

T.C.
BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ
ANABİLİM DALI

ÖĞRETİM YÖNETİM SİSTEMİ ÜZERİNDE ÜNİVERSİTE
(LİSANS) DÜZEYİNDEKİ ÖĞRENCİ HAREKETLİLİĞİNİN VERİ
MADENCİLİĞİ YÖNTEMLERİYLE ANALİZİ

Özkan ÖZBAY

YÜKSEK LİSANS TEZİ
ANKARA – 2015

T.C.
BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ
ANABİLİM DALI

ÖĞRETİM YÖNETİM SİSTEMİ ÜZERİNDE ÜNİVERSİTE
(LİSANS) DÜZEYİNDEKİ ÖĞRENCİ HAREKETLİLİĞİNİN VERİ
MADENCİLİĞİ YÖNTEMLERİYLE ANALİZİ

DANIŞMAN
Yrd. Doç. Dr. Halil ERSOY

Özkan ÖZBAY

YÜKSEK LİSANS TEZİ
ANKARA – 2015

T.C.
BAŞKENT ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ ANABİLİM DALI

ÖĞRETİM YÖNETİM SİSTEMİ ÜZERİNDE ÜNİVERSİTE (LİSANS)
DÜZEYİNDEKİ ÖĞRENCİ HAREKETLİLİĞİNİN VERİ MADENCİLİĞİ
YÖNTEMLERİYLE ANALİZİ

Özkan ÖZBAY

YÜKSEK LİSANS TEZİ

Bu tez, / /2015 tarihinde aşağıda üye adları yazılı jüri tarafından kabul edilmiştir.

Unvan	Adı Soyadı	İmza
Doç. Dr.	Ömer DELİALİOĞLU	
Yrd. Doç. Dr.	Serpil YALÇINALP	
Yrd. Doç. Dr.	Halil ERSOY	

ONAY

/ /2015

Prof. Dr. Sadegül AKBABA ALTUN

Eğitim Bilimleri Enstitü Müdürü

TEŞEKKÜR

Çalışmam süresince tez danışmanlığımı üstlenen, bilimsel katkı ve manevi desteğini hiçbir zaman esirgemeyen değerli hocam Yrd. Doç. Dr. Halil ERSOY'a,

Yüksek Lisans eğitimim boyunca benden yardımlarını, desteğini, sabrını ve bilgisini esirgemeyen değerli hocalarım Doç. Dr. Serpil CULA'ya, Yrd. Doç. Dr. Serpil YALÇINALP'e, Yrd. Doç. Dr. Filiz KALELİOĞLU'na ve Yrd. Doç. Dr. Emine CABI'ya,

Yüksek lisans eğitimini birlikte sürdürdüğüm değerli sınıf arkadaşlarım Hacer TÜRKOĞLU, Hüseyin Melih ALTIN ve Melih AŞICI'ya,

Yüksek lisans eğitimim boyunca manevi desteğini hiçbir zaman esirgemeyen sevgili aileme, çalışmamın her aşamasında katkılarını esirgemeyen arkadaşlarıma teşekkürlerimi sunarım.

Özkan ÖZBAY

ÖZET

ÖĞRETİM YÖNETİM SİSTEMİ ÜZERİNDE ÜNİVERSİTE (LİSANS) DÜZEYİNDEKİ ÖĞRENCİ HAREKETLİLİĞİNİN VERİ MADENCİLİĞİ YÖNTEMLERİYLE ANALİZİ

Özkan ÖZBAY

YÜKSEK LİSANS TEZİ

Bilgisayar teknolojisindeki gelişmeler veri depolama sistemlerinin yaygınlaşmasını sağlayarak büyük miktarlardaki verilerin manyetik ortamlarda saklanması kolay ve ucuz hale getirmiştir. Ancak depolanan bu veriler tek başlarına değersiz olup bir anlam ifade etmemektedir. Depolanan verilerin anlamlı hale getirilebilmesi için bilgisayar sistemleri kullanılarak belli bir amaca yönelik işlenerek bilgiye dönüştürülmesi gerekmektedir. Veriler arasındaki örüntülerin keşfedilerek anlamlı bilgi haline getirilmesi veri madenciliği yöntemleri ile sağlanabilmektedir.

Bu çalışmada üniversite (lisans) düzeyindeki öğrencilerin Öğretim 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.

Öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarı düzeyleri arasındaki ilişkinin belirlenmesinde disiplinler arası bir alan olan veri madenciliği teknikleri kullanılmıştır. İlk olarak veri madenciliği modellerinden tanımlayıcı model olan kümeleme modeli kullanılmış ve bu modelin oluşturulmasında K-Means algoritmasından yararlanılmıştır. Kümeleme modelinde veri kaynağı olarak 2012-2013

eđitim đretim yılı bahar dnemi đretim Teknolojileri ve Materyal Tasarımı dersini alan đrencilerin yılsonu akademik başarı notları kullanılmıştır. Kmeleme işleminin gerçekleştirilmesi sonucu đrencilerin düşük, orta ve yüksek akademik başarı grubu olmak zere 3 gruba ayrılmasına karar verilmiştir. K-Means algoritması tarafından akademik grupları için ortalama deęerler hesaplanarak đrenciler akademik başarı durumlarına gre bu akademik başarı gruplarına yerleştirilmiştir. Kmeleme işlemi sonucu oluşturulan akademik başarı gruplarının; đrenci zellikleri hakkında bilgi sağlayabileceęi, tahmin edici modeller içerisinde yer alan sınıflandırma modellerinden olan karar aęaçlarının oluşturulmasına ve elde edilen sonuçların daha kolay yorumlanabilmesine katkı sağlayacağı söylenebilir.

Kmeleme işlemi gerçekleştirildikten sonra karar aęaçlarının oluşturulmasında veri kaynaęı olarak đrencilerin Moodle YS üzerinde gerçekleştirdikleri eylemler kullanılmıştır. C5.0, CART (Classification And Regression Trees), CHAID (Chi-squared Automatic Interaction Detector), QUEST (Quick, Unbiased, Efficient Statistical Tree) algoritmaları kullanılarak oluşturulan karar aęaçlarının doęruluk oranları incelenmiş ve bu algoritmalar içerisinde en yüksek doęruluk oranını veren CART algoritmasının doęruluk oranı %85,0 olarak bulunmuştur. Karar aęaçları en yüksek doęruluk oranını veren CART algoritması kullanılarak oluşturulmuştur.

Çalışma sonucunda elde edilen bulgulara baęlı olarak đrencilerin YS üzerindeki hareketlilięi ve akademik başarı düzeyleri arasında anlamlı bir ilişkinin olduęu ortaya konulmuştur. YS üzerinde yüksek düzeyde hareketlilik sergileyen đrencilerin yüksek akademik başarı grubunda, orta düzeyde hareketlilik sergileyen đrencilerin orta akademik başarı grubunda ve düşük düzeyde hareketlilik sergileyen đrencilerin düşük akademik başarı grubunda yer aldıkları söylenebilir.

Anahtar Kelimeler: Moodle, đretim Ynetim Sistemleri, Veri Madencilięi, đrenme Analitięi, Eęitsel Veri Madencilięi

ABSTRACT

ANALYSIS OF STUDENT USE OF LEARNING MANAGEMENT SYSTEM THROUGH DATA MINING METHODS IN UNDERGRADUATE LEVEL

Özkan ÖZBAY

MASTER THESIS

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 being warehoused are worthless and mean nothing by themselves. In order to make the data being warehoused 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 are 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 associative scanning from scanning models. The research includes 40 pupils who receive in Instructional Technologies and Material Design during spring semester and reside 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 are 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, are used. Clustering model which is one of the data mining models and is descriptive is used primarily. K-Means algorithm is parlayed to constitute the model. Pupils' academic grades which are marked according to the success in Instructional Technologies and Material Design lessons are used in clustering model as data source during 2012-2013 spring semesters. After clustering process has done, students are separated into three groups which are called "low academic achievement", "medium academic

achievement” and “high academic achievement”. They are inserted into these groups according to their academic achievement level. Doing that, K-Means algorithm is used in order to be able to calculate average values. It can be said that the academic achievement groups which are 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 handily.

After clustering process, Moodle (Modular Object Oriented Dynamic Learning Environment) LMS datas are aimed to create decision trees as data source. The decision trees which are constituted by using C5.0, CART (Classification And Regression Trees), CHAID (Chi-squared Automatic Interaction Detector), QUEST (Quick, Unbiased, Efficient Statistical Tree) algorithms are 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 are created by using CART algorithm.

Depending upon the findings which acquired at the end of the study, 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.

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

İÇİNDEKİLER

TEŞEKKÜR	i
ÖZET.....	ii
ABSTRACT	iv
İÇİNDEKİLER	vi
SİMGELER VE KISALTMALAR DİZİNİ	viii
ŞEKİLLER DİZİNİ	ix
TABLolar DİZİNİ	x
1. GİRİŞ	11
1.1. Problem	11
1.2. Amaç	13
1.3. Önem	13
1.4. Sınırlılıklar.....	14
1.5. Sayıtlar	14
2. LİTERATÜR.....	15
2.1. Veri Madenciliği.....	15
2.1.1. Veri Madenciliği Modelleri	16
2.1.2. Veri Madenciliği Süreci.....	22
2.1.3. Veri Madenciliğinde Kullanılan Programlar	25
2.2. Eğitsel Veri Madenciliği	26
2.2.1. Eğitsel Veri Madenciliği Süreci.....	26
2.3. Öğrenme Analitiği.....	28
2.3.1. Öğrenme Analitiği Süreci	28
2.3.2. Öğrenme Analitiğinin Eğitimde Kullanımı	30
2.4. Öğretim Yönetim Sistemleri (ÖYS).....	31
2.4.1. Moodle	33
2.5. Yurt İçi ve Yurt Dışında Yapılan Çalışmalar	35
3. YÖNTEM.....	39
3.1. Araştırmanın Modeli	39
3.2. Araştırma Örneklemi.....	40
3.3. Veri Kaynakları	41
3.4. Verilerin Analizi.....	42

3.4.1. Problemin Tanımlanması	43
3.4.2. Verilerin Hazırlanması	43
3.4.3. Modelin Kurulması ve Değerlendirilmesi	50
3.4.4. Modelin Kullanılması	52
3.4.5. Modelin İzlenmesi	52
3.5. Araştırma Süreci	53
4. BULGULAR	55
4.1. Birinci Araştırma Sorusuna İlişkin Bulgular	55
4.1.1. Öğrenciler Tarafından Gerçekleştirilen Moodle Eylemleri ve Dağılım Oranları	55
4.2. İkinci Araştırma Sorusuna İlişkin Bulgular	62
4.2.1. Tanımlayıcı Modellerinin Kullanılması	62
4.2.2. Tahmin Edici Modellerinin Kullanılması	67
5. TARTIŞMA VE SONUÇ	77
5.1. Birinci Araştırma Sorusuna İlişkin Sonuçlar	78
5.2. İkinci Araştırma Sorusuna İlişkin Sonuçlar	78
5.3. Öneriler	81
5.3.1. Araştırmaya Dönük Öneriler	81
5.3.2. Uygulamaya Dönük Öneriler	82
KAYNAKLAR LİSTESİ	83
EKLER	93
Ek-1. Moodle Eylemleri ve İşlevleri	94
Ek-2. SPSS Clementine Kullanılarak Oluşturulan Veri Madenciliği Modeli	96
Ek-3. SPSS Clementine’de Kullanılan Düğümler ve İşlevleri	97
Ek-4. CART Algoritması Hata ve Doğruluk Oranı	99
Ek-5. C5.0 Algoritması Hata ve Doğruluk Oranı	100
Ek-6. CHAID Algoritması Hata ve Doğruluk Oranı	101
Ek-7. QUEST Algoritması Hata ve Doğruluk Oranı	102
Ek-8. İntihal Raporu	103

SİMGELER VE KISALTMALAR DİZİNİ

CART	: Classification And Regression Trees Sınıflandırma ve Regresyon Ağaçları
CHAID	: Chi-squared Automatic Interaction Detector Ki-kare Otomatik Etkileşim Detektörü
CSV	: Comma Seperated Values Virgül ile Ayrılmış Değerler
ID3	: Iterative Dichotomiser 3
MOODLE	: Modular Object Oriented Dynamic Learning Environment Modüler Nesne Tabanlı Dinamik Öğrenme Ortamı
ÖYS	: Öğretim Yönetim Sistemi
QUEST	: Quick, Unbiased, Efficient Statistical Tree Hızlı, Yansız, Etkili İstatistiksel Ağaç

ŞEKİLLER DİZİNİ

Şekil 2.1. Veri Madenciliği Modelleri	16
Şekil 2.2. Karar Ağaçlarının Yapısı	18
Şekil 2.3. Kümeleme İşlemi Sonrası Veri Kümleri	21
Şekil 2.4. Veri Madenciliği Süreci	22
Şekil 2.5. Eğitsel Veri Madenciliği Süreci	26
Şekil 2.6. Öğrenme Analitiği Süreci	29
Şekil 3.1. Öğrenci Hareketliliği Veri Madenciliği Süreci	43
Şekil 3.2. Veri Birleştirme ve Temizleme İşlemi Sonrası Veri Yığını	46
Şekil 4.1. Eylemlerin Dağılım Oranları	56
Şekil 4.2. Öğrenci Eylem Sayısı	57
Şekil 4.3. Öğrenci Eylem Dağılımı	59
Şekil 4.4. Akademik Başarı Grupları	63
Şekil 4.5. Akademik Başarı Gruplarında Yer Alan Öğrencilerin Yaptıkları Eylem Sayılarına Göre Dağılımı	65
Şekil 4.6. CART Algoritması İle Oluşturulan Karar Ağacı	70
Şekil 4.7. Tahmin Edilen ve Gerçek Akademik Başarı Düzeyleri	75

TABLÖLAR DİZİNİ

Tablo 2.1. Moodle İstatistikleri	33
Tablo 2.2. Moodle Kullanım Oranı En Yüksek İlk 10 Ülke	35
Tablo 3.1. Araştırma Örneklemine Ait Demografik Bilgiler	40
Tablo 3.2. Ders Geçme Sistemi Not Ağırlıkları	40
Tablo 3.3. Moodle Üzerindeki Eylemler	48
Tablo 3.4. Öğrenci Ders Notları	49
Tablo 4.1. Öğrenci Eylem Sayısı	58
Tablo 4.2. Öğrenci Bazında Eylemler	61
Tablo 4.3. Kullanıcı Akademik Başarı Grupları	64
Tablo 4.4. Eylemler ve Akademik Başarı Grupları Tarafından Gerçekleştirilme Sayıları	66
Tablo 4.5. Tahmin Modellerinin Doğruluk Oranları	67

1. GİRİŞ

1.1. Problem

Bilgi ve araştırma bulgularının pratik alanlara uygulanması olarak tanımlanan teknoloji; iş, eğitim, iletişim, eğlence, sağlık alanlarında kullanılarak özellikle sanayi toplumlarında insan hayatının ayrılmaz bir parçası haline gelmiştir (Czaja ve diğerleri, 2006). Teknolojinin bu denli insan hayatındaki öneminin artmasıyla birlikte teknolojik ürünler insanoğlunun vazgeçilmez birer parçası olmuştur. Bu teknolojik ürünler eğlenceden iletişime, sağlıktan bankacılığa, sanayiden eğitime kadar hayatın her aşamasında kullanılır hale gelmiş ve birçok konuda insanoğlunun hayatını kolaylaştırmıştır.

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 ve diğerleri, 2001). Yani veri madenciliği, büyük miktarlardaki veri setlerinde saklı durumda bulunan örüntü ve eğilimleri keşfetme işlemidir (Thuarisingham, 2003).

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 uzaktan 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 kullanılabilecek bilgilerin üretilmesi mümkün olabilmektedir (Bienkowski, Feng, & Means, 2012).

Öğretim Yönetim Sistemleri (ÖYS) eğitim ortamlarında eğitimciler tarafından sıklıkla başvurulanan öğretim araçları haline gelmiştir. 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 saylamak 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 öğrencilerin profillerinin oluşturulmasında, öğrenme ortamının kişiselleştirilmesinde, eğitim ortamının kalitesinin artırılmasında geri bildirimler sağlar.

Bu araştırmada üniversite (lisans) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarıları arasındaki ilişki incelenmiştir. Öğrencilerin ÖYS üzerindeki hareketliliği ve akademik başarıları bu çalışmanın çekirdek noktasını oluşturmaktadır. “Üniversite (lisans) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarı düzeyleri arasında anlamlı bir ilişki var mıdır?” sorusu problem cümlemizi oluşturmaktadır.

1.2. Amaç

Bu araştırmanın genel amacı; veri madenciliği tekniklerini kullanarak üniversite (lisans) 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?

Araştırma sorusunda yer alan öğrencilerin akademik başarılarının nasıl ve ne kadar doğruluk oranıyla tahmin edilebileceği veri madenciliğinde kullanılan algoritmalar ile sağlanmaktadır.

1.3. Önem

Literatürde ÖYS'lerinin yeterince esnek ve kolay bir kullanıma sahip olup olmadıkları, özelliklerinin neler olması gerektiği, uzaktan eğitimde nasıl kullanılabilecekleri, farklı ÖYS'lerin avantaj ve dezavantajlarının belirlendiği çalışmalar olsa da (Aydın ve Biroğlu, 2008; Aydın, 2011; Lonn ve diğerleri, 2011; Türker, 2012; Floyd ve diğerleri, 2012; Reis ve diğerleri, 2012; Yapıcı ve Akbayın, 2012; Ergül, 2013), veri madenciliği teknikleriyle sistem tarafından toplanan verilerin analiz edildiği, bu verilere dayanarak ÖYS üzerindeki öğrenci hareketliliğinin incelendiği çalışmalar nadirdir (Whitmer ve diğerleri, 2012; Leony ve diğerleri, 2012; San Diego ve diğerleri, 2012).

Teknolojinin eğitimle bütünleştirilmesi konusunda önemli bir yere sahip olan ÖYS'lerin eğitime destek amaçlı daha etkin kullanılmasına ve veri madenciliği teknikleriyle ÖYS'ler üzerindeki öğrenci hareketliliğinin analiz edilerek ÖYS'lerin etkin bir biçimde değerlendirilmesine katkı sağlayacağı düşünülmektedir.

Üniversite (lisans) düzeyinde eğitim gören öğrencilerin ÖYS üzerindeki hareketliliğinin veri madenciliği teknikleriyle incelenerek veri yığınları içerisinde elde edilecek bilgilerle öğrenci akademik başarısı gibi önemli bilgilerin ne düzeyde tahmin edilebileceğinin belirlenmesine ve öğrencilerin akademik başarısını yükseltmeye yönelik tedbirlerin alınmasına katkı sağlayacağı düşünülmektedir.

ÖYS’ler üzerindeki öğrenci hareketliliğinin incelendiği çalışmaların yeterli sayıda olmadığı göz önüne alındığında bu çalışmanın araştırmacılar tarafından konu ile ilgili yapacakları yeni çalışmalarına kaynak olması beklenmektedir.

1.4. Sınırlılıklar

Bu çalışma 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 ile sınırlıdır.

Başkent Üniversitesi tarafından 2007’den beri derslere destek amaçlı kullanılan ve ihtiyaçları doğrultusunda uyarlanan açık kaynak kodlu Moodle (Modular Object Oriented Dynamic Learning Environment) ÖYS kullanımı ile sınırlıdır.

2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersinde kullanılan Moodle ÖYS özellikleri ile sınırlıdır.

Öğretim elemanın eğitimde Moodle ÖYS’yi kullanma istek ve becerisi ile sınırlıdır.

