

EĞİTİMDE BÜYÜK VERİ TEKNOLOJİSİNİN YERİ VE EĞİTİMDE BÜYÜK VERİ ANALİZ YÖNTEMLERİ

The Place Of Big Data Technology In Education And Big Data Analysis Methods In Education

Doç. Dr. Murat BAY

Karamanoğlu Mehmetbey Üniversitesi, İktisadi ve İdari Bilimler Fakültesi Karaman/Türkiye

ORCID ID: <https://orcid.org/0000-0002-2968-2258>

Mukaddes Özlem ÇELİK

Karamanoğlu Mehmetbey Üniversitesi, Sosyal Bilimler Enstitüsü, Yönetim Bilişim Sistemleri Bölümü, Karaman/Türkiye

ORCID ID: <https://orcid.org/0000-0003-4321-5645>

ÖZET

Bu araştırmanın amacı büyük veri teknolojisinin eğitimde getirdiği yenilikler ve avantajları tartışmaktır. Büyük veri teknolojisi özellikle kâr amacı güden pazarlama örgütlerinde yaygın olarak kullanılsa da diğer alanlarda kullanımı kısıtlı kalmıştır. Toplum gelişimi ve refahı için eğitimin önemi ortadadır. Bu nedenle büyük veri teknolojisinin eğitim alanında getirebileceği yenilikler toplumların gelişimini destekleyecek niteliktedir. Bunun yanında büyük veri teknolojisi sayesinde var olan eğitim sisteminden toplanan veriler sayesinde sistemin gereksinimlerini ortaya koyarak gelişimini destekleyebilecektir. Eğitimde büyük veri analizinin avantajlarının farkında olan araştırmacılar tarafından 1995 yılında başlatılmış olan eğitimsel veri analizi algoritması bugünde geliştirilmekte olup yeni analiz yöntemleri literatürde yer almaktadır. En yaygın kullanılan eğitimde büyük veri analiz yöntemleri ise eğitimsel veri analizi ve öğretim analizidir. Bu yöntemler bu çalışmada detaylı olarak incelenmiştir.

Anahtar Kelimeler: Büyük Veri, Eğitim 4.0, Büyük Veri Analiz Yöntemleri

ABSTRACT

The aim of this research is to discuss the innovations and advantages that big data technology brings in education. Although big data technology has been widely used, especially in nonprofit marketing organizations, its use has remained limited in other areas. The importance of education for Community Development and well-being is obvious. For this reason, the innovations that big data technology can bring in the field of education are qualified to support the development of societies. In addition, thanks to the data collected from the existing education system thanks to big data technology, it will be able to support its development by revealing the requirements of the system. The educational data analysis algorithm, which was started in 1995 by researchers who are aware of the advantages of big data analysis in education, is being developed today and new analysis methods are included in the literature. The most commonly used methods of big data analysis in education are educational data analysis and instructional analysis. These methods were examined in detail in this study.

Keywords: Big Data, Education 4.0, Big Data Analysis Methods

1. GİRİŞ

Bilgi çağının gelişiminde en büyük rolü oynayan internet kullanımının artması ile bilişim teknolojileri ve mobil cihazların gelişimi paralel olarak hız kazanmıştır. Bu sayede dijital verinin oluşması ve yayılma hızı saniyenin binde biri ile ifade edilen değerlerle artarak dijital kültürün değişim ve gelişimine olanak tanımaktadır.

(Chatti vd., 2014: www.nwlink.com/~donclark/performance/understanding.html.30.01.2021). Ayrıca 2019 yılı ile tüm dünyada hızla yayılan korona virüs salgını süreci nedeniyle günlük yaşantımız dahil olmak üzere bütün alışkanlıklarımız değişime uğramıştır. Bu değişimler iş ve eğitim hayatında da hızlı ve mecburi değişimler yaşanmasına neden olmuştur. Bu değişimler temasın en aza indirgenebildiği internet teknolojileri destekli bilişim teknolojileri yardımı ile iletişim kurularak çözülmesi amaçlanmıştır. Eğitimde ise bu durum tüm dünyada çevrimiçi eğitim yöntemi ile aşılmaya çalışılmıştır.

1991 yılından bugüne hem iş dünyası hem de günlük yaşam alışkanlıklarımız oldukça farklılaşmıştır. Özellikle şirketlerin ve ülkelerin rekabet gücünü elde etme eğilimleri eğitilmiş işgücüne sahip olmaları ile doğru orantılı olarak artmaya devam etmektedir. Bu durum bilgi çağında gelişim olarak adlandırılmanın yanında “Öğrenmek kazandırır” sloganı ile açıklanmıştır (Trilling ve Fadel,2009:8). Eğitim, toplumun gelişimini dört temel aşamada bağlayıcı rol oynamaktadır. Eğitim bireyin işine ve kendisine katkıda bulunmasını, kişisel gelişimini ve yeteneklerini kullanmasını, sorumluluklarını yerine getirmesini, toplumun gelenekleri ve değerlerini ileriye taşımasını sağlar (Trilling ve Fadel,2009:12). Bu nedenle eğitimsel veri analizinin eğitim alanında kullanımının yaygınlaşması hem toplum hem de bireyin gelişimi için önem arz ettiği düşünülmektedir. Bu nedenle yapılan araştırmada büyük veri teknolojisi ve bu teknolojilerin eğitimde kullanıma biçimleri olan eğitimsel veri analiz metotları incelenmiştir.

2. BÜYÜK VERİ

İnternetin yaygınlaşmasıyla artan sosyal medya kullanımı da veri hacminin artış hızına oldukça uygun bir örnek olarak gösterilebilir. 2017 yılında yapılan bir araştırmaya göre kullanıcılara ücretsiz olarak hizmet veren Flickr sosyal medya sitesinde günde ortalama 1,63 milyon fotoğraf paylaşılmaktadır (Wikimedia Commons, 2021). Sosyal medya dışında, çevrimiçi bankacılık, elektronik mail, sensörler, internet üzerinden yapılan canlı yayınlar gibi birçok alanda kullanılmaktadır (Zikopoulos ve Eaton, 2011:4; Begoli ve Horey, 2012:1). Ayrıca sosyal medyada paylaşılan her bir fotoğrafın ortalama boyutunun 2 MB olduğu kabul edilirse günde ortalama 3.6 TB’lık bir veri akışının olduğu sonucuna ulaşılabilir ve elde edilen bu sonuç sadece Flickr sosyal medya ağı için geçerlidir. Bu nedenle büyük veri yüksek hızda değişen büyük hacimli veriler olarak ifade edilmektedir (Gosain ve Chugh, 2014:1). Dünya üzerinde ise yaklaşık 2.5 kentilyon bayt veri akışı gerçekleşmektedir (Dobre ve Xhafa, 2014:1).

