

**Öğrenme Yönetim Sistemi Üzerindeki Öğrenci  
Hareketliliğinin Veri Madenciliği Yöntemleriyle Analizi**  
**Analysis of Student Dynamism into Learning Management  
System through Data Mining Methods**

Özkan ÖZBAY<sup>1</sup>, Halil ERSOY<sup>2</sup>

<sup>1</sup>Jandarma ve Sahil Güvenlik Akademisi, Sosyal Bilimler Bölümü,  
ozkan.ozbay@msn.com

<sup>2</sup>Başkent Üniversitesi, Bilgisayar ve Öğretim Teknolojileri Öğretmenliği Programı,  
hersoy@baskent.edu.tr

**Makalenin Geliş Tarihi: 10.12.2016**

**Yayına Kabul Tarihi: 13.12.2016**

**ÖZ**

*Bu çalışmada lisans mezunu öğrencilerin Öğrenme Yönetim Sistemi (ÖYS) üzerindeki hareketliliği ile akademik başarı düzeyleri arasındaki ilişki veri madenciliği yöntemleri kullanılarak incelenmiştir. Bu bağlamda bu çalışma, tarama modellerinden ilişkisel tarama modeline uygun olarak düzenlenmiş betimsel ve nicel bir çalışmadır. Çalışmanın örneklemini, Başkent Üniversitesi Eğitim Bilimleri Enstitüsü Pedagojik Formasyon Sertifika Programı 2012–2013 eğitim - öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersini alan 40 öğrenci oluşturmaktadır. Çalışmada veri kaynağı olarak öğrencilerin ilgili eğitim öğretim yılına ait Moodle (Modular Object Oriented Dynamic Learning Environment), ÖYS üzerindeki hareketliliğini içeren log kayıtları ve yılsonu akademik başarı notları kullanılmıştır. Çalışma sonucunda elde edilen bulgulara bağlı olarak öğrencilerin ÖYS üzerindeki hareketliliği ve akademik başarı düzeyleri arasında anlamlı bir ilişkinin olduğu ortaya konulmuştur.*

**Anahtar Sözcükler:** Moodle, Öğrenme Yönetim Sistemleri, Veri Madenciliği, Öğrenme Analitiği, Eğitsel Veri Madenciliği.

**ABSTRACT**

*In this study, undergraduates were examined to discover the connection between their dynamism into Learning Management System (LMS) and their levels of achievement with using the data mining methods. In this regard, the study was designed in accordance with relational screening model and it is descriptive and quantitative. The research includes 40 pupils who took the course Instructional Technologies and Material Design during spring semester and resided in Baskent University Institute of Education Sciences Pedagogical Formation Certification Program throughout 2012-2013 academic years. Students' academic achievement grades and log*

*registrations placing in LMS which shows related year's access were used as data source. Depending upon the findings which were acquired at the end of the study, it was revealed that there is a meaningful connection between the activities on LMS and the academic achievement levels of students.*

**Keywords:** *Moodle, Learning Management Systems, Data Mining, Learning Analytics, Educational Data Mining.*

## GİRİŞ

Bilgisayar teknolojilerinin gelişmesi, veri depolama sistemlerinin yaygınlaşması ve internet erişiminin artması çok çeşitli verilerin kaydedilmesini ve bu verilerin manyetik ortamlarda saklanmasını kolay ve ucuz hale getirmiştir. Ancak bu şekilde üretilen ve depolanan veriler tek başlarına değersizdir. Bu veriler belli bir amaç doğrultusunda işlendiği zaman bir anlam ifade etmeye başlarlar (Kalikov, 2006).

Veriler, bilgisayar sistemleriyle belirli bir amaca yönelik işlenerek bilgiye dönüşmektedir (Kurt ve Erdem, 2012). Bu yüzden günümüzde büyük miktarlardaki verileri işleyip kullanılabilir hale getiren teknikler büyük önem kazanmaya başlamıştır. Ham veriyi bilgiye veya anlamlı hale dönüştürme işlemleri ise veri madenciliği ile yapılabilmektedir (Kalikov, 2006). Veri madenciliği, kullanıcılara yeni yöntemlerle anlaşılabilir ve faydalı olan verileri özetlemek ve aralarındaki beklenmeyen ilişkileri bulmak için genellikle büyük gözlemsel veri kümelerinin analiz edilmesi olarak tanımlanmaktadır (Hand, Mannila ve Smyth, 2001). Kısaca veri madenciliği, büyük miktarlardaki veri setlerinde saklı durumda bulunan örüntü ve eğilimleri keşfetme işlemidir (Thuarisingham, 2003).

Veri madenciliğinde temelde tahmin edici ve tanımlayıcı olmak üzere iki ana model bulunmaktadır. Tahmin edici modeller, eldeki veri yığınlarını kullanarak bir modelin geliştirilmesini ve oluşturulan bu modeli kullanarak sonuçları önceden bilinmeyen veri yığınları için sonuçların tahmin edilmesini amaçlamaktadır (Seven, 2009). Özellikle karar alma sürecinde önemli bir role sahiptir. Kendi içinde sınıflandırma ve eğri uydurma olmak üzere ikiye ayrılmaktadır. Tahmin edici modeller içerisinde en yaygın kullanıma sahip olan ise sınıflandırma modelidir (Han ve Kamber, 2006; Arslan, 2008). Tanımlayıcı modeller karar verme aşamasında rehberlik etmede kullanılacak veri örüntülerinin tanımlanmasını sağlamaktadır. Genelde sepet analizi olarak geçen, bir alışveriş sepetindeki ürünler arasındaki ilişkiyi ortaya çıkaran ve bir elektronik alışveriş sitesinde bir ürün seçildiğinde, müşteriye başka ürünleri de önermesi tanımlayıcı modeller ile yapılan veri madenciliğine örnek olarak verilebilir (Erten, 2015). En yaygın kullanılan

tanımlayıcı modeller kümeleme, özetleme, birliktelik kuralları ve sıra örüntüleridir (Dunham, 2003; Aynekin, 2006; Bozkır, 2009).

Veri madenciliğinde veri yığımindan bilgi elde edilmeden önce eldeki verilerin; problemin tanımlanması, verilerin hazırlanması, modelin kurulması ve değerlendirilmesi, modelin kullanılması ve modelin izlenmesi işlemlerine tabi tutulması gerekir. Veri madenciliğinde yapılan çalışmanın amaca hizmet edebilmesi için yapılması gerek ilk şey problemin açık şekilde tanımlanmasıdır. Problemin tanımlanması aşaması gerçekleştirildikten sonra verilerin hazırlanması aşamasına geçilir. Bu aşamada veri madenciliği uygulanırken oluşturulacak modelin veri kaynaklarının neler olduğunun belirlenmesi ve modelde kullanılmak için uygun hale getirilmesi gerekmektedir. Verilerin hazırlanması, toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme olmak üzere 5 aşamada gerçekleştirilir. Veriler gerekli veri hazırlama işlemlerine tabi tutulduktan sonra modelin kurulması ve değerlendirilmesi aşamasına geçilir. Veri madenciliğinde belirlediğimiz probleme uygun, bizi sonuca ulaştırabilecek en iyi modelin bulunabilmesi ancak çok sayıda modelin kurulup denenmesiyle mümkündür. Bu nedenle model kurma aşaması en iyi modele ulaşıncaya kadar yinelenen bir süreçtir. Belirlediğimiz probleme uygun olarak kurulan ve geçerliliği test edilip kabul edilen bir model, modelin kullanılması aşamasında doğrudan problemin çözümünde kullanılabilir. Modelin izlenmesi aşamasında ise zaman içerisinde problemin farklılaşması ya da verilerin değişmesi nedeniyle probleme uygun olarak kurulan ve geçerli kabul edilen bir model geçerliliğini yitireceğinden dolayı sürekli olarak izlenmesi ve yeniden düzenlenmesi gerekmektedir.

Veri madenciliği büyük miktarlarda verilerin üretilip depolandığı her alanda kullanılabilir. Veri madenciliğinin kullanıldığı bazı uygulama alanları Pazar Araştırması, Risk Analizi, Kurum Kaynaklarının Kullanımı, Sağlık, Ticaret, Alışveriş, Bankacılık, Sigortacılık ve Eğitim'dir.

Eğitim alanında veri madenciliği uygulamaları teknolojinin eğitime entegre olmasıyla birlikte artış göstermiştir. Buna bağlı olarak eğitim ortamlarından elde edilen verilerin analiz edilmesinde eğitsel veri madenciliği ve öğrenme analitiği kavramlarının ortaya

çıktığı, öğrenci, öğretmen, idari personel ve eğitim kurumlarına ait verilerin analizinde veri madenciliği tekniklerinden yararlanıldığı görülmektedir (Ünal, 2014).

Eğitimde veri madenciliği, öğrencilerin tam olarak izlenmesinin güç olduğu geleneksel eğitimde daha az kullanılırken, öğrencilerin izlenmesinin daha kolay olduğu ve öğrenme ortamlarında gerçekleştirilen pek çok öğrenci davranışının kaydedildiği internet tabanlı eğitimde daha geniş uygulama potansiyeli bulmaktadır. Eğitsel veri madenciliği ve öğrenme analitiği gibi yeni yaklaşımlar ile eğitim ortamlarından elde edilen büyük veri yığınlarının analiz edilmesi sayesinde öğrenme süreci veya öğrenci davranışları hakkında bilgi sahibi olmak mümkündür. Bu sayede öğrencilerin davranışlarının, tecrübelerinin, bilgi düzeylerinin modellenmesi, benzer öğrenci profillerinin oluşturulması, uyarlanabilir ve kişiselleştirilebilir ortamlarda kullanılacak bilgilerin üretilmesi mümkün olabilmektedir (Bienkowski, Feng ve Means, 2012).