Çalışmada tercih edilen algoritmalar ve oluşturulan modelin doğruluk oranı ile sınırlıdır.

1.5. Sayıtlar

Öğretim Teknolojileri ve Materyal Tasarımı dersini alan 40 öğrencinin ve öğretim elemanının genel bilgisayar okuryazarlığının ve Moodle ÖYS’yi kullanma becerilerinin yeterli olduğu varsayılmıştır.

2. LİTERATÜR

Bu bölümde; veri madenciliğinin tanımına, veri madenciliğinin kullanım amacına, veri madenciliğinin uygulama alanlarına, veri madenciliği modellerine, veri madenciliği sürecine, veri madenciliğinde kullanılan programlara, eğitsel veri madenciliğinin tanımına, eğitsel veri madenciliği sürecine, öğrenme analitiği tanımına, öğrenme analitiği sürecine, öğrenme analitiğinin eğitimde kullanımına, öğretim yönetim sistemlerinin tanımı ve genel özelliklerine, Moodle (Modular Object Oriented Dynamic Learning Environment) Öğretim Yönetim Sistemi (ÖYS)'ne, yurt içi ve yurt dışında yapılan çalışmalara yer verilmiştir.

2.1. Veri Madenciliği

İlk bilgisayarların üretilmesiyle başlayan ve gelişen, teknolojiyle beraber artarak devam eden manyetik ortamlarda veri saklama işlemi günümüzde büyük miktarlardaki verilerin depolanmasını kolay ve ucuz hale getirmiştir. Market alışverişlerinden bankacılık işlemlerine, sağlık işlemlerinden kamu uygulamalarına, internet kullanımından eğitim uygulamalarına kadar birçok alanda üretilen veriler manyetik ortamlarda saklanmaya başlamıştır. Ancak manyetik ortamlarda depolanan bu veri yığınları tek başlarına bir anlam ifade etmediğinden veri yığınları içerisinde keşfedilmeye değer ve işe yarar bilgilerin ortaya çıkarılması gerekmektedir.

Veri yığınlarının artan veri depolama sistemlerine bağlı olarak her geçen gün artması ve bu veriler arasındaki ilişkilerin daha karmaşık hale gelmesinden dolayı veri yığınlarının analiz edilmesinde geleneksel yöntemler yetersiz kalmıştır. Veriler arasındaki ilişkilerin keşfedilmesi ve anlamlı örüntülerin ortaya çıkarılabilmesi için yeni yöntem ve araçların geliştirilmesine ihtiyaç duyulmuştur. Buna bağlı olarak büyük miktarlardaki veri yığınlarının analiz edilmesinde bilgisayar teknolojileri, istatistik, veri tabanı teknolojileri ve diğer disiplinleri bir araya getiren veri madenciliği ortaya çıkmıştır.

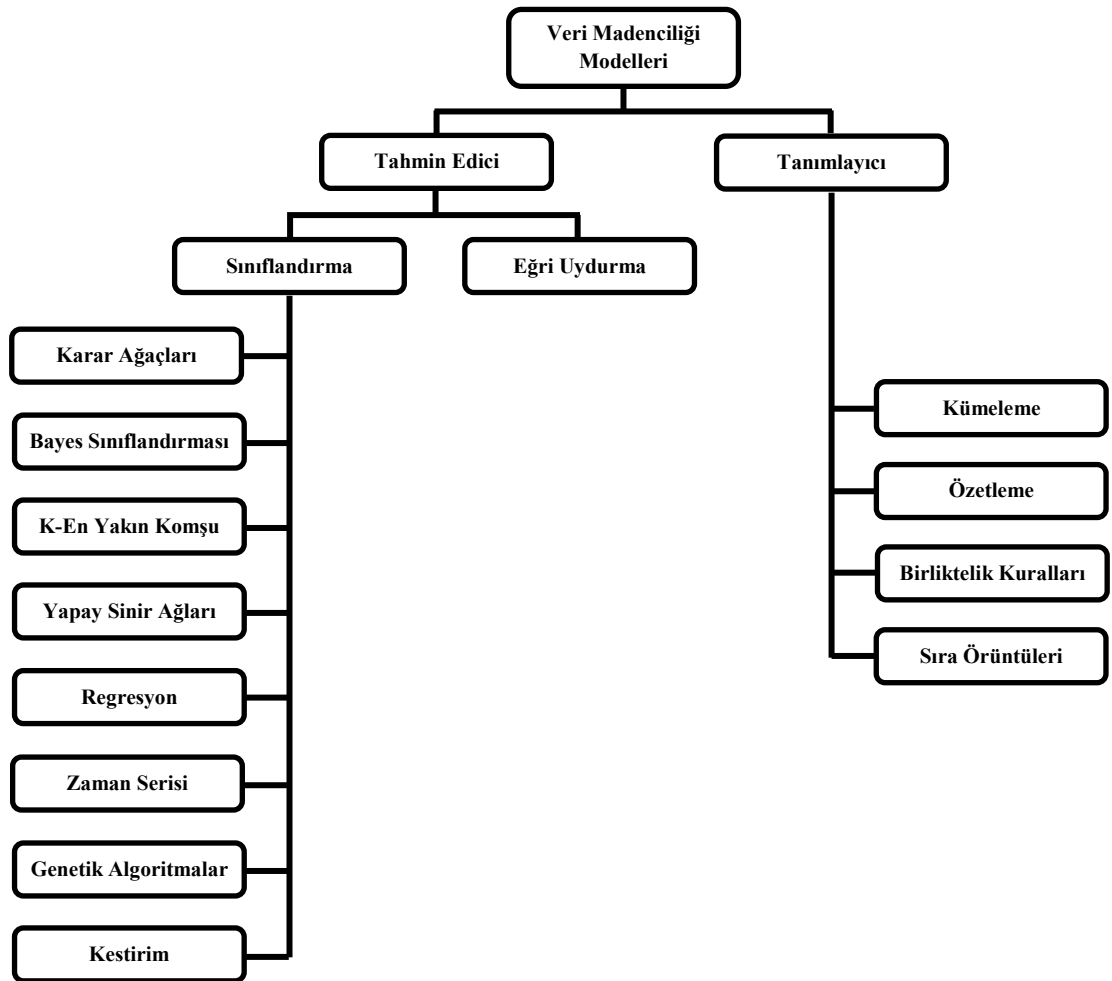
Veri yığınlarının anlamlı hale getirilmesi ve işe yarar bilgilere dönüştürülmesini sağlayan veri madenciliği, veri tabanlarındaki ham verinin tek başına ortaya koyamadığı önceden bilinmeyen, geçerli, güvenilir, potansiyel olarak kullanışlı ve anlaşılabilir örüntülerin bilgisayar programları kullanılarak ortaya çıkarılması işlemidir (Holsheimer

ve Siebes,1994; Fayyad ve diğerleri, 1996; Cabena ve diğerleri, 1998; Hand ve diğerleri, 2001; Hung, 2005).

Veri madenciliğinin kullanımı belirli uygulama alanlarıyla sınırlandırılmaz. Veri madenciliğini verinin üretilip kayıt altına alındığı her alanda kullanmak mümkündür. Sağlık, endüstri, mühendislik, pazarlama, bankacılık ve eğitim alanları veri madenciliğinin kullanıldığı başlıca uygulama alanlarıdır.

2.1.1. Veri Madenciliği Modelleri

Veri madenciliğinde tahmin edici ve tanımlayıcı olmak üzere iki ana model bulunmaktadır. Bu modeller aşağıda açıklanmıştır. Veri madenciliği modelleri Şekil 2.1’de gösterilmektedir.



Şekil 2.1. Veri Madenciliği Modelleri (Dunham, 2003)

2.1.1.1. Tahmin Edici Modeller

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. Tahmin edici modeller özellikle karar alma sürecinde önemli bir role sahiptir. Örneğin bir banka, müşterilerinin önceki dönemlerde kullanmış oldukları kredilere ilişkin tüm verilere sahiptir. Bu verilere uygun olarak kurulan model ile müşterilerinin daha sonraki kredi taleplerinde verilecek olan krediyi geri ödeyip ödeyemeyeceği konusunda tahminlerde bulunabilir (Arslan, 2008, Onat, 2008).

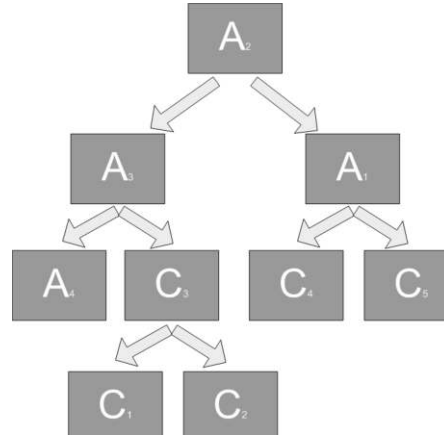
Tahmin edici modeller kendi içinde sınıflandırma ve eğri uydurma olmak üzere ikiye ayrılmaktadır. Tahmin edici modeller içerisinde en yaygın kullanıma sahip olan ise sınıflandırma modelidir.

Sınıflandırma modeli geçmiş verilerin anlaşılmasında ve yeni örneklemin nasıl davranacağını tahmin edilmesinde kullanılır. Sınıflandırma modeli bir sınıfı diğerinden ayıran örüntüleri keşfeder (Arslan, 2008). Sınıflandırma modeli örüntü tanıma, hastalık tanısı, kalite denetimi, dolandırıcılık tespiti, pazarlama, risk analizi, müşteri değerlendirmesi gibi konularda oldukça fazla kullanılmaktadır. Sınıflandırma modeli içerisinde birçok farklı model yer almaktadır.

- Karar Ağaçları (Bozkır, 2009),
- Bayes Sınıflandırması (Arslan, 2008; Bozkır, 2009; Ayık ve diğerleri, 2007),
- K-En Yakın Komşu (Kaya ve Köymen, 2008; Aynekin, 2006),
- Yapay Sinir Ağları (Sever ve Oğuz, 2002),
- Genetik Algoritmalar (Aynekin, 2006; Ayık ve diğerleri, 2007),
- Regresyon (Kudyaba ve Hoptroff, 2001),
- Zaman Serisi Analizleri (Roiger ve Geatz, 2003),
- Kestirim (Dunham, 2003) sınıflandırmada kullanılan başlıca modellerdir.

Sınıflandırma modelleri içerisinde *karar ağaçları* diğer sınıflandırma modelleriyle karşılaştırıldığında oluşturulması ve anlaşılması kolay bir yöntem olduğundan sınıflandırma modelleri içinde en çok kullanılan modeldir. Karar ağaçlarının sık kullanılan bir model olmasının bir diğer nedeni ise model şeffaflığını

sağlaması ve görsel bir sunuma sahip olmasıdır (Gürcan, 2009). Karar ağaçları üç temel unsurdan oluşur; karar düğümleri, dallar ve yapraklar. Bir karar ağacının gövde yapısını karar düğümleri oluşturur ve her bir karar düğümü dallara ayrılır, ayrılan dallarda ise yapraklar yer alır. Karar ağaçlarında kullanılan algoritmanın yapısına bağlı olarak düğümler, dallar ve yapraklar oluşturulur. Kullanılan algoritmaya göre her düğüm iki ya da daha fazla dala ayrılabilir. Karar ağacının oluşturulmasına ilk olarak en üstte yer alan kök düğümle başlanır. Karar düğümleri gerçekleştirilecek testi belirtir ve test sonucuna göre karar düğümleri dallara ayrılır. Her dal olası bir sonucu gösterir. Karar düğümlerinden ayrılan her bir dal sınıflama işlemini tamamlamaya çalışır. Dalın ucunda sınıflama işlemi gerçekleştirilirse yapraklar oluşturulur. Oluşturulan yapraklar veri üzerinde belirlenmek istenen sınıflardan biridir. Fakat dalın ucunda sınıflama işlemi gerçekleşmiyorsa yeni bir karar düğümü oluşturulur. Karar düğümleri kök düğümünden başlayarak yaprağa ulaşana kadar ardışık düğümler şeklinde oluşturulmaya devam eder. Bu yöntemde sınıflandırma işlemi için bir ağaç oluşturulmakta ve daha sonra veri tabanındaki her kayıt bu ağaca uygulanarak çıkan sonuca göre sınıflandırılmaktadır (Bozkır, 2009). Şekil 2.2’ de örnek bir karar ağacı yapısı yer almaktadır.



Şekil 2.2. Karar Ağaçlarının Yapısı (Bozkır, 2009)

Karar ağaçları oluşturulurken birçok farklı algoritmadan yararlanılmaktadır. Karar ağaçlarında kullanılan algoritmalar aralarında bir takım farklılıklar olmasına karşın büyük oranda benzerlikte göstermektedir. ID3 (Iterative Dichotomiser 3), C4.5, C5.0, CART (Classification And Regression Trees), CHAID (Chi-squared Automatic Interaction Detector) ve QUEST (Quick, Unbiased, Efficient Statistical Tree) algoritmaları karar ağaçları oluşturmada en sık kullanılanlardır.

ID3 (Iterative Dichotomiser 3) Algoritması

Quinlan (1986) tarafından geliştirilen ve sadece kategorik veriler ile çalışan entropiye dayalı bir sınıflandırma algoritmasıdır. Bu algoritmada amaç karar ağaçları oluşturulurken veri yığını içerisindeki verilerin mümkün olduğunca benzer hale getirilmesi ve oluşturulacak karar ağacı derinliğinin minimum seviyede tutulmasıdır. Böylece karmaşıklık minimum seviyeye inerken kazanç maksimum seviyede olur.

C4.5 Algoritması

C4.5 algoritması Quinlan (1993) tarafından geliştirilen ID3 algoritmasının bazı eksik ve sorunlu yönlerinin giderilmesi için yine Quinlan tarafından geliştirilmiş bir algoritmadır. ID3 algoritmasının temel özelliklerini kullanır fakat ID3 algoritmasının özelliklerine ek olarak bölünme-dağılma bilgisi, kayıp değerlerin tespit edilmesi ve sayısal verilerin hesaba katılması işlemlerini gerçekleştirir.

C5.0 Algoritması

En yaygın kullanılan karar ağaçları algoritmalarından olan C5.0 algoritması; Quinlan'ın ID3 algoritması kullanılarak elde edilmiş olan C4.5 algoritmasının geliştirilmiş halidir (Hou ve diğerleri, 2014). C4.5 algoritmasına benzer olarak ID3 algoritmasının temel özelliklerini kullanmaktadır. C4.5 algoritmasının doğruluğunu artırmak için geliştirilmiştir. C5.0 algoritması C4.5'e göre çok daha hızlı olup, hafızayı daha verimli kullanmaktadır (Çalış ve diğerleri, 2014). Bu yüzden C5.0 algoritması daha çok büyük veri yığınlarında kullanılmaktadır.

CART (Classification and Regression Trees) Algoritması

Breiman, Friedman, Olshen ve Stone (1984) tarafından geliştirilen kümedeki kayıtları kendinden daha homojen olacak şekilde alt kümelere bölen ve bu işlemi homojenlik kriterlerine ulaşıncaya kadar yinelemeli olarak tekrar eden bir sınıflandırma algoritmasıdır.

CART algoritması hem sayısal hem de nominal değerler üzerinde çalışabilir. Sınıflandırma ve regresyonu destekleyici bir yapıdadır. CART algoritması verileri her seferinde iki alt kümeye ayırmaktadır. CART algoritması maksimum büyüklükteki ağacı oluşturmaya çalışarak en saf, en iyi bölünme durumunu elde etmeye çalışır. Bu

işlemi gerçekleştirirken karar ağacı sürekli olarak bölünerek büyür ve yeni bölünme işlemi gerçekleştirilemeyinceye kadar bölünme işlemine devam edilir. Bölünme işlemi bittiğinde ise uçtan köke doğru budama işlemi başlatılır. Olası en başarılı karar ağacı her budama işleminden sonra bir test verisi kullanılarak tespit edilir (Sezer ve diğerleri, 2010).

CHAID (Chi-Squared Automatic Interaction Detector) Algoritması

Kass (1980) tarafından geliştirilen homojen değişken düzeylerini birleştirmeye ve heterojen düzeyleri ise ayrı tutarak dallandırma yapmaya yarayan bir sınıflandırma algoritmasıdır. Dalların uygulanabilmesinde en iyi ayırt edici değişkenin belirlenmesinde ki-kare istatistiğini temel alır.

QUEST (Quick, Unbiased, Efficient Statistical Tree) Algoritması

QUEST algoritması Loh ve Shih (1997) tarafından geliştirilen ön yargılı seçimi daha genel hale getiren, yanlılıktan uzak ve hızlı hesaplama yapabilen bir sınıflandırma algoritmasıdır. Büyük veri yığınlarındaki hesaplama maliyetini düşürmek amacıyla geliştirilmiştir. CART algoritması gibi ikili karar ağacı yapısını kullanmaktadır.

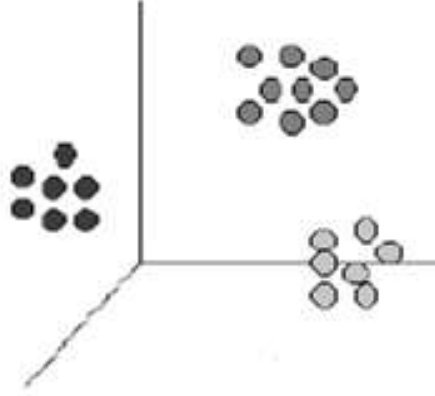
2.1.1.2. Tanımlayıcı Modeller

Tanımlayıcı modeller karar verme aşamasında rehberlik etmede kullanılacak veri örüntülerinin tanımlanmasını sağlamaktadır. En yaygın kullanılan tanımlayıcı modeller aşağıda yer almaktadır.

- Kümeleme (Kaya ve Köymen, 2008),
- Özetleme (Dunham, 2003),
- Birliktelik Kuralları (Bozkır, 2009; Ayık ve diğerleri, 2007; Aynekin, 2006),
- Sıra Örüntüleri (Bozkır, 2009)

Tanımlayıcı modeller içerisinde *kümeleme* yöntemi, soyut ve somut birbirine benzer nesnelerin bir grupta toplanması işlemini gerçekleştirmektedir. Benzer ya da birbiriyle ilişkili olan nesne grupları aynı kümeyi oluştururken, farklı ya da birbiriyle ilişkili olmayan nesneler ise başka bir küme oluşturur (Kaya ve Köymen, 2008). Kümeleme yöntemindeki amaç heterojen yapıdaki verileri benzer özelliklerine göre gruplayarak homojen sayılabilecek yapıda veri kümeleri oluşturmaktır. Kümeleme

işleminde her bir kümede yer alan veri dahil oldukları grupta yer alan diğer verilerle ortak özelliğe sahiptir. Şekil 2.3'te kümeleme işlemi sonucunda elde edilen veri kümeleri görülmektedir.



