

Eğitim Sisteminde Veri Madenciliği Uygulamaları ve Farkındalık Üzerine Bir Durum Çalışması

Abdulkadir ÖZDEMİR (*)

Rabia SAYLAM (**)

Berna Betül BİLEN (***)

Öz: Eğitimsel veri madenciliği, veri madenciliği yaklaşımında nispeten yeni bir alandır. Bu çalışma, Türkiye’de veri madenciliği teknolojisinin eğitim sisteminde kullanımı üzerine farkındalığı ortaya koymayı hedeflemektedir. Bu maksatla bir anket hazırlanmış, anket katılımcısı olan öğretmenlerin demografik karakteristikleri, teknoloji ve internete olan ilgileri, e-okul sistemine yaklaşımları ve veri madenciliği teknolojisi farkındalığı ile eğitim sisteminde veri madenciliği teknolojisinin uygulanmasına yönelik görüşleri alınmıştır. Elde edilen verilerin istatistiksel analizi için SPSS yazılımı kullanılmıştır.

Anahtar Kelimeler: Veri Madenciliği, Eğitimsel Veri Madenciliği, Veri Madenciliği Farkındalığı.

Data Mining Applications in The Education System and a Case Study On Awareness

Abstract: Educational data mining is a relatively new field in the data mining approach. This study aims to explore awareness on the use of data mining technology in Turkish education system. A survey was prepared for this purpose; the demographic characteristics of the interviewees which was teachers, their relevance to technology and the internet, approaches to e-school system, the awareness of data mining technology and the understanding of the application of data mining technology in the education system has been acquired. SPSS software was used for statistical analysis of the obtained data.

Keywords: Ekonomik Küreselleşme, İşsizlik, Panel Veri Analizi, Yükselen Piyasa Ekonomileri.

Makale Geliş Tarihi: 17.06.2018

Makale Kabul Tarihi: 16.10.2018

*) Doç.Dr. Öğretim Üyesi, Atatürk Üniversitesi İİBF Yönetim Bilişim Sistemleri Bölümü (e-posta: abdulcadir@atauni.edu.tr)

**) Öğr. Gör. Milli Savunma Üniversitesi, Hava Harp Okulu/Bilgisayar Mühendisliği Bölümü (e-posta: rsaylam@hvkk.tsk.tr)

***) Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Yönetim Bilişim Sistemleri A.B.D (e-posta: b.b_bilen@hotmail.com)

I. Giriş

Veri madenciliği, iş süreçlerinin etkinliğini ve verimliliğini artırmak için farklı kuruluş türlerinde kullanılan nispeten yeni bir teknolojidir. Veri madenciliği teknikleri, depolanan veri değerleri arasındaki yeni ilişkileri keşfetmek için büyük veri kümelerini analiz eder (Diwani ve Sam, 2014). Dijital veri miktarının artması ancak bu verileri analiz eden insan sayısında değişiklik olmaması, bu analizler için veri madenciliği tekniklerinin kullanımına zorlamıştır (Savaş vd., 2012).

İş dünyasında yaygın olarak kullanılan veri madenciliği eğitim alanında çeşitli uygulama alanlarına sahiptir (Yukselturk vd., 2014). Eğitim alanında, veri madenciliği ve analizleri de büyük bir dönüşüm potansiyeline sahiptir: insanların öğrenmesini keşfetmek, öğrenmeyi tahmin etmek ve gerçek öğrenme davranışını anlamak için kullanılabilir. Böylelikle eğitim sistemi çıktılarında önemli bir kalite düzeyi yakalanabilir. Eğitimsel veri madenciliği, bu hedeflere ulaşarak daha iyi ve daha akıllı öğrenme teknolojisi tasarlamak ve öğrenenleri ve eğitimcileri daha iyi bilgilendirmek için kullanılabilir (Baker, 2014).

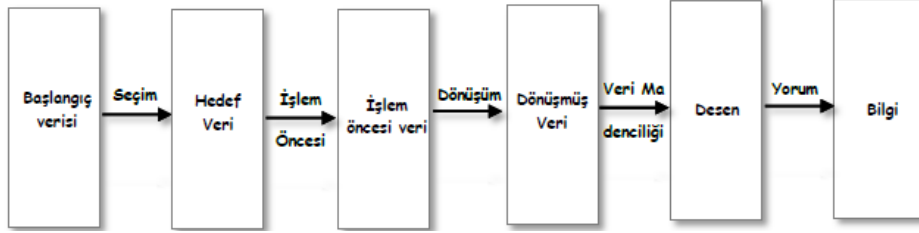
Bu çalışmada sırasıyla; veri madenciliği ve uygulama alanları kısaca tanımlanmış, eğitim sisteminde veri madenciliği uygulamaları incelenerek, öğrenme teknikleri, öğrenci performansı tahmini ve öğrencinin okulu bırakması öngörülerinin ortaya konulabildiği görülmüştür. Bir durum çalışması olarak; Erzurum ilindeki teknik liselerde öğretmenlere yapılan EK’te yer alan anket çalışması analizlerine göre özellikle eğitim alanında veri madenciliği farkındalığı keşfedilmeye çalışılmış ve son olarak çalışma sonuç bölümüyle tamamlanmıştır.