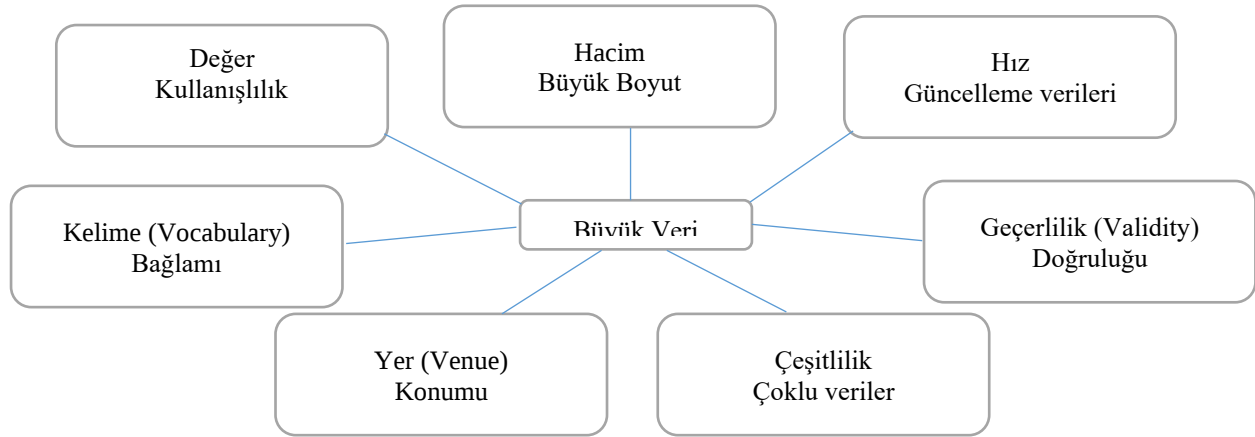
1990’ların başından bugüne kadar iş dünyasında kuruluşlar büyük hacimli verileri depolayarak ve analiz ederek, stratejik ve operasyonel yönleri hakkında daha iyi kararlar almaya çalışmaktadırlar. Ancak kullanılan bu verilerin doğası her geçen gün ile birlikte değişime uğrayarak veri yönetimi ve analizi konusunda karmaşıklığa neden olmaktadır (Daniel, 2014:4). Benzer şekilde yükseköğretim kurumları da öğrenen ve öğrenme sürecine katkı sağlamak için öğrenme analitiklerini kullanmaktadır. Çünkü yükseköğretim kurumlarında giderek artan öğrenci sayısı, nitelikli öğrenci yetiştirme ihtiyacı, öğretim kalitesinin uluslararası düzeyde olmasını sağlamak gibi artan ihtiyaçların ekonomik, sosyal ve politik değişimlere yanıt vermesi beklenmektedir (Daniel, 2014:1; Bozkurt, 2016:56).

Büyük veri; yapılan araştırmaların yaklaşımları doğrultusunda, geleneksel yöntemlerle işlenemeyecek kadar fazla verinin iş zekası, makine öğrenmesi, analitik gibi farklı yöntemler kullanılarak işlenmesidir (Hurvitz vd., 2013:23). Dolayısıyla veri istatistikleri, büyük veri analizinin yalnızca bir parçasıdır; büyük veri analizi, verileri istatistiksel olarak hesaplamak, mevcut büyük miktardaki veri işlemeye dayalı olarak gerçek ihtiyaçları karşılamak ve yüksek düzeyli analitik veriler çıkarmak için çeşitli algoritmaların kullanılmasıdır. Büyük veri, çok büyük (çoklu bilgi ve büyük veri hacmi), yüksek hız (hızlı bilgi işleme), çeşitlilik (çeşitli veri türleri) ve kullanım değeri (toplanan veri bilgilerinin kullanım değeri) ile karakterize edilmektedir (Pan, 2019:221).

Büyük Verinin ilk tanımı, Meta Group tarafından "3V" adı verilen üç özelliğini açıklayarak geliştirilmiştir. 3V: Hacim (volume), Hız (velocity), ve Çeşitlilik (variety)’tir. Veri kalitesine dayalı olarak, IBM, doğruluk (veracity) tanımını ekleyerek büyük veri tanımını 4V üzerinden açıklamıştır. Aynı şekilde Oracle, ilk yapılan tanıma atf yaparak tanıma bir başlık ekleyerek büyük verinin katma değerini vurgulayan tanıma eklediği Değer (value) adlı dördüncü bir V ile kendi 4V tanımını oluşturmuştur (Baaziz ve Quoniam, 2014:2). Literatürde büyük verinin bileşenleri konusunda fikir birliği sağlanamamıştır. Büyük veriyi oluşturan altı bileşen; çeşitlilik (variety), hız (velocity), verinin hacmi (volume), doğrulama (verification), doğruluk (veracity) ve değer (value) dir (Normandeau, 2013, <https://insidebigdata.com/er.trh.12.12.2020>).

Bu V’ lerin içeriğine bakacak olursak;

- ✓ Çeşitlilik (variety): farklı türden verilerin çapraz referans ile birleştirilerek tam olarak yararlanılması,
- ✓ Hız (velocity): verinin üretilebileceği hız,
- ✓ Verinin hacmi (volume): veri kümesinin boyutu,
- ✓ Doğruluk (veracity): toplanan verinin eksiklik ya da belirsizliği,
- ✓ Değişkenlik (volatility): toplanan ve çıkarılan verinin zamanla daha az alakalı hale gelmesi (Poulovassilis, 2016:1).



Şekil 1. Büyük Verinin Özellikleri (Bamiah vd., 2018:130)

3. EĞİTİMDE BÜYÜK VERİ KULLANIMINA GEÇİŞ: EĞİTİM 4.0

Üretim ve endüstri alanında yapılan gelişmeler ve ilerleyen teknoloji endüstri alanında belirleyici rol oynamaktadır (Öztemel, 2018:26). Teknolojinin gelişmesi ile beraber üretim yöntemleri değişmiş ve daha az maliyetle daha çok ürün üretilmesi, hızlı üretim, insansız üretim, insan makine etkileşimi gibi konular endüstriye dahil olmuştur (Brettel vd., 2014:38; Saurabh vd., 2018:234). Eğitimde de endüstri 4.0 da görülen gelişmelere benzer durumlar ile karşılaşmaktadır (Wallner ve Wagner, 2016:156).

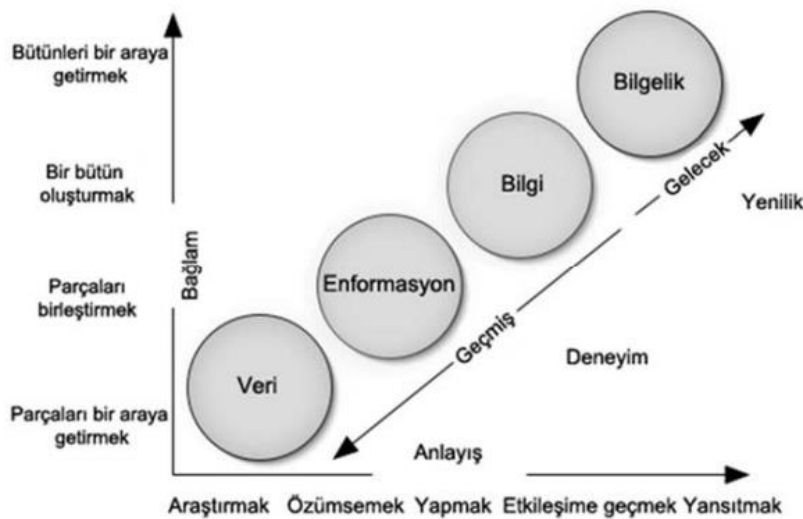
Eğitim 1.0 tarım toplumu insanların eğitim ihtiyacını karşılamıştır. Burada öğrenci öğretmenin yaptığı şeyler üzerinden izleyerek öğrenmiştir (Puncreobutr, 2016:93).