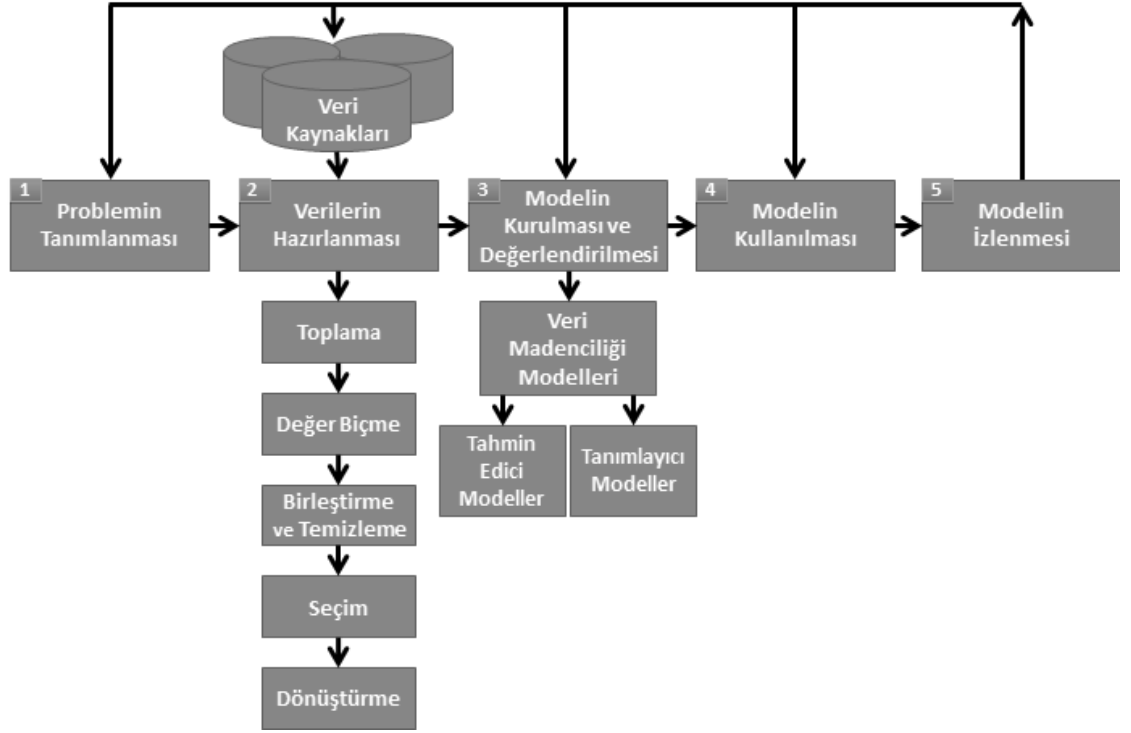
Şekil 2.3. Kümeleme İşlemi Sonrası Veri Kümeleri (Kaya ve Köymen, 2010)

Kümeleme yönteminde veri yığınlarına uygulanacak kümeleme işlemi öncesi küme sayısının ya da kümeler arasındaki maksimum minimum benzerlik ölçüsünün belirtilmesi gerekir. Böylece elde edilecek küme sayısı anlaşılabilir ölçüde olacak ve sonuçların anlaşılabilirliği artacaktır.

Kümeleme işleminin gerçekleştirilmesinde K-Means yöntemi en iyi bilinen ve en çok kullanılan yöntemdir. K-Means yöntemi kümeler içindeki nesnelerin ortalamaları üzerine kuruludur. K-Means yönteminin genel mantığı n adet veri nesnesinden oluşan bir veri kümesini, giriş parametresi olarak verilen k adet kümeye bölümlenmektir (Sarıman, 2011). Giriş parametresi olarak kullanılacak küme sayısı keyfi olarak belirlenmektedir. K-Means algoritması çalıştırıldığında n adet nesne içerisinde rastgele k adet nesne seçilmektedir. Seçilen k adet nesneden her biri bir kümenin merkezini ya da ortalamasını temsil etmektedir. Geri kalan nesneler ise oluşturulan kümelerin ortalama değerlerine olan uzaklıklarına göre en yakın oldukları kümelere dahil edilmektedir. Daha sonra her bir kümenin ortalama değeri yeniden hesaplanmakta ve yeni küme merkezleri belirlenerek tekrar nesnelerin merkez uzaklıkları tespit edilmektedir. Tüm nesneler kümelere yerleşinceye kadar bu işlem sürdürülmektedir.

2.1.2. Veri Madenciliği Süreci

Veri madenciliğinde veri yığınının bilgi elde edilmeden önce, eldeki verilerin bir takım işlemlere tabi tutulması gerekir. Bu işlemler sırasıyla aşağıda açıklanmıştır. Veri madenciliği süreci Şekil 2.4'te gösterilmektedir.



Şekil 2.4. Veri Madenciliği Süreci (Shearer, 2000)

2.1.2.1. Problemin Tanımlanması

Veri madenciliğinde yapılan çalışmanın amaca hizmet edebilmesi için yapılması gerek ilk şey problemin açık şekilde tanımlanmasıdır. Problemin doğru tanımlanması uygulanacak veri madenciliğiyle elde edilecek başarıyı arttıracak, problemin çözümü doğru ve hızlı bir şekilde gerçekleşecektir. Problemin yanlış tanımlanması ise zaman ve maliyet kaybına neden olacaktır. Bu yüzden problemin tanımlanmasında yanlış tahminlerde bulunulması halinde ortaya çıkacak maliyetin ve doğru tahminlerde kazanılacak faydalara ilişkin tahminlerin de bu aşamada yer alması gerekmektedir (Onat, 2008).

2.1.2.2. Verilerin Hazırlanması

Veri madenciliği uygulanırken oluşturulacak modelin veri kaynaklarının neler olduğunun belirlenmesi ve modelde kullanılmak için uygun hale getirilmesi bu aşamada

gerçekleştirilir. Modelde verilerden kaynaklı ortaya çıkan problemlerin meydana gelmesi bu aşamaya tekrardan dönülmesine ve verilerin yeniden düzenlenmesine neden olacaktır (Piramuthu,1998). Verilerin hazırlanması; toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme olmak üzere 5 aşamada gerçekleştirilir.

Toplama: Veri madenciliği uygulanması planlanan problem için gerekli olduğu düşünülen verilerin neler olduğu ve bu verilerin hangi kaynaklardan toplanacağını belirlendiği aşamadır. Verilerin hangi kaynaklardan toplandığı veri madenciliği uygulaması için önem arz etmektedir. Çünkü toplanan verilerin güvenilir olup olmadığı doğru sonuçlara ulaşıp ulaşılmayacağını etkileyecektir. Ayrıca yetersiz verinin toplanması da veri madenciliği sürecini etkileyeceği gibi gereğinden fazla veride zaman kaybına neden olacaktır. Bu nedenle veri madenciliği uygulamasından doğru sonuçlar elde etmek için toplanan verilerin yeterli ve güvenilir kaynaklardan olması gerekmektedir.

Değer Biçme: Veri madenciliğinde kullanılacak verilerin yeterli miktarda olması ve veri madenciliği sürecine yardımcı olmak için farklı yerlerden de veriler toplanması gerektiğinden veriler arasında kodlama, ölçü birimi farklılıkları gibi uyumsuzluklar olabilmektedir. Bu nedenle toplanan verilerin veri madenciliği sürecinde iyi sonuçlar verebilmeleri için ne ölçüde uyumlu olduklarının belirlenmesi gerekmektedir.

Birleştirme ve Temizleme: Birleştirme ve temizleme aşamasında farklı kaynaklardan toplanan veriler bir önceki aşamada uyum sorunu giderildikten sonra tek bir veri tabanında toplanır. Tek bir veri tabanında toplanan veriler içerisinde hatalı ve eksik olan veriler belirlenerek veri tabanından silinir. Temizleme aşamasında yapılacak hatalar sürecin başa dönmesine neden olabilir. Dolayısıyla birleştirme ve temizleme aşamasında amaca hizmet edecek veri kaybını önlemek için titiz davranılmalı ve sadece çıkarılması gereken veriler çıkarılmalıdır.

Seçim: Veri madenciliğinde kullanılacak modelin özelliğine göre verilerin seçiminin yapılması gerekir. Örneğin tahmin edici bir model kullanılacaksa bağımlı ve bağımsız değişkenler belirlenmeli ve veriler buna göre seçilmelidir. Ayrıca sonuca ulaşmada bir etkisi olmayan sıra numarası, kimlik numarası gibi ya da problemle ilgili olmayan diğer anlamsız verilerin kurulacak modele girmemesi gerekmektedir. Çünkü

bu tip veriler diğer verilerin modeldeki ağırlıklarını azaltacaklardır. Veri tabanındaki veri miktarının çok fazla olması durumunda seçilecek verilerin rastgereliği bozmayacak ve tüm popülasyonu temsil edecek şekilde seçilmesi gereklidir. Aynı zamanda çok miktarda verinin yer aldığı veri tabanları üzerinde çalışılırken birkaç farklı model denenerek bunlar arasından en güvenilir ve en güçlü modelin seçilmesi gerekir.

Dönüştürme: Modelde kullanılacak verilerin bazı kodlamalar kullanılarak tanımlanması veya gösterim şeklinin değiştirilmesi gerektiğinde kullanılmaktadır. Bazı veri madenciliği algoritmaları bu işlemleri otomatik olarak gerçekleştirebilmektedir ancak yine de dönüştürme işlemlerinin algoritmalara bırakılmadan yapılması daha güvenilir olacaktır.

2.1.2.3. Modelin Kurulması ve Değerlendirilmesi

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 ulaşılıncaya kadar yinelenen bir süreçtir.

Model kurulduktan sonra modelin değerlendirilmesi gerekmektedir. Modelin değerlendirilmesinde kullanılan en basit yöntem geçerlilik testidir. Geçerlilik testi büyük veri yığınları için yapılırsa basit geçerlilik, küçük veri yığınları için yapılırsa çapraz geçerlilik testi kullanılır. Basit geçerlilik testinde verilerin %5'lik ile %33'lük arasındaki bir kısmı test verisi olarak kullanılır. Çapraz geçerlilik testinde ise veriler rastgele iki eşit parçaya ayrılır. İlk olarak ayrılan birinci parça üzerinde modelin eğitimi ve ikinci parça üzerinde modelin test işlemi yapılır. Daha sonra ikinci parça üzerinde modelin eğitimi ve birinci parça üzerinde modelin test işlemi yapılır. Her iki işlem sonucunda ulaşılan hata oranlarının ortalaması kullanılır. Kurulan modelin doğruluk derecesi ne kadar yüksek olursa olsun gerçek dünyayı tam anlamıyla modelliğini garanti edebilmek mümkün değildir (Yaralıoğlu, 2004).

2.1.2.4. Modelin Kullanılması

Belirlediğimiz probleme uygun olarak kurulan ve geçerliliği test edilip kabul edilen bir model doğrudan problemin çözümünde kullanılabilir. Ayrıca başka problemlerin alt uygulaması olarak da kullanılabilir.

2.1.2.5. Modelin İzlenmesi

Belirlediğimiz probleme uygun olarak kurulan ve geçerli kabul edilen bir modelin zaman içerisinde problemin farklılaşması ya da verilerin değişmesi nedeniyle geçerliliğini yitireceğinden dolayı sürekli olarak izlenmesi ve yeniden düzenlenmesi gerekmektedir. Dolayısıyla tahmin edilen ve gözlenen değişkenler arasındaki farklılıklar takip edilerek model sonuçları izlenebilir ve modelin yeniden düzenlenip düzenlenmemesi konusunda bir karara varılabilir.

2.1.3. Veri Madenciliğinde Kullanılan Programlar

Çok büyük veri tabanlarındaki bilgilerin incelenerek, örüntü keşfi için doğru soruların sorulması, ilginç örüntülerin keşfi için hipotezlerin geliştirilmesi pratikte zor olduğundan dolayı ilgili örüntülerin keşfedilmesinde veri madenciliği programları kullanılır (Arslan, 2008).

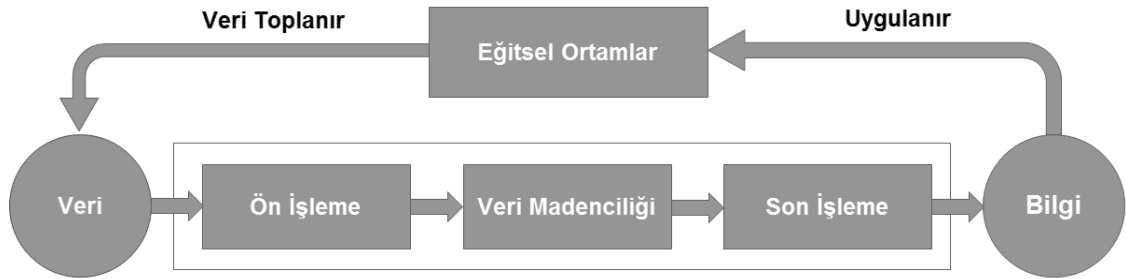
Veri madenciliği sürecinde verilerin analiz edilmesinde Analysis Manager, Clementine, DBMiner, Weka, Data Logic/R, Darwin/ODM, SAS Enterprise Miner, INLEN, KDW (Knowledge Discovery Workbench), SKICAT (Sky Image Classification & Archiving Tool), R-MINI, TASA (Telecommunication Network Alarm Sequence Analyzer), GCLUTO (Graphical Clustering Toolkit) programları kullanılmaktadır.

2.2. Eğitsel Veri Madenciliği

Eğitsel veri madenciliği; eğitim ortamlarından elde edilen verilerin veri madenciliği sürecine tabi tutularak elde edilen sonuçların öğrencilerin ve öğrenme ortamlarının daha iyi anlaşılmasında kullanılması ile ilgilenen bir disiplin olarak tanımlanmaktadır (Siemens & Baker, 2012).

2.2.1. Eğitsel Veri Madenciliği Süreci

Veri madenciliği sürecine benzer olarak eğitsel veri madenciliğinde de veri yığınlarından bilgi elde edilmeden önce eğitim ortamlarından elde edilen verilerin bir takım işlemlere tabi tutulması gerekir. Bu işlemler sırasıyla aşağıda açıklanmıştır. Eğitsel veri madenciliği süreci Şekil 2.5'te gösterilmektedir.



Şekil 2.5. Eğitsel Veri Madenciliği Süreci (García, Romero, Ventura ve de Castro, 2011)

Eğitsel Ortamlar: Eğitsel veri madenciliğinde kullanılan verilerin büyük bir bölümü öğretim yönetim sistemleri, zeki öğrenme ortamları, web 2.0 araçları, çevrimiçi değerlendirme sistemleri, geleneksel eğitim ortamlarından ya da araştırmacılar tarafından özel olarak geliştirilen öğrenme ortamlarından elde edilmektedir (Márquez-Vera, Cano, Romero, & Ventura, 2013; Romero & Ventura, 2013; Peña-Ayala, 2014).

Veri: Eğitsel veri madenciliğinde kullanılan veriler genellikle çevrimiçi öğrenme ortamlarından elde edilen log kayıtları, etkileşim verileri, ölçek verileri ve kontrollü deneylerden elde edilen verilerden oluşmaktadır (Cristóbal, Sebastián, Mykola, & Ryan, 2010). Ayrıca bunlara ek olarak gözlem yoluyla elde edilen veriler ya da öğrencilerin kendi ifadeleriyle yazdıkları mesajlarda veri kaynağı olarak kullanılabilir (Cavalcanti, Pires, Cavalcanti, & Pires, 2012).

Ön İşleme: Veri madenciliği sürecinde olduğu gibi eğitsel veri madenciliği sürecinde de veri yığınının bilgi elde edilmeden önce, eldeki verilerin bir takım ön işleme aşamasına tabi tutulması gerekir. Ön işleme aşamasının amacı veri kalitesini artırmak ve analiz aşaması için ideal değişkenleri elde etmektir (Hämäläinen & Vinni, 2010). Ön işleme aşaması veri madenciliği sürecinde yer alan verilerin hazırlanması aşamasıyla birebir benzerdir. Verilerin hazırlanması aşamasında olduğu gibi veri toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme adımları ön işleme aşamasında gerçekleştirilir.

Veri Madenciliği: Ön işleme aşaması gerçekleştirildikten sonra elde edilen veriler eğitsel veri madenciliğinin temelini oluşturan veri madenciliği yöntemlerine göre analiz edilir. Burada veri madenciliği sürecinde de açıklandığı gibi seçilen problem durumuna uygun olarak belirlenen veri madenciliği yöntemi kullanılarak analiz işlemi gerçekleştirilir. Yapılan çalışmalar incelendiğinde eğitsel veri madenciliğinde en çok kullanılan veri madenciliği yöntemlerinin sınıflama, kümeleme, regresyon ve ilişki madenciliği olduğu görülmektedir (Peña-Ayala, 2014; Romero & Ventura, 2010).

Son İşleme: Veri madenciliği sonunda elde edilen bilgilerin, model ya da örüntülerin yorumlanması işlemi gerçekleştirilir (Romero & Ventura, 2013).

Bilgi: Sonuç olarak elde edilen bilgiler karar verme sürecinde ya da eğitsel ortamların iyileştirilmesinde kullanılmaktadır.

2.3. Öğrenme Analitiği

Eğitimde teknoloji entegrasyonunun artmasıyla birlikte “öğrenme analitiği” kavramı ortaya çıkmıştır (Elias, 2011). Eğitsel veri madenciliği, akademik analitikler, sosyal ağ analizi gibi farklı tekniklerin devşirilmesiyle ortaya çıkan bu kavram yararlı bilgi elde etmek, kişinin önceki öğrenmelerini yansıtmak ve eğitim-öğretimi geliştirmek için kullanılır (Dyckhoff ve diğerleri, 2012; Akt: Yıldız ve Bahçeci, 2014).

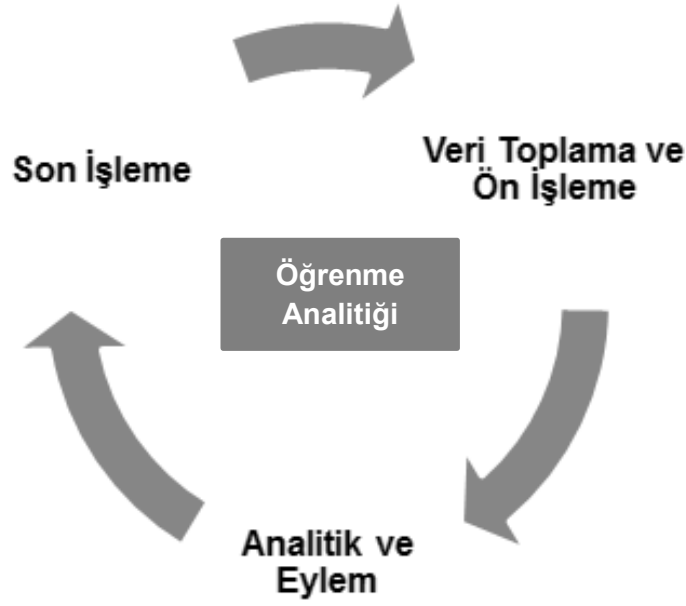
Öğrenme analitiği kavramı ilk kez Siemens (2010) tarafından “öğrenme üzerinde tahmin ve tavsiye yapabilmek için akıllı veri, öğrenenin ürettiği veri, bilgi ve sosyal bağlantıları keşfetmek için analizlerin kullanılması” olarak tanımlanmıştır. Johnson ve arkadaşları (2011) tarafından ise öğrencilerin ürettiği ve onlar adına toplanan verilerin büyük kısmının akademik ilerlemelerini, gelecekteki performanslarını ve potansiyel sorunlarını zamanında tahmin etmek için yorumlanması olarak tanımlanmıştır.

Öğrenme analitiği genel anlamda ise; öğrenci ve öğrencinin bulunduğu bağlam ile ilgili verilerin öğrenmenin ve öğrenmenin gerçekleştiği ortamın daha iyi anlaşılması ve optimize edilmesi amacıyla; ölçülmesi, toplanması, analiz edilmesi ve raporlanması olarak tanımlanmaktadır (Siemens & Baker, 2012).

