

Derleme Makale/Review Article

Türkiye’de eğitim alanında yapılan veri madenciliği ve yapay zekâ çalışmaları

Emine Aruğaslan¹, Hanife Çivril¹

¹Isparta Uygulamalı Bilimler Üniversitesi, Uzaktan Eğitim Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, 32200, Isparta, Türkiye

Anahtar Kelimeler

Eğitimde yapay zekâ
Uyarlanabilir öğrenme ortamları
Zeki öğrenme sistemleri
Eğitsel veri madenciliği
Öğrenme analitikleri

Makale geçmişi:

Geliş Tarihi: 13.11.2020
Kabul Tarihi: 21.09.2021

Öz: Veri madenciliği ve yapay Zekâ teknikleri birçok alanda olduğu gibi eğitim alanında da uygulanmakta ve eğitim-öğretim faaliyetlerinde yenilikler sunmaktadır. Veri madenciliğinin ve yapay Zekânın eğitimde kullanımının yaygınlaşmasıyla birlikte bilimsel araştırmalar da yapılmaktadır. Bu çalışmanın amacı, Türkiye’de yapılan ve eğitimde veri madenciliği ve yapay Zekâ kullanımını ele alan çalışmaları incelemek ve bu konudaki eğilimleri tespit etmektir. Bu doğrultuda alanyazın taraması yapılmış ve veriler betimsel analizi yöntem ile analiz edilmiştir. Çalışmada kullanılacak makaleleri belirlemek amacıyla TR Dizin, Dergipark ve Google Akademik tarama ve indeksleme motorlarında; eğitim, veri madenciliği, eğitsel veri madenciliği, öğrenme analitikleri, yapay zekâ, zeki öğrenme sistemleri, uyarlanabilir öğrenme ortamları gibi anahtar kelimeler temel alınarak taramalar yapılmış ve konuyla ilgili tam metin olarak ulaşılan makaleler incelemeye alınmıştır. Yapılan taramalar neticesinde toplam 112 adet makale çalışmaya dahil edilmiş ve bu makaleler, hazırlanan makale inceleme formu çerçevesinde analiz edilmiştir.

Atıf için/To Cite:

Aruğaslan E. Çivril H. Türkiye’de Eğitim Alanında Yapılan Veri Madenciliği ve Yapay Zekâ Çalışmaları. Uluslararası Teknolojik Bilimler Dergisi, 13(2), 81-89, 2021.

Data mining and artificial intelligence studies in the field of education in Turkey

Keywords

Artificial intelligence in education
Adaptive learning environments
Intelligent learning environments
Educational data mining
Learning analytics

Article history:

Received: 13.11.2020
Accepted: 21.09.2021

Abstract: Data mining and artificial intelligence techniques are applied in the field of education as well as in many areas and offer innovations in education and training activities. With the widespread use of data mining and artificial intelligence in education, scientific research has been carried out on these issues. The aim of this study to examine studies on educational data mining and artificial intelligence use and to identify trends in this field. In this context, a literature review was made and the data were analyzed by descriptive analysis method. In order to determine the articles to be used in the study, researches were conducted with keywords such as education, data mining, educational data mining, learning analytics, artificial intelligence, intelligent learning systems, and adaptive learning environments in TR Index, Dergipark and Google Scholar searches and indexing engines and full text articles on the subject were examined. As a result of the searches, a total of 112 articles were included in the study and these articles were analyzed within the framework of the Article Review Form.

1. Giriş

Ağ teknolojilerinin ve güçlü veri toplama ve depolama araçlarının hızlı gelişiminin bir sonucu olarak her geçen gün pek çok alanda devasa büyüklükte verilerin toplandığı bir çağda yaşamaktayız [1][2]. Bu inanılmaz derecede büyüyen veri kümesini analiz etmek ve bu verileri bilgiye dönüştürmek günümüzde bir ihtiyaç

haline gelmekte ve bunun için ise güçlü ve çok yönlü araçlara gereksinim duyulmaktadır. Bu gereksinim, veri madenciliğinin doğmasını sağlamıştır [1]. Veri madenciliği, basitçe büyük ve karmaşık veri kümelerindeki yapıların ve modellerin keşfi olarak tanımlanmaktadır [3]. Veri madenciliğinin en kapsamlı tanımı Gartner Inc. tarafından “Veri madenciliği, örüntü tanıma teknolojilerinin yanı sıra istatistiksel ve

matematiksel teknikleri kullanarak, depolarda saklanan büyük miktarda veriyi eleyerek anlamlı yeni korelasyonları, modelleri ve eğilimleri keşfetme sürecidir.” şeklinde yapılmıştır [4].

Veri madenciliğinde bir dizi görev bulunmaktadır. Veri madenciliğinin bu görevleri, tahmin edici ve tanımlayıcı olarak iki kategoride gruplandırılabilir. Tahmin edici görevler, sınıflandırma, regresyon, zaman serisi analizi ve kestirimken, tanımlayıcı görevler kümeleme, özetleme, birliktelik kuralı ve sıra örüntüleridir [5]. Tanımlayıcı veri madenciliği görevleri, verilerin genel özelliklerini karakterize ederken; tahmin edici veri madenciliği görevleri, yeni bir veri kümesinin nasıl davranacağını tahmin etmek için mevcut veri kümesi üzerinde çıkarımlar gerçekleştirmektedir.

Veri madenciliğinin bankacılık, sigortacılık, pazarlama, tıp ve biyoloji, e-ticaret gibi pek çok alanın yanı sıra eğitim alanında da son yıllarda kullanıldığı görülmektedir. Eğitimde veri madenciliğinin kullanılması, eğitsel veri madenciliği ve öğrenme analitiği kavramlarını ortaya çıkarmıştır. Eğitsel veri madenciliği disiplinler arası bir araştırma alanıdır [6]. Farklı eğitim seviyelerindeki eğitim sürecinde yer alan kabul sistemleri, kayıt sistemleri, öğrenci bilgi sistemleri, öğrenme yönetim sistemleri gibi öğrenmeyi veya eğitimi destekleyen her türlü sistemden gelen verileri incelemeye ve faydalı bilgiler keşfetmeye odaklanmaktadır [6][7][8]. Öğrenme analitikleri ise keşfedilen bilgilerin öğrenme sürecine ilişkin iyileştirmelerde ve öğretim tasarımında kullanılması ile ilgilidir [9]. Eğitsel veri madenciliğindeki ve öğrenme analitiklerindeki analiz alanı, veriler, süreç ve hedefler oldukça benzerdir. Her iki alan da eğitim alanına odaklanır. Bununla birlikte, öğrenme analitiği için kullanılan teknikler eğitsel veri madenciliğinde kullanılan tekniklerden oldukça farklı olabilir. Eğitsel veri madenciliği, tipik veri madenciliği tekniklerinin (kümeleme, sınıflandırma, birliktelik kuralı gibi) uygulanmasına odaklanır. Öğrenme analitikleri ise veri madenciliği tekniklerine ek olarak, istatistiksel ve görselleştirme araçları veya sosyal ağ analizi teknikleri gibi diğer yöntemleri de içerir [10].