Eğitim 2.0'da ise öğrenci gelişen sanayi toplumuna çözüm üretmek için öğrenmeyi amaçlamış ve teknolojiyi bir öğrenme aracı olarak kullanmıştır (Puncreobutr, 2016:93; Harkins, 2008:2-3).

Eğitim 3.0'da teknoloji artık bireylerin gelişmesi ve bilgiyi yeniden üretmesi için bir araç olmuştur. Teknolojinin kullanılması ile birlikte bilgiye ulaşma hızı artmış ve sanal öğrenme ortamı kavramı ortaya çıkmıştır (Harkins, 2008:2-3).

Eğitim 4.0'da bireyler endüstri 4.0'a benzer şekilde dinamik şartlar altında çalışabilecek ve yeni girişimler yapabilecek kişiler olması amaçlanmaktadır (Harkins, 2008:2-3).

Dünyadaki teknolojik gelişmeler ile üretim ve eğitim konuları da hızla gelişmiş ve artık eğitim, eğitim 4.0 endüstri ise endüstri 4.0 adını almıştır.



Şekil 2. Veri, Enformasyon, Bilgi Ve Bilgelik Kavramları Arasındaki İlişki
(Clark, www.nwlink.com/~donclark/performance/understanding.html, 30.01.2021).

Veriler araştırma, oluşturma, toplama ve keşif yoluyla elde edilen bilgi parçacıdır. Verilerin düzenlenmesi ile bilginin oluşturulması kolayca sonuçlara varabilmemizi sağlar. Aynı zamanda görsel veya işitsel verilerin birleştirilerek bilgiye dönüştürülür. Ancak bilgi, farklı bakış açıları ile deneyimlenmesi ile deneyim karmaşıklığı ortaya çıkar. Bir kişinin bilgisinin diğerine aktarılması sırasında eğitimi alanın bakış açısını anlamak zordur. Bu nedenle eğitim ve öğretimin başarısını anlamak güçtür. Bilgelik, anlayışın nihai yani son seviyesidir. Bilgide olduğu gibi, bilgelikte insanın iç görüşü ve deneyimleri ile alakalıdır (Clark,2021,www.nwlink.com/~donclark/performance/understanding.html.30.01.2021).

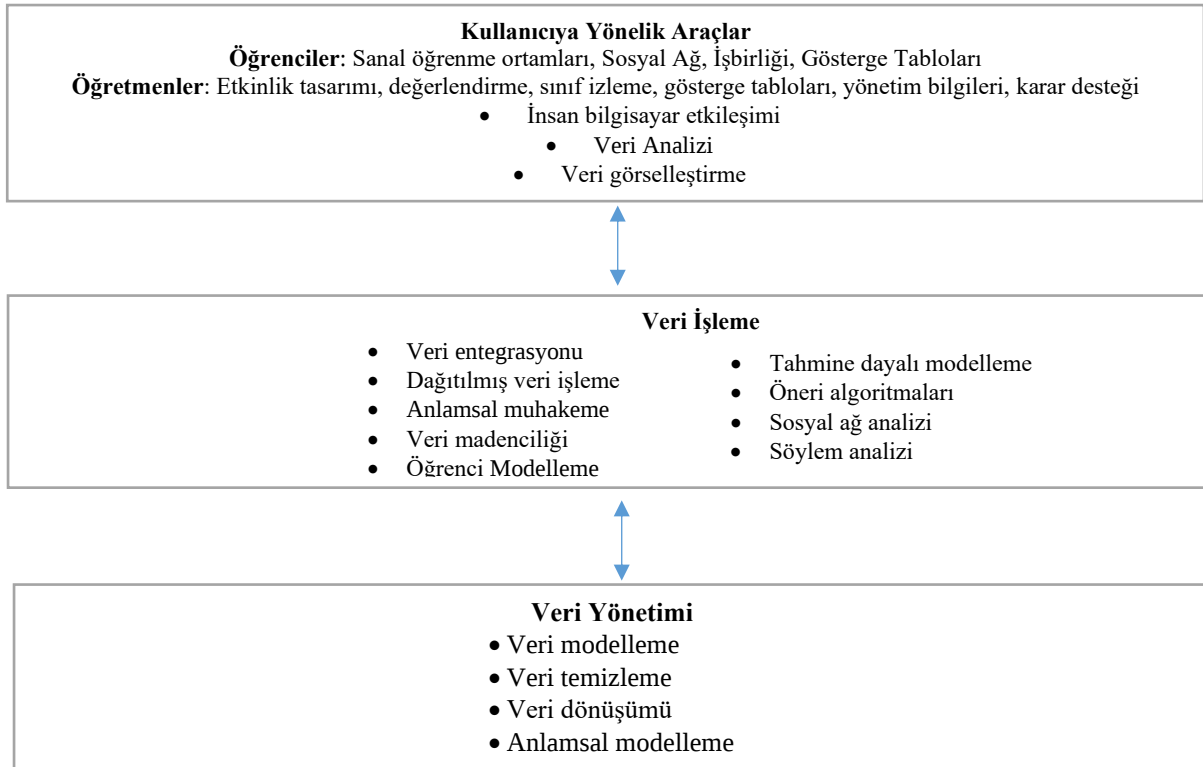
Şekil 2’de ise verinin bilgelige yolculuğu aşamalı olarak gösterilmektedir. İşte tam burada öğrenme analitiğinin gücü devreye girmektedir. Öğrenenin sağladığı veriler yoluyla mümkün olan en doğru sonuca ulaşılması ve öğrenenin yönlendirilmesinde temel taşlardır diyebiliriz.

4. EĞİTİMDE BÜYÜK VERİ ANALİZ YÖNTEMLERİ

4.1. Öğrenme Analizi

Büyük verinin oluşması ile birlikte veriyi işleme yöntemleri de yenilik kazanmıştır. Çünkü elde edilen verinin büyüklüğü bu verinin geleneksel yöntemler ile işlenmesine olanak sağlamamaktadır. Dolayısıyla veri işleme ve analiz etme yöntemleri de yenilik kazanmıştır. Büyük verinin ortaya çıkması ve gelişimiyle birlikte gelen bu verinin işlenmesi ve farklı alanlarda kullanılabilmesi gerekliliği ortaya çıkmıştır. Bu gereklilik doğrultusunda büyük verinin kullanılmasını sağlayacak yeni analitik yaklaşımlar doğmuştur. Kullanımı yaygınlaşmakta olan büyük verinin işlenmesinde soyut (soft) ve somut (hard) teknolojilerce anlamlandırılabilir olması ve bu veriden bir sonuca ulaşılabilir olmasının oldukça önemlidir. Verinin toplanması, düzenlenmesi ve analiz edilmesi sürecinde kullanılan yazılımlar soyut teknolojilere örnek gösterilebilir. Ayrıca elde edilen verinin ihtiyaç duyulan alanda kullanılabilir hale getirilmesi sürecinde kullanılan bilişim teknolojileri somut teknolojilere örnek gösterilmektedir. Günümüzde bahsi geçen bu soyut ve somut sistemler ile her geçen saniye veriler toplanmakta ve analiz edilip anlamlandırılmaktadır. Bir anlama sahip olan bu veri raporlandırılıp ihtiyaç duyulan alanlarda kullanılmaktadır. Öğrenme analitiği ise eğitim süreçlerinde kullanılan soyut ve somut teknolojilerin kullanılarak büyük verinin eğitim alanında verimli kullanılma amacıyla işlenmesidir (Bozkurt, 2016:60).