II. Veri Madenciliği Uygulama Alanları ve Veri Madenciliği Modelleri

Madencilik terimi genellikle, mali/ekonomik değeri olan herhangi bir maddeyi toprağın altından/içerisinden yeryüzüne çıkarıp, o maddeyi işleyerek ya da işlemeden paraya dönüştürme işlemi olarak bilinir. Benzer şekilde veriler de, veri tabanlarından çıkarılmayı bekleyen değerli madenler gibidir. Günümüzde mobil iletişim araçlarının kullanımının yaygınlaşması İnternet kullanımını daha da yaygın hale getirmiştir. Bilgi ve iletişim teknolojilerindeki ilerlemeler verilere erişimi ve miktarını da artırmıştır. Teknolojinin sunduğu imkânlar sayesinde büyük miktarda ve farklı türlerdeki veriler farklı özelliklerdeki veri tabanlarında depolanmaktadır. Veri tabanlarında depolanan bu verilerden anlamlı sonuçlar üretmek için veriler üzerinde çeşitli yöntem ve teknikler kullanarak birtakım işlemler yapılması gerekmektedir. Veri madenciliği daha önceden bilinmeyen, geçerli ve uygulanabilir bilgilerin geniş veri tabanlarından elde edilmesi ve bu bilgilerin işletme kararlarında kullanılmasıdır (Silahtaroglu, 2013).

Veri madenciliği, aynı zamanda bir süreçtir. Diwani ve Sam (2014) ise veri madenciliğini büyük ve çoğunlukla organize edilmemiş veri tabanlarından faydalı bilgi çıkarma süreci olarak tanımlamıştır. Veri yığınları arasından çeşitli işlemler yaparak veriyi ortaya çıkarmanın yanı sıra, bilgi keşfi sürecinde örüntüleri ayrıştırarak süzmek ve bir sonraki adıma hazır hale getirmek de bu sürecin bir parçasıdır. Bu süreç Şekil 1’de

gösterilmiştir.



Şekil 1 : Veri Madenciliğinde Bilgi Keşfetme Süreci

Çıkarılan bilgiler daha sonra bir tahmin veya sınıflandırma modeli oluşturmak, mevcut bir modeli rafine etmek veya ambar haline getirilen veri tabanının bir özetini sağlamak için kullanılabilir (Diwani ve Sam, 2014).

Veri madenciliğinde tahmin etmeye ve tanımlamaya yönelik iki tür model bulunmaktadır. Tahmin edici modeller, mevcut verilerden hareketle sonucu bilinmeyen veriler üzerinde tahmin üretmeyi amaçlarken tanımlayıcı modeller ise veri tabanındaki verinin genel özelliklerini tanımlarlar (Özbay, 2015).

Günümüzde veri madenciliği uygulamalarıyla birçok alanda karşılaşmak mümkündür. Otomatik veri toplama araçları, yaygın bilgi teknolojileri ile birlikte veri ambarlarından büyük verileri anlamlandırmak daha kolay hale gelmiştir. Aşağıda yer alan alanlar veri madenciliği tekniklerinin kullanıldığı alanlardan bazılarıdır (Silahtaroglu, 2013).

Pazarlama Yönetimi:

Müşterilerin satın alma örüntülerinin belirlenmesi: Müşteriler herhangi bir ürünü aldıktan sonra, anlamlı bir sıklıkla başka bir ürünü alıyor mu? DVD satın alanların %36'sı 2 hafta sonra boş DVD ve stabilo kalem ve DVD zarfı alıyor. Evet bu değerli bir bilgi ancak veri madenciliği, %7.7'si 15 gün içinde, mağazaya tekrar gelip elektronik dışında bir ürün alıyorlar gibi kestirilmesi güç bilgileri getirmelidir.

Müşterilerin demografik özellikleri arasındaki bağlantıların bulunması: Müşterilerin, yaşı, eğitim durumu, medeni hali vs. ile satın aldıkları ürünler arasında herhangi bir korelasyon var mı? Yaşı 25-33 arasında olup DVD yazıcı alan evli erkeklerin %36'sı süt alır.

Posta kampanyalarında cevap verme oranının artırılması: Kampanyaya sadece katılması en muhtemel kişileri dâhil etmek bir çözüm olabilir. Amaç geçmiş deneyimlere (verilere) bakarak olası genel eğilimi belirlemektir.

Mevcut müşterilerin elde tutulması, yeni müşterilerin kazanılması: Telekomünikasyon şirketleri tarafından sıklıkla kullanılmaktadır. Hangi kullanıcının ne zaman hat değiştireceğini kestirmek veri madenciliğinin işidir.