Yapay Zekâ ise, öğrenme, problem çözme ve örüntü tanıma gibi insan Zekâsıyla yaygın olarak ilişkili bilişsel problemleri çözmeye odaklanan bilgisayar bilimi alanı olarak tanımlanmaktadır. Diğer bir deyişle insanlar gibi çalışabilen akıllı makineler yaratmak için yapılan çalışmadır. Eğitim alanı da yapay Zekâdan önemli ölçüde etkilenmektedir. Günümüzde farklı eğitim seviyelerinde öğretim robotları, akıllı destek sistemleri, uyarlanabilir öğrenme sistemleri gibi çeşitli yapay Zekâ uygulamaları kullanılmaktadır [11].

Son yıllarda dünyada eğitim alanında veri madenciliği ve yapay zekâ konusunda yapılan çalışmaların arttığı bilinmektedir. Bu çalışmada, Türkiye’de eğitimde veri madenciliği ve yapay Zekâ kullanımını konusundaki mevcut durumu ortaya koymak amaçlanmış ve alanyazın taraması yapılarak bu konuları ele alan çalışmalar incelenerek eğilimler tespit edilmeye çalışılmıştır.

1.1. Eğitsel Veri Madenciliği ile İlgili Yapılan Benzer Çalışmalar

Eğitsel veri madenciliği ile ilgili alanyazın taraması yöntemini kullanarak mevcut durumu ortaya koymak amacıyla yapılan bazı çalışmalar mevcuttur. Baker ve Yasef (2009), çalışmalarında eğitim amaçlı veri madenciliğinin nasıl geliştiğini ve eğitimde veri madenciliği araştırmalarındaki bazı ana eğilimlerin neler olduğunu araştırmıştır [12]. Mohamad ve Tasir (2013), çalışmalarında veri madenciliğinin önceki akademisyenler tarafından nasıl ele alındığını ve eğitim araştırmalarında veri madenciliğindeki en son eğilimleri incelemiştir [13]. Papamitsiou ve Economides (2014), öğrenme analitiği ve eğitsel veri madenciliği konusundaki mevcut durumu ve bunların öğrenme üzerindeki etkisini anlamak amacıyla 2008-2014 yılları arasında yayınlanan deneysel durum çalışmalarını incelemiştir. Yayınlanan çalışmaların araştırma sorularını, yöntemini ve bulgularını analiz ederek çalışmaları gruplandırmıştır [14]. Shahiri ve AbdulRashid (2015), öğrenci performansını analiz etmede kullanılan değişkenleri tanımlamak ve öğrencilerin performansını tahmin etmek için kullanılan tahmin yöntemlerini incelemek amacıyla sistematik ilişkisel alanyazın taraması yapmıştır [15]. Sin ve Muthu (2015), çalışmalarında eğitimde büyük veri teknolojilerinin uygulamalarını incelemekte ve eğitsel veri madenciliği ve öğrenme analitiği ile ilgili mevcut literatürün bir incelemesini sunmaktadır [16]. Sukhija, Jindal ve Aggarwal (2015), çalışmalarında 2001-2015 yılları arasında yayınlanan çeşitli çalışmaların sonuçlarını inceleyerek eğitsel veri madenciliğinin gelişimini resmetmiştir [17]. Dutt, Ismail ve Herawan (2017), kümeleme algoritması ve kümeleme algoritmasının eğitsel veri madenciliği bağlamında uygulanabilirliği ve kullanılabilirliği konusunu ele alan ve 1983-2016 arasında yayınlanan çalışmaları inceleyerek sistematik alanyazın taraması yapmıştır [18]. Kumar, Singh ve Handa (2017), eğitimde veri madenciliğinin farklı tahmin teknikleri arasındaki farkı anlamak ve analiz etmek, öğrenci performansını tahmin etmek için kullanılan farklı öğrenci özelliklerini ve farklı tahmin tekniklerini belirlemek ve anlamak için sistematik alanyazın taraması yapmıştır [19]. Tekin ve Öztekin (2018), çalışmalarında 2006-2016 yılları arasında eğitsel veri madenciliği ile ilgili yayınlanmış olan çalışmaları incelemiş ve eğitim alanındaki veri

madenciliği konusunda güncel eğilimleri araştırmışlardır [20].

1.2. Yapay Zekâ ile İlgili Yapılan Benzer Çalışmalar

Eğitimde yapay Zekânın kullanımı ile ilgili alanyazın taraması yöntemini kullanarak mevcut durumu ortaya koymak amacıyla yapılan bazı çalışmalar mevcuttur. Chassignola, Khoroshavin ve Klimova (2018), çalışmalarında yapay teknolojilerinin çalışma sürecine olası etkisini belirlemek ve eğitim ortamındaki olası değişiklikleri tahmin etmek amacıyla alanyazın taraması yapmış ve yapılan çalışmaları özelleştirilmiş eğitim içeriği, yenilikçi öğretim yöntemleri, teknolojiyle geliştirilmiş değerlendirme, öğrenci ve öğretim görevlisi arasındaki iletişim olmak üzere dört kategoride incelemiştir [21]. Zawacki-Richter vd. (2019), 2007-2018 yılları arasında yayımlanan ve yükseköğretimde yapay Zekâ uygulamaları üzerine yapılan çalışmaları incelemiştir. Sonuçları profil oluşturma ve tahmin, ölçme ve değerlendirme, uyarlanabilir sistemler ve kişiselleştirme, akıllı eğitim sistemleri başlıkları altında sunmuştur [22]. Chen, Chen ve Lin (2020), yapay Zekânın eğitim üzerindeki etkisini değerlendirmek amacıyla alanyazın taraması yapmıştır [23].