Şekil 3’de bilgi işlem tekniklerinin kullanabileceği katmanları, öğrenme süreçleri ve öğrencilerle ilgili verileri yöneterek kullanmak için genel bir mimari modellenmiştir (Poulovassilis, 2016:3).



Şekil 3: Veri Mimarisini ve Hesaplama Teknikleri (Poulovassilis, 2016:3)

Öğrenme sistemlerinden sağlıklı büyük veri alabilmek için hem sistem içerisinde hem de kullanıcı yani öğrenci ve öğretmenden gelen verilerin sağlıklı olması beklenir. Bunun için kurulan sistemde dikkat edilmesi gereken önemli noktalar bulunmaktadır. Öncelikle, öğrencilerin sistem deneyimleri oluştuktan sonra bir ölçme aracı yardımıyla veriler toplanmalıdır. Bu veriler ışığında verilen derslerin sırası, öğrenim modelinin ne kadar faydalı olduğu gibi sorulara cevap bulunarak, öğrenci akademik performansının artması sağlanırken, öğretmen etkililiği artırılabilir (Daniel, 2014:12). Bunun yanında öğrenci bilgi modeli oluşturulması gerekmektedir. Bu model öğrencinin sisteme dair sorduğu sorular, yaşadığı zorluklar gibi verileri içermelidir. Bunun yanında öğrenci dosyaları oluşturularak kişiselleştirilmiş öğrenme ortamı sağlanabilir. Bu dosyalarda temel öğrenme bilgileri, temel bilgi verileri, aynı öğrenme özelliklerine sahip öğrencilerin verileri, veri madenciliği ile öğrencilerin öğrenme özelliklerine göre gruplanarak, öğrenme analizi ve makine öğrenimi algoritmaları ile analiz edilerek kişiselleştirilmiş öğrenme ortamları sunulabilir (Li ve Zhai, 2018:342; Lukarov vd., 2014:93). Son olarak, öğrenci davranış modeli; öğrencilerin öğrenme davranışlarını gözlemlemek, dersi tamamlama, öğrenme süresi, performans puanları gibi veriler incelenerek oluşturulmalıdır (Eynon, 2013:238). Aşamaları belirtilen sistemlerden eğitim alanında kullanılan en yaygın olanı öğrenme analizi veya analitiği olarak tanımlanmaktadır.

Geleneksel eğitim siteminde kullanılan öğrenci bilgi sistemi, öğrencinin akademik verileri, demografik bilgileri gibi temel bilgileri depolamaktadır. Bu kısıtlı veri ile analiz yapılabilirken günümüzde kullanılan eğitim yönetim sistemleri ile öğrencinin ihtiyaç duyduğu materyal, gönderdiği ödevlerin yönetimi gibi verilerin desteği ile artık öğrencinin dijital ayak izlerine ulaşılabilir ve bu büyük miktarda veri sağlayabilir. Bu durum dijital öğrenme ortamlarının sağladığı yenilikler ve zenginleştirilebilir eğitim olanakları olarak görülmektedir (Fisher vd., 2020:131). Teknolojik gelişmelerin, günümüz dünyasında üretilen bilginin ve çevrimiçi ortamlarda büyük veriyi oluşturan dijital izlerin sürekli bir biçimde arttığı günümüzde öğrenme analitiğinin daha sık kullanılacağı düşünülmektedir (Bozkurt, 2016:74).

En genel tanımıyla öğrenme analizi “öğrenmeyi ve bunun gerçekleştiği ortamları anlamak ve optimize etmek amacıyla öğrenciler ve bağlamları hakkındaki verilerin ölçülmesi, toplanması, analizi ve raporlanmasıdır”(Solaresearch, 2021). Büyük veri teknolojisinin eğitim üzerinde özelleştirildiği alan birçok araştırmacı tarafından öğrenme analitiği veya analizi olarak tanımlanmaktadır. Ancak öğrenme analizi veya analitiği tanımı çok net ve belirgin şekilde yapılmamıştır. Buna rağmen öğrenme analitiğinin avantajları ve dezavantajları konusunda farklı düşünceler mevcuttur.

Öğrenme analitiği, büyük bir toplum için algoritmik genişlemenin geçişinin spesifik olarak şekillenmiş halidir, ayrıca eğitim üzerindeki geniş etkileri toplumlara yön verebilir (Knight ve Shu, 2017:17). Bunun yanında öğrenme analitiği, veri analizi sırasında gecikmeyi en aza indirgeyerek, öğrenci ve öğretmenler için kurs veya bireysel düzeyde yardımcı olacak verilerin seçilmesi ve işlenmesinin yanında kişiselleştirilmiş hizmetler oluşturulmasına odaklanır (Elias, 2011:4). Bu nedenle büyük veri işleme modellerinin eğitim alanında daraltılarak analizi (öğrenme analitiği) sonucu öğrencilerin öğrenme yöntemleri, görsel bilgi kaynakları, öğrenme birimleri, kurs ya da sınıf içerisinde davranışlar incelenerek veri toplanır. Bu sayede öğrencinin akademik başarısı ve öğretmenin öğretim yeterliliği önemli ölçüde artırılabilir (Li ve Zhai, 2018:342). Ayrıca öğrenme analitiği, eğitimin teknolojik bir sisteme dönüştürülmesi yoluyla öğrenci ve öğretmenleri marjinalleştirme, “öğrenmek” tanımının altında bahsedilenlerin sınıflandırılmasına alternatifler oluşturabilir (Knight ve Shu, 2017:19).

Avantajları olduğu gibi büyük veri teknolojisinin eğitimde kullanılmasının getirdiği dezavantajları da vardır. İnternet üzerinde bulunan her sistem için risk teşkil eden güvenlik, gizlilik, etik, uzman eksikliği, veri depolama, işleme ve verilerin çeşitliliğinin oluşturduğu birlikte analizlerinden elde edilen verilerin güvenilirliği gibi birçok dezavantaj oluşturabilecek durum söz konusudur (Bamiah vd., 2018:234-235). Örneğin, eğitim veri setlerinin uyumsuzluğu veya algoritma ya da teknolojik yaklaşımların yeterliliği gibi teknik sorunlar veri analizinin etkililiği, açıklığı, yorumlama gücü gibi sorunlar yaratırken suistimalin önünü açmaktadır. Bu nedenle öğrenme analizinden doğru faydalanmanın altı kritik boyutu şekilde belirtilmiştir (Drachsler ve Greller, 2012:122).