Pazar sepeti analizi: A ürününü alan müşteri yanında başka bir B ve C gibi ürün alıyor mu? B ürününü alan kişinin aynı gün C ürününü alma olasılığı nedir?

Risk Yönetimi ve Dolandırıcılık Saptama: Kredi kartı dolandırıcılığı, İnternet işlemleri; e-nakit dolandırıcılığı, Sigorta dolandırıcılığı, Bilgisayar sistemleri ve bilgisayar ağlarına girme gibi başlıca örnekler göz önüne alındığında; kişinin deneyimlerinden çok kurumun deneyimleri yani verileri dev-reye sokmak gerekmektedir.

Diğer Uygulamalar: Biyoloji (DNA sıra analizi) : Veri madenciliğiyle geliştirilen sıralama örnek analizi ve benzerlik arama yöntemleri DNA (100.000 gen) verisi üzerinde analiz yapmayı kolaylaştırır.

Tıp: Gebelik esnasında çocukta oluşabilecek herhangi bir down sendromu riskinin belirlenmesinde dış bulgular ve veri madenciliği teknikleri devreye girmektedir.

Veri Madenciliği Modelleri: Veri madenciliğinde kullanılan modellerden (Çalış vd., 2014);

Tanımlayıcı modellerde daha çok veriler arasındaki gizli kalmış ilişki, örüntülerin tanımlanması sağlanır.

Tahminleyici modellerde sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlanmaktadır.

İşlevlerine göre Veri madenciliği modellerini gruplarsak (Çalış vd., 2014);

Sınıflama ve Regresyon: Sınıflama, kategorik değerleri tahmin ederken, regresyon, süreklilik gösteren değerlerin tahmin edilmesinde kullanılır. Sınıflandırma, bir veri ögesini, önceden tanımlı sınıflardan birine tasnif ederken, regresyon veri ögesini gerçek değerli bir tahmini değışkene eşler. Örneğin bir sınıflama modeli, banka kredi uygulamalarının güvenli veya riskli olmalarını kategorize etmek amacıyla kurulurken, regresyon modeli, geliri ve mesleği verilen potansiyel müşterilerin bilgisayar ürünleri alırken yapacakları harcamaları tahmin etmek için kurulabilir.

Sınıflama ve regresyon modellerinde kullanılan başlıca teknikler: Yapay Sinir Ağları, Genetik Algoritmalar, K-En Yakın Komşu (K-Nearest Neighbour), Naive-Bayes sınıflayıcısı, Lojistik Regresyon ve Karar Ağaçlarıdır.

Kümeleme: Elde bulunan veriler incelenerek birbirine benzeyenler bir kümeye, benzemeyenler ise bir başka kümeye toplanır. Önceden belirlenmiş bir sınıflandırma yoktur. Olsaydı kümeleme değil sınıflandırma olurdu.

Bağlantı Analizi: Veri tabanındaki her bir kayıt veya kayıtlar grubu arasında bir bağlantı ilişki yaratılmaya çalışılır. Birliktelik kuralları (bir ürün satın alındığında yanında bir ürün satın alınması), Örüntü tanıma (El, yüz, resim, metin tanıma), Ardışık zaman örüntüleri (Ardışık zamanlarda yapılan alışveriş ya da paternler) ve Benzer zaman

keşfi bu başlık altında yer alır (Silahtaroglu, 2013).

III. Eğitim Sisteminde Veri Madenciliği Uygulamaları

Eğitimsel Veri Madenciliği (EDM), eğitsel bağlamdan gelen benzersiz veri türlerini keşfetme yöntemlerini geliştirme ile ilgilenen gelişmekte olan bir disiplindir. EDM'nin yöntemleri, büyük veri koleksiyonlarından örtülü ve ilginç kalıpların otomatik çıkartılmasıdır. Veri madenciliği yararlı modelleri belirlemek, öğretmenlerin öğretimine yönelik daha objektif geribildirim almak ve öğrencilerin nasıl öğrendiği hakkında daha fazla bilgi edinmek için web etkinliğini değerlendirmek için e-öğrenme verilerini keşfetmek, görselleştirmek ve analiz etmek için uygulanabilir (Magdin, 2015).

Eğitim alanındaki veri madenciliği gelişimi, diğer alanlarla karşılaştırıldığında oldukça hızlı olmuştur. Eğitimde veri madenciliği ile ilgili ilk makaleler ve çalıştaylar sadece on yıl önce başladı; 2008'de ilk konferans serisi (Uluslararası Eğitimsel Veri Madenciliği Konferansı) ve 2009'da ilk dergi (Eğitim Veri Madenciliği Dergisi) yayınlanmaya başladı. Eğitimde veri madenciliğinin daha sonra ortaya çıkmasının nedenlerinden biri kayıtların genel olarak kağıt formlarda dolapların içinde depolanmasıdır. Çevrimiçi öğrenme sistemlerinden gelen verilerde ise çoğunlukla zor kullanılan biçimlerde depolanıyor ya da önemli bilgiler hiç saklanmıyordu (Baker, 2014).