2. Yöntem

Bu çalışmada, eğitimde veri madenciliği ve eğitimde yapay Zekâ konularını ele alan çalışmaları inceleyerek mevcut durumu ortaya koymak ve eğilimleri tespit etmek amacıyla betimsel analiz yöntemi kullanılmıştır. Yıldırım ve Şimşek (2011)’e göre; betimsel analiz nitel verilerin işlenerek elde edilen bulguların tanımlanması ve yorumlanması adımlarından oluşur [24]. Alanyazın taraması yapılarak ulaşılan çalışmaların seçimi için birtakım ölçütler belirlenmiştir ve bu ölçütleri sağlayan çalışmalar betimsel analize tabii tutulmuştur. Bu çalışma, Türkiye’deki mevcut durum ve eğilimleri incelediği için yayınların çalışmaya dahil edilebilmesi için Türkiye’de yayınlanan dergilerde bulunan ve tam metin olarak ulaşılabilen makale olması bir ölçüt olarak belirlenmiştir. Bir diğer ölçüt ise yayınlanan çalışmaların doğrudan eğitim alanında veri madenciliği ve yapay Zekâ konularına odaklanıyor olmasıdır.

2.1. Veri Toplama Araçları

Eğitimde veri madenciliği ve yapay Zekâ konularını ele alan çalışmaları incelemek için kapsamlı bir alanyazın taraması yapılmıştır. Çalışmanın amacı doğrultusunda belirlenen ölçütler göz önünde bulundurularak elektronik ortamda yayınlanan makalelere ulaşmak için TR Dizin, Dergipark ve Google Akademik tarama ve indeksleme motorları kullanılmıştır. Tarama yapılırken eğitim, veri madenciliği, eğitsel veri madenciliği,

öğrenme analitikleri, yapay zekâ, zeki öğrenme sistemleri, uyarlanabilir öğrenme ortamları gibi anahtar kelimeler kullanılmıştır. Tarama sonucunda çalışmaya 112 makale dahil edilmiştir. Makalelerin analiz edilebilmesi için araştırmacılar tarafından Microsoft Excel paket programı kullanılarak bir “Makale İnceleme Formu” hazırlanmıştır. Bu formda “basım yılı”, “yazarların unvanları”, “araştırmacının kurumu”, “çalışma konusu”, “çalışma grubu (örneklem)”, “eğitim kademesi”, “veri toplama araçları” ve “anahtar kelimeler” sütunları oluşturulmuştur. Eğitimde veri madenciliği çalışmalarının analizi için ise forma “modelleme yöntemleri”, “veri analiz yöntemleri” ve “kullanılan araçlar” sütunları eklenmiştir.

2.2. Veri Analizi

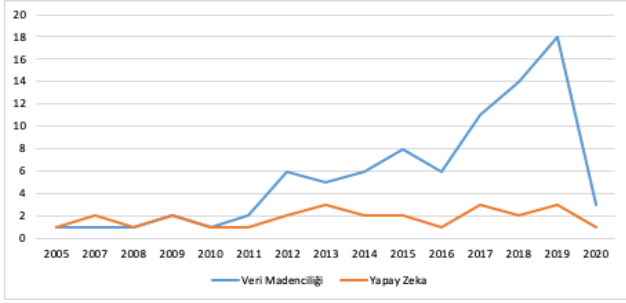
Bu çalışmada verilerin analizi için betimsel analiz kullanılmıştır. Hazırlanan Makale inceleme formu çerçevesinde çalışmaya dahil edilen bütün makaleler her iki araştırmacı tarafından ayrı ayrı analiz edilmiştir. Ortaya çıkan bulgular karşılaştırılarak incelenmiş ve araştırmacılar arasında tutarlılık sağlanmıştır. Analizlerin güvenilirliğini sağlamak amacıyla bir uzman tarafından çalışmaların değerlendirilmesi istenilmiştir. Analiz sonucunda elde edilen bulgular, tanımlayıcı istatistikler (yüzde ve frekans) kullanılarak sunulmuştur. Bulgular; tablo, grafik ve şekillerle görselleştirilmiştir.

3. Bulgular

Bu çalışmada yapılan taramalar neticesinde veri madenciliği (85) ve yapay Zekâ (27) ile bağlantılı olduğu düşünülen toplam 112 adet makaleye yer verilmiş ve tüm incelemeler bu makaleler üzerinden gerçekleştirilmiştir.

3.1. Genel Bilgiler

Bu bölümde çalışmada yer alan makalelere ait basım yılları incelendiğinde aşağıdaki gibi bir tablo ile karşılaşılmıştır. Grafik 1 incelendiğinde yıllara göre özellikle veri madenciliği konularını içine alan çalışmaların 2010 yılı itibari ile giderek arttığı söylenebilir.



Grafik 1. Çalışmada yer alan makalelerin konularına göre basım yıllarının dağılımı*

*Veriler 2020 yılının Şubat ayına kadar olan çalışmaları kapsamaktadır.

112 adet çalışmayı yapan araştırmacıların çalıştığı kurumlara bakıldığında neredeyse tamamına yakını yükseköğretim kurumları oluşturmuştur. Yükseköğretim kurumları dışında Akademiler, MEB'e bağlı okullar ve verilerin Türkiye'den toplandığı yurtdışındaki eğitim kurumları oluşturmaktadır. Toplamda 241 araştırmacı bu makalelerde görev almış ve en fazla 6 yazarlı olacak şekilde çalışmalar yapılmıştır.

3.2. Eğitimde Veri Madenciliği İle İlgili Çalışmalar

Eğitimde Veri Madenciliği konusunu ele alan 85 adet makale incelendiğinde aşağıdaki gibi bir tablo ile karşılaşmıştır (Tablo 1). Çalışmaların yapılma amaçlarında en çok göze çarpan konunun %49,4 ile "öğrenci başarıları" olduğu görülmüştür. İkinci sırada veri madenciliği konuları içerisinde yine önemli yer tutan "Öğrenme analitikleri" gelmektedir.