Şekil 4. Öğrenme Analitiği Çerçevesi (Drachslar ve Greller, 2012:122).

Paydaşlar: Öğrenme analitiğinden yaralananlar ve katkıda bulunanlardır. Paydaş boyutunda veri sağlayıcı ve veri analiz sonucunu kullanan aynı özne olabilir (Drachslar ve Greller, 2012:121). Örneğin bir öğretmen veri sağlayıcı durumundayken aynı zamanda öğrencilere dair işlenmiş veriyi kullanabilir.

Enstrümanlar (Yöntem): Şekil 4'te belirtildiği gibi öğrenme analitiğinin uygulandığı teknolojiler, algoritmalar, teori ve diğer yorumlama birimlerini (eğitimsel veri madenciliği veya klasik istatistiksel teknikler) temsil eder (Drachslar ve Greller, 2012:122).

İç Sınırlılıklar (Yetenekler): Öğrenme analitiği tek başına yeterli sonuç vermeyecektir. Sonuçların anlaşılır ve yorumlanabilir olması için algoritmalarla faydalanmak gerekir (Reffay ve Chanier, 2003:343). Bu nedenle öğrenme analizi sonuçlarını yorumlamak için üst düzey yeterliliklere ihtiyaç vardır. Paydaşların bu yeterliliklere sahip olmaması iç sınırlılıklara yol açacaktır (Drachslar ve Greller, 2012:122).

Dış Kısıtlamalar: Eğitimde öğrenme analizi uygulanırken yeni etik ve güvenlik sorunları meydana gelmektedir (Hildebrandt, 2006:43). Verinin kullanımı sırasında doğabilecek güvenlik tehditlerine yönelik politikalar, verinin korunumu kanununa ek konular eklenmelidir (Drachslar ve Greller, 2012:122).

Amaçlar (Hedefler): Ulaşmak istenen hedeflerin belirlenmesi gerekliliğini tanımlar. Ayrıca bilgi akışını, sosyal etkileşimleri izleyerek karşılaştırılması yoluyla örgütsel etkililiği ve verimliliği artırabilir (Govaerts vd., 2010:93). Hedefler, eğitim sürecinin evrimine katkıda bulunmak için her zaman eğitim ufuklarının genişletilmesine ve ilerlemesine yöneliktir (Rodrigues vd., 2018:1704).

Veri: Eğitim veri kümelerinin oluşturuldukları ve paylaşıldıkları ortamları tanımlar. Öğrenme analitiği ise farklı eğitim sistemlerinden elde edilen veri kümelerinden yararlanır. Kurumlar öğrenci ilerlemesini gözlemlemek, yönetmek ve belirli fonlardan faydalanmak amacıyla rapor vermek en yaygın analiz sonuçları kullanım yöntemleridir. Ayrıca eğitim veri kümelerinin birbirine bağlanması, öğrenci odaklı eğitim hizmetlerine ve dolayısıyla eğitimi kişiselleştirmeye yönelik karma uygulamaların geliştirilmesini sağlayacaktır (Moon vd., 2020:16).

Dijital çağda internet ile iletişim teknolojilerinin devrimi öğretmenlerin öğrencilerine hızlı ve kişiselleştirilmiş öğrenme gibi birçok fırsat tanıyabilirler. Bunun yanı sıra öğreneninde bilgiye erişimde herhangi bir sorun yaşamaması bir avantajdır ancak bu durum öğrenen açısından başka bir kafa karışıklığına neden olmaktadır. İnternet üzerinde bulunan devasa bilgi yığnında ulaşılan bilgi ne kadar

güvenli veya hangi yöntemi kullanarak bilgiye en hızlı ve doğru yolla ulaşılır, gibi sorular öğrencinin kafasını hayli karıştıracaktır. Tam olarak burada öğrenme analizi ile elde edilmiş veriler öğretmene öğrencinin öğrenme stili, uygun materyal seçimi gibi öğrenciye özgü bilgi vererek öğrencinin motivasyon kaybetmeden devam edebilmesi için destek olacaktır (Hamada ve Nishikawa, 2010:101).

4.2. Eğitimsel Veri Analizi

1960'lı yıllarda başlayan veri kaydetme yöntemleri sonucu 1980'lerde bu kaydedilen bilgilere ulaşmayı kolaylaştıran veri tabanları geliştirilmiştir. Bu verilerin çoğalıp birikmesi ile verinin anlamlı bilgiye dönüştürülmesi önem kazanmıştır. Veri madenciliği, tanımlayıcı ve tahmin edici modeller meydana getiren veri işleme yöntemlerinden birisidir (Timor ve Künc, 2021:2). Veri madenciliği, çevrimiçi ticarete para miktarı, müşteri sayısı ve sadakati ile ölçülebilen karı artırmakken, eğitimde öğrenmeyi geliştirmeyi amaçlamaktadır (Romero ve Ventura, 2007:136). Eğitimsel veri madenciliği; eğitim ortamlarından gelen benzersiz veri türleri içinde keşifler yapmaya yönelik yöntemlerin geliştirilmesi ve bu yöntemleri öğrencilerin daha iyi öğrenmesi için kullanımına odaklanan bilimsel araştırma alanı olarak tanımlanmaktadır (Baker, 2010:112).

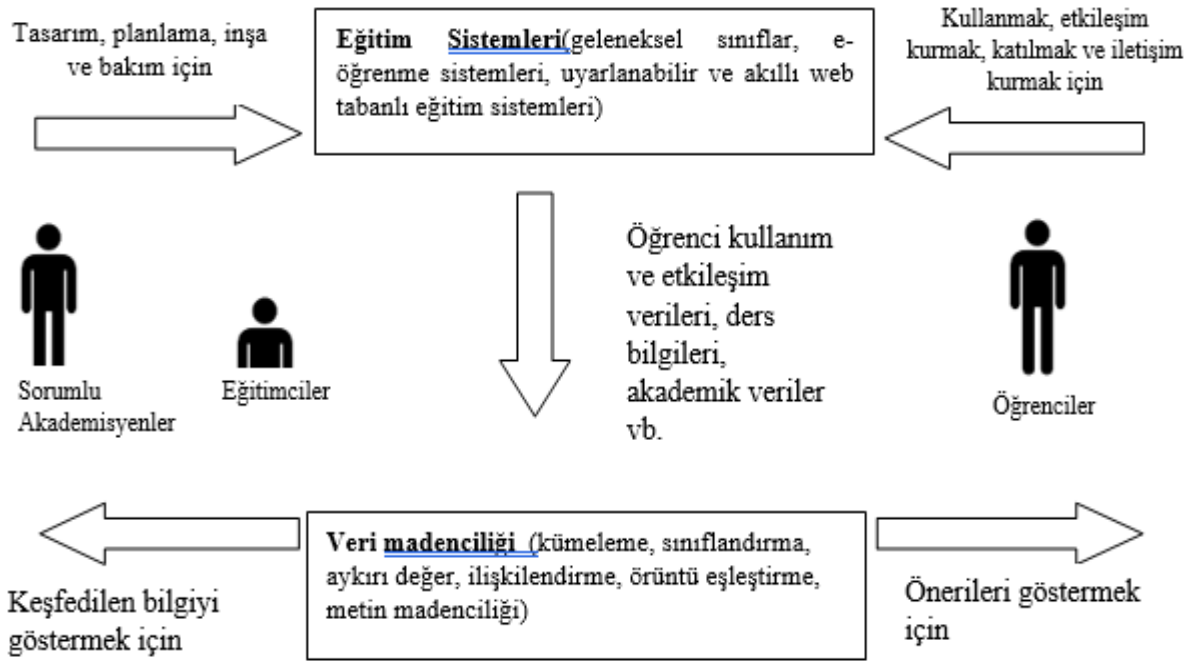
Eğitimsel veri madenciliği (EDM), öğrenmenin kardeş topluluğu olarak görülmektedir. Bu topluluk 2005'te başlayan ve 2008'de her yıl düzenlenen konferans haline gelen, 2011 yılında ise uluslararası eğitimsel veri madenciliği sosyal topluluğu isimli bir atölye kurulması ile her gün gelişen bir topluluk haline gelmiştir (Baker ve Inventado, 2014:61). Bu süreç detaylı olarak Şekil 5'te incelenmiştir.