Öğrenme Yönetim Sistemleri (ÖYS) eğitim ortamlarında eğitimciler tarafından yaygın olarak kullanılır hale gelmişlerdir. Birçok farklı ÖYS olmasına rağmen ÖYS'lerin ortak kullanım amaçları öğretimi desteklemek, öğrencinin bilgiyi kendisinin yapılandırmasını olanak tanımak, eğitim kalitesini arttırmak ve kalıcılığı arttırmak olarak sıralanabilir (Yıldız ve Bahçeci, 2014). Ancak ÖYS'lerden en üst düzeyde fayda sağlamak sadece sistemin kullanılması ile mümkün değildir. ÖYS'ler gerçekleştirilen tüm kullanıcı davranışlarını kayıt altına almaktadır. Eğitsel veri madenciliği ve öğrenme analitikleri ise bu noktada devreye girer ve öğrenci profillerinin oluşturulmasında, öğrenme ortamının kişiselleştirilmesinde, eğitim ortamının kalitesinin artırılmasında geri bildirimler sağlar.

Literatürde ÖYS'lerin yeterince esnek ve kolay bir kullanıma sahip olup olmadıkları, özelliklerinin neler olması gerektiği, uzaktan eğitimde nasıl kullanılacakları, farklı ÖYS'lerin avantaj ve dezavantajlarının belirlendiği çalışmalar olsa da (Aydın ve Büroğlu, 2008; Aydın, 2011; Ergül, 2013; Floyd, Schultz ve Fulton, 2012; Lonn, Teasley ve Krumm, 2011; Reis, Baktır, Çelik, Erkoç, Özçakır, Özdemir ve Şahin, 2012; Türker, 2012; Yapıcı ve Akbayın, 2012) veri madenciliği teknikleriyle sistem tarafından toplanan verilerin analiz edildiği, bu verilere dayanarak ÖYS üzerindeki öğrenci hareketliliğinin incelendiği çalışmalar nadirdir (Leony, Pardo, Valentin, Quinones ve Kloos, 2012; San

Diego, Ballard, Hatzipanagos, Webb, Khan, Blake, Dore, Konstantinidis ve Barrett, 2012; Whitmer, Fernandes ve Allen, 2012).

Bu araştırmanın genel amacı, veri madenciliği tekniklerini kullanarak üniversite (lisans mezunu) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile yılsonu akademik başarıları arasındaki ilişkiyi belirlemektir. Bu ilişkinin ortaya çıkarılmasında veri yığnında yer alan verilerin veri madenciliği teknikleriyle analiz edilmesi, anlaşılabilir hale getirilmesi ve daha önceden ortaya çıkarılmayan ilişkilerin ortaya çıkarılması amaçlanmıştır.

Araştırmanın genel amacı doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. Öğrencilerin ÖYS üzerindeki hareketliliği incelendiğinde öğrenciler tarafından gerçekleştirilen eylemler nasıl bir dağılım göstermektedir?
2. Öğrenciler tarafından ÖYS üzerinde gerçekleştirilen eylemler analiz edildiğinde öğrencilerin akademik başarıları nasıl ve ne kadar doğruluk oranıyla tahmin edilebilir?

## YÖNTEM

Bu çalışmada, üniversite (lisans mezunu) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarı düzeyleri arasındaki ilişki incelenmiştir. Bu bağlamda bu araştırma, tarama modellerinden ilişkisel tarama modeline uygun olarak düzenlenmiş betimsel ve nicel bir çalışmadır. Tarama modelleri, geçmişte ya da halen var olan bir durumu var olduğu şekliyle betimlemeyi amaçlayan araştırma yaklaşımlarıdır (Karasar, 2012). Tarama modelleri kendi içinde ikiye ayrılır. Bunlar, genel tarama ve örnek olay tarama modelleridir. Genel tarama modeli tekil ya da ilişkisel tarama olarak gerçekleştirilebilir. İlişkisel tarama modelleri, iki ve daha çok sayıdaki değişken arasında birlikte değişim varlığını ve/veya derecesini belirlemeyi amaçlayan araştırma modelidir (Karasar, 2012). İlişkisel tarama modelinde çalışmaya konu olan olay, birey veya nesne kendi koşulları içinde değiştirilmeye çalışılmadan tanımlanmaya çalışılır. Değişkenlere müdahale edilmemesi nedeniyle ilişkisel tarama modelleri nedensel karşılaştırma araştırmalarına benzer. Ancak nedensel karşılaştırma araştırmalarında bir bağımlı

değişkeni etkileyen bağımsız değişkenler neden-sonuç ilişkisi içinde belirlenmeye çalışılırken ilişkisel tarama modelinde sadece değişkenlerin birlikte değişimleri incelenir (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2012).

### Çalışma Grubu

Araştırmanın çalışma grubunu Başkent Üniversitesi Eğitim Bilimleri Enstitüsü 2012-2013 eğitim öğretim yılı bahar dönemi Pedagojik Formasyon Sertifika Programı, Öğretim Teknolojileri ve Materyal Tasarımı dersini alan 10'u erkek ve 30'u kız olmak üzere toplam 40 öğrenci oluşturmaktadır. Öğrencilerin tamamı Fen Edebiyat Fakültesi Matematik Bölümü mezunu öğretmen adaylarıdır. Çalışma grubunun demografik özellikleri ile ilgili veriler Tablo 1'de yer almaktadır:

**Tablo 1.** Çalışma Grubuna Ait Demografik Bilgiler

Cinsiyet	F	%
Erkek	10	25.00
Kız	30	75.00
Toplam	40	100.00

Öğretim Teknolojileri ve Materyal Tasarım dersi 5 ayrı şubede verilmekte olup tüm şubelerde ortak müfredat, ortak kitap kullanılmakta ve ortak hedefler güdülmektedir. Ancak araştırmada 1 şubeden elde edilen veriler kullanılmıştır. Derse ait notlar, klasik ve çoktan seçmeli şekilde yapılan vize ve final sınavları ile öğretim elemanı tarafından belirlenen belirli bir zaman diliminde gönderilmesi gereken 5 ödev ve öğretim elemanının kanaat notundan oluşmaktadır. Derse ait geçme sisteminin not ağırlıkları Tablo 2'de yer almaktadır:

**Tablo 2.** Ders Geçme Sistemi Not Ağırlıkları

Ödev 1	Ödev 2	Ödev 3	Ödev 4	Ödev 5	Kanaat	Vize	Final
%5	%5	%5	%5	%5	%5	%30	%40

Derslerde açık kaynak kodlu ÖYS olan Moodle kullanılmaktadır. ÖYS'nin derslerde kullanım teknik ve yöntemi ders sorumlusu öğretim elemanına aittir. ÖYS, dersi veren öğretim elemanları tarafından ders ile ilgili dosya, kaynak ve materyallerinin eklenmesi, ödevler ile ilgili açıklamaların verilmesi, ödevlerin gönderilmesi, ödev notlarının görüntülenmesi, kursa kayıtlı diğer öğrencilerle iletişim kurulması amacıyla kullanılmaktadır. Ancak ÖYS'nin derslerde kullanımı dersi veren öğretim elemanının ÖYS'yi kullanma istek ve becerisine bağlı olarak değişmektedir.

### Veri Kaynakları

Araştırma problemleri doğrultusunda öğrencilerin ilgili derse ait yılsonu akademik başarıları notları ve ÖYS üzerindeki hareketliliklerini içeren log kayıtları veri kaynağı olarak kullanılmıştır. Derse ait yılsonu akademik başarı notları ve öğrenci hareketliliğine ait log kayıtları derse giren öğretim elemanı tarafından ÖYS üzerinden indirilerek araştırmacıya verilmiştir.

### Verilerin Analizi

ÖYS üzerindeki öğrenci eylemlerinin analizinde IBM firması tarafından geliştirilmiş olan veri madenciliği uygulamalarında kullanılan görsel modelleme aracı SPSS Clementine 12.0 programı kullanılmıştır.

Çalışmanın gerçekleştirilmesinde öğrencilerin ÖYS üzerindeki hareketliliğini içeren log kayıtları ve yılsonu akademik başarı notları kullanılmıştır. Veri madenciliği uygulanırken oluşturulacak modelin veri kaynaklarının neler olduğunun belirlenebilmesi, modelin kurulabilmesi ve verilerin analiz edilebilmesi için veri madenciliği sürecinde yer alan verilerin hazırlanması adımları uygulanmıştır. Veriler gerekli veri hazırlama aşamalarına tabi tutulduktan sonra veri madenciliği sürecinin diğer adımı olan modelin kurulması ve değerlendirilmesi aşamasına geçilmiştir. Bu aşamada bizi sonuca ulaştırabilecek en iyi



modelin bulunabilmesi için çok sayıda model kurulup denenerek en iyi modele ulaşıncaya kadar yinelenmiştir. Bu aşamadan sonra probleme uygun olarak kurulan ve geçerliliği test edilip kabul edilen modelin kullanılması aşamasına geçilmiş ve elde edilen veriler analiz edilmiştir. Veri analizi sürecinde gerçekleştirilen işlemler detaylı olarak aşağıda açıklanmıştır.

### **Problemin Tanımlanması**

Bu araştırmada “Üniversite (lisans mezunu) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarı düzeyleri arasında anlamlı bir ilişki var mıdır?” sorusu problem cümlemizi oluşturmaktadır.

### **Verilerin Hazırlanması**

Bu aşamada veri toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme adımları gerçekleştirilmiştir.

1. Toplama: Veri madenciliği için gerekli olduğu düşünülen verilerin hangi kaynaklardan toplanacağına karar verilmiş ve ilgili kaynaklardan veriler elde edilmiştir. Çalışma kapsamında öğrencilerin ÖYS üzerindeki hareketliliğini içeren veriler ÖYS log kayıtlarından elde edilmiştir. Ayrıca öğrencilere ait yılsonu akademik başarı notları dersi veren ilgili öğretim elemanından temin edilmiştir.

2. Değer Biçme: Veri madenciliğinde kullanılacak olan veriler arasındaki kodlama, ölçü birimi farklılıkları gibi uyumsuzluklar kontrol edilerek verilerin ne ölçüde uyumlu oldukları belirlenmiştir. Veri madenciliği sürecinde kullanılacak olan öğrenci hareketliliğine ait veriler sadece log kayıtları aracılığıyla elde edildiğinden dolayı veriler arasında kodlama, ölçü birimi gibi farklılıklar bulunmamaktadır.