Öğrenme stilleri geçmişten itibaren geniş bir araştırma konusu olmuştur. Magdin (2015) veri madenciliği yöntemlerini uygulayarak, öğrenci için çalışma materyalinin kişiselleştirilmesi üzerine bir inceleme yapmıştır. Her birey için uygun bir öğrenme tekniği seçildiği ve çalışma materyalinin her öğrencinin imkân ve kabiliyetlerine göre ayarlandığı durumda öğrencilerin dersi akıcı bir şekilde geçtiği görülüyor.

Eğitim sistemini süreçlerden ve öğelerden oluşan bir sistem olarak ele alan bir tanımda eğitim, belli amaçlar ve sonuçlara ulaşmak üzere, planlanmış ve düzenlenmiş, eğitim ortamı, eğitici, eğitilen ve eğitim materyalinden oluşan bir bütün şeklinde tanımlanmaktadır (Keskinkılıç vd., 2006).

Süreç iyileştirmeye dayalı kalite anlayışı ile eğitim sisteminin kalitesini ve dolayısıyla eğitim sistemi çıktılarının kalitesini artırmak için eğitim süreçlerinin iyileştirilmesi gerekmektedir. Bu doğrultuda eğitim süreçlerini iyileştirerek eğitim sisteminin ve çıktılarının kalitesini artırmak için; öğrencilerin davranışlarını analiz ederek, performansı öngörerek ve ders yazılımı geliştiricilerine eksikliklerini tespit etme konusunda yardımcı olacak ve yazılım süreçlerinin iyileştirecek geribildirim sağlanmalıdır. Bu noktada önemli olan öğrencilerin şu ana kadar karar vericiler tarafından bilinmeyen tutum ve davranışlarına odaklanmak gerekmektedir (Sorour vd., 2017).

Osmanbegović ve Suljić (2012)'un çalışmasında, 2010-2011 öğretim yılında İktisat Fakültesi öğrencileri üzerinde yapılan anket ile; demografik veriler ve geçmiş başarıları ve üniversitedeki başarısı ile ilgili veriler toplanmıştır. Bu analiz, algoritmaların eğitilmesinden ve test edilmesinden sonra, öğrencilerin muhtemel başarısı üzerine

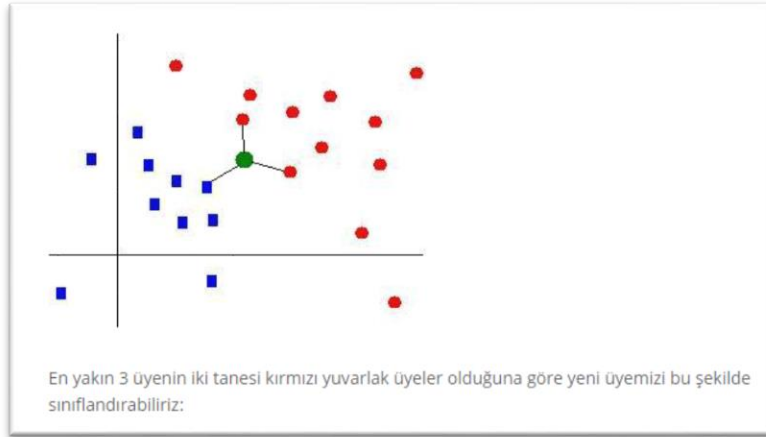
sonuçlar çıkarılmasını mümkün kılmıştır.

Yurdakul ve Topal (2015)'ın öğrenci başarı ya da başarısızlığını etkileyen etmenleri belirleyerek, birbirleri arasındaki ilişkileri ortaya koydukları çalışmada; öğrenci-öğretmen ve öğrenci-öğrenci ilişkisinin iyi olması, öğrencinin bir çalışma odasının olması arasında başarıyı artıran bir ilişki bulunduğu ortaya konmuştur. Ayrıca öğrencinin çok kardeşi olmaması, annenin üniversite mezunu olması da başarıyı artıran etmenler arasında bulunmaktadır.

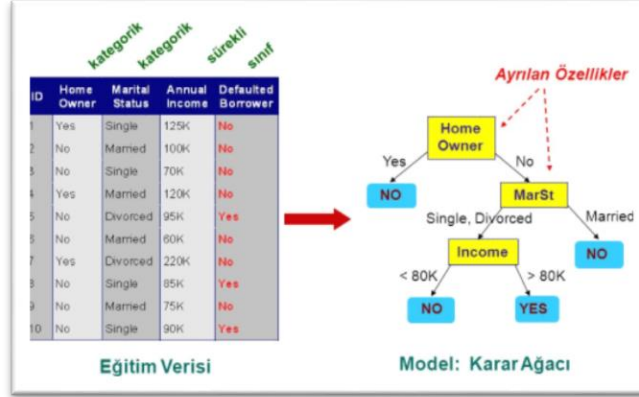
Öğrencinin serbest yorum verilerini her dersin ardından toplayarak öğrenci performansını tahmin etmeye çalışan Sorour ve arkadaşları (2017) öğrenme analitiği ve veri madenciliği yaklaşımlarını sentezlemiştir.