• 1995	Corbett&Anderson, Bayesian Bilgi İzleme Modeli – Günümüzde hala geçerliliğini koruyan ilk eğitimsel veri madenciliği algoritması
• 2000	İlk eğitimsel veri madenciliği ile ilgili atölye çalışması
• 2001	Zaiane eğitimsel veri madenciliği yöntemlerinin potansiyeli hakkında teorik makale
• 2005	"Eğitim amaçlı veri madenciliği" terimini kullanan ilk atölye çalışması
• 2006	Eğitimsel veri madenciliği hakkında yayınlanan ilk kitap: "E-öğrenmede veri madenciliği", Romero ve Ventura
• 2008	Eğitimsel Veri Madenciliği üzerine ilk uluslararası konferans
• 2009	Eğitimsel Veri Madenciliği Dergisi ilk sayısını yayınladı
• 2010	Eğitimsel Veri Madenciliği ile ilgili ilk el kitabı yayımlandı; Romero, Ventura, Pechenizkiy ve Baker
• 2011	İlk Öğrenme Analitiği ve Bilgi Konferansı düzenlendi
• 2011	IEDMS kuruldu (Entegre Elektronik Veri Yönetim Sistemi)
• 2012	SoLAR kurdu
• 2013	İlk öğrenme analizi yaz enstitüsü kuruldu
• 2022	Tüm eğitim araştırmaları analitik ve veri madenciliğini içermeye başlayacaktır.

Şekil 5: Eğitimsel Veri Madenciliğindeki (EDM) Önemli Kilometre Taşı Çalışmalarının Zaman Çizelgesi (Baker ve Inventado, 2014:62)

Şekil 6' da görebileceğimiz gibi, sorumlu eğitimciler ve akademisyenler eğitim sistemlerini tasarlamaktan, planlamaktan, inşa etmekten ve sürdürmekten sorumludur. Öğrenciler kurulan sistemi kullanır ve sorumlularla etkileşimde bulunurlar. Kurslar, öğrencilerin kullanımı ve etkileşimiyle ilgili mevcut tüm bilgilerden yararlanarak, e-öğrenme sürecini iyileştirmeye yardımcı olan bilgileri keşfetmek için farklı veri madenciliği teknikleri uygulayabilirler. Keşfedilen bilgi sadece sağlayıcılar (eğitimciler) tarafından değil aynı zamanda kendi kullanıcıları (öğrenciler) tarafından da kullanılabilir. Bu nedenle, eğitim sistemlerinde

veri madenciliği uygulaması, her bir bakış açısı ile farklı aktörlere yönlendirilebilir (Zorrilla vd.,2005:206-207).



Şekil 6: Eğitim Sistemlerinde Veri Madenciliği Uygulama Döngüsü (Romero ve Ventura, 2007:136).

Eğitimsel veri analizinin temel amacı; öğrencileri ve öğrenmeyi daha iyi anlamak için veri havuzlarını kullanarak öğrencilerin yararına olacak şekilde uygulamaya dönüştürmek için bu verileri teori ile birleştiren hesaplama yaklaşımları geliştirmektir (Romero ve Ventura, 2010:601).

Mevcut veri madenciliği araçlarının çoğu, eğitimcilerin kullanamayacağı kadar karmaşıktır (Romero ve Ventura, 2007:140) ve özellikleri bir eğitimcinin ihtiyaç duyabileceği kapsamın çok ötesindedir. Bu nedenle eğitim yönetim sistemleri yöneticisi tarafından analizlerin gerçekleştirilmesi gerekmektedir. Yöneticilerin uyguladığı veri madenciliği teknikleri sonucu elde ettiği raporların eğitimciler için öğrencinin öğrenmesini ve çevrimiçi eğitimi nasıl iyileştireceklerine ilişkin kararlar almada yardımcı olmalıdır. Bu nedenle Eğitimsel Veri Madenciliği, eğitim bağlamından gelen benzersiz veri türlerini keşfetmek için yöntemler geliştirmeyle ilgilenen yeni bir disiplindir (Romero vd., 2008:369).

Veri madenciliği tekniklerinin uygulandığı ana e-öğrenme problemlerinden veya konularından bazıları öğrencinin öğrenme performansının değerlendirilmesiyle ilgileniyor, ders uyarlaması ve Öğrencilerin öğrenme davranışı, öğrenme materyallerinin değerlendirilmesi ve web tabanlı eğitim kursları, e-öğrenme kurslarının hem öğretmenlerine hem de öğrencilerine geri bildirim sağlar ve tipik olmayan öğrencilerin öğrenme davranışını tespit eder (Romero vd., 2008:369).

Romero ve Ventura (2007), eğitimsel veri madenciliği alanındaki çalışmaları kategorilere ayırmaktadır: *İstatistik ve görselleştirme*, öğrencilerden elde edilen öğrenme verileri eğitimsel veri madenciliğinin başlangıç noktası olsa da tek başına yeterli değildir. Bu yaklaşımda veri madenciliği, aksine önce hipotezler oluşturulur sonra bu hipotezler test edilir. Bu nedenle keşif odaklı bir iştir diyebiliriz. Ayrıca eğitim sistemlerine uygulanan farklı *web madenciliği* teknikleri vardır, ancak hemen hemen hepsi sonraki üç taneden birinde gruplandırılabilir: *kümeleme*, *sınıflandırma*, *aykırı değer tespiti*; *ilişki kuralı madenciliği*, *sıralı model madenciliği*; *metin madenciliği*dir (Romero ve Ventura, 2007:140). Bununla birlikte, sıralı model madenciliği (Sequential pattern mining), metin madenciliği (text mining), aykırı değer analizi (outlier analysis) ve sosyal ağ analizi (social network analysis) gibi çevrimiçi öğrenmede de kullanılan başka veri madenciliği teknikleri de vardır (Romero vd., 2008:380).