3. Birleştirme ve Temizleme: Veri madenciliği sürecinde kullanılacak olan veriler birleştirilerek hatalı ve eksik olan veriler veri tabanından silinmiştir. Log kayıtlarında 5788 satır ve 8 sütundan (Ders Adı, Tarih, Saat, IP Adresi, Kullanıcı Adı, Eylem, Eylem Bağlantıları ve Bilgi Alanı) oluşan 46304 veri kaydına rastlanmıştır. Çalışmada kullanılacak olan bu veriler veri temizleme işlemine tabi tutulmuştur. Veri temizleme

işlemi sırasında eksik ve veri madenciliği sürecinde kullanılmayacak olan veriler veri yığınının çıkarılmıştır. Veri temizleme işlemi sonrası 430 satırlık 3440 veri kaydı (Sistem Yöneticisi ve Ders Öğretim Elemanına ait veriler) veri yığınının çıkarılmış ve 5358 satır ve 8 sütundan oluşan 42864 veri kaydı veri madenciliği sürecinde kullanılmak üzere değerlendirmeye tabi tutulmuştur. Ayrıca öğrencilerin yılsonu akademik başarı notları veri birleştirme işlemine tabi tutularak veri temizleme işlemi sonucu elde edilen veri yığına eklenmiştir. Veri birleştirme ve temizleme işlemi sonrasında SPSS Clementine programı kullanılarak elde edilen veri yığına ait bilgiler Şekil 1’de yer almaktadır.

Şekil 1’de yer alan veri yığına ait bilgiler, birbirinden farklı her bir veri sütununda kaç farklı değişkenin olduğunu ve bu değişkenlere ait toplamdaki veri kaydını göstermektedir. Ders adı sütununda 1 farklı ders adının, tarih sütununda 105 farklı tarihin, kullanıcı adı sütununda 40 farklı kullanıcının, eylem sütununda 16 farklı eylemin, eylem bağlantıları sütununda 86 farklı eylem bağlantısının, bilgi sütununda 57 farklı bilginin yer aldığı ve her bir sütunda 5358 tane kaydın olduğu görülmektedir.

Field	Sample Graph	Type	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
Ders Adı		Flag	--	--	--	--	--	1	5358
Tarih		Set	--	--	--	--	--	105	5358
IP Adresi		Set	--	--	--	--	--	--	5358
Kullanıcı Adı		Set	--	--	--	--	--	40	5358
Öğrenci Akademik Başarı		Range	59.000	86.000	77.999	5.736	-0.472	--	5358
Eylem		Set	--	--	--	--	--	16	5358
Eylem Bağlantıları		Set	--	--	--	--	--	86	5358
Bilgi		Set	--	--	--	--	--	57	3729

Şekil 1. Veri Birleştirme ve Temizleme İşlemi Sonrası Veri Yığını

*4. Seçim:* Veri madenciliği sürecinde kullanılacak olan modelin özelliğine göre verilerin seçim işlemi yapılmıştır. Derse ait Ders Adı, Tarih, Saat, IP Adresi, Kullanıcı Adı, Eylem, Eylem Bağlantıları, Bilgi Alanı ve Öğrenci Akademik Başarısı alanlarından problemle ilgisi ve sonuca ulaşmada bir etkisi olmayan Ders Adı, Tarih, Saat, IP Adresi, Eylem Bağlantıları ve Bilgi Alanı alanları modele dâhil edilmemiştir. Çünkü problemle ilgisi olmayan bu tip veriler diğer verilerin modeldeki ağırlıklarını azaltacaklardır. Seçim aşamasında Kullanıcı Adı, Eylem ve Öğrenci Akademik Başarısı alanları modelde kullanılacak olan veriler olarak belirlenmiştir. Bu alanlara ait bilgiler aşağıda detaylı olarak açıklanmıştır.

*Kullanıcı Adı:* Kullanıcı Adı alanında yer alan öğrenci isimleri kodlama işlemi gerçekleştirilerek kullanılmıştır. Kodlama işlemi öğrenci isimlerinin alfabetik sıraya göre sıralanarak ID\_1'den ID\_40'a kadar kod verilmesi suretiyle oluşturulmuştur.

*Eylem:* ÖYS, sistem yöneticisi, ders öğretmeni ve öğrenciler tarafından gerçekleştirilen tüm eylemleri kayıt altına almakta ve bu kayıtlar ders adı, tarih, saat, IP adresi, kullanıcı adı, eylem, eylem bağlantıları ve bilgi alanı şeklinde 8 farklı bölümden oluşmaktadır. ÖYS üzerinde sistem yöneticisi, ders öğretmeni ve öğrenci olma duruma bağlı olarak farklı eylemler gerçekleştirilebilmekte ve tüm bu kullanıcı yetkilerine ait eylemler farklı miktarlarda olmak üzere toplamda 449 eylem kayıt altına alınabilmektedir. Log kayıtları incelendiğinde öğrenciler tarafından toplamda 16 adet farklı eylemin gerçekleştirildiği tespit edilmiştir. Bu eylemler Tablo 3'te yer almaktadır:

**Tablo 3.** ÖYS Üzerindeki Öğrenci Eylemleri

Eylem (İngilizce)	Eylem (Türkçe)
Assignment view all	Tüm gönderilere bak
Assign upload	Ödev yükle
Assign view	Ödev gönderisine bak
Blog view	Blog görüntüle
Course enrol	Derse kaydol
Course recent	Ders geçmişi

Courseuser report	Ders kullanıcı raporu
Course view	Ders sayfasına bak
Forum user report	Forum kullanıcı raporu
Forum view forum	Forum sayfasına bak
Forum view forums	Tüm forumları listele
Resource view	Ders kaynağını görüntüle
Resource view all	Tüm ders kaynaklarını görüntüle
Upload upload	Dosya yükle
User view	Kullanıcı görüntüle
User view all	Tüm kullanıcıları listele

*Öğrenci Akademik Başarısı:* Seçim aşaması sonrası veri madenciliği sürecinde kullanılacak olan öğrenci yılsonu akademik başarı notları Tablo 4'te gösterilmektedir:

**Tablo 4.** Öğrenci Ders Notları

Kullanıcı ID	Cinsiyet	100'lük Notu
ID_1	Erkek	76
ID_2	Kız	83
ID_3	Kız	84
ID_4	Kız	83
ID_5	Kız	74
ID_6	Kız	83
ID_7	Kız	74
ID_8	Erkek	72
ID_9	Kız	75
ID_10	Kız	75
ID_11	Kız	85
ID_12	Erkek	59
ID_13	Kız	73
ID_14	Kız	84

---

ID_15	Kız	79
ID_16	Erkek	73
ID_17	Erkek	83
ID_18	Kız	67
ID_19	Kız	62
ID_20	Kız	77
ID_21	Kız	75
ID_22	Kız	75
ID_23	Kız	74
ID_24	Kız	72
ID_25	Kız	84
ID_26	Erkek	73
ID_27	Erkek	79
ID_28	Kız	66
ID_29	Kız	75
ID_30	Kız	69
ID_31	Erkek	86
ID_32	Kız	67
ID_33	Kız	86
ID_34	Kız	63
ID_35	Erkek	72
ID_36	Kız	78
ID_37	Kız	83
ID_38	Kız	72
ID_39	Kız	81
ID_40	Erkek	77

---

5. *Dönüştürme*: Modelde kullanılacak verilerin kodlamalar kullanılarak tanımlanması, veri madenciliği için uygun formlara dönüştürülmesi ya da gösterim şeklinin değiştirilmesi gerekmektedir. Birçok algoritma bu işlemleri otomatik olarak gerçekleştirmektedir. Öğrenci hareketliliğine ait veriler veri madenciliği sürecinde kullanılabilir uygun formlarda olduğundan dönüştürme işlemine tabi tutulmamıştır.

#### *Modelin Kurulması ve Değerlendirilmesi*

Öğrenci başarı tahminlerinin modellenmesinde tahmin edici modellerden sınıflandırma modeli olan karar ağaçlarından, öğrenci akademik başarı gruplarının oluşturulmasında tanımlayıcı modellerden olan kümeleme yönteminden yararlanılmıştır.

Veri madenciliği sürecinde öğrenci akademik başarı gruplarının oluşturulmasında kümeleme yönteminden yararlanılmıştır. Kümeleme işleminde benzer niteliklere sahip soyut ya da somut nesnelere bir grubu oluştururken bununla bağlantılı olmayanlar diğer soyut ya da somut nesnelere başka bir grubu oluşturmaktadır. Bu çalışmada kümeleme modelinin kullanılmasının amacı öğrenci profillerinin belirlenmesi, benzer özelliklere sahip öğrencilerin gruplanması ve oluşan grupların genel olarak özelliklerinin karşılaştırılabilmesidir. Ayrıca kümeleme işleminin gerçekleştirilmesi öğrenciler hakkındaki sayısal verilerin kategorik sınıflara çevrilmesini sağlayarak verilerin tahmin modellerinde kullanılmasını daha uygun hale getirecektir. Kümeleme işleminin gerçekleştirilmesinde K-Means yöntemi en iyi bilinen ve en çok kullanılan yöntemdir (Sarıman, 2011). Bu yüzden kümeleme işlemi K-Means algoritması kullanılarak gerçekleştirilmiştir. Kümeleme işlemi gerçekleştirilirken elde edilebilecek küme sayısının anlaşılabilir ölçüde olması ve sonuçların anlaşılabilirliğinin artırılabilmesi için küme sayısının ya da kümeler arasındaki benzerlik ölçüsünün belirtilmesi gerekmektedir. Öğrenci yılsonu akademik başarı notlarına göre kümeleme işleminin gerçekleştirilmesinde, öğrencilerin, geçme notunun altında kalma ihtimali olanlar, orta seviyede başarı gösterenler ve üst seviyede başarı gösterenler şeklinde 3 gruba ayrılmasına karar verilmiştir. Oluşturulacak küme sayısı belirlendikten sonra K-Means algoritması kullanılarak öğrenci notlarının birbirine olan yakınlıklarına göre gruplama işlemi gerçekleştirilmiştir. Kümeleme işlemi tamamlandıktan sonra elde edilen gruplar,

grup ortalamaları dikkate alınarak Düşük Akademik Başarı, Orta Akademik Başarı ve Yüksek Akademik Başarı olmak üzere isimlendirilmiştir.

