, sensör, bilgisayar, depolama sistemleri gibi teknolojik alt yapıların oluşturulması.

2-Yazılımların temin edilmesi veya geliştirilmesi: büyük verinin analizi için kullanılacak R, Mplus, SAS vs. gibi yazılımların temin edilmesi veya yeni yazılımların geliştirilmesi.

3- Mevcut personelin bu konuda eğitilmesi: eğitimden sorumlu tüm personel için büyük veriyi tüm yönleriyle ele alan önemli eğitimlerin tasarlanması.

4-Bu alana yönelik yeni personel istihdam edilmesi ve yetiştirilmesi: büyük verinin analiz edilmesi için psikometrist, istatistikçi, bilgisayar sistemleri uzmanı, yazılımcı gibi farklı alanlardan uzmanların ortak olarak çalışması gerekmektedir. Eğitim verisi alanında uzmanlaşmış bu grup bir personelin yetiştirilmesi önemli bir gerekliliktir.

5-Her ilin ya da bölgenin büyük veri analizini, gerçekleştirecek merkezlerin kurulması: örneğin; Milli Eğitim Bakanlığı, tüm il merkezlerinde ölçme ve değerlendirme birimleri kurmuştur. Bu birimler, il bazında büyük verinin kontrol ve denetimini üstlenebilir.

6-Merkezi sistem oluşturularak tüm verilerin merkezi sistemde birleştirilmesi: merkezi eğitim sistemine bağlı tek bir merkez kurularak tüm il ya da bölge verilerinin bu merkezde toplanması,

7-Son olarak büyük verinin kullanımına ilişkin etik ilkelerin korunması ve denetlenmesinden sorumlu özel birimlerin kurulması aşamaları uygulanarak eğitimde büyük verinin kullanımına ilişkin ciddi gelişmeler sağlanabilir.

Kaynakça

- Akçapar, B. (2020). M. Şeker, A. Özer ve C. Korkut, (Edt.), *Küresel salgının anatomi: insan ve toplumun geleceği* (s. 320-342). Türkiye Bilimler Akademisi Yayınları.
- Aksu, G., ve Güzeller, C. O. (2016). PISA 2012 matematik okuryazarlığı puanlarının karar ağacı yöntemiyle sınıflandırılması: Türkiye örneklemini. *Eğitim ve Bilim*, 41 (185), 101-122. <https://dx.doi.org/10.15390/EB.2016.4766>
- Baykal, A. (2006). Veri madenciliği uygulama alanları. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 7 (2006), 95-107.
- Bousbia, N., and Belamri, I. (2014). Which contribution does EDM provide to computer-based learning environments. In A. Pena-Ayala (ed.), *Educational data mining* (pp. 3-28). Springer.
- Buhl, H. U., Röglinger, M., Moser, F., and Heidemann, J. (2013). Big data. *Business & Information Systems Engineering*, 5(2), 65-69. <https://doi.org/10.1007/s12599-013-0249-5>
- Cope, B., and Kalantzis, M. (2016). Big data comes to school: Implications for learning, assessment, and research. *American Educational Research Association (AERA) Open*, 2 (2), 1-19. <https://doi.org/10.1177/2332858416641907>
- Çetintav, G., Tot, E. A., ve Yılmaz, R. (2022). Derste teknoloji kullanımının TIMSS 2019 sonuçlarına etkisinin veri madenciliği ile analizi. *Bilgi ve İletişim Teknolojileri Dergisi*, 4 (1), 26-43. <https://doi.org/10.53694/bited.876229>
- Doğan, K., ve Arslantekin, S. (2016). Büyük veri: önemi, yapısı ve günümüzdeki durum. *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi*, 56 (1), 15-36.
- Gibson, D. C., Webb, M., and Ifenthaler, D. (2015, October). *Challenges of big data in educational assessment* [Poster presentation]. 12th International Conference on Cognition and Exploratory Learning in Digital Age, Ireland.
- Hadi, H. J., Shnain, A. H., Hadishaheed, S., and Ahmad, A. H. (2015). Big data and five V's characteristics. *International Journal of Advances in Electronics and Computer Science*, 2 (1), 16-23.

- Har Carmel, Y. (2016). Regulating 'big data education' in europe: lessons learned from the US. *Internet Policy Review*, 5 (1), 1-17. <https://doi.org/10.14763/2016.1.402>
- Hotaman, D. (2020). Öğrenci başarısının değerlendirilmesinde eğitsel veri madenciliğinin kullanımı. *Ulakbilge Sosyal Bilimler Dergisi*, 8 (48), 577-587. <https://doi.org/10.7816/ulakbilge-08-48-08>
- INOXOFT, (2022, 18 Ekim). *Use of big data in education industry: history, benefits and examples*. <https://inoxoft.com/blog/impact-of-big-data-on-education-history-benefits-and-examples/>
- Işıklı, Ş. (2014). Büyük veri, epistemoloji ve etik tartışmalar. *AJIT-e: Academic Journal of Information Technology*, 5 (17), 89-122. <https://doi.org/10.5824/1309-1581.2014.4.006.x>
- Kutlu, Ö., Doğan, C. D. ve Karakaya, İ. (2014). Ölçme ve değerlendirme performans dayalı durum belirleme. Pegem Akademi Yayıncılık.
- Martínez-Abad, F., Gamazo, A., and Rodríguez-Conde, M. J. (2020). Educational data mining: identification of factors associated with school effectiveness in PISA assessment. *Studies in Educational Evaluation*, 66 (2020), 1-10. <https://doi.org/10.1016/j.stueduc.2020.100875>
- Özen, Z., Kartal, E., ve Emre, İ. E. (2017). H. F. Odabaşı, B. Akkoyunlu, A. İşman (Edt.) *Eğitim teknolojileri okumaları* (s. 106-118). TOJET.
- Özer, M. ve Suna, H.E. (2020). M. Şeker, A. Özer ve C. Korkut, (Edt.), *Küresel toplumun anatomisi: İnsan ve toplumun geleceği* (s. 171-192). Türkiye Bilimler Akademisi Yayınları.
- Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal Of Asynchronous Learning Networks*, 16 (3), 9-20.
- Romero, C., and Ventura, S. (2013). Data mining in education. *WIREs Data Mining Knowl Discov*, 3 (1), 12-27. <https://doi.org/10.1002/widm.1075>
- Sin, K., and Muthu, L. (2015). Application of big data in education data mining and learning analytics a literature review. *ICTACT Journal On Soft Computing*, 5 (4), 1035-1049.
- Toptaş, O., ve Şen, A. (2021). Eğitimde ölçme değerlendirmede büyük verinin önemi. *Düşünce ve Toplum Sosyal Bilimler Dergisi*, 3 (4), 223-243.
- Tosunoğlu, E., Yılmaz, R., Özeren, E., ve Sağlam, Z. (2021). Eğitimde makine öğrenmesi: araştırmalardaki güncel eğilimler üzerine inceleme. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, 3 (2), 178-199.
- Williamson, B. (2015, December). *Smarter learning software: Education and the big data imaginary*. [Paper presentation]. Big Data-Social Data Conference, University of Warwick, Coventry.
- Zikopoulos, P., Eaton, C., de Roos, D., Deutch, T., and Lapis, G. (2012). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media.