Baker (2010) eğitsel veri madenciliği üzerine farklı bir bakış açısı olarak aşağıdaki şekilde sınıflandırmıştır.

1. Tahmin
 - ✓ Sınıflandırma
 - ✓ Regresyon
 - ✓ Yoğunluk tahmini
2. Kümeleme
3. İlişki madenciliği
 - ✓ Birlik kuralı madenciliği
 - ✓ Korelasyon madenciliği
 - ✓ Sıralı model madenciliği
 - ✓ Nedensel veri madenciliği
4. İnsan yargısı için verilerin damıtılması
5. Modellerle keşif

Tahmin: Amaç, belirlenen tahmin değişkenlerinin (verilerin diğer yönlerinin) bazı kombinasyonlardan yola çıkılarak verilerin tek bir yönünü yani tahmin edilen asıl değişkeni çıkarabilen bir model geliştirmektir (Baker, 2010:115). Ayrıca E-öğrenme sistemlerini kullanırken öğrencilerin davranışlarının ve performansının tahmin edilmesi, genel olarak e-öğrenme ortamlarının yanı sıra sanal derslerin de iyileştirilmesini kolaylaştırma potansiyeline sahiptir (Castro vd.,2007:192).

Kümeleme: Bu modelin amacı veri kümesinin daha önceden net olarak kategorilere ayrılmadığı durumlarda analiz sonucu doğal olarak bir araya gelen verileri kümeleyerek kategorilere ayırmaktır (Amershi ve Conati, 2006:466).

İlişki madenciliğinde temel amaç çok sayıda değişken arasında yapılan analiz sonuçlarında aralarında en güçlü ilişki bulunan değişkenleri tespit etmektir. İlişki madenciliğinin genel olarak kabul edilen dört türü bulunmaktadır. Bunlar: ilişki kuralı madenciliği, korelasyon madenciliği, sıralı model madenciliği ve nedensel veri madenciliğidir(Baker, 2010:117).

İnsan yargısı için verilerin damıtılması: Veri analizi sonrası elde edilen sonuçlar sunulduğunda önemli çıkarımların yanında farklı ve istenmeyen yorumların yapılması mümkündür. Bu durumun önüne geçebilmek amacıyla veri analizi sonrası sunulan verilerin görselleştirilerek anlam kazandırılması olayıdır (Baker, 2010:117-118).

Dördüncü kategori ise evrensel olarak veri madenciliği olarak görülmesi de, Romero ve Ventura, (2007)'nın *istatistik ve görselleştirme* kategorisiyle uyumludur (Baker ve Yacef,2009:4).

Modellerle Keşif: Bir modelle keşifte, tahmin, kümeleme veya bazı durumlarda veri mühendisliği yoluyla bir model geliştirilir. Bu model daha sonra tahmin veya ilişki madenciliği gibi başka bir analizde bir bileşen olarak kullanılır. Tahmin durumunda, oluşturulan modelin tahminleri, yeni bir değişkeni tahmin etmede yordayıcı değişkenler olarak kullanılır. İlişki madenciliği durumunda ise oluşturulan modelin tahminleri ile ek değişkenler arasındaki ilişkiler incelenir. Bu, bir araştırmacının karmaşık ve gizli bir yapı ile çok çeşitli gözlemlenebilir yapılar arasındaki ilişkiyi incelemesini sağlayabilir (Baker, 2010:117). Ayrıca bu veri analiz yöntemi Romero ve Ventura, (2007)'nın web madenciliği kategorisi ile ilişkili görünmektedir.

5. SONUÇ ve ÖNERİLER

Çalışma temel olarak eğitimde büyük veri teknolojisinin kullanım yöntemlerini irdelemiştir. Eğitimde veri analizi yöntemleri adına birçok farklı yöntem araştırmacılar tarafından öne sürüldüğü görülmektedir. Bu durum temelde ihtiyaç duyulan sonuca odaklı olarak değişmektedir. İstenilen sonuç değişkeninin belirlenmesi yöntemin belirlenmesinde ilk adım olacaktır. Veri analizi öncesi öğrencinin akademik başarı durumunu etkileyen değişkenlerin tahmini, yaşanan coğrafyanın öğrenen üzerinde etkisi, derslerin işleniş sırasının eğitim kalitesine etkisi gibi birçok farklı değişken arasındaki ilişki veya karşılıklı ilişkileri

irdelenebilir. Bu nedenle daha önce önemle bahsettiğimiz öncelikle veri analizi sonrası elde edilmek istenen bilginin ne olduğuna karar verilmelidir.

Eğitim ülkeler ve toplumlar açısından yaşadığımız yüzyılın temel yapı taşı olduğunu söylemek mümkündür. Eğitim seviyesi düşük toplumların gelişimi de o seviyede düşük olduğunu gözlemlemek zor olmasa gerek. Bu nedenle çağımız bilgi çağı olarak adlandırılıyor. Bilgi çağında gelişimi yakından takip ederek gelişimi yönlendirebilmek amacını taşıyan ülke ve toplumların eğitimsel analiz yöntemlerini günümüzde aktif ve faydalı bir şekilde kullanıyor olması beklenir. Çalışma boyunca faydaları ve yöntemleri tartışılan bu veri analiz yöntemleri eğitim kalitesi, öğrenenin öğrenme yöntemleri, eğitimin kendine katması gereken özellikleri vb. birçok eğitimin kalitesini artırma yönünde yapılması gerekenler konusunda net bilgi sağlayabildiği görülmektedir. Bu nedenle özellikle çevrimiçi eğitime mecbur kaldığımız bu dönemde eğitimsel veri analizi yapılabilmesi adına sağlıklı veri toplanarak analiz edilebilir. Bu sayede eğitim sisteminin, eğitimcilerin ve öğrencilerin gelişimi adına atılabilecek adımlar net olarak belirlenebilir.

KAYNAKÇA

Amershi, S.,&Conati, C. (2006, June). Automatic recognition of learnergroups in exploratory learning environments. In International Conference on IntelligentTutoringSystems (pp. 463-472). Springer, Berlin, Heidelberg.

Baaziz, A.,&Quoniam, L. (2014). How to use Big Data Technologies to optimize operations in UpstreamPetroleumIndustry. In 21st World PetroleumCongress, Moscow.

Baker, R. S. J. D. (2010). Data miningforeducation. International encyclopedia of education, 7(3), 112-118.

Baker, R. S.,&Inventado, P. S. (2014). Educational datamining and learning analytics. In Learning analytics (pp. 61-75). Springer, New York, NY.

Baker, R. S.,&Yacef, K. (2009). Thestate of educational datamining in 2009: A review and future visions. JEDM| Journal of Educational Data Mining, 1(1), 3-17.

Bamiah, S. N.,Brohi, S. N., &Rad, B. B. (2018). Bigdata technology in education: Advantages, implementations, andchallenges. Journal of Engineering Science and Technology, 13, 229-241.

Begoli, E.,&Horey, J. (2012, August). Design principles for effective knowledge discovery from big data. In 2012 JointWorking IEEE/IFIP Conference on Software Architecture andEuropean Conference on Software Architecture (pp. 215-218). IEEE.