Okulu bırakma eğitimde ciddi bir sorundur. Öngörüsü ve çözülmesi kolay değildir. Márquez-Vera ve arkadaşları (2016), Meksika'da okul bırakma oranı en yüksek olan lise eğitiminde Meksikalı öğrencilerin verilerini kullanarak bir vaka çalışması yapmıştır. Okulu bırakma konusunu ele alan ilk adımda IF-THEN kuralları, öğrencinin okula devam etmesine veya okuldan ayrılmasına neden olan göstergeleri ve ilişkileri ortaya koymuştur. Bir sonraki adımda, bırakma tehlikesiyle karşı karşıya olan her öğrencinin kendine has ihtiyaçlarını ve sorunlarını belirlemek ve daha sonra etkili ve uygun bir bırakma önleme stratejileri sağlamak için programlar uygulamaktır.

Yükseltürk ve arkadaşları (2014)'nın çalışmasında ise çevrimiçi öğrenmede terk sorununun çözümünde veri madenciliği kullanımının yararlı olup olmayacağı incelenmiştir. Okulu bırakan öğrencileri sınıflandırmak için, en yaygın veri madenciliği yaklaşımları, Şekil 2 k-En Yakın Komşu (k-NN), Şekil 3 Karar Ağacı (DT), Naive Bayes (NB) ve Sinir Ağları (NN) temel alınarak uygulanmıştır. Bu dört farklı sınıflandırma algoritması, 10 kat çapraz doğrulama tekniği kullanılarak eğitilmiş ve test edilmiştir.



Şekil 2: k-En Yakın Komşu



Şekil 3: Karar Ağacı

IV. Türk Eğitim Sisteminde Veri Madenciliği Farkındalığı - Bir Durum Çalışması

A. Amaç

Milli Eğitim Bakanlığına bağlı okullarda yaklaşık 18 milyon öğrenci eğitim görmektedir. Bu zengin öğrenci potansiyeli göz önüne alındığında, öğrencilere ait kayıtlı verilerin işlenmeye hazır madenler olduğu düşünülebilir. Bu kapsamda çalışmanın amacı, veri madenciliği tekniklerini kullanarak eğitim sisteminin temel unsurlarından biri olan öğretmenlerin bu potansiyelle ilgili kullanım düzeylerini tespit etmek ve farkındalıklarını ortaya koymaktır.

B. Veri Toplama

Eğitim sisteminin temel unsurlarından biri olan öğretmenlerin Veri Madenciliği tekniklerini kullanma potansiyelleri ve farkındalıklarını ortaya koymak amacıyla yürütülen bu çalışmada veriler, 2016-2017 öğretim yılında Mart ayı itibariyle Erzurum'da teknik liselerde çalışan öğretmenler arasında yapılan anketler ile toplanmıştır. Toplam 15 maddeden oluşan anket soruları genel anlamda Demografik, Teknoloji, Eğitimde Teknoloji ve Veri madenciliği kapsamaları altında gruplanmıştır.

C. Araştırmanın Sınırlılıkları

Bu çalışmada kolay ulaşılabilir örneklem kullanılmıştır. Çalışma Erzurum Palandöken ilçesinde bulunan Kazım Karabekir Mesleki ve Teknik Anadolu Lisesi ve Yakutiye ilçesinde bulunan Atatürk Mesleki ve Teknik Anadolu Lisesi ile Erzurum Mesleki ve Teknik Anadolu Lisesi öğretmenleriyle sınırlıdır.

D. Analiz ve Bulgular

Ankette tanımlayıcı istatistik analizler SPSS aracı ile yapılmıştır. Ankete katılan 100 öğretmene sorulan tüm soruların frekans değerleri çıkarılmıştır. Katılımcıların %46'sı

kadın %54'ü erkek, %51'inin hizmet süresi 10 yılın altında %26'sı 11-20 yıl arasında %23'ü 20 yılın üzerindedir.

Katılımcıların %89'u teknolojinin günlük çalışma faaliyetleri üzerinde olumlu etkisi olduğunu düşünmekte, %97'si eğitimle ilgili faaliyetlerin internet üzerinden takip edilmesini öğrenci öğretmen ve veliler için faydalı bulmaktadır. Veri madenciliği kavramını biliyor musunuz sorusuna ise katılımcıların %53'ü bilmiyor cevabını verirken, %30 duymuştum ve yalnızca %17'si biliyorum cevabını vermiştir. Veri madenciliği farkındalığını tespit etmek için veri madenciliği kavramı da açıklanarak (geçmiş verilere dayalı tahminler, analizler yapılabilir vb.) sorulan doğrudan ve dolaylı sorulara verilen cevaplar incelendiğinde çoğunlukla ortada yani 'kısmen' cevabı verilmiştir.