Öğrenci başarı tahminlerinin modellenmesinde, veri madenciliği modeli olarak tahmin edici modeller içerisinde yer alan sınıflandırma modellerinden olan karar ağaçları seçilmiştir. Karar ağaçları diğer sınıflandırma modellerine göre oluşturulması basit, anlaşılması kolay, başarı oranı yüksek ve sınıflandırma modelleri içinde en çok tercih edilen model olduğundan veri madenciliği analizinde tercih edilmiştir.

Öğrencilerin başarı durumlarını tahmin etmeyi amaçlayan karar ağacı modelleri oluşturulmadan önce modellerde yer alacak girdi değişkenleri belirlenmiştir. Girdi değişkeni olarak kümeleme yöntemiyle oluşturulan akademik başarı grupları ve veri hazırlığı adımlarına tabi tutulan değişkenler kullanılmıştır. Modellerin oluşturulması sırasında bu girdi değişkenleri veri madenciliği algoritmalarından yararlanılarak analiz edilmiştir.

Karar ağaçlarının oluşturulmasında birçok algorithmadan yararlanılmaktadır. Bu çalışmada veri madenciliği algoritmalarından en sık kullanılan C5.0, CART, CHAID, QUEST algoritmaları kullanılmıştır. Birden fazla algoritma kullanılarak her bir algoritmanın başarı tahminindeki doğruluk oranları hesaplanmıştır.

#### *Modelin Kullanılması*

Bu aşamada veri madenciliği algoritmalarından C5.0, CART, CHAID, QUEST algoritmaları kullanılarak oluşturulan modellerden en yüksek doğruluk oranını veren CART algoritması kullanılarak karar ağaçları oluşturulmuştur.

#### *Modelin İzlenmesi*

Belirlediğimiz probleme uygun olarak kurulan, geçerliliği test edilip, kabul edilen ve kullanılmaya başlayan model zaman içerisinde problemin farklılaşmasına bağlı olarak ya da verilerin değişmesi nedeniyle geçerliliğini yitirebilir bundan dolayı modelin sürekli olarak izlenmesi ve bu değişen durumlara göre yeniden düzenlemesi gerekmektedir.

## BULGULAR

### Birinci Araştırma Sorusuna İlişkin Bulgular

Araştırma Sorusu: Öğrencilerin ÖYS üzerindeki hareketliliği incelendiğinde öğrenciler tarafından gerçekleştirilen eylemler nasıl bir dağılım göstermektedir?

Birinci araştırma sorusu kapsamında öğrencilerin hareketlilikleri geleneksel istatistiksel yöntemler kullanılarak incelenmiştir. ÖYS üzerinde gerçekleştirilen eylemler ve bu eylemlerin öğrenci bazında dağılım oranlarına ait bulgulara ulaşılmıştır.

Öğrenciler tarafından gerçekleştirilen 16 adet farklı eylemin toplam gerçekleştirilen 5358 adet eylem içindeki dağılım oranları incelendiğinde %56,03 ile ders kaynağını görüntüle, % 29,99 ile ders sayfasına bak eylemlerinin öğrenci hareketliliğinin büyük bir kısmını oluşturduğu görülmektedir. Şekil 2’de 16 adet farklı eylemin toplamda gerçekleştirilen eylemler içerisindeki dağılımları ayrıntılı olarak görülmektedir:

Value	Proportion	%	Count
Ders kaynağını görüntüle	56,03	56,03	3002
Ders sayfasına bak	29,99	29,99	1607
Ödev gönderisine bak	4,4	4,4	236
Tüm kullanıcıları listele	3,04	3,04	163
Kullanıcı görüntüle	2,18	2,18	117
Tüm gönderilere bak	1,57	1,57	84
Derse kaydol	0,75	0,75	40
Forum sayfasına bak	0,65	0,65	35
Ders geçmiş	0,49	0,49	26
Tüm ders kaynaklarını görüntüle	0,43	0,43	23
Forum kullanıcı raporu	0,17	0,17	9
Blog görüntüle	0,13	0,13	7
Tüm forumları listele	0,09	0,09	5
Ders kullanıcı raporu	0,04	0,04	2
Ödev yükle	0,02	0,02	1
Dosya yükle	0,02	0,02	1

Şekil 2. Eylemlerin Dağılım Oranları

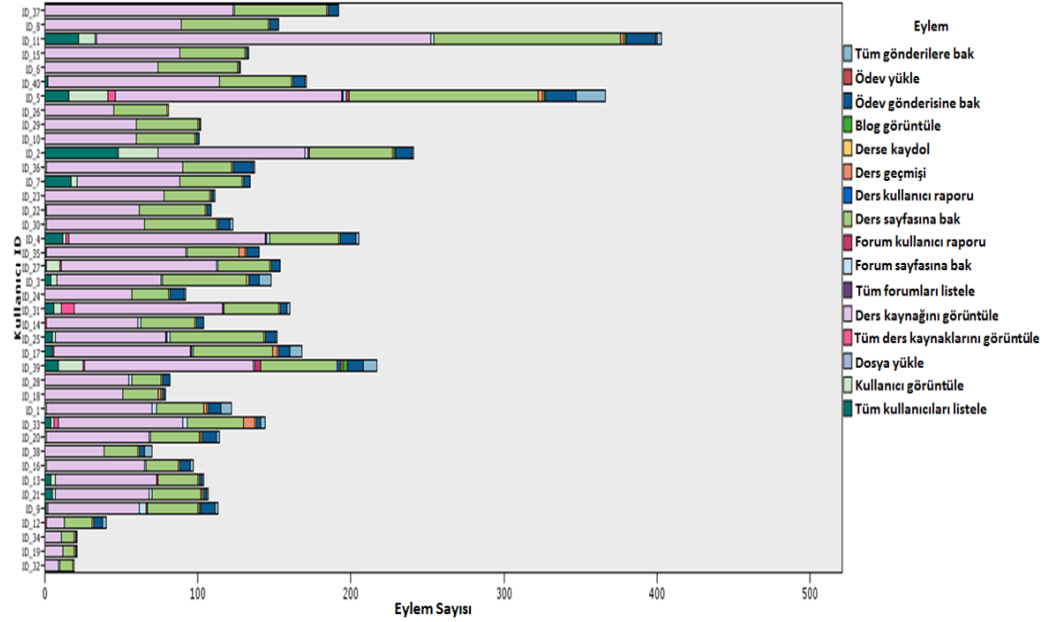
Toplam eylem sayısının (5358 Adet Eylem) öğrenci bazında dağılımı incelendiğinde minimum eylem sayısının (19 Adet Eylem) ID\_32 adı ile kodlu öğrenci ve maksimum eylem sayısının (403 Adet Eylem) ID\_11 adı ile kodlu öğrenci tarafından gerçekleştirildiği tespit edilmiştir. Öğrenci bazında gerçekleştirilen toplam eylem sayılarının büyükten küçüğe doğru sıralanışı Tablo 5’te ayrıntılı olarak yer almaktadır.



**Tablo 5.** Öğrenci Eylem Sayısı

Kullanıcı_ID	Toplam Eylem Sayısı
ID_11	403
ID_5	366
ID_2	241
ID_39	217
ID_4	205
ID_37	192
ID_40	171
ID_17	168
ID_31	160
ID_27	154
ID_8	153
ID_25	152
ID_3	148
ID_33	144
ID_35	140
ID_36	137
ID_7	134
ID_15	133
ID_6	128
ID_30	123
ID_1	122
ID_20	114
ID_9	113
ID_23	111
ID_22	109
ID_21	107
ID_14	104
ID_13	104
ID_29	102
ID_10	101
ID_16	97
ID_24	92
ID_28	82
ID_26	81
ID_18	79
ID_38	70
ID_12	40
ID_34	21
ID_19	21
ID_32	19

ÖYS üzerinde gerçekleştirilen 5358 adet eylem içerisindeki 16 adet farklı eylemin öğrenci bazında dağılımları tespit edilmiştir. Öğrenci eylem dağılımları Şekil 3'te görülmektedir.



Şekil 3. Öğrenci Eylem Dağılımı

Şekil 3 incelendiğinde, öğrencilerin ders içeriklerine erişebilmeleri, öğretim elemanı tarafından paylaşılan dosya ve ödev gönderilerini görebilmeleri için derse kaydol eylemini gerçekleştirdikleri görülmektedir. Bu eylem ilgili derse kaydolabilmek için öğrenciler tarafından mutlaka bir defa gerçekleştirilmektedir. Öğretim elemanı tarafından öğrencilerden tartışma forumu, blog, ödev gönderisi yükleme ve dosya yükleme işlemlerini gerçekleştirmeleri istenmemesine rağmen öğrencilerin bir kısmının bu işlemlerle ilgili olan forum sayfasına bak, forum kullanıcı raporu, tüm forumları listele, blog görüntüle, ödev yükle ve dosya yükle eylemlerini gerçekleştirdikleri görülmektedir. Derse kayıtlı bazı öğrencilerin diğer arkadaşlarını görüntülemek için tüm kullanıcıları listele eylemini gerçekleştirdikleri ve derse kayıtlı bir arkadaşının profilini incelemek için kullanıcı görüntüle eylemini gerçekleştirdikleri görülmektedir. Ders geçme sisteminin

%25'lik kısmını oluşturan ders öğretmeni tarafından belirlenen ödevlere ait ödev gönderi durumunu, not ve son teslim tarihini görüntülemek için öğrencilerin ödev gönderisine bak eylemini gerçekleştirdikleri görülmektedir. Ayrıca öğrencilerin ÖYS'ye en son giriş yaptıkları tarihe kadar olan tüm ödev gönderilerini görüntülemek için tüm gönderilere bak eylemini gerçekleştirdikleri görülmektedir. Öğrencilerin derse kaydedildikleri tarihten itibaren gerçekleştirdikleri ders aktivitelerini, ders kaynak ve aktivite güncellemelerini, forum ve ödev gönderilerini görüntülemek için ders geçmişi eylemini ve ders içerisinde kendilerine ait gerçekleştirdikleri aktiviteleri görüntülemek için ders kullanıcı raporu eylemini gerçekleştirdikleri görülmektedir. Ayrıca öğrencilerin derse ait tüm kaynakları görüntülemek için tüm ders kaynaklarını görüntüle eylemini gerçekleştirdikleri görülmektedir.