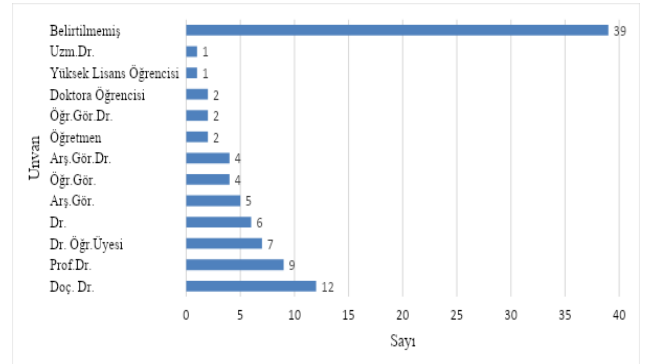
Tablo 1. Eğitimde veri madenciliği çalışma konuları

Eğitimde Veri Madenciliği Çalışma Konuları	f	%
Öğrenci başarıları	42	49,4
Öğrenme analitikleri	7	8,2
Eğitmen performansları	5	5,9
Teorik çalışmalar	4	4,7
Öğrenci verilerinin sınıflandırılması	3	3,5
Öğrenme stilleri	3	3,5
E-öğrenme riskleri	3	3,5
Devam/devamsızlık	2	2,4
Sosyal ağlar	2	2,4
Web Tabanlı Eğitim	2	2,4
Profillerinin belirlenmesi	2	2,4
Öğrenme ortamı hakkında görüşlerin incelenmesi	2	2,4
Program müfredatlarının karşılaştırılması	2	2,4
Siber zorbalık	1	1,2
Sanal laboratuvar	1	1,2
OBS sisteminin kullanılabilirliği	1	1,2
Eylem öğrenme sürecinin okul yöneticileri karar vermelerine etkisi	1	1,2
Öğrenci velilerinin güvenli İnternet kullanımı farkındalığı	1	1,2
Büyük veri	1	1,2
Genel Toplam	85	100

Öğrenci başarılarını konu alan eğitimde veri madenciliği çalışmaları detaylandırıldığı; öğrenci başarılarının değerlendirilmesi, öğrenci başarılarına etki eden faktörler, öğrenci başarılarının artırılması, öğrenci başarılarına etki eden ÖYS (Öğrenme Yönetim Sistemleri) etkinlikleri, ÖSYM verilerinin öğrenci başarısına etkileri, bilgisayar okuryazarlığı, lise mezuniyetlerinin fakültelerle olan ilişkisi gibi konular karşımıza çıkmaktadır.

Öğrenme ortamlarındaki hareketlerin analiz edilerek incelenmesini hedef alan ve son yıllarda popüler hale gelen "öğrenme analitikleri" çalışmaları da veri madenciliği ile ilgili araştırılan makalelerde diğer çalışma konularına oranla yüksek bir paya sahiptir. Teorik çalışma konuları araştırmacıların veri madenciliğini temel alarak yaptıkları belirli bir alandaki araştırmaları inceledikleri konulardan oluşmaktadır.

Çalışmaları yapan araştırmacıların unvanları incelendiğinde 85 çalışmada, 94 kişi görev almış ve makalelerdeki dergi formatı gereği unvan bilgisi çoğunlukla paylaşılmamış ve grafikte de "belirtilmemiş" olarak verilmiştir (Grafik 2). Unvan belirtilen çalışmalarda en çok "Doç.Dr." unvanına sahip öğretim elemanları yer almıştır.



Grafik 2. Araştırmacıların unvanları

Eğitimde veri madenciliği üzerine yapılan çalışmalara bakıldığında %89,7'lik dilimi öğrenci gruplarının oluşturduğu görülmektedir (Tablo 2). Bu sonuç eğitim çalışmalarında beklenen bir durumdur. Eğitim seviyelerine göre akademisyenleri ve öğretmenleri hedef alan çalışmaların oldukça az sayıda olduğu görülmüştür. Akademisyenleri içeren çalışmalar genellikle eğitmen performansları konularında yapılmış olup, veli ve yöneticileri ilgilendiren çalışmalar ise bir konudaki farkındalık tespitinin veri madenciliği yöntemleri ile yapılması üzerinedir.

Tablo 2. Çalışma grupları

Çalışma Grubu	f	%
Öğrenci	61	89,7
Akademisyen	2	2,9
Genel (Web taramalarına katılanlar)	1	1,5
Üstün Yetenekli Öğrenciler	1	1,5
Öğretmen	1	1,5
Veli	1	1,5
Yönetici	1	1,5
Genel Toplam	68	100

Eğitimde yapılan çalışmaların konu içerikleri ve çalışma gruplarının yanı sıra eğitimin hangi kademesinde bu sürecin tamamlandığı bilgisi de önem arz etmektedir. Bu çalışmadaki eğitim kademeleri incelendiğinde aşağıdaki gibi bir sonuçla karşılaşmıştır (Tablo 3). Yükseköğretim kurumları, araştırma yapan kurumların başında gelmektedir. Bu sebeple yükseköğretim kurumlarının elde edilen sonuçlar içerisinde en üstünde yer alıyor olması beklenen bir durumdur. Bunun yanı sıra MEB bünyesinde de çalışmalar çoğunlukla yükseköğretim kurumlarındaki akademisyenler tarafından yürütülmüştür.