Teknolojinin ve eğitim faaliyetlerinin internetten takip edilmesini çok büyük oranda faydalı bulan öğretmenlerin, veri madenciliği ile ilgili sorulara büyük oranda kararsız cevaplar vermesi bu alanda farkındalığın yetersiz olduğuna işaret etmektedir.

Demografik özelliklerden cinsiyetin, teknoloji ve veri madenciliği soruları ile ilişkisi Tablo 1, 2, 3 ve 4'te t-testi ile incelendiğinde; kadınların erkeklere göre hem teknoloji hem de veri madenciliği yaklaşımında istatistiksel olarak %95 (p-değeri 0.05'ten küçük) güven aralığında biraz daha az ilgili olduğu söylenebilir. Ancak bu farklılık Eğitimde teknoloji sorularında ortaya çıkmamıştır.

Tablo 1: Teknoloji Grup İstatistikleri

Group Statistics					
	Cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
Teknoloji_ortalama	Kadin	46	1.5163	.36667	.05406
	Erkek	54	1.3519	.38399	.05225

Tablo 2 : Teknoloji t-testi

Independent Samples Test										
		Levene's Test for Equality of Variances			t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Teknoloji_ortalama	Equal variances assumed	.060	.806	2.179	98	.032	.16445	.07547	.01469	.31422
	Equal variances not assumed			2.187	96.698	.031	.16445	.07519	.01522	.31369

Tablo 3 : Veri Madenciliği Grup İstatistikleri

Group Statistics					
	Cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
VeriMaden_ortalama	Kadin	46	1.9620	.43769	.06453
	Erkek	54	1.7824	.43450	.05913

Tablo 4 : Veri madenciliği t-test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
VeriMaden_ortalama	Equal variances assumed	.012	.914	2.053	98	.043	.17955	.08747	.00596	.35314
	Equal variances not assumed			2.051	95.264	.043	.17955	.08753	.00579	.35330

Veri madenciliği farkındalığının hangi özelliklere göre değişiklik gösterdiğini tespit etmek amacıyla farklı değişkenler ele alınarak One-Way ANOVA testleri yapılmıştır. Bu testler sonucunda; veri madenciliği yaklaşımı ile öğretmenlerin yaş ve hizmet süreleri (Yaş ve Hizmet süresi orantılı olduğundan çok farklı bir sonuç çıkması beklenmezdi) arasındaki ilişki Tablo 5 ve 6'da ortaya konmuş ve istatistiksel olarak %95 güven aralığında (p-değeri 0.05'ten büyük) bir ilişki olmadığı görülmüştür.

Tablo 5 : Yaş-Veri Madenciliği ANOVA Testi

ANOVA					
VeriMaden_ortalama	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.446	2	.223	1.141	.324
Within Groups	18.981	97	.196		
Total	19.428	99			

Tablo 6: Hizmet Yılı-Veri Madenciliği ANOVA Testi

ANOVA					
VeriMaden_ortalama	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.633	2	.316	1.633	.201
Within Groups	18.795	97	.194		
Total	19.428	99			

Veri madenciliği ile ilişkili çarpıcı bir tespit ise branşlardan çıkmıştır. İstatistiksel olarak %95 güven aralığında (p-değeri $0.01 < 0.05$) faktörler arasındaki ilişki Tablo 7’de gösterilmiştir. Tablo 8’de yer alan %95’lik güven aralığı değerleri incelendiğinde; Fen, Sosyal ve Sağlık branşlarında görev yapan öğretmenlerin Mesleki ve Teknik Branşlarda (bilgisayar, elektronik vb.) görev yapan öğretmenlere göre veri madenciliği yaklaşımına daha uzak oldukları istatistiksel olarak söylenebilir.

Tablo 7: Hizmet Yılı-Branş ANOVA Testi

ANOVA					
VeriMaden_ortalama	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.033	3	1.011	5.920	.001
Within Groups	16.395	96	.171		
Total	19.428	99			

Tablo 8 : Branş Tanımlayıcılar

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Fen	16	2.0000	.41833	.10458	1.7771	2.2229	1.25	2.75
Sosyal	27	2.0000	.42743	.08226	1.8309	2.1691	1.00	2.50
Mesleki ve Teknik	44	1.6705	.39903	.06016	1.5491	1.7918	1.00	2.50
Sağlık	13	2.0769	.42555	.11803	1.8198	2.3341	1.25	2.75
Total	100	1.8650	.44299	.04430	1.7771	1.9529	1.00	2.75

Teknoloji ve Eğitimde teknolojiye ilişkin sorularda ise yaş, hizmet süresi ya da branş değişkenleri ile aralarında anlamlı bir ilişki ortaya çıkmamıştır.