2.3.1. Öğrenme Analitiği Süreci

Öğrenme analitiği farklı tekniklerin bir arada kullanılmasından dolayı eğitsel veri madenciliğinden daha fazla akademik disiplin içermekle birlikte, bilgisayar bilimleri, istatistik, psikoloji, sosyoloji, eğitim bilimleri alanlarından da kavram ve teknikler içermektedir (Bienkowski ve diğerleri, 2012).

Öğrenme analitiğinin amacı öğretmen ve okulların eğitsel olanaklarının her bir öğrencinin ihtiyaç ve yeteneğine göre uyarlanmasını olanaklı kılmaktır (Johnson, Smith, Willis, Levine, & Haywood, 2011; Akt: Akçapınar, 2014). Bu doğrultuda öğrenme analitiği sürecinin gerçekleştirilmesinde veri madenciliği ve eğitsel veri madenciliği süreçlerine benzer olarak veri yığınlarından bilgi elde edilmeden önce eğitim ortamlarından elde edilen verilerin bir takım işlemlere tabi tutulması gerekir. Öğrenme analitiği süreci; veri toplama ve ön işleme, analitik ve eylem, son işleme adımlarından oluşan tekrarlı bir süreçtir. Bu işlemler sırasıyla aşağıda açıklanmıştır. Öğrenme analitiği süreci Şekil 2.6’da gösterilmektedir.



Şekil 2.6. Öğrenme Analitiği Süreci (Chatti ve diğerleri, 2012).

Veri Toplama ve Ön İşleme: Veri madenciliği ve eğitsel veri madenciliği süreçlerinde olduğu gibi öğrenme analitiği sürecinde de ilk ve en önemli aşama verilerin güvenilir kaynaklardan elde edilmesi ve ön işleme aşamasına (veri toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme) tabi tutulmasıdır. Veriler bu işlemler gerçekleştirilerek bir sonraki aşama olan analitik ve eylem aşamasında kullanılmak üzere hazır hale getirilir.

Analitik ve Eylem: Veri toplama ve ön işleme aşaması gerçekleştirildikten sonra elde edilen verilere; eğitsel veri madenciliğinde olduğu gibi daha etkili öğrenme yaşantılarının sağlanmasına yardımcı olmak ve gizli örüntülerin ortaya çıkarılmasını sağlamak amacıyla farklı veri madenciliği teknikleri uygulanır. Analiz sonucunda elde edilen bilgilere dayanarak; izleme ve analiz, tahmin ve müdahale, mentörlük ve tutorluk, değerlendirme ve geri bildirim, uyarılma, kişiselleştirme ve tavsiye şeklinde eylemler gerçekleştirilir (Akçapınar, 2014).

Son İşleme: Öğrenme analitiği tekrar eden bir süreç olduğundan zaman içerisinde problemin farklılaşması ya da verilerin değişmesi nedeniyle; farklı kaynaklardan yeni verilerin eklenmesi, veri yığınının tekrar gözden geçirilmesi, mevcut değişkenlerin düzenlenmesi ve yeni analitik yönteminin seçilmesi gibi işlemler gerçekleştirilir (Beal, Qu, & Lee, 2006).

Eğitsel veri madenciliği ve öğrenme analitiği kavramları hedef ve ilgi alanları açısından birbirine benzer iki alan olmalarına karşın kullandıkları teknik ve yöntemler açısından bir takım farklılıklar göstermektedirler. Eğitsel veri madenciliği daha etkili eğitim ortamlarının oluşturulmasında eğitim ortamlarından elde edilen verileri veri madenciliği teknikleriyle analiz ederken, öğrenme analitiği yine aynı amaçla veri madenciliği tekniklerine ek olarak istatistik, görselleştirme, sosyal ağ analizi, duygu analizi, söylem analizi gibi yöntemleri de kullanmaktadır (Chatti ve diğerleri, 2012; Peña-Ayala, 2014; Romero & Ventura, 2007; Romero & Ventura, 2010).

2.3.2. Öğrenme Analitiğinin Eğitimde Kullanımı

Günümüzde her eğitim kademesinde eğitim uygulamalarının önemli bir parçası olan öğrencilerin; kişisel bilgileri, notları, başarılı ve başarısız olduğu dersler gibi birçok bilgi geniş veri tabanlarında tutulmaktadır. Anlamli ilişkilerin araştırılabileceği ve önemli bilgilerin elde edilebileceği bu veri yığınları; eğitimde aksaklıklara neden olan problemlerin tespitinde, öğrenci başarılarının tahmin edilmesinde, öğrenme ortamlarının öğrenci özelliklerine göre uyarlanabilmesinde, dersi bırakma eğilimi olan öğrencilerin tespitinde ve bu doğrultuda eğitim kalitesinin artırılmasında kullanılabilir (Kotsiantis, 2009; Özpolat & Akar, 2009; Romero ve diğerleri, 2010; Ting-Wen ve diğerleri, 2013). Öğrenme analitiğinin eğitimde kullanılması ilişkin yapılan bazı çalışmalar aşağıda özetlenmiştir.

Lopez ve diğerleri (2012) tarafından gerçekleştirilen çalışmada, öğrencilerin Moodle forum kullanım verilerinin ders başarıları üzerinde önemli bir etken olup olmadığı araştırılmıştır. Forum kullanım verileriyle ilişkili 8 değişken belirlenmiş ve bu değişkenler kullanılarak öğrencilerin ders başarıları geçti-kaldı şeklinde tahmin edilmeye çalışılmıştır. Ders başarılarının tahmin edilmesinde farklı sınıflama algoritmaları kullanılarak doğru sınıflama oranlarına göre performansları karşılaştırılmıştır. Çalışma sonucunda öğrencilerin Moodle üzerindeki forum kullanım verilerinin ders başarısının tahmin edilmesinde iyi bir değişken olduğu gösterilmiştir. Romero ve diğerleri (2008) tarafından gerçekleştirilen benzer bir çalışmada ise öğrencilerin Moodle forum kullanım verilerine ek olarak Moodle kullanımı ile ilgili olduklarını düşündükleri 9 farklı değişken belirlenmiş ve farklı sınıflama algoritmaları kullanılarak öğrencilerin ders başarıları tahmin edilmeye çalışılmıştır. Çalışma

sonucunda sınıflama algoritmaları kullanılarak elde edilen kuralların öğrencilerin sınıflandırılmasında, öğrencilere zamanında müdahale edilmesinde ya da karar verme sürecinde kullanılabileceği belirtilmiştir.

Macfadyen ve Dawson (2010) tarafından gerçekleştirilen çalışmada, öğretim yönetim sisteminden elde edilen ve öğrencilerin akademik başarılarıyla istatistiksel olarak anlamlı ilişki gösteren 15 adet değişken belirlenmiş ve lojistik regresyon yöntemi kullanılarak öğrencilerin akademik başarıları tahmin edilmeye çalışılmıştır. Çalışma sonucunda öğrenci etkileşimleriyle ilgili öğretim yönetim sisteminden elde edilen verilerden pedagojik olarak anlamlı bilgiler üretilebileceği belirtilmiştir. Benzer olarak Hu, Lo ve Shih (2014) tarafından gerçekleştirilen çalışmada, öğretim yönetim sistemi üzerindeki öğrenci etkileşim verileri analiz edilerek öğrencilerin derste sergileyebilecekleri performanslarını tahmin etmeye yarayan bir erken uyarı sistemi geliştirilmeye çalışılmıştır.

Bahçeci (2015) tarafından gerçekleştirilen çalışmada, alan yazındaki tüm çalışmalar derlenerek özellikle öğretim yönetim sistemleri üzerinde kullanılabilen Excel Pivot Tables, eLAT (Exploratory Learning Analytics Toolkit), SNAPP (The Social Networks Adapting Pedagogical Practice), CosyLMSA, ViMoodle ve LAe-R (Learning AnalyticsenrichedRubric) öğrenme analitiği araçları incelenmiştir. Çalışmada öğrenme analitiği araçlarını kullanan sistemlerin daha etkin olduğu ve öğrenci başarısını, etkileşim ve topluluk hissini arttırdığı sonucuna ulaşılmıştır.

2.4. Öğretim Yönetim Sistemleri (ÖYS)

Uzaktan eğitimin yanı sıra örgün eğitimde de sağladıkları katkılardan dolayı günümüzde birçok eğitim kurumunda kullanılan ÖYS'ler eğitimde yapılabilecek tüm uygulamaları üzerinde barındıran ve bunları web ortamına taşıyan web tabanlı sistemlerdir. ÖYS, öğrenim sürecini planlamayı, değerlendirmeyi, uygulamayı sağlayan bir yazılım ya da web tabanlı bir teknoloji olarak da tanımlanabilir (Aydın ve Biroğul, 2008). ÖYS'ler; eğitimcilere bilgi sunumu, öğrenme materyallerinin yönetimi ve paylaşımı, çoklu iletişim imkânları ve sınıf yönetimi konusunda sağladıkları olanaklarla derse yardımcı öğrenme ortamları olarak da tercih edilen sistemlerdir (Lonn ve Teasley, 2009).

ÖYS'ler ticari amaçlı ve açık kaynak kodlu olmak üzere iki şekilde kullanıma sunulmaktadır.

Ticari amaçlı ÖYS'ler kaynak kodları gizli ve belli bir lisans ücreti karşılığında kullanıcılarına kullanma imkânı sunan yazılımlardır. Türkiye'de Enocta ticari olarak bu alanda hizmet veren şirketlerden biridir. Ayrıca kurumlar ve üniversiteler, kendi bünyelerinde geliştirdikleri ÖYS'leri diğer kurum ya da üniversitelere pazarlayabilmektedirler (Reis ve diğerleri, 2012). Ticari amaçlı bazı ÖYS'ler şunlardır; Blackboard, CCNet, eCollege, Fedena, GeoLearning, HotChalk, Informetica, It's learning, Learn.com, Meridian Knowledge Solutions, Plateau Systems, SharePointLMS, SSLearn, Thinking Cap LMS ve Vitalect.

Açık kaynak kodlu ÖYS'ler kaynak kodları gizli olmayan ve kullanıcılardan belli bir lisans ücreti talep edilmeyen yazılımlardır. Açık kaynak kodlu ÖYS'ler kullanıcılara yazılımı sınırsız kullanma özgürlüğü sunmasının yanında yazılımın kaynağına rahatça erişebilme, yazılımı değiştirme, geliştirme, kopyalama ve dağıtma izni de veren sistemlerdir (Stallman, 2012). Açık kaynak kodlu ÖYS'lerin sahip oldukları genel özellikleri aşağıda yer almaktadır (Duran ve diğerleri, 2006).

- Açık kaynak kodlu ÖYS'ler öğretim elemanlarının özel bir eğitim almasına gerek kalmadan kullanabilecekleri şekilde tasarlanmış sistemlerdir.
- Geniş bir kitlenin eğitim amaçlarına göre kullanabileceği şekilde yapılandırıldıklarından dolayı dil desteği içerirler ve kullanıcılar sistemi kendi ana dillerinde kullanım imkânına sahiptirler.
- Öğrencilerin bireysel farklılıklarını dikkate alan öğrenme yaklaşımlarına uygun pedagojik gerekliliklere göre biçimlendirilmişlerdir.
- Standartlaştırılmış ilkelere (SCORM standartları) uygunluk sağlayan içerik yönetimi sunarlar.
- Yöneticilere, yetkilendirmede esneklik sağlarlar.
- Öğretim elemanlarına, derse katılan öğrencilerin ÖYS'deki hareketlerini gösteren istatistikler sunarlar.
- Öğrenenlerin gruplar halinde çalışmasına imkân sunarak işbirlikli öğrenmeye yardımcı olurlar.

Moodle, ATutor, Dokeos, Bodington, Fle3 Learning Environment, Claroline, Docebo, eStudy, Drupal, DotLRN, eFront, Sakai, OLAT, en çok kullanılan açık kaynak kodlu ÖYS'lerdir (Altınparmak ve diğerleri, 2011). Moodle ve ATutor yazılımları, kurulum, çevrimiçi işbirliği ve iletişim, tasarım ilkeleri, verimlilik araçları, sistem, içerik, kurs ve değerlendirme yönetimi, destek, güvenlik ve küçük-büyük ölçekli kurumlara uygunluk açılarından diğer ÖYS yazılımlarına göre daha kullanışlı ve yönetilebilir niteliktedirler (Reis ve diğerleri, 2012).

2.4.1. Moodle

Başkent Üniversitesi tarafından açık kaynak kodlu ve üniversite tarafından kendi ihtiyaçları doğrultusundan uyarlanan Moodle kullanılmaktadır. Moodle, 1999'da Martin Dougiamas tarafından geliştirilmeye başlanan, ücretsiz olarak indirilebilen, PHP programlama dilini destekleyen ve Windows, Linux, Mac OS X gibi farklı işletim sistemlerinde kullanılabilen açık kaynak kodlu bir öğretim yönetim sistemidir. Kelime olarak Moodle "Modular Object Oriented Dynamic Learning Environment" yani "Esnek Nesne Yönelimli Dinamik Öğrenme Ortamı" anlamındadır (Dougiamas, 2014). İlk sürümü 2002'de kullanılmaya başlanmış ve son olarak 2.8.2 sürümü 2015 Ocak ayında kullanıma sunulmuştur. Moodle, 220 ülkede yaklaşık 53 bin kayıtlı siteyle 1,5 milyon eğitimci ve 70 milyon kullanıcı tarafından kullanılmaktadır. Moodle ait istatistikler Tablo 2.1'de yer almaktadır.

Tablo 2.1. Moodle İstatistikleri

(<https://moodle.net/stats/> sayfası 2015 yılı verilerinden uyarlanmıştır.)

Moodle İstatistikleri	
Kayıtlı Site	52.966
Ülkeler	220
Kurslar	7.801.045
Kullanıcılar	70.722.240
Forum Mesajları	143.625.147
Kaynaklar	71.685.726
Sınav Soruları	302.197.966

Moodle, açık kaynak kodlu olmasından dolayı özelleştirilebilir bir yapıya sahiptir. Aynı zamanda geleneksel eğitimle verilen dersleri ve çevrimiçi verilen dersleri desteklemek için öğrenme nesneleri oluşturma imkânı da sunmaktadır. Budan dolayı birçok ÖYS’de olmayan pedagojik özelliklere sahiptir. Moodle’ın bu özellikleri, öğrencilerin aktif birer katılımcı olmasına ve öğrenme sorumluluğunu üstlenmesine yardımcı olmaktadır (Lakhan ve JhunJhunwala, 2008). Moodle’nin en önemli özelliği ise tüm kullanıcılar tarafından kolay kullanılabilmesidir. Moodle’nin diğer genel özellikleri aşağıda yer almaktadır.

- Moodle ücretsizdir.
- Windows, Linux ve Mac OS X işletim sistemleri altında kullanılabilir.
- Ticari ÖYS’lerle (Blackboard, CCNet, eCollege, Fedena, GeoLearning, HotChalk, Informetika, It’s learning, Learn.com, Meridian Knowledge Solutions, Plateau Systems, SharePointLMS, SSLearn, Thinking Cap LMS, Vitalelect) yarışmakta olup kullanım açısından büyük bir paya sahiptir.
- 220 ülkede 120 dil desteği mevcuttur.
- İstenilen dil paketi indirilip sisteminize yüklenebilir ve dil paketi sizin tarafınızdan geliştirilebilir.
- Kullanıcılar hiç bir programlama ve veri tabanı bilgisine sahip olmadan kullanabilmektedir.
- Sistemdeki eksiklikler ya da sorunlar sizin tarafınızdan çözülebildiği gibi geliştirici kitle tarafından çok hızlı bir şekilde giderilebilmektedir.
- Ücretsiz olduğundan kullanıcı kitlesi çok geniştir.
- Geliştirici kitle tarafından yeni özellikler sürekli olarak eklenmekte ve ücretsiz olarak yayınlanmaktadır.

Moodle 220 ülkede üniversiteler, liseler, ilköğretim okulları, devlet daireleri, sağlık örgütleri, askeri kurumlar, havayolları, petrol şirketleri, özel eğitimciler başta olmak üzere birçok alanda çeşitli kurum ve bireyler tarafından kullanılmaktadır. Moodle sistemine kayıtlı 52.966 site arasından ülkelere göre site sayısı oranı en fazla olan ilk 10 ülke Tablo 2.2’de yer almaktadır.

Tablo 2.2. Moodle Kullanım Oranı En Yüksek İlk 10 Ülke
(<https://moodle.net/stats/> sayfası 2015 yılı verilerinden uyarlanmıştır.)

Ülke	Kayıtlı Site
Amerika Birleşik Devletleri	8.467
İspanya	5.782
Brezilya	3.589
Birleşik Krallık	2.812
Meksika	2.127
Almanya	2.082
İtalya	1.573
Kolombiya	1.487
Avustralya	1.370
Rusya Federasyonu	1.284

Moodle bilgi aktarımı, içerik oluşturma, ödev ve sınav hazırlama, tartışma platformu, uzaktan sınıf yönetimi, dosya depolama gibi işlemleri gerçekleştirebilmekte ve sistem üzerinde gerçekleştirilen tüm bu hareketliliği kaydedebilmektedirler (Mostow, 2004). Ancak kaydedilen bu verilerden öğrenci sayısının artması durumunda eğitimciler tarafında anlamlı bilgiler çıkarılması güçleşmektedir. Bu nedenle anlamlı bilgilerin ortaya çıkarılmasında, veri örüntülerinin tanımlanmasında, verilerin görselleştirilmesi ve analiz edilmesinde eğitsel veri madenciliği ve öğrenme analitiğinin temelini oluşturan veri madenciliğinden faydalanılabilir (Talavera ve Gaudioso, 2004).

2.5. Yurt İçi ve Yurt Dışında Yapılan Çalışmalar

Birçok araştırmacı veri madenciliği, eğitsel veri madenciliği ve öğrenme analitiklerinin çeşitli uygulamalarını araştırmakta ve bu uygulamaların önemini anlamak, ilgili alandaki kaliteyi artırmak için birçok çalışma yapmaktadır. Bu bağlamda gerek yurt içinde gerek yurt dışında yapılan çalışmalar gün geçtikçe artmaktadır.

Eğitim alanında yapılan çalışmalar geleneksel ve uzaktan eğitim ortamlarından elde edilen veriler kullanılarak gerçekleştirilmektedir. Geleneksel eğitimde yapılan çalışmalar; inceleme ve tanıtmaya (Barahate, 2012; Siemens ve Baker, 2012; Romero ve Ventura, 2013; Kumar ve Vijayalakshimi, 2013; Sharma ve Singh, 2013; Ali, 2013;

Bhise, Thorat ve Supekar, 2013), öğrencilerin akademik başarı ve başarısızlıkları ile bunları etkileyen faktörlerin belirlenmesi (Superby, Vandamme ve Meskens, 2006; Ayık, Özdemir ve Yavuz, 2007; Tosun, 2007; Halees, 2008; Bozkır ve diğerleri, 2009; Baker, Gowda ve Corbett, 2011; Birtıl, 2012; Taşdemir, 2012; Tiwari, Singh ve Vimal, 2013; Bilen, Hotaman, Aşkın ve Büyüklü, 2014), öğrenci özelliklerini belirleme ve bunlara göre öğrencilerin gruplanması (Kelly ve Tangney, 2005; Bresfelean ve diğerleri, 2008; Gaafar ve Khanmis, 2009; Zhang ve diğerleri, 2010; Koç ve Karabatak, 2012; Şengür ve Tekin, 2013; Gülen ve Özdemir, 2013; Hark, 2013) üzerine yoğunlaşmış durumdadır. Uzaktan eğitimde de benzer amaçlar doğrultusunda çalışmalar gerçekleştirilmiştir (Lee, 2007; Kayri, 2008; Şuşnea, 2009; Whitmer, Fernandes & Allen, 2012; Leony, Pardo, Valentin, Quinones & Kloos, 2012; San Diego ve diğerleri, 2012; Sevindik ve Kayışlı, 2012; Jiang ve Cai, 2013; ALMazroui, 2013). Bu alanda yapılan çalışmalardan bazıları aşağıda özetlenmiştir.