Tablo 3. Çalışmaların yürütüldüğü eğitim kademesi

Eğitim Kademesi	f	%
Yükseköğretim (Lisans)	47	55,3
MEB (Ortaokul)	7	8,2
MEB (Genel)	6	7,1
MEB (Lise)	7	8,2
Yükseköğretim (Lisansüstü)	2	2,4
Diğer (Teorik Çalışmalar)	16	18,8
Genel Toplam	85	100

Veri toplama araçlarına ait veriler Tablo 4’te verilmiştir. Bu tabloda yazılım geliştirmeyi konu alan çalışmalarda veri toplama araçları kullanılmamış ve ayrıca teorik çalışmaların çoğunda hazır veriler kullanılmış ve literatür taraması yapılmıştır. Tablo 4 incelendiğinde kullanılan veri toplama araçlarından en çok tercih edileni %34,3 ile “OBS Verileri” olmuştur. OBS verilerinin toplanarak yapıldığı Veri madenciliği çalışmalarında öğrencilere ait veriler alınmış ve araştırma verisi olarak kullanılmıştır. Bu veriler çoğunlukla yaş, cinsiyet, medeni durum, yılsonu ve arasınan notları, fakülte, bölüm bilgileri, doğum tarihi, adres vb. bilgileri içermektedir. Veri madenciliği ile ilgili en çok araştırma konusunun öğrenci başarısı olduğu düşünüldüğünde OBS üzerinden en çok toplanan verinin “yıl sonu notları” olması rastlantı değildir. Bunun dışında ikinci sırada %25,5 ile “anket” gelmektedir. Eğitimde Veri madenciliği çalışmalarında Anket ile veri toplamak ve bu verilerin Veri madenciliği modelleme yöntemleri ile analiz edilmesi sıkça rastlanan bir durum olmuştur. Öğrencilerin başarılarının analizi, karşılaştırması veya tahmininde kullanılan ÖSYM, PISA gibi veri kaynaklarına da sıklıkla başvurulduğu söylenebilir.

Tablo 4. Veri toplama araçları

Veri Toplama Araçları	f	%
OBS Verileri	35	34,3
Anket	26	25,5
Log Kayıtları	11	10,8
Tez/Makale Arama Motoru	10	9,8
ÖSYM Verileri	8	7,8
LMS Verileri	4	3,9
PISA Notları	3	2,9
Başarı Testi	2	2,0
Görüşme	1	1,0
Deney Verileri	1	1,0
TIMMS Notları	1	1,0
Genel Toplam	102	100

Veri toplama araçları ile oluşturulan veri setleri içerisindeki değerlerin arasında farklılıklar olduğu için veri setinin uygun hale getirilmesi amacıyla 9 çalışmada verilerin normalizasyon işlemine tabii tutulduğu belirtilmiştir. 5 tane çalışmada normalizasyon için kullanılan matematiksel yöntem yer verilmemiştir. 1 tane çalışmada normalizasyon için D Min Max Normalizasyonu, 3 tane çalışmada ise Z-Score Normalizasyonu tekniğinin kullanıldığı görülmüştür.

Tablo 5’te eğitimde veri madenciliği ile ilgili çalışmaya dahil edilen makalelerde en çok kullanılan Veri madenciliği modelleme yöntemleri verilmiştir. Teorik çalışmalar ve veri madenciliği tekniklerinin kullanılmadığı çalışmalar, aşağıda verilen modelleme, yöntem ve veri toplama araçları tablolarında yer almamıştır. Tahmin edici model %57,4 ile en çok kullanılan yöntem olmuştur. Veri madenciliğini temel alan araştırmalar incelendiğinde anahtar kelimelerinde ve konu içeriğinde Veri madenciliği ifadesi geçen ancak özet ve özellikle yöntem bölümlerinde veri madenciliği ile ilgili detaylı ve olması gereken bilgilendirmelerin yapılmamasından dolayı %11,8’lik bir oranla “belirtilmemiş” bilgisi tabloda yer almıştır. Ancak bu durumda çoğu çalışmanın modelleme yöntemini yazdığı sonucu çıkarılabilir (%89,2).

Tablo 5. Veri madenciliği modelleme yöntemleri

Veri Madenciliği Modelleme Yöntemleri	f	%
Tahmin edici model	39	57,4
Tanımlayıcı model	13	19,1
Tahmin edici model, Tanımlayıcı model	8	11,8
Belirtilmemiş	8	11,8
Genel Toplam	68	100

Bu çalışmadaki eğitimde veri madenciliği alanındaki yapılan araştırmalarda en çok kullanılan analiz yöntemleri %50,6 ile sınıflandırma analizi olmuştur (Tablo 6). Yapılan araştırmalarda tabloda verilen yöntemlerin birlikte kullanıldığı da görülmüştür. Yapılan tüm çalışmalardan veri analiz yöntemi “belirtilmemiş” olanlar çıkartıldığında çalışmaların çoğunun tek yöntemle, bir kısmının da birden fazla yöntemle analiz edildiği görülmüştür. Çalışmalarda en

fazla 3 yöntem birlikte kullanılmış ve çalışmalarda hem tahmin edici modelin farklı yöntemlerinin hem tanımlayıcı yöntemin farklı yöntemlerinin hem de tahmin edici modelle tanımlayıcı modelin farklı yöntemlerinin birlikte kullanıldığı görülmüştür.

Tablo 6. Veri analiz yöntemleri

Veri Analiz Yöntemleri	f	%
Sınıflandırma Analizi	44	50,6
Birliktelik Kuralı	13	14,9
Kümeleme	12	13,8
Regresyon	10	11,5
Belirtilmemiş	8	9,3
Genel Toplam	87	100

Çalışmalar incelendiğinde en çok kullanılan ana analiz yöntemlerinden olan sınıflandırma analizinde en çok kullanılan algoritmalar Karar ağaçları, Yapay Sinir Ağları, Naive Bayes, K-en Yakın Komşu, CN2 Rules olurken; Birliktelik Kuralında ise en çok kullanılan algoritma Apriori olmuştur.

Veri madenciliğinde verilerin analiz edilmesi için kullanılan birçok yazılım bulunmaktadır. Çalışmalar incelendiğinde Tablo 7’de verilen yazılımların kullanıldığı görülmüştür. WEKA ve SPSS analiz programları veri madenciliği çalışmalarında %20,5 ile en çok kullanılan yazılımlar olmuştur. Çalışmalarda yine aynı oranda kullanılan yazılımların belirtilmediği görülmüştür. Kullanılan yazılım çalışmalarda belirtilmese de kullanılan analiz yöntemleri çalışmalarda paylaşılmıştır. Verilerin analizlerinde aynı çalışmada en fazla 3 adet yazılım kullanılmıştır.