Teknoloji, Eğitimde teknoloji ve Veri madenciliği soruları arasındaki ilişkileri belirlemek amacıyla yapılan korelasyon testi Tablo 9’da verilmiştir. Bu sonuçlara göre %90 güven aralığında Veri Madenciliği soruları ile Teknoloji ve Eğitimde teknoloji soruları arasında pozitif korelasyon gözlemlenmiştir.

Tablo 9 : Sorular Arasındaki Korelasyon Analizi

Correlations				
		VeriMaden_ortalama	EgitimTek_ortalama	Teknoloji_ortalama
VeriMaden_ortalama	Pearson Correlation	1	.179	.273**
	Sig. (2-tailed)		.075	.006
	N	100	100	100
EgitimTek_ortalama	Pearson Correlation	.179	1	.089
	Sig. (2-tailed)	.075		.377
	N	100	100	100
Teknoloji_ortalama	Pearson Correlation	.273**	.089	1
	Sig. (2-tailed)	.006	.377	
	N	100	100	100

V. Sonuç

İş dünyasında yaygın olarak kullanılan veri madenciliği ve analizleri eğitim alanında da büyük bir dönüşüm potansiyeline sahiptir.

Bu çalışmada, eğitim sisteminde veri madenciliği uygulamalarına ait literatürde yer alan çalışmalar incelenmiş ve Türk eğitim sisteminde veri madenciliği farkındalığını ortaya koymak amaçlanmıştır. Farkındalığı tespit etmek amacıyla hazırlanan, 15 sorudan oluşan anket Erzurum ilinde yer alan teknik okullarda görev yapan 100 öğretmen tarafından doldurulmuştur. Ankette yer alan sorular genellediğinde; Demografik, Teknoloji, Eğitimde Teknoloji ve Veri Madenciliği başlıklarının bulunduğu söylenebilir. Anket sonuçları SPSS istatistik programına aktarılmış ve analiz edilmiştir.

Katılımcıların %53'ünün veri madenciliği kavramını bilmediği, %30'unun daha önce duyduğu ve yalnızca %17'sinin bu kavramı tam olarak bildiği ortaya çıkmıştır. İlişkili sorularda da durum çok farklı değildir. Buradan çıkarımla birçok eğitimcinin, veri madenciliği teknolojisinin farkında olmadığı ve daha önce hiç kullanmadığı söylenebilecektir.

Yapılan t-testte istatistiksel olarak kadınların erkeklere göre hem teknoloji hem de veri madenciliği yaklaşımında biraz daha az ilgili olduğu, One-Way ANOVA testlerinde veri madenciliği yaklaşımı ile öğretmenlerin hizmet süreleri arasında anlamlı bir ilişki bulunmadığı ancak branşlar ile veri madenciliği arasında anlamlı bir ilişki bulunduğu, bu ilişkide Mesleki ve Teknik Branş (bilgisayar, elektronik vb) öğretmenlerinin diğer öğretmenlere göre veri madenciliği yaklaşımına daha yakın olduğu sonucu çıkarılmıştır. Ayrıca Veri Madenciliği soruları ile Teknoloji ve Eğitimde teknoloji soruları arasında pozitif korelasyon gözlemlenmiştir.

Eğitim Sisteminde Veri Madenciliği Uygulamaları başlığı altında yer alan çalışmalar da göstermiştir ki bu alanda veri madenciliği uygulanması hem eğitim tekniklerinin geliştirilmesi hem de öğrenci performans veya davranışlarının önceden tahmin edilmesi anlamında katkılar sağlamaktadır. Ülkemizde de eğitim sisteminde veri madenciliği farkındalığının artırılması ve eğitimde veri madenciliği uygulamalarının geliştirilmesinin büyük fayda sağlayacağı değerlendirilmektedir.

Kaynaklar

Baker, R.S., (2014). "Educational Data Mining: An Advance for Intelligent Systems in Education", *IEEE Intelligent Systems*, 29, 3.

Çalış, A.,Kayapınar S. Ve Çetinyokuş T. (2014). "Veri Madenciliğinde Karar Ağacı Algoritmaları ile Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama", *Endüstri Mühendisliği Dergisi*, 25 (3-4), (2-19).

Diwani, S. A. ve Sam, A., (2014). "Data Mining Awareness and Readiness in Healthcare Sector: a case of Tanzania", *Advances in Computer Science: an International Journal*, 3(1), No.7.