Kayri (2008) tarafından gerçekleştirilen çalışmada, öğrenci performans göstergelerinin sürekli izlenebildiği ve ürünler arasındaki örüntünün bilgisayar sistemleri tarafından oldukça kolay yapılabilirdiği e-portfolio değerlendirmeleri için veri madenciliğinde kullanılan yöntemlerin alternatif bir ölçme yaklaşımı olarak kullanımı önerilmiştir.

Halees (2008) tarafından gerçekleştirilen çalışmada, Gaza Islamic Üniversitesinde 2007-2008 eğitim öğretim döneminde 151 öğrenciye ait kişisel, akademik ve e-öğrenme sistemi ile ilgili kayıtlar kullanılarak öğrenci davranışları veri madenciliği ile değerlendirilmeye çalışılmış ve ulaşılan sonuçlara göre öğrencilerin performanslarının iyileştirilmesi amaçlanmıştır. Benzer olarak Bresfelean ve diğerleri (2008) tarafından gerçekleştirilen çalışmada veri madenciliğindeki sınıflandırma ve veri kümeleme yöntemlerini kullanarak öğrenci profilleri ortaya çıkarılmaya çalışılmış ve akademik başarısızlık nedenlerinin saptanarak öğrenci başarılarının iyileştirilmesi amaçlanmıştır.

Gaafar ve Khanmis (2009) tarafından Kahire Amerikan Üniversitesinde gerçekleştirilen çalışmada, farklı veri tabanlarından elde edilen veriler kullanılarak oluşturulan veri yığnında farklı veri madenciliği yöntemleri kullanılarak başarılı yani mezun olabilecek, başarısız yani okulu bırakabilecek şeklinde iki farklı öğrenci

profilinin modellenmesi amaçlanmıştır. Zhang ve diğerleri (2010) tarafından gerçekleştirilen çalışmada, veri madenciliğinin risk altındaki öğrencilere nasıl yardım edebileceğini, dersin uygunluğunun nasıl değerlendirilebileceğini ve elde edilen sonuçların öğrencilere nasıl uyarlanabileceğini araştırmışlardır. Birtıl (2012) tarafından gerçekleştirilen çalışmada, öğrencilerin başarısızlık nedenlerini belirlemek için uygulanan anket veri madenciliği yöntemlerinden kümeleme yöntemi kullanılarak incelenmiş, öğrencileri başarısızlığa iten etkenlerin hangilerinin aynı anda görüldüğün tespit edilmesi ve bunlar arasındaki ilişkilerin belirlenmesi amaçlanmıştır. Şengür ve Tekin (2013) tarafından gerçekleştirilen çalışmada ise 2011 yılında mezun olan 127 Bilgisayar ve Öğretim Teknolojileri Öğretmenliği öğrencisinin 4 yıl süresince almış oldukları 49 kültür ve mesleki dersin yılsonu notları kullanılarak mezuniyet notlarının tahmin edilmesi amaçlanmıştır.

Hark (2013), tarafından gerçekleştirilen çalışmada, 905 öğrenciden anket yoluyla elde edilen verilerden veri madenciliği yöntemlerinden birliktelik kuralları kullanılarak etkileşimli tahta kullanımına dair öğrenci tutumları incelenmeye çalışılmıştır.

Gülen ve Özdemir (2013) tarafından gerçekleştirilen çalışmada, Ankara Bilim ve Sanat Merkezine devam eden 12 ve üstü yaşındaki üstün yetenekli öğrencilere Akademik Benlik Kavramı Ölçeği ve Boş Zamanları Değerlendirme Anketi uygulanarak elde edilen verilerden Üstün yetenekli öğrencilerin sıklıkla bir arada ilgi duydukları alanlar Apriori birliktelik algoritması kullanılarak tespit edilmeye çalışılmıştır. Çalışmadan elde edilen bulgulardan Bilim ve Sanat Merkezlerindeki üstün yetenekli öğrencilerin eğitiminin bireysel ihtiyaçlarına göre farklılaştırılması ve ders programlarının daha etkin olarak düzenlenmesi gibi konularda pek çok fayda sağlayacağı öngörülmüştür.

Bozkır ve diğerleri (2009) tarafından gerçekleştirilen çalışmada, 2008 yılı ÖSS dönemine ilişkin ÖSYM'nin internet sitesinde yer alan bilgi anketinden elde edilen veriler veri madenciliği kullanılarak analiz edilmiş ve ÖSS'de başarıyı temsil etmekte kullanılan sözel, sayısal, eşit ağırlık puanları ve bu puan türlerinde Türkiye'deki başarı sırasıyla birlikte Ortaöğretim Başarı Puanı ve Ağırlıklı Ortaöğretim Başarı Puanı gibi sonuçları etkileyen en önemli faktörlerin bulunması amaçlanmıştır. Bu çalışmaya benzer

olarak Taşdemir (2012) tarafından gerçekleştirilen çalışmada, ÖSYM tarafından gönderilen öğrenci verilerinden ve öğrencilerin öğrenim sırasındaki ders başarılarından yararlanılarak öğrencilerin akademik başarısına etki eden etmenler veri madenciliği yöntemlerinden regresyon kullanılarak bulunmaya çalışılmıştır. Bilen ve diğerleri (2014) tarafından gerçekleştirilen çalışmada ise İstanbul ilinde 2011 yılında LYS sınavına giren 42 farklı lise türünün sınavdaki başarı performanslarına göre kümelene mesi amaçlanmıştır. Kümeleme işleminin gerçekleştirilmesinde ve hangi test türlerinin kümelere ayrışmada daha etkili olduğunun belirlenmesinde veri madenciliği yöntemlerinden kümeleme ve karar ağaçları teknikleri kullanılmıştır.

Yapılan çalışmalara bakıldığında eğitimde veri madenciliği çalışmalarının birçoğunun geleneksel eğitim ortamlarından elde edilen veriler kullanılarak gerçekleştirildiği görülmektedir. Uzaktan eğitim sistemlerinden elde edilen veriler kullanılarak gerçekleştirilen veri madenciliği çalışmalarının geleneksel eğitim ortamlarına göre daha az olduğu dikkat çekmektedir. Yapılan çalışmalarda, öğrenci akademik başarısı ve başarısızlıkları ile bunlara etki eden etmenlerin belirlenmesi, öğrenci profillerinin ortaya çıkarılarak bunlara göre öğrencilerin gruplandırılması, akademik başarısı düşük olan risk altındaki öğrencilere nasıl yardım edilebileceğinin belirlenmesi, öğrencilerin olası mezuniyet notlarının tahmin edilmesi gibi sonuçlara ulaşmaya çalışılarak eğitimdeki eksikliklerin tespit edilmesi ve buna bağlı olarak eğitim kalitesinin artırılmaya çalışıldığı görülmektedir.

3. YÖNTEM

Bu bölümde; araştırmanın modeli, araştırma örnekleme ve özellikleri, veri analizi aşaması ve araştırma sürecine yer verilmiştir.

Bu araştırmanın amacı; veri madenciliği tekniklerini kullanarak üniversite (lisans) düzeyindeki öğrencilerin Öğretim Yönetim Sistemi (ÖYS) üzerindeki hareketliliği ile yılsonu akademik başarıları arasındaki ilişkiyi belirlemektir. Bu amaç 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?

3.1. Araştırmanın Modeli

Bu çalışmada, üniversite (lisans) 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 varolan bir durumu varolduğu 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 araştırmaya konu 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 ve diğerleri, 2012).

3.2. Araştırma Örneklemi

Araştırmanın örneklemini 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. Araştırmanın örneklemini oluşturan öğrencilerin tamamı Fen Edebiyat Fakültesi Matematik Bölümü mezunu öğretmen adaylarıdır. Araştırmanın örneklemini oluşturan öğrencilerin demografik özellikleri ile ilgili tanımlayıcı istatistikler Tablo 3.1'de yer almaktadır.

Tablo 3.1. Araştırma Örneklemine Ait Demografik Bilgiler

Cinsiyet	f	%
Erkek	10	25,00
Kız	30	75,00
Toplam	40	100,00

Pedagojik Formasyon Sertifika Programına ait dersler akşam 17.00-20.00 saatleri arasında işlenmekte olup bu program kapsamında verilen Öğretim Teknolojileri ve Materyal Tasarım dersi 5 ayrı şubede verilmekte ve bütün şubelerde ortak müfredat, ortak kitap ve ortak hedefler güdülmektedir. Dersi veren tüm öğretim elemanları tarafından 5 şubede de aynı stratejiler kullanılmaktadır. Ancak araştırmaya diğer şubeler dahil edilmemiştir. Derse ait notlar; klasik ve çoktan seçmeli şekilde yapılan vize ve final sınavlarından elde edilen notlar ve öğretim elemanı tarafından belirlenen, Moodle üzerinden detaylı olarak açıklanan ve belirli bir zaman diliminde gönderilmesi gereken 5 ödev ile öğretim elemanının kanaat notundan oluşmaktadır. Derse ait geçme sisteminin not ağırlıkları Tablo 3.2'de yer almaktadır.

Tablo 3.2. 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

Başkent Üniversitesi Pedagojik Formasyon Sertifika Programı kapsamında Moodle ÖYS'nin derslerde kullanım teknik ve yöntemi ders sorumlusu öğretim elemanına aittir. Moodle, 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 Moodle'nin derslerde kullanımı dersi veren öğretim elemanının Moodle'yi kullanma istek ve becerisine bağlı olarak değişmektedir.

Araştırmanın örneklemini oluşturan öğrencilere dersi veren öğretim elemanı 2007 yılından beri Moodle ÖYS'yi bütün derslerde etkin olarak kullanmaktadır. Öğretim elemanı Başkent Üniversitesinde Moodle yöneticiliği yapmaktadır. Ayrıca ÖYS'ler ve video konferans sistemleri üzerine çalışmalar yapmaktadır. Araştırmanın örneklemini oluşturan öğrencilerin tamamı ise Moodle ÖYS'yi daha önce hiç kullanmamış olup, 2012-2013 eğitim öğretim yılı bahar dönemi Pedagojik Formasyon Sertifika Programı, Öğretim Teknolojileri ve Materyal Tasarımı dersinde Moodle ÖYS'yi ilk defa kullanmışlardır.

3.3. Veri Kaynakları

Araştırma problemleri doğrultusunda öğrencilerin 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersine ait yılsonu akademik başarıları notları ve öğrencilerin Moodle üzerindeki hareketliliklerine ait log kayıtları olmak üzere iki farklı veri kaynağı kullanılmıştır.

Öğrencilerin 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersinde ait yılsonu akademik başarı notları ve öğrenci hareketliliğine ait log kayıtları derse giren öğretim elemanı tarafından Moodle üzerinden indirilerek araştırmacıya verilmiştir.

Moodle ÖYS'ye erişimde herhangi bir tarayıcı kısıtlaması bulunmamaktadır. Öğretim elemanları ve öğrenciler tarafından sisteme giriş işlemi kullanıcı adı ve şifre ile gerçekleştirilmektedir. Moodle sistemine giriş yapıldıktan sonra öğretim elemanı ve öğrenciler tarafından gerçekleştirilen hareketlilik yapılan fare tıklamalarına göre kayıt altına alınmaktadır. Moodle tarafından, dersin adı, tarih ve saat, IP adresi, giriş yapan kullanıcının adı, gerçekleştirdiği eylem, gerçekleştirdiği eyleme ait bağlantı adresi, gerçekleştirilen eylem bağlantısına ait bilgi alanı kayıt altına alınmaktadır. Ancak kayıt

altına alınan bu veriler Moodle tarafından herhangi bir veri madenciliği işlemine tabi tutulmamakta sadece ham veri şeklinde isteyen öğretim elemanları tarafından indirilebilmektedir.

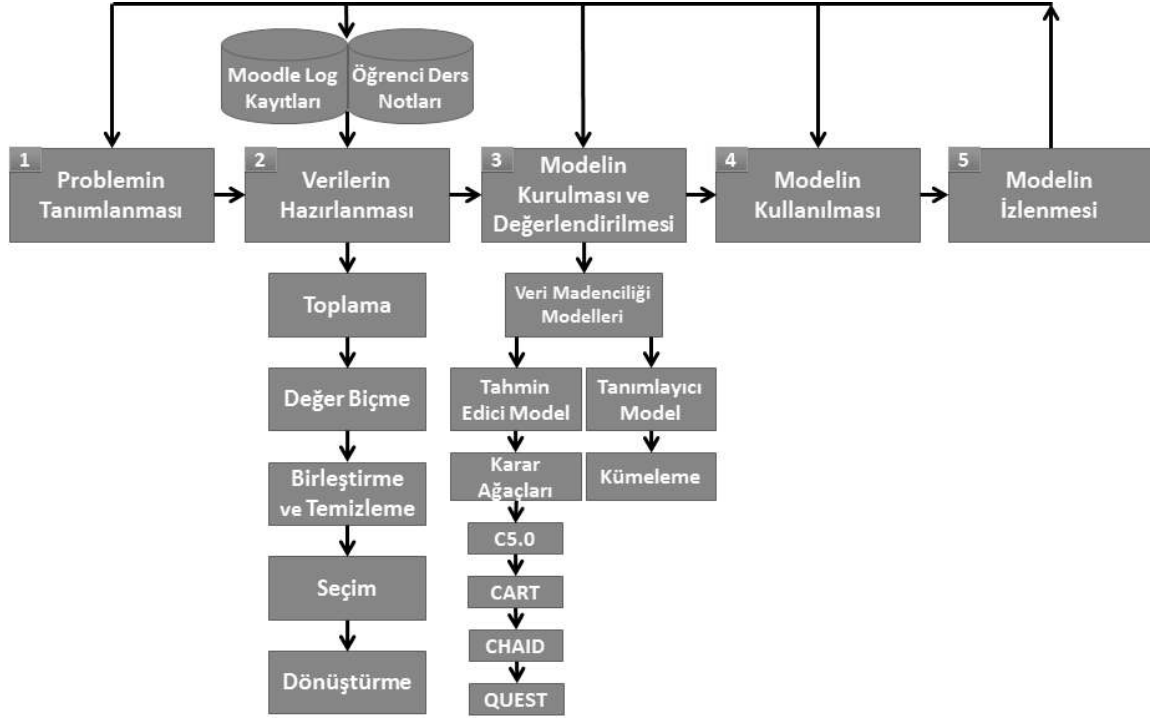
3.4. Verilerin Analizi

Öğretim Teknolojileri ve Materyal Tasarımı dersi için Moodle ÖYS üzerindeki öğrenci hareketliliği verilerinin analiz aşamasında kullanım yaygınlığı, kolay uyum, algoritma çeşitliliği, ara yüz kolaylığı, veri madenciliği ve bilgi keşfi konusunda yayın ve araştırma yapan KDnuggets dergisinin sektör analiz çalışmaları sonucunda 3 yıldır üst üste “en çok tercih edilen Veri Madenciliği çözümü” seçilmesi gibi üstünlüklerinden dolayı SPSS Clementine 12.0 kullanılmıştır. SPSS Clementine 12.0 IBM firması tarafından geliştirilmiş veri madenciliği uygulamalarında kullanılan görsel modelleme aracıdır.

SPSS Clementine 12.0 programı ilgili firmanın internet sitesinden demo sürümü (30 günlük süreyle programın tüm özelliklerini kullanma imkanı vermektedir.) indirilerek temin edilmiştir.

Çalışmanın gerçekleştirilmesinde 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersini alan 40 öğrencinin Moodle ÖYS üzerindeki hareketliliğini içeren log kayıtları ve yılsonu akademik başarı notları kullanılmıştır. Veri madenciliği uygulanırken oluşturulacak modelinin veri kaynaklarının neler olduğunun belirlenebilmesi, modelin kurulabilmesi ve verilerin analize edilebilmesi için veri madenciliği sürecinde yer alan verilerin hazırlanması adımları (veri toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme) uygulanmıştır. Veriler gerekli veri hazırlama aşamalarına (veri toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme) 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.

Şekil 3.1’de Moodle üzerinde gerçekleştirilen öğrenci hareketliliğinin veri analizi süreci gösterilmektedir. Veri analizi sürecinde gerçekleştirilen işlemler detaylı olarak aşağıda açıklanmıştır.



Şekil 3.1. Öğrenci Hareketliliği Veri Madenciliği Süreci (Shearer, 2000’den uyarlanmıştır.)

3.4.1. Problemin Tanımlanması

Veri madenciliğinde yapılan çalışmanın amaca hizmet edebilmesi için ilk olarak problemin açık bir şekilde tanımlanması gerekmektedir. Bu aşamada problemin doğru tanımlanması uygulanacak olan veri madenciliğinden elde edilecek olan başarıyı arttıracak, problemin çözümünün doğru ve hızlı bir şekilde gerçekleşmesini sağlayacaktır.

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

3.4.2. Verilerin Hazırlanması

Verilerin hazırlanması aşamasında; veri toplama, değer biçme, birleştirme ve temizleme, seçim, dönüştürme adımları gerçekleştirilmiştir.

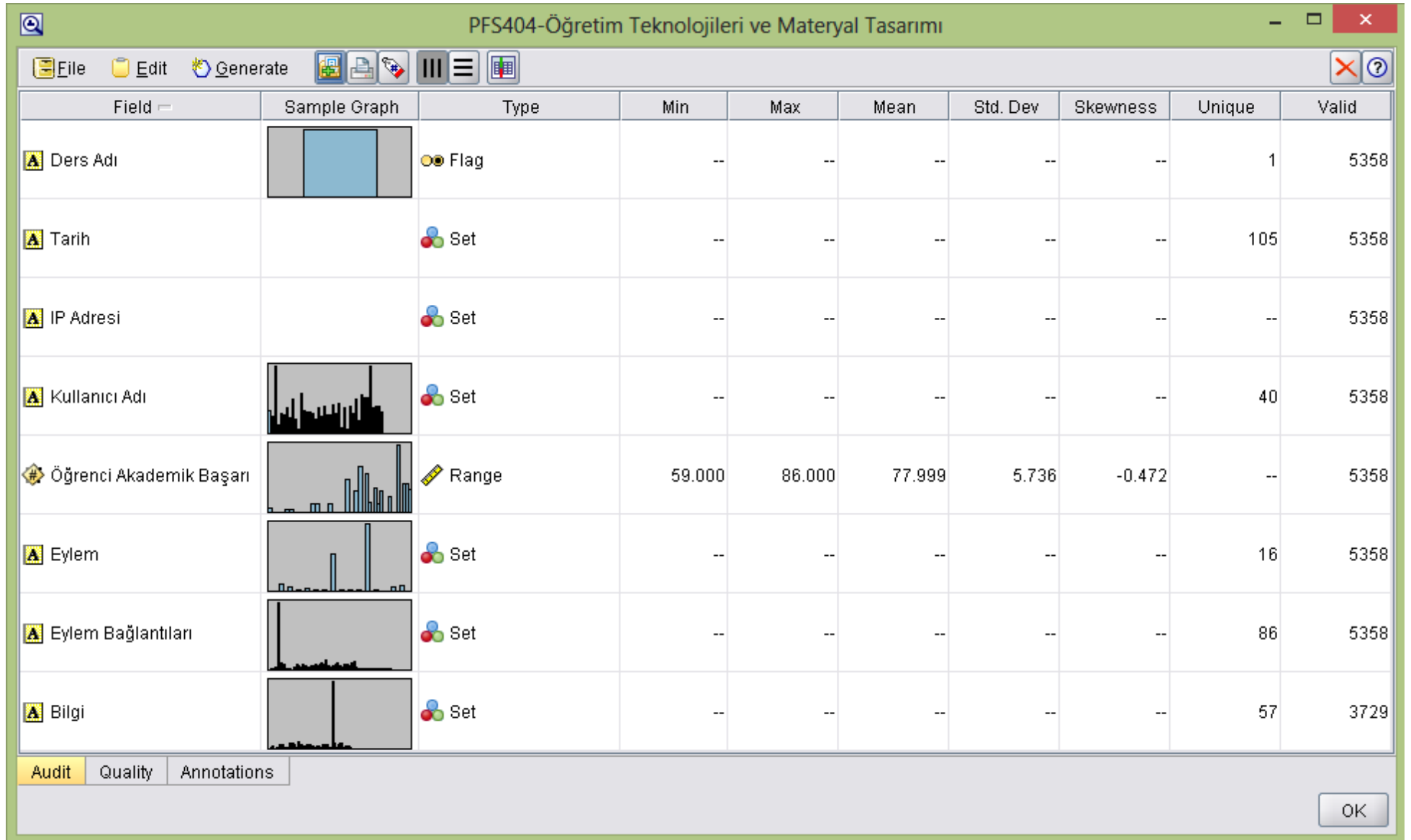
1. *Toplama:* Verilerin hazırlanması aşamasında ilk olarak veri toplama işlemi gerçekleştirilmiştir. Bu aşamada veri madenciliği için gerekli olduğu düşünülen verilerin hangi kaynaklardan toplanacağına karar verilmiş ve ilgili kaynaklardan veriler elde edilmiştir. Çalışma kapsamında kullanılan Öğretim Teknolojileri ve Materyal Tasarımı dersini alan 40 öğrencinin ÖYS üzerindeki hareketliliğini içeren veriler Moodle aracılığıyla kaydedilmiş log kayıtlarından elde edilmiştir. Ayrıca öğrencilere ait 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersi yılsonu akademik başarı notları dersi veren ilgili öğretim elemanından temin edilmiştir.