Tablo 7. Veri analizinde kullanılan araçlar

Veri Analiz Yazılımları	f	%
WEKA	16	20,5
Belirtilmemiş	16	20,5
SPSS (Clementine, Modeler)	16	20,5
Microsoft SQL Server Analysis Services	6	7,7
MATLAB	5	6,4
Rapid Miner Studio	4	5,1
Orange Data Mining	2	2,6
Microsoft Office Excel	2	2,6
KNIME	2	2,6
Visual Studio	2	2,6
Business Intelligence Development Studio	1	1,3
D-Melt	1	1,3
Purple Insight MineSet	1	1,3
Tanagra	1	1,3
SAS Enterprise Miner	1	1,3
PASW Statistics	1	1,3
R	1	1,3
Genel Toplam	78	100

Veri madenciliği ile ilgili 85 çalışmada toplam 345 anahtar kelime bulunmaktadır. En çok kullanılan anahtar kelimelerden ilk 10 tanesi Tablo 8’de verilmiştir.

Tablo 8. Veri madenciliği çalışmalarındaki anahtar kelimeler

Anahtar Kelimeler	f
Veri madenciliği (data mining)	35
Eğitsel veri madenciliği (educational data mining)	25
Öğrenme analitikleri (learning analytics)	10
Sınıflandırma algoritmaları	7
Yapay sinir ağları	6
Makine öğrenmesi	5
Birliktelik kuralı	5
WEKA	4
Karar ağaçları	4
Web tabanlı eğitim	4

Bu anahtar kelimeler incelendiğinde çoğunlukla veri madenciliği kavramının kullanıldığı görülmüştür. Diğer öne çıkan ve çok kullanılan kelimeler; öğrenme, eğitimsel, sınıflandırma, öğrenci analizi gibidir. Araştırmacılar veri madenciliği yöntemlerini, modellerini, platformları kullanılan yazılımları, algoritmaları anahtar kelimelerinde kullanmışlardır. Şekil 1’de veri madenciliği ile ilgili detaylı bir şekil oluşturmak için anahtar kelime verileri kelime bulutuna sokulmuş ve aşağıdaki gibi bir görsel oluşturulmuştur [25].



Şekil 1. Eğitimde veri madenciliği çalışmalarında anahtar kelimeler için kelime bulutu

3.3. Eğitimde Yapay Zekâ ile İlgili Çalışmalar

Eğitimde yapay Zekâ ile ilgili 27 adet makaleye ulaşılmıştır. Yapay Zekâ çalışmalarına ait konu içerikleri; uyarlanabilir öğrenme ortamı, zeki öğrenme sistemleri, bireyselleştirilmiş öğrenme, robot öğretmen, sanal gerçeklik, bulut bilişim ve yapay zekâ, yazılım değerlendirilmesi, eğitimsel yazılım geliştirme, sanal gerçeklik, bulut bilişim, etkileşimli eğitim yazılımı, dağıtık yapay zekâ teknikleri şeklindedir. Yapılan çalışmaların yıllara göre dağılımına bakıldığında 2005-2020 yılları arasında olduğu görülmüştür.

Çalışmalarda toplam 46 araştırmacı görev almış ve genellikle öğretim üyeleri tarafından yapılmıştır (Tablo 9).

Tablo 9. Araştırmacı unvanları

Unvanlar	f	%
Belirtilmemiş	14	30,4
Dr. Öğr. Üyesi	11	23,9
Prof. Dr.	7	15,2
Doç. Dr.	4	8,7
Arş. Gör.	3	6,5
Öğr. Gör.	2	4,3
Öğrenci	2	4,3
Dr.	2	4,3
Arş.Gör. Dr.	1	2,2
Genel Toplam	46	100

Eğitimde yapay Zekâ üzerine yapılan çalışma konuları Tablo 10’da verilmiştir. Çalışmalarda en çok %28,6 ile yapay Zekâ yazılımının/ortamının öğretime etkisinin değerlendirilmesi yapılmıştır.

Tablo 10. Eğitimde yapay Zekâ araştırmalarındaki çalışma konuları

Çalışma Konuları	f	%
Yazılımın/ortamın öğretime etkisinin değerlendirilmesi	10	28,6
Teorik	9	25,7
Yazılım/ortam geliştirme	9	25,7
Yazılım/ortam değerlendirmesi	8	22,9
Genel Toplam	35	100

Eğitimde Yapay Zekâ konusundaki araştırmaların çalışma grupları incelendiğinde teorik ve yazılım/ortam geliştirme üzerine yapılan çalışmalarda bu kapsamda bilgi yer almamaktadır (Tablo 11). Ancak bu çalışmaların çoğunlukla öğrenciler üzerinde yapıldığı söylenebilir.

Tablo 11. Çalışma grupları

Çalışma Grubu	f	%
Öğrenci	14	70
Öğretmen	3	15
Akademisyen	2	10
Uzman	1	5
Genel Toplam	20	100

Çalışmalar eğitim kademelerine göre incelendiğinde daha çok yükseköğretim düzeyinde çalışmalar yapılmıştır (Tablo 12).

Tablo 12. Çalışmalardaki eğitim kademesi

Eğitim Kademesi	f	%
Yükseköğretim	10	66,6
MEB	3	20
Ortaokul	2	13,3
Genel Toplam	15	100

Veri toplama araçları olarak anket, ölçek, başarı testi ve görüşme formları en çok kullanılan araçlar olmuştur (Tablo 13).

Tablo 13. Veri toplama araçları

Veri Toplama Araçları	f	%
-----------------------	---	---

Anket/Ölçek	8	34,8
Başarı testi	6	26,1
Görüşme	4	17,4
Tez/Makale arama motoru	2	8,7
Log Kayıtları	1	4,3
Delphi yöntemi	1	4,3
Ses Kayıtları	1	4,3
Genel Toplam	23	100

Yapay Zekâ ile ilgili 27 çalışmada toplam 110 anahtar kelime bulunmaktadır. En çok kullanılan anahtar kelimelerden ilk 5 tanesi Tablo 14’te verilmiştir.

Tablo 14. Yapay Zekâ çalışmalarındaki anahtar kelimeler

Anahtar Kelimeler	f
Yapay Zekâ (artificial intelligence)	8
Zeki öğretim sistemleri	6
Uyarlanabilir öğrenme ortamları	4
Web tabanlı öğretim	4
Bilgisayar destekli öğretim	3

Yapay Zekâ konularındaki anahtar kelimelere ait bilgiler aşağıdaki şekilde verilmiştir (Şekil 2). Anahtar kelimeler içerisinde en çok göze çarpanının “uyarlanabilir” ifadesinin olduğu görülmüştür.