- Keskinkılıç, M. ve Işıldak, M., S., (2006). “Bilgisayarlı Muhasebe Eğitiminin Sistem Yaklaşımıyla İncelenmesi”, *Mali Çözüm Dergisi*, İSMMMO Yayın Organı, 2006, 75, Nisan-Mayıs, 131-140.
- Magdin, M., (2015). “Personalization of Student in Course Management Systems on the Basis Using Method of Data Mining”, *The Turkish Online Journal of Educational Technology*, 14 1.
- Márquez-Vera, C., Cano, A., Romero, C., Noaman, A.Y.M., Fardoun, H.M. and Ventura, S., (2016). “Early Dropout Prediction Using Data Mining: A Case Study With High School Students”, *Expert Systems*, 33, 1.
- Osmanbegović, E. ve Suljić, M., (2012). “Data Mining Approach for Predicting Student Performance”, *Economic Review – Journal of Economics and Business*, X, 1.
- Özbay, Ö. (2015). “Veri Madenciliği Kavramı ve Eğitimde Veri Madenciliği Uygulamaları”, *Uluslararası Eğitim Bilimleri Dergisi*.2, 5.
- Savaş, S., Topaloğlu, N. ve Yılmaz, M. (2012). “Veri Madenciliği ve Türkiye’deki Uygulama Örnekleri”, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 11 21.
- Silahtaroglu, G., (2013). Veri madenciliği, Papatya Yayınları, İstanbul.
- Sorour, S.E., Goda, K. ve Mine, T., (2017). “Comment Data Mining to Estimate Student Performance Considering Consecutive Lessons”, *Educational Technology & Society*, 20 (1), 73–86.
- Yukselturk, E. Ozekes, S. ve Türel, Y.K., (2014). “Predicting Dropout Student: An Application of Data Mining Methods in an Online Education Program”, *European Journal of Open, Distance and e-Learning*, 17 / No. 1.
- Yurdakul, S. ve Topal T., (2015). Veri Madenciliği ile Lise Öğrenci Performanslarının Değerlendirilmesi, 17. Akademik Bilişim Konferansı.

EK

ANKET SORULARI

1.Cinsiyetiniz:

- a. Kadın
- b. Erkek

2.Yaşınız:

- a. 35'in altında
- b. 35-45
- c. 45'in üzerinde

3.Hizmet süreniz:

- a. 1-10 yıl
- b. 11-20 yıl
- c. 20 yıl üzeri

4.Branřınız:

- a. Fen Bilimleri (Fizik, Matematik...)
- b. Sosyal Bilimler(Tarih,Cođrafya,...)
- c. Mesleki ve Teknik B. (Bilgisayar,Elektronik,...)
- d. Sađlık Bilimleri

5.Teknolojinin günlük çalıřma faaliyetleriniz üzerinde olumlu etkisi var mı?

- a. Evet, kesinlikle var.
- b. Kararsızım.
- c. Hayır, yok.

6.Yeni teknolojileri kullanmaya ilgili misiniz?

- a. Evet, çok ilgiliyim.
- b. Çok fazla meraklı deđilim.
- c. Hiç ilgili deđilim.

7.Kaç yıldır internet kullanıyorsunuz?

- a. 10 yıldan fazladır.
- b. 5-10 yıldır.
- c. 5 yıldan az zamandır.

8.Günde kaç saat internette geçiriyorsunuz?

- a. 3 saatten fazla
- b. 1-3 saat
- c. 1 saatten az

9.Eđitimle iliřkili faaliyetlerin internet üzerinden takip edilmesi hakkında ne düşünöyorsunuz?

- a. Öđrenci, öđretmen ve veliler için faydalı buluyorum.
- b. Veliler için faydalı buluyorum.
- c. Zaman kaybı olduđunu düşünöyorum.

- 10. Eğitimle ilgili kayıtların dosyada evrak olarak saklanması yerine sistemlere girilmesi, veri tabanlarında tutulması gerekli midir?**
- Evet, kesinlikle.
 - Kısmen gereklidir.
 - Gereksizdir.
- 11. e-okul sistemini hangi amaçla kullanıyorsunuz?**
- Öğrenci bilgilerini edinmek ve tüm işlemleri yapmak için.
 - Not girişi ve ödev vermek için
 - Not girişi için.
- 12. Veri madenciliği kavramını biliyor musunuz?**
- Evet, biliyorum.
 - Daha önce duymuştum.
 - Hayır, bilmiyorum.
- 13. Mevcut eğitim yöntemlerine ek olarak öğrenci verilerinin analizinde veri madenciliği teknolojisi kullanmanın faydalı olacağını düşünüyor musunuz?**
- Faydalı olacağını düşünüyorum.
 - Ne gibi bir faydası olabilir bilmiyorum.
 - Faydalı olacağını düşünmüyorum.
- 14. Öğrencilerin geçmiş bilgilerine dayanarak performansı tahmin edilebilir mi?**
- Evet, kesinlikle edilebilir.
 - Kısmen edilebilir.
 - Hayır öğrenci heran değişebilir.
- 15. Veri madenciliği teknikleri kullanarak, öğrencilerin hangi alana yönelmeleri (fen, sosyal vb.) konusunda öneriler verilebilir mi?**
- Evet, kesinlikle verilebilir.
 - Kısmen verilebilir.
 - Verilemez.