2. *Değer Biçme:* Verilerin hazırlanması aşamasında veri toplama işlemi tamamlandıktan sonra değer biçme işlemi gerçekleştirilmiştir. Bu aşamada 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 Moodle kayıtları aracılığıyla elde edildiğinden dolayı veriler arasında kodlama, ölçü birimi gibi farklılıklar bulunmamaktadır. Ancak SPSS Clementine 12.0 programına verilerin bağlanabilmesi, veri birleştirme ve temizleme işlemlerinin daha kolay gerçekleştirilebilmesi için CSV (Comma Seperated Values) kayıt türünde olan log kayıtları Excel'e çevrilmiş ve veri madenciliği için uygun hale getirilmiştir.

3. *Birleştirme ve Temizleme:* Verilerin hazırlanması aşamasında değer biçme işlemi tamamlandıktan sonra bir sonraki işlem olan birleştirme ve temizleme işlemi gerçekleştirilmiştir. Bu aşamada veri madenciliği sürecinde kullanılacak olan veriler birleştirilerek hatalı ve eksik olan veriler veri tabanından silinmiştir. Öğretim Teknolojileri ve Materyal Tasarımı dersine ait Moodle 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 Öğretim Teknolojileri ve Materyal Tasarımı dersine ait yılsonu akademik başarı notları veri birleştirme işlemine tabi tutularak veri temizleme işlemi sonucu elde edilen veri yığınınına eklenmiştir. Veri birleştirme ve temizleme işlemi sonrasında SPSS Clementine programı kullanılarak elde edilen veri yığınınına ait bilgiler Şekil 3.2’de yer almaktadır.

Şekil 3.2’de yer alan veri yığınınına ait bilgiler; birbirinden farklı her bir veri sütununda (Ders Adı, Tarih, Saat, IP Adresi, Kullanıcı Adı, Eylem, Eylem Bağlantıları ve Bilgi Alanı) kaç farklı değişkenin olduğunu ve bu değişkenlere ait toplamdaki veri kaydı 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.



Şekil 3.2. Veri Birleştirme ve Temizleme İşlemi Sonrası Veri Yığını

4. *Seçim:* Verilerin hazırlanması aşamasında birleştirme ve temizleme işlemi tamamlandıktan sonra seçim işlemi gerçekleştirilmiştir. Bu aşamada veri madenciliği sürecinde kullanılacak olan modelin özelliğine göre verilerin seçim işlemi yapılmıştır. Veri madenciliği sürecinde kullanılacak olan Öğretim Teknolojileri ve Materyal Tasarımı dersine 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 modelin oluşturulmasında modele dahil edilmemiştir. Çünkü problemle ilgisi olmayan bu tip veriler diğer verilerin modeldeki ağırlıklarını azaltacaklardır. Seçim aşamasında veri madenciliği sürecinde Kullanıcı Adı, Eylem ve Öğrenci Akademik Başarısı alanları modelde kullanılacak olan veriler olarak belirlenmiştir. Modelde kullanılacak olan bu alanlara ait bilgiler aşağıda detaylı olarak açıklanmıştır.

Kullanıcı Adı: Çalışmada kullanılacak olan 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: Moodle; 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. Moodle 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. Moodle tarafından kaydedilen log kayıtları incelendiğinde Öğretim Teknolojileri ve Materyal Tasarımı dersini alan 40 öğrenci tarafından toplamda 16 adet farklı eylemin gerçekleştirildiği tespit edilmiştir. Moodle üzerinde gerçekleştirilen bu eylemler Tablo 3.3'te yer almaktadır. Eylemlerin işlevlerine ait detaylı açıklamalar Ek-1'de sunulmuştur.

Tablo 3.3. Moodle Üzerindeki Eylemler

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
course user 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 Öğretim Teknolojileri ve Materyal Tasarımı dersine ait öğrenci yılsonu akademik başarı notları Tablo 3.4'te gösterilmektedir.

Tablo 3.4. 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*: Verilerin hazırlanması aşamasında seçim işlemi tamamlandıktan sonra son aşama olan dönüştürme işlemine geçilmiştir. Bu aşamada veri madenciliği sürecinde 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. Öğretim Teknolojileri ve Materyal Tasarımı dersine ait öğrenci hareketliliği verileri veri madenciliği sürecinde kullanılabilecek uygun formlarda olduğundan dönüştürme işlemine tabi tutulmamıştır.

3.4.3. Modelin Kurulması ve Değerlendirilmesi

Veri madenciliği modelleri tahmin edici ve tanımlayıcı modeller olmak üzere ikiye ayrılmaktadır. Bu araştırmada her iki modelden de yararlanılmıştır. 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 nesneler bir grubu oluştururken bununla bağlantılı olmayanlar diğer soyut ya da somut nesneler başka bir grubu oluşturmaktadır. Kümeleme işlemi yaparken birkaç özelliğe bakarak gruplama yapmak gayet basitken, gruplamada kullanılacak nitelik sayısının artmasıyla birlikte bu işlem zorlaşmaktadır. Bu durumda veri madenciliği yöntemlerinden tanımlayıcı model olan kümeleme yöntemi devreye girmektedir. 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 algoritmasından yararlanılmıştır. 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 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. Kümeleme işlemine ait grupların oluşturulmasına bulgular kısmında detaylı olarak yer verilmiştir.

Veri madenciliği sürecinde 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 (Veri Toplama, Değer Bıçme, Birleştirme ve Temizleme, Seçim, Dönüştürme) tabi tutulan değişkenler kullanılmıştır. Modellerin oluşturulması sırasında bu girdi değişkenleri veri madenciliği algoritmalarından yararlanılarak analiz edilmiştir.

Veri madenciliği sürecinde karar ağaçlarının oluşturulmasında birçok algoritmadan yararlanılmaktadır. Bunlar içerisinde ID3, C4.5, C5.0, CART, CHAID ve QUEST algoritmaları en sık kullanılanlarıdır. Bu nedenle bu çalışmada veri madenciliği algoritmalarından en sık kullanılan C5.0, CART, CHAID, QUEST algoritmaları kullanılmıştır. ID3 ve C4.5 algoritmalarının dahil edilmemesinin nedeni ise C5.0 algoritmasının bu algoritmaların geliştirilmiş hali olmasından dolayıdır. Birden fazla algoritma kullanılarak her bir algoritmanın başarı tahminindeki doğruluk oranları hesaplanmıştır.

Modellerin kurulması, doğruluk ve geçerliliğin test edilmesinde Jackknife, Bootstrap ve Çapraz Geçerlilik olmak üzere üç farklı yöntem kullanılmaktadır. Bu

yöntemler korelasyonlu yapıyı bozup, yanlışlıkları ortadan kaldırabilmek için matematiksel model kurulumu ve kurulan modelin geçerliliğinin test edilmesinde sıklıkla kullanılmaktadır (Temel ve diğerleri, 2012). Bu çalışmada öğrenci başarı düzeyinin tahmin edilmesi için kurulan modellerin doğruluğu ve geçerliliğinin tespit edilmesinde modeli oluşturan tüm verilerin kullanılmasına olanak sağlayan çapraz geçerlilik yöntemlerinden olan çift çapraz geçerlilik yöntemi kullanılmıştır. Çift çapraz geçerlilik yöntemi veri miktarının sınırlı olması durumunda tercih edilen bir yöntemdir. Modelin doğruluğunun test edildiği çift çapraz geçerlilikte veri yığını ilk olarak test ve eğitim verisi olmak üzere tesadüfi olarak iki eşit parçaya ayrılır. Daha sonra model oluşturularak veriler üzerinde basit geçerlilik yöntemindeki gibi hata ve doğruluk oranları hesaplanır. Hata ve doğruluk oranları hesaplandıktan sonra test ve eğitim verileri yer değiştirilerek aynı işlem tekrarlanır. Elde edilen iki doğruluk değerinin ortalaması alınarak modelin doğruluk oranı hesaplanır.

3.4.4. Modelin Kullanılması

Veri madenciliği sürecinde modelin kurulması ve değerlendirilmesi aşamasında probleme uygun olarak kurulan ve geçerliliği test edilip kabul edilen model oluşturulduktan sonra modelin kullanılması aşamasına geçilmiştir. 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. Bu veriler ışığında öğrencilerin ÖYS üzerindeki hareketliliğinin akademik başarıya olan etki düzeyi hakkında bilgiler elde edilmiştir. Öğrenci başarı tahmininde kullanılan veri madenciliği tahmin modeli Ek-2’de ve bu modelde yer alan düğümler ve işlevleri Ek-3’te yer almaktadır.

3.4.5. Modelin İzlenmesi

Veri madenciliği sürecinde 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.

3.5. Araştırma Süreci

Araştırmacı tarafından ilk olarak yapılan çalışmanın amaca hizmet edebilmesi için veri madenciliği sürecinin ilk adımı olan problemin tanımlanması işlemi gerçekleştirilmiştir. Problemin doğru şekilde tanımlanması veri madenciliğiyle elde edilecek başarıyı arttıracığından, zaman ve maliyet kaybına engel olacağından bu aşama üzerinde titizlikle durulmuştur.

Problemin tanımlanmasından sonra bir sonraki adım olan veri madenciliği uygulanırken oluşturulacak modelin veri kaynaklarının neler olduğunun belirlendiği ve modelde kullanılmak için verilerin uygun hale getirildiği verilerin hazırlanması işlemi gerçekleştirilmiştir. Bu doğrultuda araştırmada Moodle ÖYS log kayıtları ve derse kayıtlı olan öğrencilerin yılsonu akademik başarı notları veri kaynağı olarak belirlenmiştir. Veri kaynakları belirlendikten sonra verilerin hazırlanması işleminin ilk basamağı olan veri toplama işlemi gerçekleştirilmiştir. Araştırmacı tarafından Moodle ÖYS'yi 2007 yılından beri tüm derslerin etkin olarak kullanan Pedagojik Formasyon Sertifika Programı kapsamında Öğretim Teknolojileri ve Materyal Tasarımı dersini veren ilgili öğretim elemanından derse ait Moodle log kayıtları talep edilmiştir. Öğretim elemanı tarafından kullanıcı adı ve şifre ile sisteme giriş yapılarak 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersine ait Moodle log kayıtlarına ulaşılmıştır. Moodle log kayıtları CSV dosya türünde indirildikten sonra araştırmacıya teslim edilmiştir. Ayrıca veri kaynağı olarak belirlenen öğrenci yılsonu akademik başarı notları da dersi veren öğretim elemanı tarafından araştırmacıya teslim edilmiştir. Veri toplama işlemi gerçekleştirildikten sonra elde edilen veri yığınının verilerin hazırlanması işleminin diğer basamakları olan değer biçme, birleştirme ve temizleme, seçim, dönüştürme işlemleri uygulanmıştır.

Verilerin hazırlanması işlemi tamamlandıktan sonra bir sonraki adım olan modelin kurulması ve değerlendirilmesi işlemi gerçekleştirilmiştir. Bu aşamada probleme uygun bizi sonuca götürebilecek en iyi modelin bulunabilmesi için çok sayıda model kurulup denenerek doğruluk oranı en yüksek modele ulaşılmıştır.

Belirlendiğimiz probleme uygun olarak kurulan ve test edilen model problemin çözümünde kullanılmıştır.

Arařtırmacı tarafından veri madencilięi sreci boyunca yapılan tm iřlemler SPSS Clementine 12.0 programı kullanılarak gerekleřtirilmiř ve gerek duyulan tm bulgulara ulařılmıřtır.

4. BULGULAR

Bu bölümde; yapılan uygulama sonucunda elde edilen bulgular araştırmanın genel amacı ve bu genel amaç doğrultusunda belirlenen alt amaçlara göre ele alınmıştır. Araştırmanın genel amacı; veri madenciliği tekniklerini kullanarak üniversite (lisans) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile yılsonu akademik başarıları arasındaki ilişkiyi belirlemektir. Bu doğrultuda belirlenen alt amaçlara ilişkin bulgular aşağıdaki gibidir.

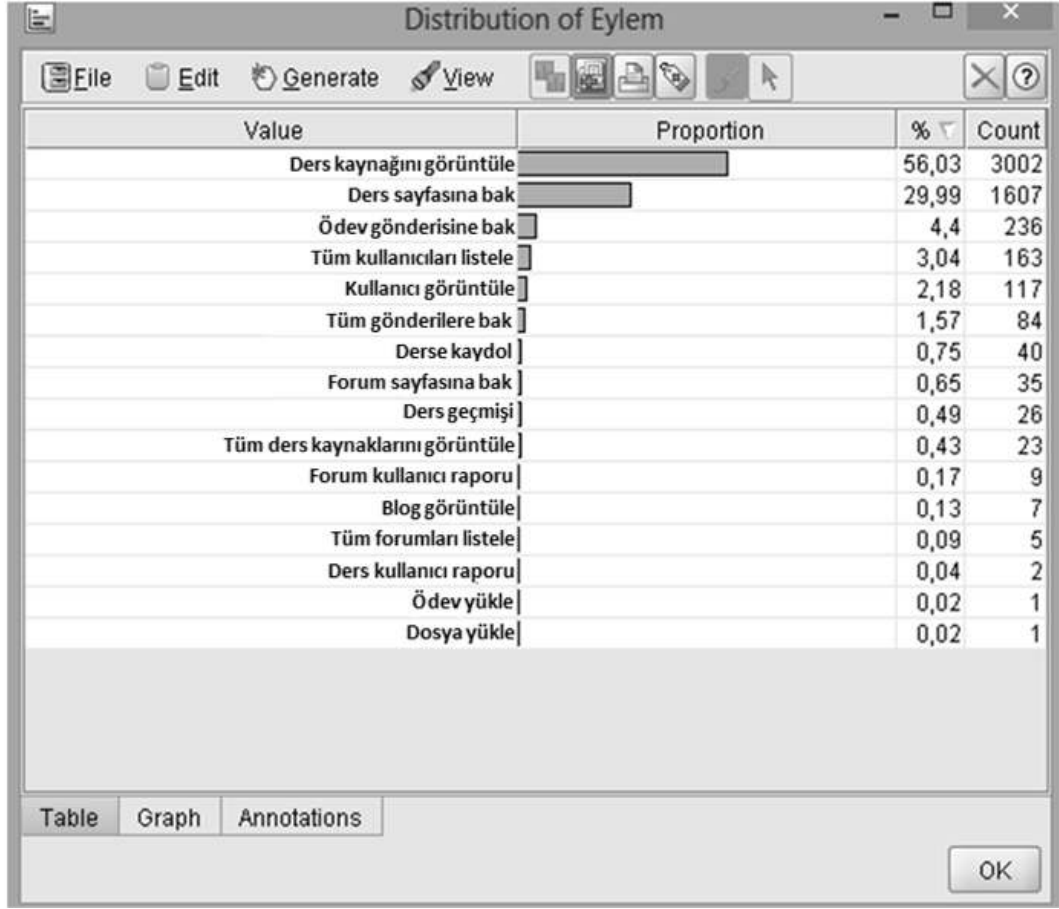
4.1. Birinci Araştırma Sorusuna İlişkin Bulgular

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

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

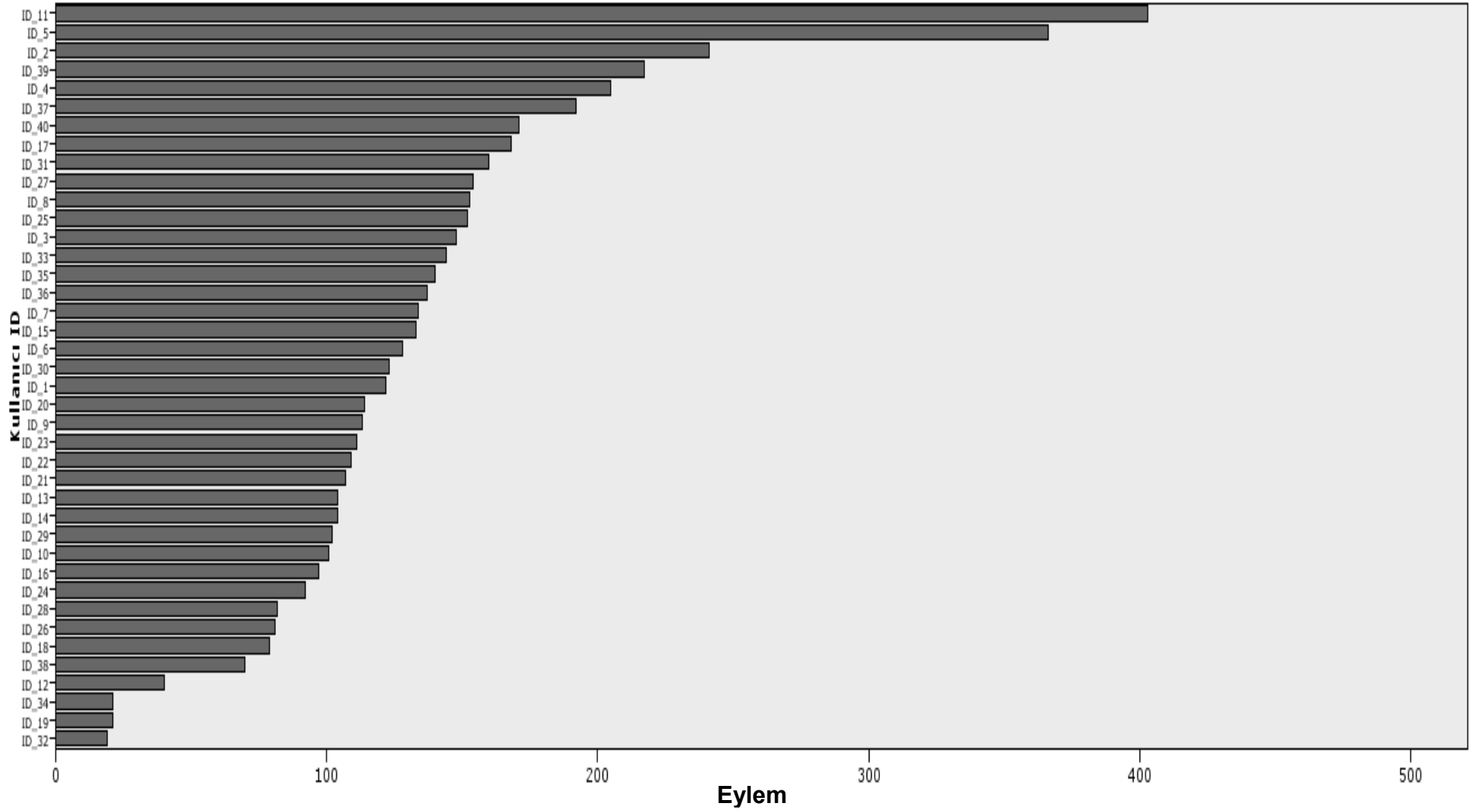
4.1.1. Öğrenciler Tarafından Gerçekleştirilen Moodle Eylemleri ve Dağılım Oranları

Moodle üzerinde 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 4.1’de 16 adet farklı eylemin toplam gerçekleştirilen eylemler içerisindeki dağılımları ayrıntılı olarak görülmektedir.