Şekil 2. Eğitimde yapay Zekâ çalışmalarında anahtar kelimeler için kelime bulutu

4. Sonuç ve Tartışma

Bu çalışmada, Türkiye’de yapılan eğitimde veri madenciliği ve eğitimde yapay zekâ kullanımını ile ilgili 112 makale betimsel analiz yöntemi ile analiz edilmiştir. Bu çalışmanın sonuçlarına göre eğitimde veri madenciliği kullanımı konusunda (eğitsel veri madenciliği veya öğrenme analitikleri) yapılan çalışmaların daha fazla olduğu ve 2010 yılından sonra bu konuda yapılan çalışmaların arttığı görülmektedir. Bu durumun nedeni öğrenci bilgi sistemleri, öğrenme yönetim sistemleri, sosyal ağlar gibi bilgi ve iletişim teknolojilerinin sunduğu ortamların kullanımının artması ve bu ortamlardan verilerin kolay toplanabilir olması olarak düşünülebilir. Araştırma konularına bakıldığında ise yapılan çalışmaların büyük oranda öğrenci başarısı üzerine odaklandığı görülmektedir.

Eğitimde veri madenciliği kullanımındaki eğilimin daha çok öğrencilerin başarısının değerlendirilmesi ve tahmin edilmesi yönünde olduğu söylenebilir. Eğitim sürecinin önemli bir parçası olan öğrencilerin başarısını tahmin etmek, eğitimcilere ve öğrencilere öğrenme ve öğretme süreçlerini geliştirmelerine yardımcı olmak için oldukça yararlı olduğundan araştırmacıların daha çok bu konu üzerine eğildiği söylenebilir. Öğrenme analitikleri, araştırılan konular arasında ikinci sırada yer alsa da küçük bir orana sahip olduğu görülmektedir. Günümüzde dünyada özellikle uzaktan eğitime karşı artan ilgi ve öğrenme yönetim sistemlerinin kullanımının artması sonucunda öğrenme analitikleri dikkat çeken önemli eğilimlerden birisi haline gelmiştir [16]. Ancak Türkiye’de öğrenme yönetim sistemlerinin kullanımı yaygın olmasına rağmen öğrenme analitiklerinin nispeten daha az çalışıldığı görülmektedir. Bu ortamlardan daha üst düzey verim sağlanabilmesi için sistem içinde kayıtlı olan kullanıcı verilerinin analiz edilmesi önemlidir. Böylece öğrenci profillerinin oluşturulmasında, öğrenme ortamının kişiselleştirilmesinde, eğitim ortamının kalitesinin artırılmasında geri bildirimler sağlanacaktır [26].

Türkiye’de Eğitimde veri madenciliği ve eğitimde yapay Zekâ üzerine odaklanan çalışmalar, daha çok öğrencilere yöneliktir. Aktif öğrenci verilerine ulaşmanın diğer iç (öğretmen, akademisyen, yönetici vb) ve dış paydaş (veli, mezun vs.) verilerine ulaşmaktan daha kolay olması bu durumu ortaya çıkarmaktadır. Eğitim sürecinde önemli bir yere sahip olan diğer iç ve dış paydaşlara yönelik çalışmaların artırılması gerektiği düşünülmektedir. Çalışmaların yürütüldüğü eğitim kademesine bakıldığında yükseköğretim daha çok tercih edildiği ve buna paralel olarak da veri toplama ortamı olarak yükseköğretimde kullanılan öğrenci bilgi sistemlerinin başı çektiği görülmektedir. Diğer veri toplama ortamlarına bakıldığında internet ortamlarının kullanıldığı görülmektedir. Bunun yanı sıra bazı araştırmacıların kendi eğitim verilerini anket, ölçek, başarı testi ve görüşme formu gibi araçları kullanarak topladığı görülmektedir.

Veri madenciliğinde verilerin analizi için çeşitli teknikler kullanılmaktadır. Analiz sonucunda yapılan çalışmalarda sınıflandırma analizinin daha çok kullanıldığı ve diğer analiz yöntemlerinin eşit oranlarda dağıldığı görülmektedir. Sin ve Muthu (2015), eğitsel veri madenciliğinde kullanılan en popüler teknikleri regresyon, en yakın komşu, kümeleme ve sınıflandırma şeklinde listelemiştir [16]. Veri madenciliğinde verilerin analiz edilmesinde çeşitli yazılımların kullanıldığı görülmektedir. En çok kullanılan yazılımlar, bir makine öğrenme yazılımı olan WEKA ve analiz süreçlerinde araştırmacılar tarafından en çok tercih

edilen analiz yazılımı SPSS olmuştur. WEKA’nın açık kaynak kodlu bir yazılım olmasının, araştırmacılar tarafından yoğunlukla tercih edilme sebebi olduğunu düşünebiliriz.