Öğrenciler tarafından gerçekleştirilen tüm bu eylemler dikkate alındığında ders kaynağını görüntüle, ders sayfasına bak ve derse kaydol eylemlerinin tüm öğrenciler tarafından gerçekleştirildiği diğer eylemlerin ise tüm öğrenciler tarafından gerçekleştirilmediği görülmektedir.

### **İkinci Araştırma Sorusuna İlişkin Bulgular**

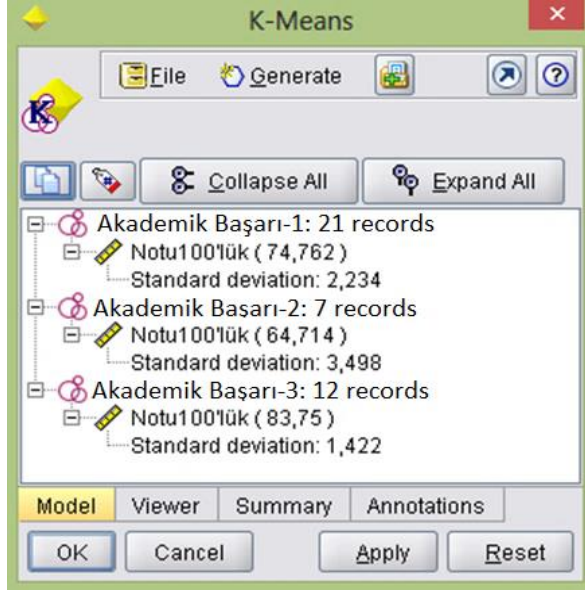
*Araştırma Sorusu: Öğrenciler tarafından ÖYS üzerinde gerçekleştirilen eylemler analiz edildiğinde öğrencilerin akademik başarıları nasıl ve ne kadar doğruluk oranıyla tahmin edilebilir?*

*İkinci araştırma sorusu kapsamında öğrencilerin gerçekleştirdikleri eylemler disiplinler arası bir alan olan veri madenciliği teknikleri kullanılarak incelenmiştir. Tanımlayıcı ve tahmin edici veri madenciliği modelleri kullanılarak öğrencilerin akademik performanslarına ilişkin bulgulara ulaşılmıştır.*

#### *Tanımlayıcı Modellerinin Kullanılması*

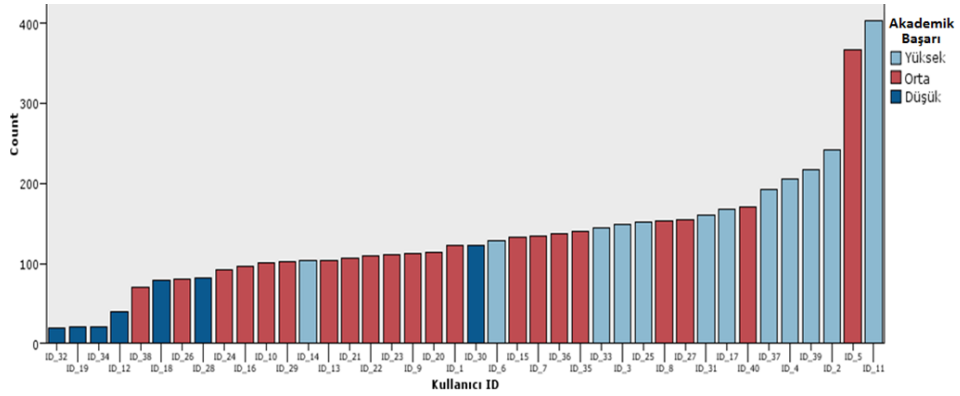
Öğrenci akademik başarı gruplarının oluşturulmasında kümeleme algoritmalarından olan K-Means algoritmasından yararlanılmıştır. K-Means algoritmasının kullanılması sonucunda elde edilen 3 grubun ortalamaları dikkate alınarak Akademik Başarı-1 isimli grup ortalaması 74,762 olan 21 kayıt Orta Akademik Başarı, Akademik Başarı-2 isimli

grup ortalaması 64,714 olan 7 kayıt Düşük Akademik Başarı, Akademik Başarı-3 isimli grup ortalaması 83,75 olan 12 kayıt Yüksek Akademik Başarı grubu olarak adlandırılmıştır. Şekil 4'te K-Means algoritması tarafından oluşturulan akademik başarı grupları gösterilmektedir.



Şekil 4. Akademik Başarı Grupları

Yüksek Akademik Başarı, Orta Akademik Başarı ve Düşük Akademik Başarı grubunda yer alan öğrencilerin yaptıkları eylem sayılarına göre dağılımı Şekil 5'te görülmektedir:



**Şekil 5.** Akademik Başarı Gruplarında Yer Alan Öğrencilerin Yaptıkları Eylem Sayılarına Göre Dağılımı

Öğrencilerin yer aldıkları akademik başarı grupları ile gerçekleştirdikleri eylem ilişkileri incelendiğinde yüksek akademik başarı grubunda yer alan öğrencilerin analizde yer alan 16 eylemi de gerçekleştirdikleri ve özellikle ders kaynağını görüntüle ile ders sayfasına bak eylemlerini daha çok gerçekleştirdikleri, orta akademik başarı grubunda yer alan öğrencilerin analizde yer alan 13 eylemi gerçekleştirdikleri ve yüksek akademik başarı grubuna benzer olarak özellikle ders kaynağını görüntüle ile ders sayfasına bak eylemlerini daha çok gerçekleştirdikleri, düşük akademik başarı grubunda yer alan öğrencilerin ise sadece 9 eylemi gerçekleştirdikleri tespit edilmiştir. Tablo 6'da eylemler ve akademik başarı grupları tarafından gerçekleştirilme durumları görülmektedir.

**Tablo 6.** Eylemler ve Akademik Başarı Grupları Tarafından Gerçekleştirilme Sayıları

Eylemler	Akademik Başarı						Toplam	
	Yüksek		Orta		Düşük		f	%
	F	%	f	%	f	%	f	%
Tüm gönderilere bak	37	0.70	43	0.80	4	0.07	84	1.57
Ödev yükle	1	0.02	-	0.00	-	0.00	1	0.02
Ödev gönderisine bak	91	1.70	122	2.27	23	0.43	236	4.4
Blog görüntüle	5	0.09	2	0.04	-	0.00	7	0.13
Derse kaydol	12	0.22	21	0.40	7	0.13	40	0.75
Ders geçmişi	13	0.25	11	0.20	2	0.04	26	0.49
Ders kullanıcı raporu	2	0.04	-	0.00	-	0.00	2	0.04
Ders sayfasına bak	659	12.3	818	15.3	130	2.39	1607	29.99
Forum kullanıcı raporu	5	0.09	4	0.08	-	0.00	9	0.17
Forum sayfasına bak	17	0.31	15	0.28	3	0.06	35	0.65
Tüm forumları listele	4	0.07	1	0.02	-	0.00	5	0.09
Ders kaynağını görüntüle	1216	23.0	1572	29.0	214	4.03	3002	56.03
Tüm ders kaynaklarını görüntüle	16	0.30	6	0.11	1	0.02	23	0.43
Dosya yükle	1	0.02	-	0.00	-	0.00	1	0.02
Kullanıcı görüntüle	68	1.26	48	0.90	1	0.02	117	2.18
Tüm kullanıcıları listele	115	2.14	48	0.90	-	0.00	163	3.04
<b>Toplam</b>	<b>2262</b>	<b>42.51</b>	<b>2711</b>	<b>50.3</b>	<b>385</b>	<b>7.19</b>	<b>5358</b>	<b>100</b>

Öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarıları arasındaki ilişkinin belirlenmesinde kullanılan C5.0, CART, CHAID, QUEST algoritmaların doğruluk oranları Tablo 7’de yer almaktadır.

**Tablo 7.** Tahmin Modellerinin Doğruluk Oranları

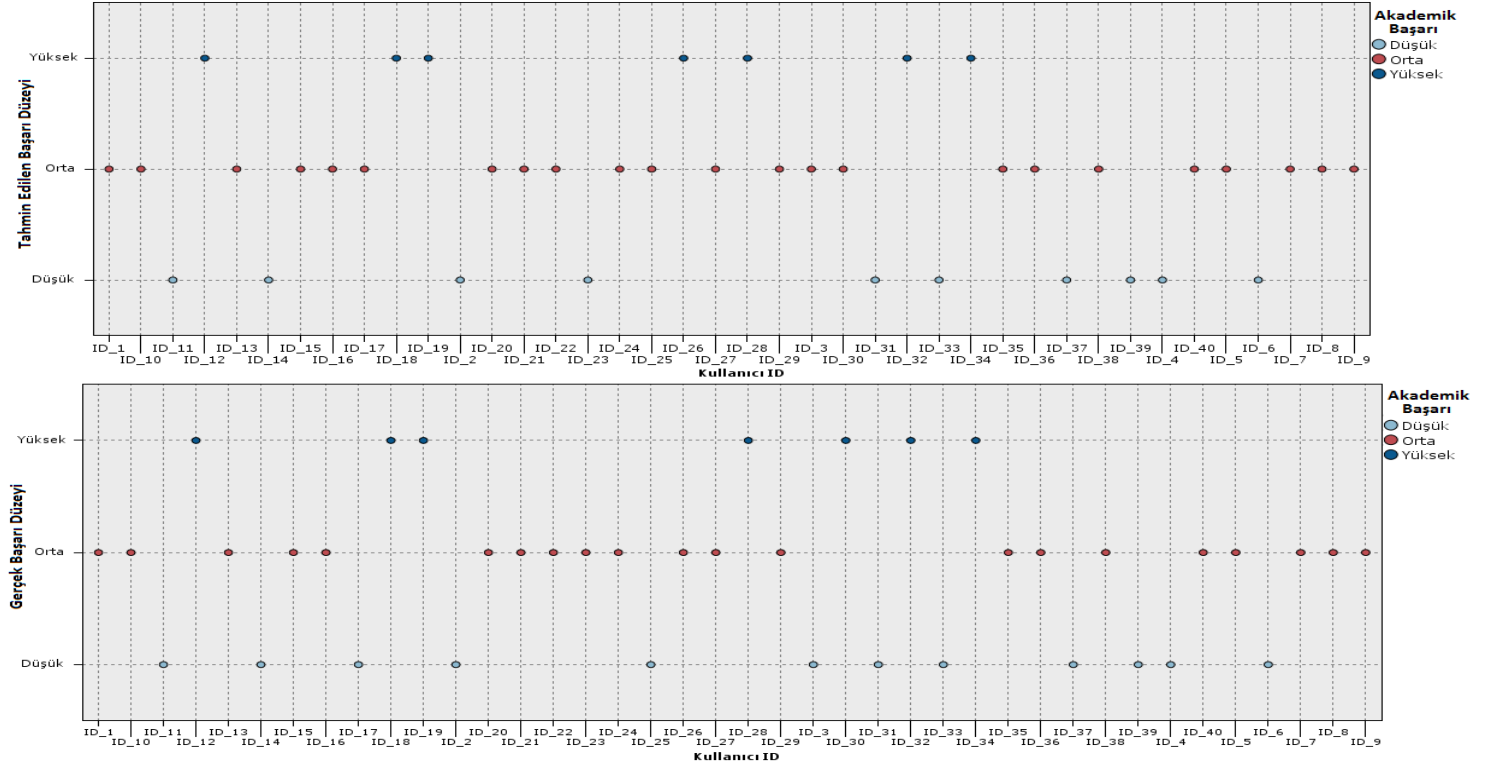
Algoritma	Doğruluk Oranı	
	Doğru	Yanlış
CART	%85.0	%15.0
C5.0	%82.5	%17.5
CHAID	%65.0	%35.0
QUEST	%65.0	%35.0