Şekil 4.1. Eylemlerin Dağılım Oranları

Moodle üzerinde gerçekleştirilen 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. Şekil 4.2’de öğrenci bazında gerçekleştirilen toplam eylem sayılarının büyükten küçüğe doğru dağılımı görülmektedir.



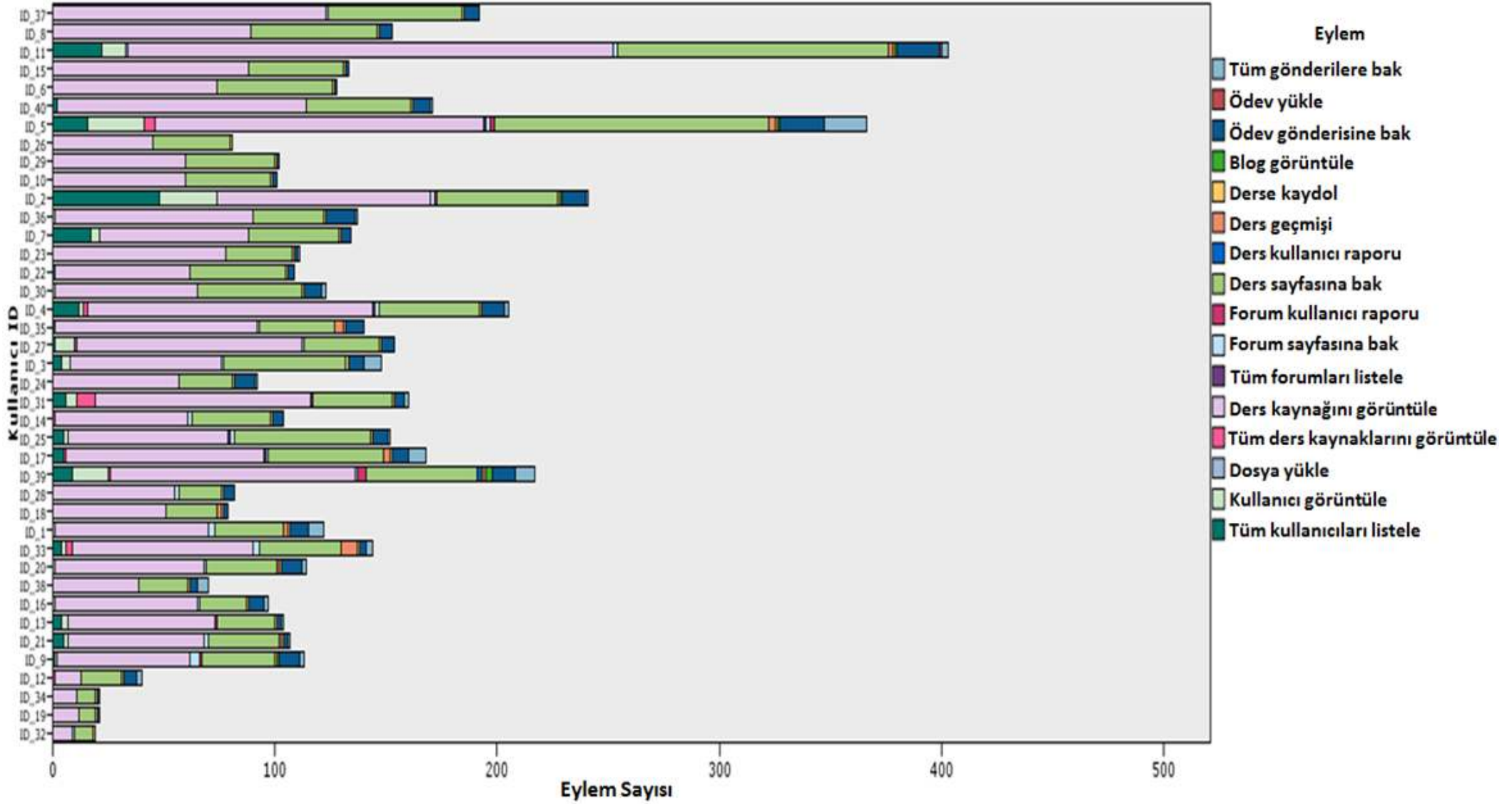
Şekil 4.2. Öğrenci Eylem Sayısı

Öğrenci bazında gerçekleştirilen toplam eylem sayılarının büyükten küçüğe doğru sıralanışı Tablo 4.1’de ayrıntılı olarak yer almaktadır.

Tablo 4.1. 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

Moodle üzerinde gerçekleştirilen 5358 adet eylem içerisindeki 16 adet farklı eylemin öğrenci bazında dağılımları tespit edilmiştir. Öğrenci eylem dağılımları Şekil 4.3’de görülmektedir.



Şekil 4.3. Öğrenci Eylem Dağılımı

Şekil 4.3 incelendiğinde; Öğrencilerin ÖYS’de yer alan 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 ÖYS üzerinde 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.

Moodle üzerinde gerçekleştirilen 16 adet farklı eylemden hangilerinin öğrenci bazında kaç öğrenci tarafından gerçekleştirilip gerçekleştirilmediği Tablo 4.2’de yer almaktadır.

Tablo 4.2. Öğrenci Bazında Eylemler

Eylem	N
Derse kaydol	40
Ders sayfasına bak	40
Ders kaynağını görüntüle	40
Ödev gönderisine bak	38
Tüm gönderilere bak	23
Forum sayfasına bak	19
Kullanıcı görüntüle	19
Tüm kullanıcıları listele	18
Ders geçmişi	10
Tüm ders kaynaklarını görüntüle	9
Blog görüntüle	5
Forum kullanıcı raporu	5
Tüm forumları listele	5
Ders kullanıcı raporu	1
Ödev yükle	1
Dosya yükle	1

Tablo 4.2 incelendiğinde; ders kaynağını görüntüle, ders sayfasına bak, derse kaydol eylemlerinin tüm öğrenciler tarafından, ödev gönderisine bak eyleminin 38 öğrenci tarafından, tüm gönderilere bak eyleminin 23 öğrenci tarafından, kullanıcı görüntüle eyleminin 19 öğrenci tarafından, forum sayfasına bak eyleminin 19 öğrenci tarafından, tüm kullanıcıları listele eyleminin 18 öğrenci tarafından, ders geçmişi eyleminin 10 öğrenci tarafından, tüm ders kaynaklarını görüntüle eyleminin 9 öğrenci tarafından, tüm forumları listele eyleminin 5 öğrenci tarafından, forum kullanıcı raporu eyleminin 5 öğrenci tarafından, blog görüntüle eyleminin 5 öğrenci tarafından, dosya yükle eyleminin 1 öğrenci tarafından, ders kullanıcı raporu eyleminin 1 öğrenci tarafından, ödev yükle eyleminin 1 öğrenci tarafından gerçekleştirildiği görülmektedir.

4.2. 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?

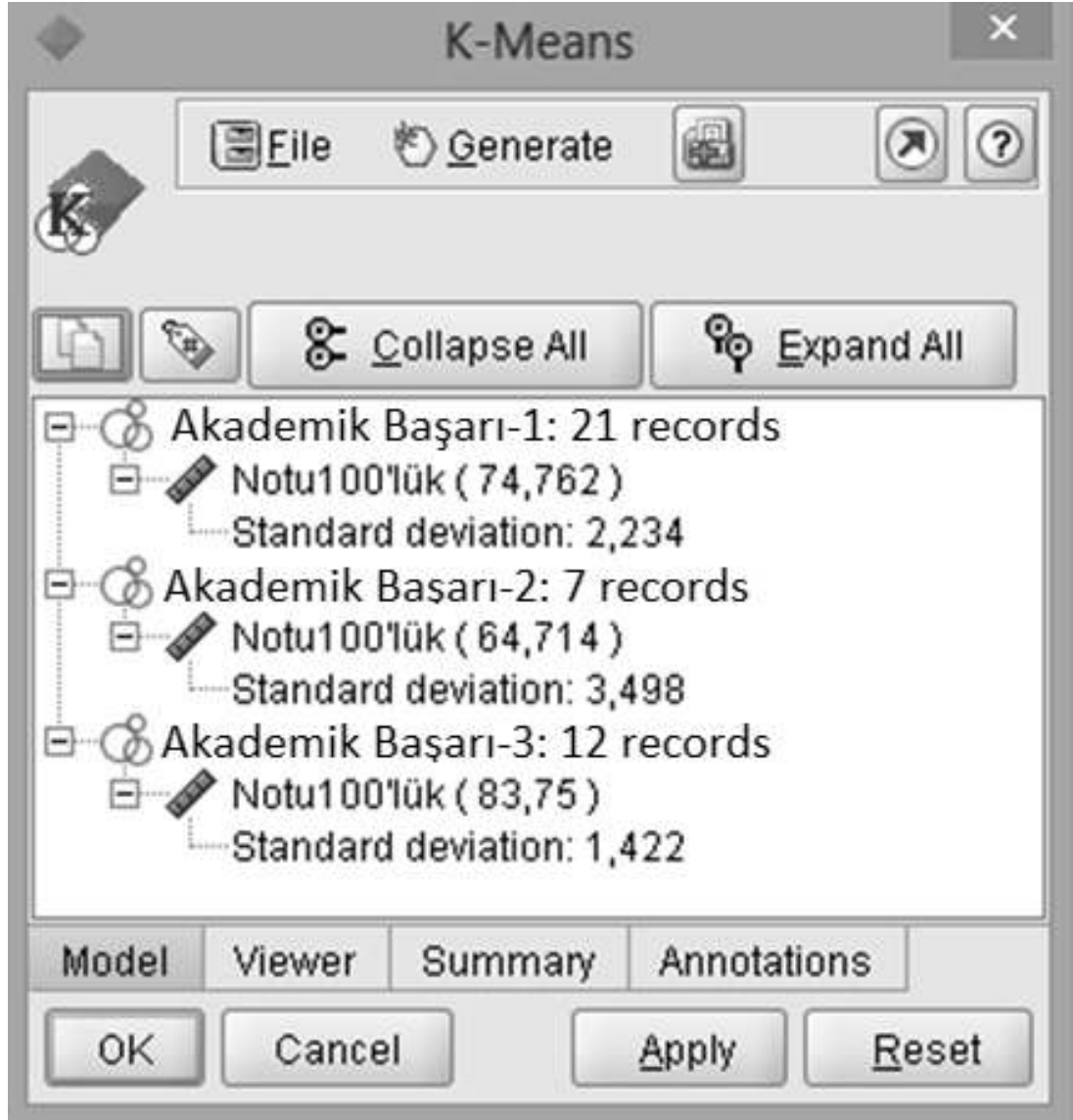
İkinci araştırma sorusu kapsamında öğrencilerin Moodle ÖYS üzerinde gerçekleştirdikleri eylemler disiplinler arası bir alan olan veri madenciliği teknikleri kullanılarak incelenmiştir. Tanımlayıcı ve tahmin edici veri madenciliği modelleri kullanılarak öğrencilerin akademik performanslarına ilişkin bulgulara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

4.2.1. Tanımlayıcı Modellerinin Kullanılması

Veri madenciliği modellerinden tanımlayıcı model olan kümeleme modeli kullanılmış ve bu modelin oluşturulmasında K-Means algoritmasından yararlanılmıştır. Kümeleme modelinde veri kaynağı olarak 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersini alan öğrencilerin yılsonu akademik başarı notları kullanılmıştır.

4.2.1.1. Kümeleme Modelinin Oluşturulması

Öğrenci akademik başarı gruplarının oluşturulmasında kümeleme algoritmalarından olan K-Means algoritmasından yararlanılmıştır. Öğrenci yılsonu akademik başarı notları K-Means algoritmasına eklenerek elde edilebilecek küme sayısının anlaşılabilir ölçüde olması ve sonuçların anlaşılabilirliğinin artırılabilmesi için akademik başarı grupları geçme notunun altında kalma ihtimali olanlar, orta seviyede başarı gösterenler ve üst seviyede başarı gösterenler olmak üzere 3 gruba ayrılmıştır. 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 4.4'te K-Means algoritması tarafından oluşturulan akademik başarı grupları gösterilmektedir.



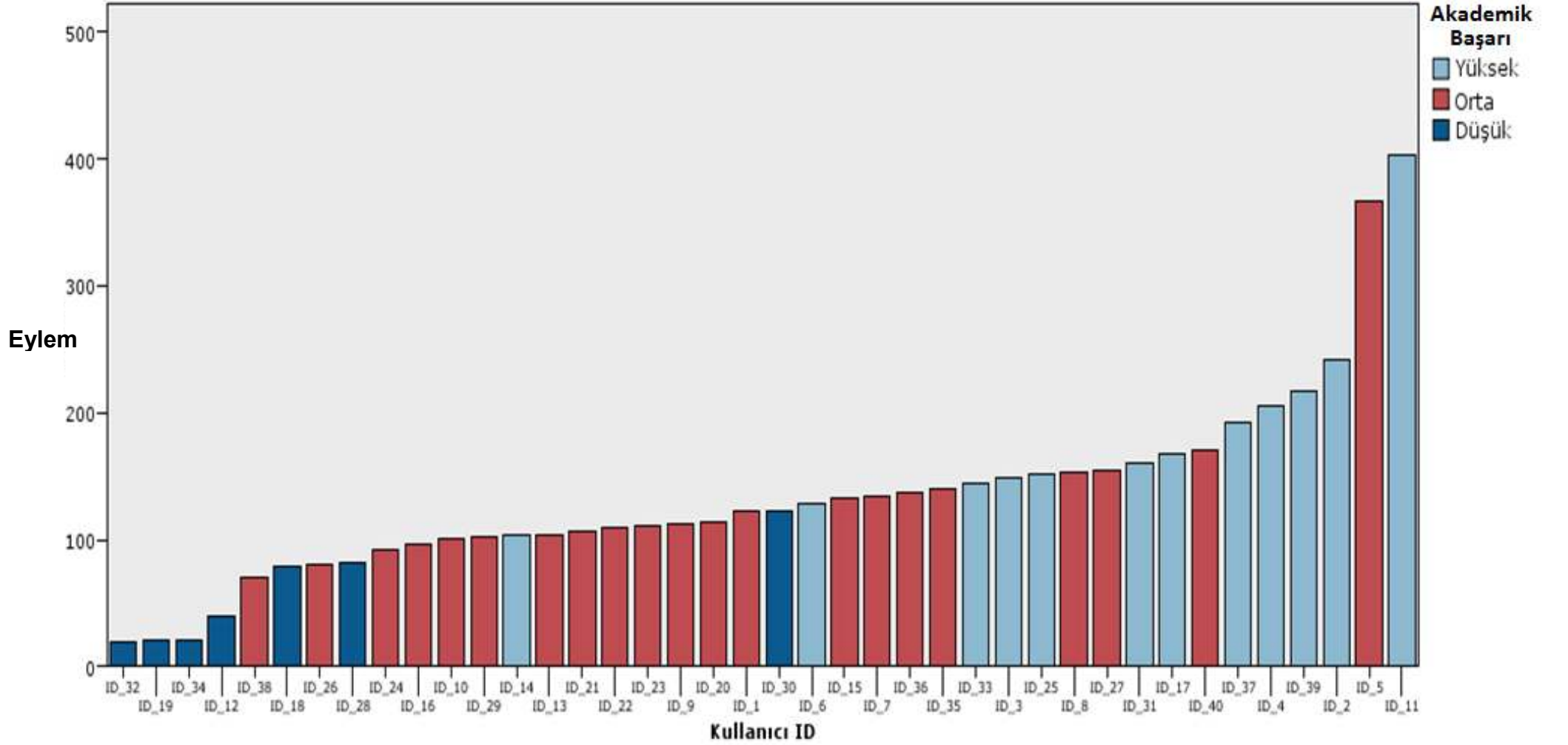
Şekil 4.4. Akademik Başarı Grupları

Kümeleme yöntemi kullanılarak oluşturulan akademik başarı gruplarında 12 öğrencinin Yüksek Akademik Başarı, 21 öğrencinin Orta Akademik Başarı, 7 öğrencinin Düşük Akademik Başarı grubunda yer aldığı ve bu öğrencilerin hangi akademik başarı gruplarında bulundukları Tablo 4.3'te görülmektedir.

Tablo 4.3. Kullanıcı Akademik Başarı Grupları

Akademik Başarı Düzeyi		
Yüksek	Orta	Düşük
ID_2	ID_1	ID_12
ID_3	ID_5	ID_18
ID_4	ID_7	ID_19
ID_6	ID_8	ID_28
ID_11	ID_9	ID_30
ID_14	ID_10	ID_32
ID_17	ID_13	ID_34
ID_25	ID_15	
ID_31	ID_16	
ID_33	ID_20	
ID_37	ID_21	
ID_39	ID_22	
	ID_23	
	ID_24	
	ID_26	
	ID_27	
	ID_29	
	ID_35	
	ID_36	
	ID_38	
	ID_40	

Yüksek Akademik Başarı, Orta Akademik Başarı ve Düşük Akademik Başarı grubunda yer alan öğrencilerin yaptıkları eylem sayılarına göre dağılımı Şekil 4.5'te görülmektedir.



Şekil 4.5. 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 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 4.4'te eylemler ve akademik başarı grupları tarafından gerçekleştirilme durumları görülmektedir.

Tablo 4.4. Eylemler ve Akademik Başarı Grupları Tarafından Gerçekleştirilme Sayıları

Eylemler	Akademik Başarı						Toplam	
	Yüksek		Orta		Düşük			
	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
Toplam	2262	42,51	2711	50,3	385	7,19	5358	100

4.2.2. Tahmin Edici Modellerinin Kullanılması

Tahmin modellerinin oluşturulmasında sınıflandırma algoritmalarından olan karar ağaçları oluşturmaya yarayan C5.0, CART, CHAID, QUEST algoritmaları kullanılmıştır. Kullanılan algoritmalar ile oluşturulan modellerin doğruluk oranları incelenmiş ve bu algoritmalar içinden en yüksek doğruluk oranına sahip olan CART algoritması kullanılarak karar ağacı oluşturma işlemi gerçekleştirilmiştir.