Eğitimde yapay Zekâ ile ilgili çalışmalar, veri madenciliği çalışmalarına göre daha sınırlı kalmıştır. Türkiye’de bu alandaki çalışmalarda bir boşluk olduğu söylenebilir. Yapay Zekânın eğitimde uygulama alanları olan robot öğrenme, uyarlanabilir öğrenme ortamları ve zeki öğrenme sistemleri ile ilgili yapılan çalışmaların neredeyse yarısının teorik ve yazılım geliştirme/ortam geliştirme boyutunda olduğu görülmektedir. Diğer çalışmalar ise geliştirilen ya da varolan yazılımların/ortamların öğretmenler veya öğrenciler tarafından değerlendirilmesi ve öğrenmenin etkililiği ve verimliliği (öğrenci başarısı, öğrencilerin bilişsel doyumu vb.) üzerine odaklanmaktadır. Yazılımların/ortamların değerlendirilmesi için kullanılan veriler anket, görüşme, sistemdeki log kayıtları ve ses kayıtlarından elde edilmiştir. Yazılımların/ortamların öğretim sürecindeki etkisinin değerlendirilmesinde ise başarı testleri ve ölçekler kullanılmıştır. Çalışmalarda ilköğretim/ortaöğretim öğrencileri ve öğretmenler olsa da örneklem grubu çoğunlukla yükseköğretim öğrencilerinden oluşmaktadır. Araştırmacıların çoğunlukla akademisyen olması ve akademisyenler için de uygun örneklem olan yükseköğretimdeki öğrenci gruplarına erişimin daha kolay olması bu durumu ortaya çıkarmaktadır. Öğrenciler dışındaki örneklem gruplarının yer aldığı çalışmaların da yapılması, eğitimde yapay Zekâ alanına katkı sağlayacaktır. Eğitimde yapay Zekâ, öğrencilere gelişmiş öğrenme deneyimleri sunarak öğrenme materyallerinin öğrencilerin ihtiyaçlarına ve yeteneklerine göre bireyselleştirmesini sağlamaktadır. Geleneksel e-öğrenme ortamları her kullanıcıya aynı içeriği sunmaktadır. E-öğrenmenin yaygınlaşmasıyla beraber bu ortamların, bireysel farklılıkları dikkate alma yönündeki eksiklikleri de daha belirgin hale gelmiştir. Bu yüzden e-öğrenme ortamlarında öğretim ve öğrenmeyi geliştirmek için yapay Zekâ teknolojilerinin kullanılması önemlidir. Bu açıdan bakıldığında Türkiye’de yapay Zekâ uygulama ortamlarının geliştirilmesi ve bu teknolojilerin çeşitli gruplar açısından değerlendirilmesi araştırmacılara çalışma konusu olarak önerilebilir.

Kaynaklar

- [1] Hand DJ, Adams NM *Data Mining*. Editörler: Balakrishnan N, Colton T, Everitt B, Piegorisch W, Ruggeri F, Teugels JL. Major Reference Works. Wiley StatsRef: Statistics Reference Online. 2015.
- [2] Wu X, Zhu X, Wu G, Ding W. Data mining with big data. *IEEE Transactions on Knowledge and Data*

- Engineering*, 26(1), 97-107, 2014. doi: 10.1109/TKDE.2013.109.
- [3] Han J, Kamber M, Pei J. *Data Mining Concepts and Techniques*. Morgan Kaufmann Publishers, Burlington, 2012.
- [4] Gartner Glossary. <https://www.gartner.com/en/information-technology/glossary> (Erişim Tarihi: 23.05.2020).
- [5] Dunham MH. *Data mining introductory and advanced topics*. Upper Saddle River, NJ: Pearson Education, Inc., 2003.
- [6] Romero C, Ventura S. Data mining in education. *WIRES Data Mining Knowl Discov*, 3, 12-27, 2013. Doi:10.1002/widm.1075
- [7] Al Mazidi A, Abusham E. Study of general education diploma students’ performance and prediction in Sultanate of Oman, based on data mining approaches. *International Journal of Engineering Business Management*, 10, 1-11, 2018.
- [8] Saa AA. Educational Data Mining & Students’ Performance Prediction. *International Journal of Advanced Computer Science and Applications*, 7(5), 212-220, 2016.
- [9] Keskin S, Aydın F, Yurdugül H. Eğitsel Veri Madenciliği ve Öğrenme Analitikleri Bağlamında E-Öğrenme Verilerinde Aykırı Gözlemlerin Belirlenmesi. *EĞİTİM TEKNOLOJİSİ Kuram ve Uygulama*, 9(1), 292-309, 2019.
- [10] Chatti MA, Dyckhoff AL, Schroeder U, Thüs H. A Reference Model for Learning Analytics. *International Journal of Technology Enhanced Learning*, 4(5), 318-331, 2012. doi: 10.1504/IJTEL.2012.051816
- [11] Chassignol M, Khoroshavin A, Klimova A, Bilyatdinova A. Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science*, 136, 16-24, 2018.
- [12] Baker RSJD, Yacef K. The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1(1), 2009.
- [13] Mohamad SK, Tasir Z. Educational Data Mining: A Review. *Procedia - Social and Behavioral Sciences*, 97, 320-324, 2013.
- [14] Papamitsiou Z, Economides AA. Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Journal of Educational Technology & Society*, 17(4), 49-64, 2014.
- [15] Shahiri AM, Husain W, Rashid NA. A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 72, 414-422, 2015.
- [16] Sin K, Muthu L. Application of big data in educational data mining and learning analytics-a literature review. *ICTAC Journal of Soft Computing*, 5(4), 1035-1049, 2015.
- [17] Sukhija K, Jindal M, Aggarwal N. The recent state of educational data mining: A survey and future visions. *2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)*, Amritsar, 1-2 October 2015.
- [18] Dutt A, Ismail MA, Herawan T. A Systematic Review on Educational Data Mining. *IEEE Access*, 5, 15991-16005, 2017.
- [19] Kumar M, Singh AJ, Handa D. Literature Survey on Student’s Performance Prediction in Education using Data Mining Techniques. *IJ. Education and Management Engineering*, 6, 40-49, 2017.
- [20] Tekin A, Öztekin Z. Eğitsel Veri Madenciliği ile İlgili 2006-2016 Yılları Arasında Yapılan Çalışmaların İncelenmesi. *Eğitim Teknolojisi Kuram ve Uygulama*, 8(2), 108-124, 2018. doi: 10.17943/etku.351473
- [21] Chassignol M, Khoroshavin A, Klimova A, Bilyatdinova A. Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science*, 136, 16-24, 2018.
- [22] Zawacki-Richter O, Marin VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators?. *International Journal of Educational Technology in Higher Education*, 16(39), 2019.
- [23] Chen L, Chen P, Lin Z. Artificial Intelligence in Education: A Review. *IEEE Access*, 8, 75264-75278, 2020. doi: 10.1109/ACCESS.2020.2988510.
- [24] Yıldırım A, Şimşek H. *Sosyal Bilimlerde Nitel Araştırma Yöntemleri*. 8th Ed. Seçkin Yayınevi, Ankara, Türkiye, 2011.
- [25] Worldcloud Generator. <https://www.wordclouds.com> (Erişim Tarihi: 23.05.2020).
- [26] Özbay Ö, Ersoy H. Öğrenme Yönetim Sistemi Üzerindeki Öğrenci Hareketliliğinin Veri Madenciliği Yöntemleriyle Analizi. *GEFAD / GUGEF*, 37(2), 523-558, 2017.