*Tahmin Edici Modellerinin Kullanılması*

CART algoritması için doğruluk oranının %85, hata oranının %15, C5.0 algoritması için doğruluk oranının %82.5, hata oranının %17.5, CHAID algoritması için doğruluk oranının %65, hata oranının %35, QUEST algoritması için doğruluk oranının %65, hata oranının ise %35 olduğu görülmektedir. Sınıflandırma algoritmalarıyla kurulan modellerin hata ve doğruluk oranları göz önüne alındığında %85 doğruluk oranı ile en yüksek doğruluk oranına CART algoritması kullanılarak ulaşılmıştır. Bu nedenle karar ağaçlarının oluşturulmasında ve başarı tahmininin gerçekleştirilmesinde CART algoritmasından yararlanılmıştır.

İlk olarak başarı düzeylerine göre gruplanan (Yüksek Akademik Başarı, Orta Akademik Başarı ve Düşük Akademik Başarı) öğrencilerin gerçekleştirdikleri eylemlerle olan ilişkileri belirlenmiştir. Öğrenciler tarafından gerçekleştirilen 16 adet farklı eylemden toplam eylem yüzdesi içerisinde %86.02 oranla ve tüm öğrenciler tarafından gerçekleştirilen ders sayfasına bak ve ders kaynağını görüntüle eylemleri girdi değişkeni olarak belirlenmiştir.

Veri yığınına CART algoritması uygulandıktan sonra öğrencilerin gerçekleştirdikleri eylem ilişkilerine göre akademik başarı düzeyleri belirlenmiştir. Şekil 6'da öğrencilerin tahmin edilen ve gerçek akademik başarı düzeyleri ayrıntılı olarak görülmektedir.



Şekil 6. Tahmin Edilen ve Gerçek Akademik Başarı Düzeyleri



Şekil 6 incelendiğinde CART algoritması kullanılarak tahmin edilen akademik başarı düzeyi ve öğrencilerin gerçek akademik başarı düzeyleri karşılaştırıldığında, ID\_17 adı ile kodlu öğrencinin gerçekte yüksek akademik başarı grubunda yer alırken tahmin edilende orta akademik başarı grubunda, ID\_23 adı ile kodlu öğrencinin gerçekte orta akademik başarı grubunda yer alırken tahmin edilende yüksek akademik başarı grubunda, ID\_25 adı ile kodlu öğrencinin gerçekte yüksek akademik başarı grubunda yer alırken tahmin edilende orta akademik başarı grubunda, ID\_26 adı ile kodlu öğrencinin gerçekte orta akademik başarı grubunda yer alırken tahmin edilende düşük akademik başarı grubunda, ID\_3 adı ile kodlu öğrencinin gerçekte yüksek akademik başarı grubunda yer alırken tahmin edilende orta akademik başarı grubunda, ID\_30 adı ile kodlu öğrencinin gerçekte düşük akademik başarı grubunda yer alırken tahmin edilende orta akademik başarı grubunda yer aldığı görülmektedir.

### **TARTIŞMA ve SONUÇ**

Veri madenciliği veri yığını içerisinde gizli kalmış anlamlı bilgilere ulaşmayı sağlayan bir süreçtir. Veri madenciliğinin eğitim alanında kullanımına baktığımızda eğitsel veri madenciliği ve öğrenme analitiği kavramlarının ortaya çıktığı dikkat çekmektedir. Eğitim alanında yer alan öğelere (öğrenci, öğretmen, ders) yönelik yapılacak veri analizleri ile eğitimde kalite artışı sağlanabilmektedir (Ünal, 2014). Eğitim ortamlarında eğitsel veri madenciliği çalışmaları geleneksel eğitim ortamlarından elde edilen veriler ve uzaktan eğitim sistemleri üzerinden elde edilen verilerle olmak üzere iki şekilde gerçekleştirilmektedir. Geleneksel eğitim ortamlarında öğrenci verilerinin toplanması daha zor iken uzaktan eğitim sistemlerinde öğrenci hareketleri kayıt altına alındığından öğrenci verilerinin toplanması ve analiz edilmesi daha kolaydır. Geleneksel eğitim ve uzaktan eğitim sistemlerinde yaygın olarak kullanılan ÖYS'ler gerçekleştirilen tüm öğrenci hareketlerini kayıt altına aldıklarından eğitsel veri madenciliği çalışmalarına önemli bir veri kaynağı sağlamaktadır.

Bu çalışmada veri madenciliği teknikleri kullanılarak üniversite (lisans mezunu) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile yılsonu akademik başarıları

arasındaki ilişkinin belirlenmesi amaçlanmıştır. Çalışma kapsamında öğrencilerin ÖYS log kayıtları ve yılsonu akademik başarı notları veri kaynağı olarak kullanılmıştır. Elde edilen veri kaynakları veri madenciliği sürecine tabi tutulmuştur. Bu süreçte belirlenen değişkenler veri madenciliği teknikleri kullanılarak analiz edilmiştir.

### **Birinci Araştırma Sorusuna İlişkin Sonuçlar**

Öğrencilerin ÖYS üzerinde gerçekleştirdikleri eylemlerin dağılımının belirlenmesinde geleneksel istatistiksel yöntemlerden yararlanılmıştır. Gerçekleştirilen 16 adet farklı eylem içerisinde en çok gerçekleştirdikleri eylemlerin ders kaynağını görüntüle ve ders sayfasına bak eylemleri olduğu görülmüştür. Bu eylemler derse kayıtlı tüm öğrenciler tarafından farklı oranlarda gerçekleştirilmiştir.

### **İkinci Araştırma Sorusuna İlişkin Sonuçlar**

Öğrencilerin akademik başarı durumlarının belirlenmesinde disiplinler arası bir alan olan eğitsel veri madenciliği ve öğrenme analitiğinin temelini oluşturan veri madenciliği tekniklerinden yararlanılmıştır. Bu doğrultuda öğrenci performanslarına ilişkin bir tahmin modelinin oluşturulması hedeflenmiştir.

Veri madenciliği uygulamalarında ilk olarak kümeleme modeli oluşturulmuştur. Kümeleme modelinin oluşturulmasında K-Means algoritması kullanılarak öğrencilerin düşük, orta ve yüksek akademik başarı grubu olmak üzere 3 gruba ayrılmasına karar verilmiştir. Bu işlemin gerçekleştirilmesi sonucu K-Means algoritması tarafından her akademik başarı grubu için hesaplanan ortalama değer temel alınarak öğrencilerin akademik başarı durumlarına göre bu gruplara yerleştirildikleri görülmektedir. Kümeleme analizi sonucu oluşturulan grupların öğrenci özellikleri hakkında bilgi sağlayabileceği düşünülmektedir. Ayrıca karar ağaçlarının oluşturulmasına ve elde edilen sonuçların daha kolay yorumlanabilmesine katkı sağlayacağı söylenebilir. Halees (2008) tarafından gerçekleştirilen çalışmada ise öğrencilere ait kişisel kayıtlar, akademik kayıtlar, kurs kayıtları ve e-öğrenme sistemiyle ilgili kayıtlar kullanılarak öğrenci davranışlarının değerlendirilmesi ve performanslarının iyileştirilmesi amaçlanmıştır. Veriler veri madenciliği teknikleri kullanılarak analiz edilmiştir. Kümeleme algoritmaları

kullanılarak öğrenciler davranışlarına göre kümelendi, bu kümeleme sonucu çok aykırı olan davranışlar ve bir arada olan davranışlar tespit edilmiştir. Amershi ve Conati (2009) tarafından gerçekleştirilen çalışmada ise çevrimiçi öğrenme ortamında yer alan öğrencilerin etkileşim verileri kullanılarak kümeleme algoritmaları yardımıyla öğrenciler yüksek öğrenme ve düşük öğrenme şeklinde iki kümeye ayrılmıştır.

Öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarıları arasındaki ilişkinin belirlenmesi ve başarı düzeylerinin tahmin edilmesinde veri madenciliği tekniklerinden olan karar ağaçları kullanılmıştır. Karar ağaçlarının oluşturulmasında en iyi sonucu verecek algoritmanın seçilmesi zor ve büyük öneme sahip bir işlemdir (Osmanbegović ve Suljić, 2012; Romero ve Ventura, 2013). Bu amaçla en çok kullanılan yöntem farklı sınıflama algoritmalarını kullanarak en yüksek doğruluk oranını veren algoritmanın seçilmesidir (Aydın, 2007; Romero, Ventura, Hervás, ve Gonzales 2008; Lopez, Luna, Romerove Ventura, 2012; Osmanbegović ve Suljić, 2012). Ancak karar ağaçlarının oluşturulmasında kullanılacak algoritmanın doğruluk oranının problemin çözümünde yeterli seviyede olup olmadığına da bakılması gerekmektedir. Veri madenciliğinde belirlenen problemde kullanılacak algoritmaların doğruluk oranları için genel bir kriter belirlemek zordur. Örneğin bazı problemlerde %60 doğruluk oranı çok iyi sayılabilirken bazılarında ise %99.9 beklenmedik bir durum olabilmektedir. Hatta bazı durumlarda 0,001 altındaki bir olasılıkla var olan hatalar yakalanmak istenebilir. Dolayısıyla problem çözümünde kullanılacak olan algoritmalarının doğruluk oranının yeterli olup olmadığının belirlenmesinde önemli olan şey problemin tanımıdır. Ayrıca kullanılan algoritmaların istenilen çıktıyı doğru vermesi ve benzer problemlerde elde edilen başarı oranının üzerinde ya da buna yakın seviyede olması gerekmektedir (Şeker, 2013).

Bu çalışmada da benzer bir yaklaşım izlenerek tahmin modelinin oluşturulmasında C5.0, CART, CHAID, QUEST algoritmaları kullanılarak doğruluk oranları incelenmiştir. Bu algoritmalar içerisinde en yüksek doğruluk oranını veren CART algoritmasının doğruluk oranı %85.0 olarak bulunmuştur. Yani gerçekte var olan durum ile karar ağacı kullanılarak oluşturulan tahmin edilen durumun birbirine benzerlik oranının %85 olduğunu göstermektedir. ÖYS üzerindeki öğrenci hareketliliği kullanılarak

gerçekleştirilen benzer çalışmalarda kullanılan algoritmaların doğruluk oranları incelendiğinde, Lopez ve diğerleri (2012) tarafından gerçekleştirilen ÖYS forum kullanım verilerinin öğrenci ders başarısının önemli bir göstergesi olup olmadığının test edildiği çalışmada BayesNet algoritması kullanılarak %87.5 doğruluk oranı elde edilmiştir. Romero, Ventura, Herv'as ve Gonzales (2008) tarafından gerçekleştirilen ÖYS forum kullanım verilerine ek olarak öğrencilerin ÖYS kullanımlarını yansıttığı düşünülen 9 farklı değişkende seçilerek öğrenci ders başarısını tahmin etmek amacıyla kullanılan sınıflandırma algoritmalarında %60 - %70 arasında doğruluk oranı elde etmişlerdir. Macfadyen ve Dawaon (2010) tarafından gerçekleştirilen Blackboard öğrenme ortamındaki etkileşim verileri kullanılarak ders başarısını tahmin etmek için kurulan regresyon modelinde %89 doğruluk oranı elde etmişlerdir.

Çalışma sonucunda elde edilen bulgulara bağlı olarak öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarı düzeyleri arasındaki ilişki ortaya çıkarılmaya çalışılmıştır. Şekil 5'te de görüldüğü üzere ÖYS üzerinde yüksek düzeyde hareketlilik sergileyen öğrencilerin yüksek akademik başarı grubunda, orta düzeyde hareketlilik sergileyen öğrencilerin orta akademik başarı grubunda ve düşük düzeyde hareketlilik sergileyen öğrencilerin düşük akademik başarı grubunda yer aldıkları söylenebilir. Yapılan benzer çalışmalarda da öğrencilerin ÖYS kullanımı ve akademik performansları arasında güçlü bir ilişkinin olduğu görülmüştür. Whitmer, Fernandes ve Allen (2012) tarafından Kaliforniya Eyalet Üniversitesinde gerçekleştirilen çalışmada ÖYS üzerinde daha çok zaman harcayan ve daha yüksek ziyaret sayısına sahip olan öğrencilerin daha yüksek akademik başarıya sahip oldukları belirtilmiştir. Beer, Clark ve Jones (2010) tarafından Queensland Üniversitesinde yapılan diğer bir çalışmada ise çevrimiçi eğitim alan lisans öğrencilerinin kurs sayfasını ziyaret etme sayısı ile final notları arasında anlamlı bir ilişki olduğu belirtilmiştir. Kurs sayfası ziyaret sayısı daha yüksek olanların daha yüksek final notlarına sahip oldukları söylenmiştir. Yine benzer olarak Akçapınar (2014) tarafından gerçekleştirilen çalışmada çevrimiçi öğrenme ortamlarında daha az aktivite gösterenlerin derste daha düşük başarı sergiledikleri, yüksek düzeyde aktivite gösterenlerin derste daha yüksek başarı sergiledikleri sonucuna ulaşılmıştır.

## Öneriler

Teknolojinin eğitim alanındaki etkisine bağlı olarak eğitimde yer alan öğelere ait birçok veri kayıt altına alınmaya başlamıştır. Özellikle çevrimiçi öğrenme ortamlarının artması bu verilerin daha kolay kaydedilmesini sağlamıştır. Kaydedilen bu verilerin daha kolay analiz edilmesinde ve öğrenme ortamlarının düzenlenmesinde eğitsel veri madenciliği ve öğrenme analitiği gelecekte en fazla kullanılacak uygulamalar olacaktır (Johnson, Smith, Willis, Levine ve Haywood, 2011). Yapılan çalışma sonucunda elde edilen bilgiler ışığında ileriki araştırmalara yön verebilecek araştırmaya ve uygulamaya dönük öneriler aşağıda sıralanmıştır:

### *Araştırmaya Dönük Öneriler*

- Çalışma kapsamında ÖYS üzerinde öğrenciler tarafından gerçekleştirilen 16 adet farklı eylemden 2 eylem girdi değişkeni olarak belirlenmiştir. Yapılacak diğer çalışmalarda farklı eylemler girdi değişkeni olarak veri madenciliği modelinde kullanılabilir ve elde edilecek sonuçlar incelenebilir.
- ÖYS tarafından tutulan öğrenci kayıtları 8 farklı alandan (ders adı, tarih, saat, IP adresi, kullanıcı adı, eylem, eylem bağlantıları ve bilgi alanı) oluşmaktadır. Çalışma kapsamında öğrencilere ait eylemler kullanılarak akademik başarı düzeyleri arasındaki ilişki belirlenmeye çalışılmıştır. Yapılacak diğer çalışmalarda farklı alanlar ile öğrencilerin akademik başarı düzeyleri arasındaki ilişkiler incelenebilir.
- Çalışma kapsamında veri madenciliği modelleri içerisinde yer alan tahmin edici modellerden olan sınıflandırma tekniklerinden karar ağaçları kullanılarak öğrenci başarıları tahmin edilmeye çalışılmıştır. Yapılacak diğer çalışmalarda diğer sınıflandırma teknikleri (Bayes, K-En Yakın Komşu, Yapay Sınır Ağları, Regresyon, Zaman Serisi Analizleri, Genetik Algoritmalar, Kestirim) kullanılarak öğrenci başarıları tahmin edilmeye çalışılabilir.

*Uygulamaya Dönük Öneriler*

- Veri madenciliği sürecinin herkes tarafından daha kolay ve hızlı gerçekleştirilebilmesi için ÖYS üzerinde çalışacak eklentiler geliştirilerek kaydedilen verilerin ÖYS üzerinde veri madenciliği sürecine tabi tutulması sağlanabilir. Öğretim elemanları tarafından buradan elde edilecek bilgiler kullanılarak öğrencileri anlık olarak izleme ve dönüt verme işlemleri gerçekleştirilebilir. Öğrenciler tarafında da bu bilgiler kendi öğrenme süreçlerini izlemek amacıyla kullanılabilir.
- ÖYS üzerinde geçmiş dönemlere ait öğrenci hareketliliği analiz edilerek bir sonraki eğitim uygulamaları için öğrencilere gerekli dönüt ve yönlendirme işlemleri sağlanabilir.
- ÖYS üzerinde dönem içerisindeki öğrenci hareketliliği analiz edilerek öğrenci profilleri belirlenebilir, öğrencilere gerekli dönüt ve dönem bitmeden yönlendirme işlemleri sağlanabilir.

**KAYNAKLAR**

- Akçapınar, G. (2014). *Çevrimiçi Öğrenme Ortamındaki Etkileşim Verilerine Göre Öğrencilerin Akademik Performanslarının Veri Madenciliği Yaklaşımı ile Modellenmesi*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi, Ankara.
- Amershi, S., & Conati, C. (2009). Combining Unsupervised And Supervised Machine Learning To Build User Models For Exploratory Learning Environments. *Journal of Educational Data Mining*, 1(1), 71-81.
- Arslan, H. (2008). *Sakarya Üniversitesi Web Sitesi Erişim Kayıtlarının Web Madenciliği ile Analizi*. Yayınlanmış Yüksek Lisans Tezi. Sakarya Üniversitesi, Sakarya.
- Aydın, S. (2007). *Veri Madenciliği ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama*. Yayınlanmış Doktora Tezi. Anadolu Üniversitesi, Eskişehir.
- Aydın, D. (2011). *Malzeme Bilgisi Dersinin Moodle Açık Kaynak Kodlu Öğretim Yönetim Sisteminde İşlenişinin Öğrenci Üzerindeki Etkinliğinin Değerlendirilmesi*. Yayınlanmış Yüksek Lisans Tezi. Bahçeşehir Üniversitesi, İstanbul.
- Aydın, C.Ç. ve Biroğul, S. (2008). E- Öğrenmede Açık Kaynak Kodlu Öğretim Yönetim Sistemleri ve Moodle. *Bilişim Teknolojileri Dergisi*, 1( 2), 31-36.
- Aynekin, G. (2006). *İnternet İçerik Madenciliğinde Yapay Sinir Ağları ve Bir Uygulama*. Yayınlanmış Yüksek Lisans Tezi. Uludağ Üniversitesi, Bursa.
- Beer, C., Clark, K. & Jones, D. (2010). *Indicators of Engagement*. Australasian Society for Computers in Learning in Tertiary Education, Proceedings, s.75-86, Sydney, Australia.
- Bienkowski, M., Feng, M. & Means, B. (2012). Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics-An Issue Brief. Erişim Tarihi: 15 Eylül 2015, <http://tech.ed.gov/learning-analytics>
- Bozkır, A. S. (2009). *OLAP ve Veri Madenciliği Teknolojilerinden Yararlanılarak Web Tabanlı Bir Karar Destek Sisteminin Gerçekleştirilmesi*. Yayınlanmış Yüksek Lisans Tezi. Hacettepe Üniversitesi, Ankara.
- Büyüköztürk, Ş., Çakmak, K.E., Akgün, Ö.E, Karadeniz, Ş. ve Demirel F. (2012). *Bilimsel Araştırma Yöntemleri*(13. baskı).Ankara: Pegem Akademi.