Bozkurt, A. (2016). Öğrenme analitiği: e-öğrenme, büyük veri ve bireyselleştirilmiş öğrenme. Açıköğretim Uygulamaları ve Araştırmaları Dergisi, 2(4), 55-81.

Brettel, M.,Friederichsen, N., Keller, M., &Rosenberg, M. (2014). How virtualization, decentralization and network building change the manufacturing landscape: An Industry 4.0 Perspective. International journal of mechanical, industrialscienceandengineering, 8(1), 37-44.

Castro, F.,Vellido, A., Nebot, A., &Mugica, F. (2007). Applying datamining techniqueto e-learning problems. InEvolution of teaching and learningparadigms in intelligent environment (pp. 183-221). Springer, Berlin, Heidelberg.

Chatti, M. A.,Lukarov, V., Thüs, H., Muslim, A., Yousef, A. M. F., Wahid, U., ... &Schroeder, U. (2014). Learning analytics: Challenges and future researchdirections. eled, 10(1).

Clark, D. (2021, Ocak 30). Big Dog and Little Dog's Bowl of Biscuits. www.nwlink.com/~donclark/performance/understanding.html adresinden alındı

Daniel, B. (2014). Big Data and analytics in higher education: Opportunities and challenges. British journal of educationaltechnology, 46(5), 2-17.

Dobre, C.,&Xhafa, F. (2014). Intelligent services for bigdata science. Future generation computer systems, 37, 267-281.

Drachsler, H.,&Greller, W. (2012, April). The pulse of learning analytics understandings and expectationsfromthetakeholders. InProceedings of the 2nd internationalconference on learninganalyticsandknowledge (pp. 120-129).



Elias, T. (2011). Learning analytics. Learning, 1-22.

Eynon, R. (2013). Therise of Big Data: whatdoes it mean for education, technology, andmediaresearch?. 38(3),237-240.

Gosain, A.,&Chugh, N. (2014). New Design Principles for Effective Knowledge Discovery from Big Data. International Journal of Computer Applications, 96(17).

Govaerts, S.,Verbert, K., Klerkx, J., &Duval, E. (2010, December). Visualizing activities for self-reflection and awareness. In International conference on web-basedlearning (pp. 91-100). Springer, Berlin, Heidelberg.

Hamada, M.,&Nishikawa, K. (2010, December). Web-BasedEnhanced Learning Style Index with Integration into an e-Learning System. In International Conference on Web-Based Learning (pp. 101-110). Springer, Berlin, Heidelberg.

Harkins, A. M. (2008). Leapfrog principles and practices: Corecomponents of education 3.0 and 4.0. FuturesResearchQuarterly, 24(1), 19-31.

Hildebrandt, M. (2006). Privacyandidentity. Privacyandthecriminallaw, 43. Eriřim adresi: <https://books.google.com.tr/>, Eriřim Tarihi: 12.12.2020.

Knight, S.,&Shum, S. B. (2017). Theory and learning analytics. Handbook of learning analytics, 17-22.

Li, Y.,&Zhai, X. (2018). Review and prospect of modern education usingbigdata. ProcediaComputerScience, 129, 341-347.

Lukarov, V.,Chatti, M. A., Thüs, H., Kia, F. S., Muslim, A., Greven, C., &Schroeder, U. (2014, September). Data Models in Learning Analytics. InDeLFIWorkshops (Vol. 1014, pp. 88-95).

Moon, J., Do, J., Lee, D., &Choi, G. W. (2020). A conceptual framework for teaching computational thinking in personalizedOERs. Smart Learning Environments, 7(1), 1-19.

Normandeau, K. (2013). Beyond Volume, Variety and Velocity is theIssue of Big DataVeracity.<https://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/> Eriřim Tarihi: 12.12.2020.

Öztemel, E. (2018). Eğitimde yeni yönelimlerin değerdendirilmesi ve Eğitim 4.0. Üniversite Arařtırmaları Dergisi, 1(1), 25-30.

Pan, L., (2019). Paper-A Big Data-Based Data Mining Tool for Physical Education and Technica andTactical Analysis. International Journal of Emerging Technologies in Learning (iJET). 14 (22), 220-231.

Poulovassilis, A. (2016). Big Data and education. BirBeck Üniversitesi of London: <http://www.dcs.bbk.ac.uk/oldsite/research/techreps/2016/bbkcs-16-01.pdf>.(23.04. 2019).

Puncreobutr, V. (2016). Education 4.0: New challenge of learning. St. TheresaJournal of Humanities and Social Sciences, 2(2), 92-97.

Reffay, C.,&Chanier, T. (2003). How social network analysis can help tomeasurecohesion in collaborative distance-learning. In Designing forchange in networked learning environments (pp. 343-352). Springer, Dordrecht.

Rodrigues, M. W.,Isotani, S., &Zarate, L. E. (2018). Educational Data Mining: A review of evaluationprocess in the e-learning. TelematicsandInformatics, 35(6), 1701-1717.

Romero, C.,&Ventura, S. (2007). Educational datamining: A surveyfrom 1995 to 2005. Expertswithapplications, 33(1), 135-146.

Romero, C.,&Ventura, S. (2010). Educational datamining: a review of thestate of the art. IEEE Transactions on Systems, Man, andCybernetics, Part C (Applications andReviews), 40(6), 601-618.

Romero, C.,Ventura, S. ve García, E. (2008). Ders yönetim sistemlerinde veri madencilği: Moodle vaka çalışması ve öğretici. Bilgisayarlar ve Eğitim, 51 (1), 368-384.

Saurabh, V., Ambadand, P., & Bhosle, S. (2018). 2nd International Conference on Materials Manufacturing and Design Engineering. (pp. 233-238)

Solaresearch. (2021, Ocak 15). <https://www.solaresearch.org/about/what-is-learning-analytics/> adresinden alındı

Timor, M., & Künç, G. Y. (2021). Ekonomik Gelişmişliği Etkileyen Bilgi Ekonomisi Değişkenlerinin Veri Madenciliği ile Belirlenmesi. Optimum Ekonomi ve Yönetim Bilimleri Dergisi, 8(1), 1-18.

Trilling, B., & Fadel, C. (2009). 21st Century Skills, Enhanced Edition: Learning for Life in Our Times. John Wiley & Sons.

Wallner, T., & Wagner, G. (2016, June). Academic Education 4.0. In International Conference on Education and New Developments (Vol. 2016, pp. 155-159).

Wikimedia Commons. (2021, Ocak 30). [https://commons.wikimedia.org:https://commons.wikimedia.org/wiki/File:How_many_public_photos_are_uploaded_to_Flickr_every_day,_month,_year%3F_\(6855169886\).png](https://commons.wikimedia.org:https://commons.wikimedia.org/wiki/File:How_many_public_photos_are_uploaded_to_Flickr_every_day,_month,_year%3F_(6855169886).png) adresinden alındı

Zikopoulos, P., & Eaton, C. (2011). Understanding bigdata: Analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media.

Zorrilla, M. E., Menasalvas, E., Marin, D., Mora, E., & Segovia, J. (2005, February). Web usage mining Project for improving web-based learning sites. In International conference on computer aided system theory (pp. 205-210). Springer, Berlin, Heidelberg.