- Dunham, M.H. (2003). *Data Mining Introductory and Advanced Topics* (1st edition). New Jersey: Prentice Hall, Pearson Education Inc.
- Ergül, E. (2013). *Bilişim Teknolojileri Öğretmen Adaylarının Moodle İle Ders İşlenmesi Hakkındaki Görüşleri*. Yayımlanmış Yüksek Lisans Tezi. Süleyman Demirel Üniversitesi, Isparta.
- Erten, H. (2015). *Veri Madenciliği Teknikleri İle Organ Nakli İçin Uygun Donör Oranının Hesaplanması*. Yayımlanmış Yüksek Lisans Tezi. Gazi Üniversitesi, Ankara.
- Floyd, C., Schultz, T. and Fulton, S. (2012). *Security Vulnerabilities in The Open Source Moodle E-learning System*. Proceedings of the 16th Colloquium for Information Systems Security Educations. 42-47, Lake Buena Vista, Florida.
- Halees, A. (2008). *Mining Students Data to Analyze Learning Behavior: A Case Study*. The 2008 International Arab Conference of Information Technology (ACIT2008), Conference Proceedings, University of Sfax, Tunisia.
- Han, J. and Kamber, M. (2006). *Data Mining: Concepts and Techniques* (2nd edition). San Francisco: Morgan Kaufmann.
- Hand, D., Mannila, H., and Smyth, P. (2001). *Principles of Data Mining* (1st edition). London: The MIT Press.
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). The 2011 Horizon Report. Erişim Tarihi: 14 Eylül 2015, <http://www.nmc.org/pdf/2011-Horizon-Report.pdf>
- Kalikov, A. (2006). *Veri Madenciliği ve Bir E-Ticaret Uygulaması*. Yayımlanmış Yüksek Lisans Tezi. Gazi Üniversitesi, Ankara.
- Karasar, N. (2012). *Bilimsel Araştırma Yöntemi* (24. baskı). Ankara: Nobel Yayın Dağıtım.
- Kurt, Ç. ve Erdem, O.A. (2012). Öğrenci Başarısını Etkileyen Faktörlerin Veri Madenciliği Yöntemleriyle İncelenmesi. *Politeknik Dergisi*, 15( 2), 111-116.
- Leony, D., Pardo, A., Valentin, L. F., Quinones, I. & Kloos, C.D. (2012). Learning Analytics In The LMS: Using Browser Extensions To Embed Visualizations Into A Learning Management System. CEUR Workshop Proceedings. Erişim Tarihi: 11 Kasım 2015, <http://ceur-ws.org/Vol-894/paper6.pdf>.



- Lonn, S., Teasley, S.D. & Krumm, A. E. (2011). Who Needs To Do What Where?: Using Learning Management Systems On Residential vs. Commuter Campuses. *Computers & Education*, 56(1), 642–649.
- Lopez, M. I., Luna, J. M., Romero, C., & Ventura, S. (2012). *Classification Via Clustering for Predicting Final Marks Based On Student Participation in Forums*. Paper presented at the 5th International Conference on Educational Data Mining, EDM 2012, Chania, Greece.
- Macfadyen, L. & Dawson, S. (2010). Mining LMS Data to Develop An Early Warning System For Educators: A Proof of Concept. *Computers & Education*, 54( 2), 588-599.
- Osmanbegović, E., & Suljić, M. (2012). Data Mining Approach for Predicting Student Performance. *Economic Review*, 10(1), 3-12.
- Reis, A.Z., Baktır, H.Ö., Çelik, B., Erkoç, M.F., Özçakır, F.C., Özdemir, Ş. ve Şahin, K. (2012). Açık Kaynak Kodlu Öğrenme Yönetim Sistemleri Üzerine Bir Karşılaştırma Çalışması. *Eğitim ve Öğretim Araştırmaları Dergisi*, 1(2), 42-58.
- Romero, C. & Ventura, S. (2013). Data Mining In Education. *Wiley Data Mining and Knowledge Discovery*, 3(1), 12-27.
- Romero, C., Ventura, S., Hervás, C., & Gonzales, P. (2008). *Data Mining Algorithms to Classify Students*. Paper presented at the Proc. Int. Conf. Educ. Data Mining, Montreal, Canada.
- San Diego, J.P., Ballard, J., Hatzipanagos, S., Webb, M., Khan, E., Blake, P., Dore, T., Konstantinidis, A., & Barrett, I. (2012). *Do Moodle analytics have a role to play in learning design, assessment and feedback?* 1st Moodle Research Conference, September, 14-15, Heraklion, Greece.
- Sarıman, G. (2011). Veri Madenciliğinde Kümeleme Teknikleri Üzerine Bir Çalışma: K-Means ve K-Medoids Kümeleme Algoritmalarının Karşılaştırılması. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 15(3), 192-202.
- Seven, M. F. (2009). *Veri tabanlarından Bilgi Keşfi: Veri Madenciliği ve Bir Sağlık Uygulaması*. Yayımlanmış Yüksek Lisans Tezi. İstanbul Üniversitesi, İstanbul.
- Şeker, Ş.E. (2013). *İş Zekası ve Veri Madenciliği* (1.baskı). İstanbul: Cinius Yayınları.
- Thuraisingham, B.M. (2003). *Web Data Mining and Applications in Business Intelligence and Counter Terrorism* (1st edition). Florida: CRC Press LLC.

- Türker, A.Y. (2012). *Uzaktan Eğitim Öğretim Yönetim Sisteminin Bulanık Çok Kriterli Karar Verme Yöntemleri İle Seçimi*. Yayımlanmış Yüksek Lisans Tezi. Kocaeli Üniversitesi, Kocaeli.
- Ünal, T.A. (2014). *Büyük Veri ve Eğitimsel Veri Madenciliğinin Eğitim Alanına Katkılarının İncelenmesi*. 8th International Computer & Instructional Technologies Symposium, Trakya Üniversitesi, Edirne.
- Whitmer, J., Fernandes, K. & Allen, W.R. (2012). Analytics in Progress: Technology Use, Student Characteristics, and Student Achievement. Erişim Tarihi: 30 Aralık 2015 <http://www.educause.edu/ero/article/analytics-progress-technology-use-student-characteristics-and-student-achievement>
- Yapıcı, Ü.İ. ve Akbayın, H. (2012). Harmanlanmış Öğrenme Ortamında Moodle Kullanımı. *Eğitim ve Öğretim Araştırmaları Dergisi*, 1(2), 2146-9199.
- Yıldız, E., Bahçeci, F. (2014). *Öğrenme Yönetim Sistemlerinde Kullanılan Öğrenme Analitikleri Araçlarının İncelenmesi*. 8th International Computer & Instructional Technologies Symposium, Trakya Üniversitesi, Edirne.

## SUMMARY

*The developments in computer technology render that a huge amount of data can conceal in magnetic mediums easily and cheaply by making data storage systems become widespread. However, the data warehoused are worthless and mean nothing by themselves. In order to make the warehoused data become meaningful, they ought to be transformed into knowledge purposefully. Discovering the pattern of data and making them meaningful, data mining methods are used.*

*In this study, undergraduates were examined to discover the connection between their dynamism into Learning Management System (LMS) and their levels of achievement with using the data mining methods. In this regard, the study is descriptive and quantitative. Hence, it is appropriate for relational screening model. The research includes 40 pupils who took Instructional Technologies and Material Design course during spring semester and resided in Baskent University Institute of Education Sciences Pedagogical Formation Certification Program throughout 2012-2013 academic years. Students' academic achievement grades and log registrations placing in LMS which shows related year's access were used as data source.*

*To determine the correlation between academic success level and access to the system, data mining techniques, which occur in interdisciplinary areas, were used. Clustering model which is one of the data mining models and is descriptive was used primarily. K-Means algorithm was parlayed to constitute the model. Pupils' academic grades which are marked according to the success in Instructional Technologies and Material Design lessons were used in clustering model as data source during 2012-2013 spring semesters. After clustering process had done, students were separated into three groups which are called "low academic achievement", "medium academic achievement" and "high academic achievement". They were inserted into these groups according to their academic achievement level. Doing that, K-Means algorithm was used in order to be able to calculate average values. It can be said that the academic achievement groups which were formed with clustering process can contribute to get knowledge about pupils'*

feature, to create decision trees which place in predictive models and to interpret the results easily.

After clustering process, Moodle (Modular Object Oriented Dynamic Learning Environment) LMS data were aimed to create decision trees as data source. The decision trees which were constituted by using C5.0, CART (Classification and Regression Trees), CHAID (Chi-squared Automatic Interaction Detector), QUEST (Quick, Unbiased, Efficient Statistical Tree) algorithms were examined to see their accuracy rate. As a result of that, CART algorithm has the highest accuracy rate among all the algorithms above-referred and it is %85.0. Correspondingly, decision trees were created by using CART algorithm.

Depending upon the findings which were acquired at the end of the study, it was revealed that there is a meaningful connection between the activities on LMS and the academic achievement levels of students. It is quite obvious that pupils who are active on LMS place in high academic achievement group, who are intermediate user of LMS are found in the medium academic achievement group, who are inactive on LMS locate low academic achievement group.