4.2.2.1. Kullanılan Algoritmalar ve Doğruluk Oranları

Öğrencilerin ÖYS üzerindeki hareketliliği ile akademik başarıları arasındaki ilişkinin belirlenmesinde kullanılan C5.0, CART, CHAID, QUEST algoritmaların doğruluk oranları Tablo 4.5'te yer almaktadır.

Tablo 4.5. 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

CART algoritması hem sayısal hem de nominal değerler üzerinde çalışabilen, her bir düğümü iki dala ayıran ve kendinden bir önceki kümeden daha homojen iki alt küme elde etmeye çalışan bir sınıflandırma algoritmasıdır. CART algoritması için hata ve doğruluk oranları hesaplandığında doğruluk oranının %85, hata oranının ise %15 olduğu görülmektedir. CART algoritması kullanılarak oluşturulan modelin hata ve doğruluk oranı Ek-4'te yer almaktadır.

C5.0 algoritması ID3 algoritmasının temel özelliklerini kullanmakla beraber C4.5 algoritmasının doğruluk oranını artırmak için geliştirilmiş bir sınıflandırma algoritmasıdır. C5.0 algoritması için hata ve doğruluk oranları hesaplandığında doğruluk oranının %82,5, hata oranının ise %17,5 olduğu görülmektedir. C5.0 algoritması kullanılarak oluşturulan modelin hata ve doğruluk oranı Ek-5'te yer almaktadır.

CHAID algoritması en iyi bölmeyi hesaplamak için ki-kare istatistiğini kullanarak karar ağaçları oluşturulmasını sağlayan bir sınıflandırma algoritmasıdır.

CHAID algoritması için hata ve doğruluk oranları hesaplandığında doğruluk oranının %65, hata oranının ise %35 olduğu görülmektedir. CHAID algoritması kullanılarak oluşturulan modelin hata ve doğruluk oranı Ek-6'da yer almaktadır.

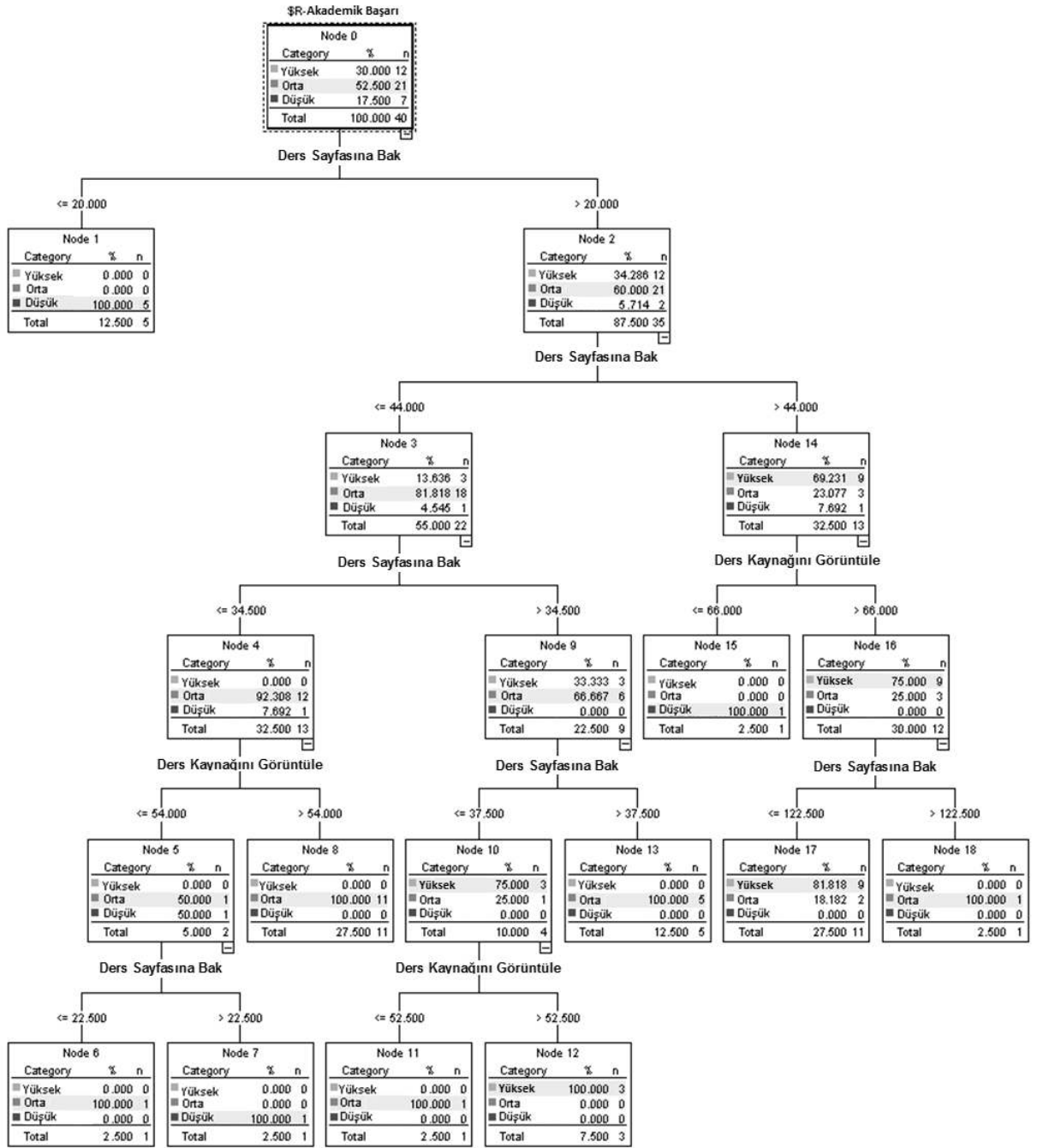
QUEST algoritması hızlı, yansız, verimli istatistiksel ağaç anlamına gelen, CART algoritmasına benzer olarak ikili karar ağacı oluşturan ve veri yığınının çok büyük olduğu durumlarda hızlı sonuçlar almak için kullanılan bir sınıflandırma algoritmasıdır. QUEST algoritması için hata ve doğruluk oranları hesaplandığında doğruluk oranının %65, hata oranının ise %35 olduğu görülmektedir. QUEST algoritması kullanılarak oluşturulan modelin hata ve doğruluk oranı Ek-7'de yer almaktadır.

4.2.2.2. CART Algoritması Kullanılarak Karar Ağaçlarının Oluşturulması

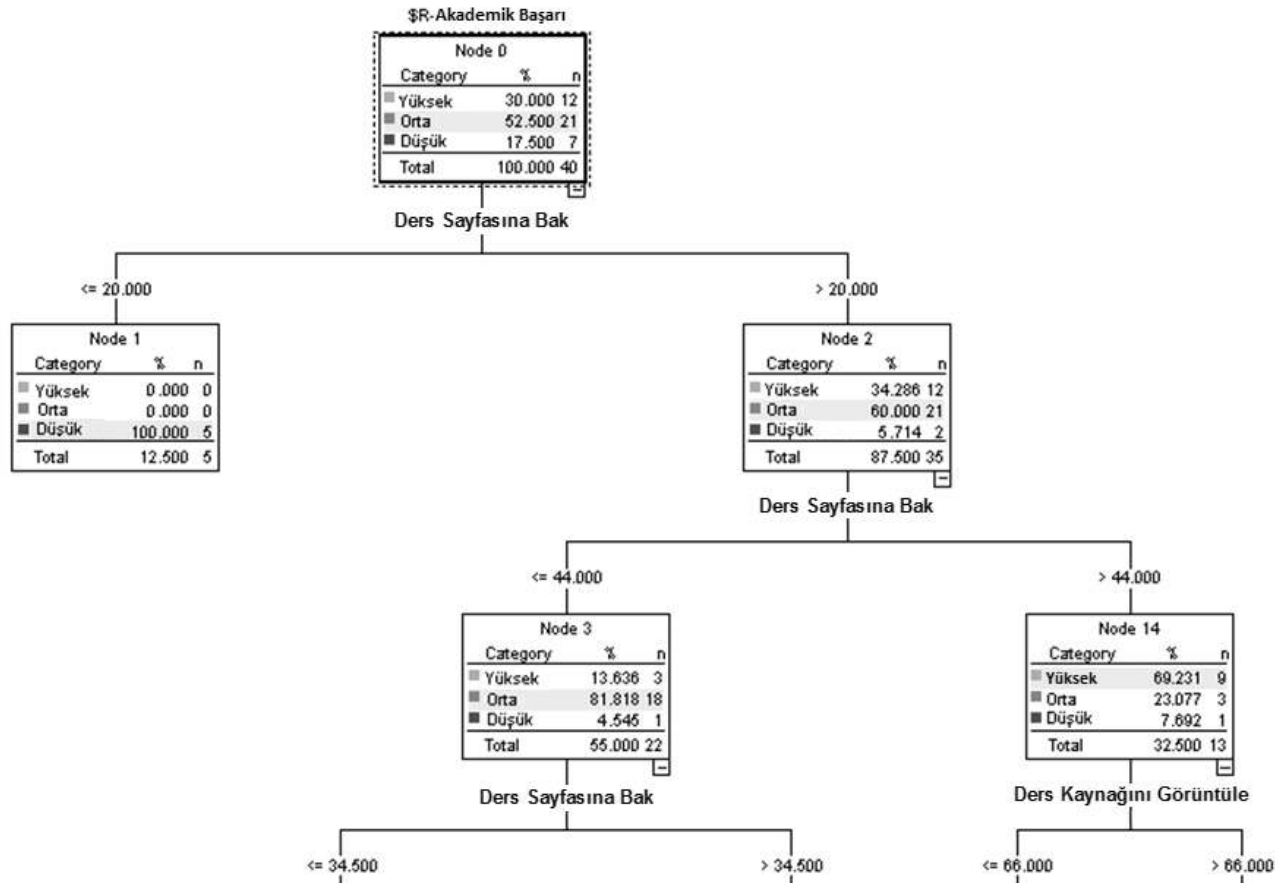
Sınıflandırma algoritmalarıyla kurulan modellerin hata ve doğruluk oranları göz önüne alındığında %85 doğruluk oranı ile en yüksek doğruluk oranına CART algoritması kullanılarak ulaşılmıştır. Bu nedenle karar ağaçlarının oluşturulmasında ve başarı tahmininin gerçekleştirilmesinde CART algoritmasından 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. Ders sayfasına bak eylemi; öğrencilerin kayıtlı olduğu dersler içerisinde seçtiği derse ait ders içeriğinin (Kurs Planı, Konular, Dokümanlar vb.) listelendiği ilk giriş sayfasını görüntülemesi işlemidir. Ders kaynağını görüntüle eylemi; öğrencilerin derse ait ders içeriğini, ekli dosyaları, derse içeriğiyle ilgili paylaşılan linkleri, ders dokümanlarını, ders sunumları ve ödev açıklamalarını görüntülenmesi işlemidir. Bu girdi değişkenleri kullanılarak oluşturulan CART algoritması sonucu oluşan karar ağacının derinliği 6 olarak bulunmuştur. Karar ağaçlarında ilk dallanma sınıfların oluşmasında en etkili değişkeni ifade etmektedir. Şekil 4.6 incelendiğinde ilk dallanmanın ders sayfasına bak eylemine göre gerçekleştiği görülmektedir. Ders sayfasına bak eylemine göre birinci (Nod 1), ikinci (Node 2), üçüncü (Node 3), dördüncü (Node 4), altıncı (Node 6), yedinci (Node 7), dokuzuncu

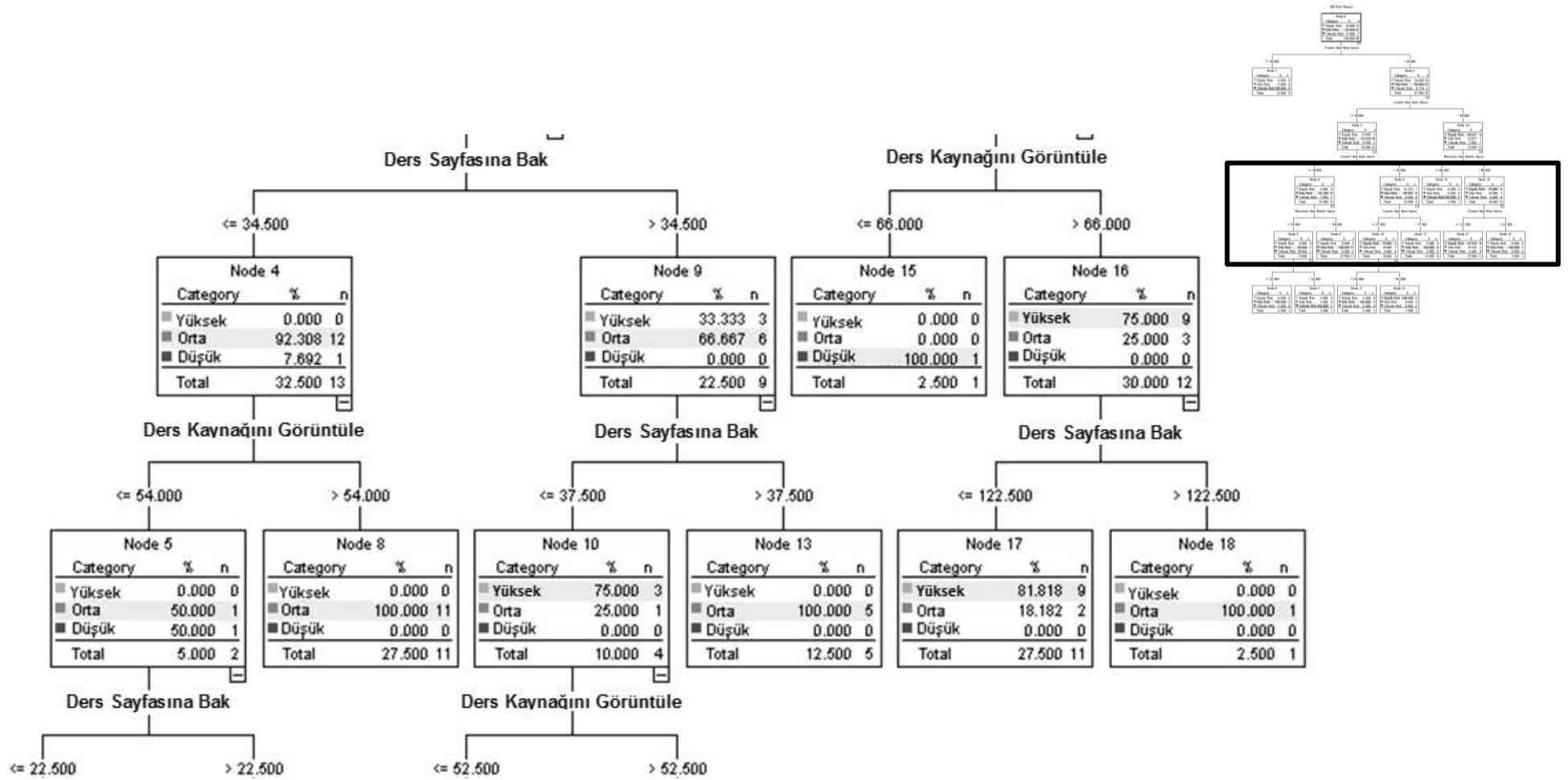
(Node 9), onuncu (Node 10), on üçüncü (Node 13), on dördüncü (Node 14), on yedinci (Node 17) ve on sekizinci (Node 18) düğümler oluşturulmuştur. Ders kaynağını görüntüle eylemine göre ise beşinci (Node 5), sekizinci (Node 8), on birinci (Node 11), on ikinci (Node 12), on beşinci (Node 15) ve on altıncı (Node 16) düğümler oluşturulmuştur. Sınıflandırma işleminin gerçekleştirilmesi akademik başarı gruplarında yer alan kişilerin eylemleri gerçekleştirme oranlarına göre algoritma tarafından hesaplanarak oluşturulmuştur. İlk düğüm (Node 1) incelendiğinde birinci düğümde (Node 1) düşük akademik başarı grubunda yer alan 5 kişi, ikinci düğümde (Node 2) yüksek akademik başarı grubundan 12, orta akademik başarı grubundan 21, düşük akademik başarı grubundan 2 kişinin yer aldığı görülmektedir. Gerçekte başarı notu düşük ve düşük akademik başarı grubun yer alan 2 öğrencinin birinci düğümde (Node 1) yer alması gerekirken bu öğrencilerin eylem sayılarının yüksek akademik başarı grubu veya orta akademik başarı grubunda yer alan öğrencilere yakın ya da fazla olması nedeniyle ikinci düğümde (Node 2) yer aldığı görülmektedir. Oluşturulan diğer düğümler de öğrencilerin gerçekleştirdikleri ders sayfasına bak ve ders kaynağını görüntüle eylem sayıları göz önüne alınarak sınıflamaya tabi tutulmuş ve dallanma işlemi gerçekleşmiştir. CART algoritmasının karar ağacı çıktısı Şekil 4.6'da görülmektedir. Ayrıca Şekil 4.6'da yer alan karar ağacının dallarında ve yapraklarında yer alan verilerin daha iyi görülebilmesi için Şekil 4.6 tekrardan 3 parça halinde gösterilmiştir.



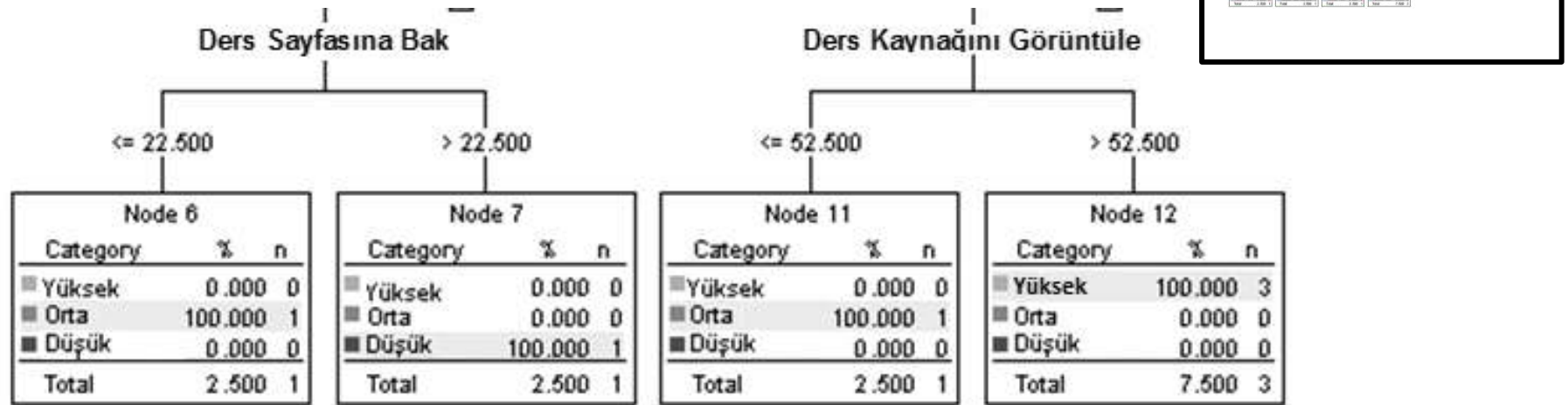
Şekil 4.6. CART Algoritması İle Oluşturulan Karar Ağacı



Şekil 4.6. CART Algoritması İle Oluşturulan Karar Ağacı - 1

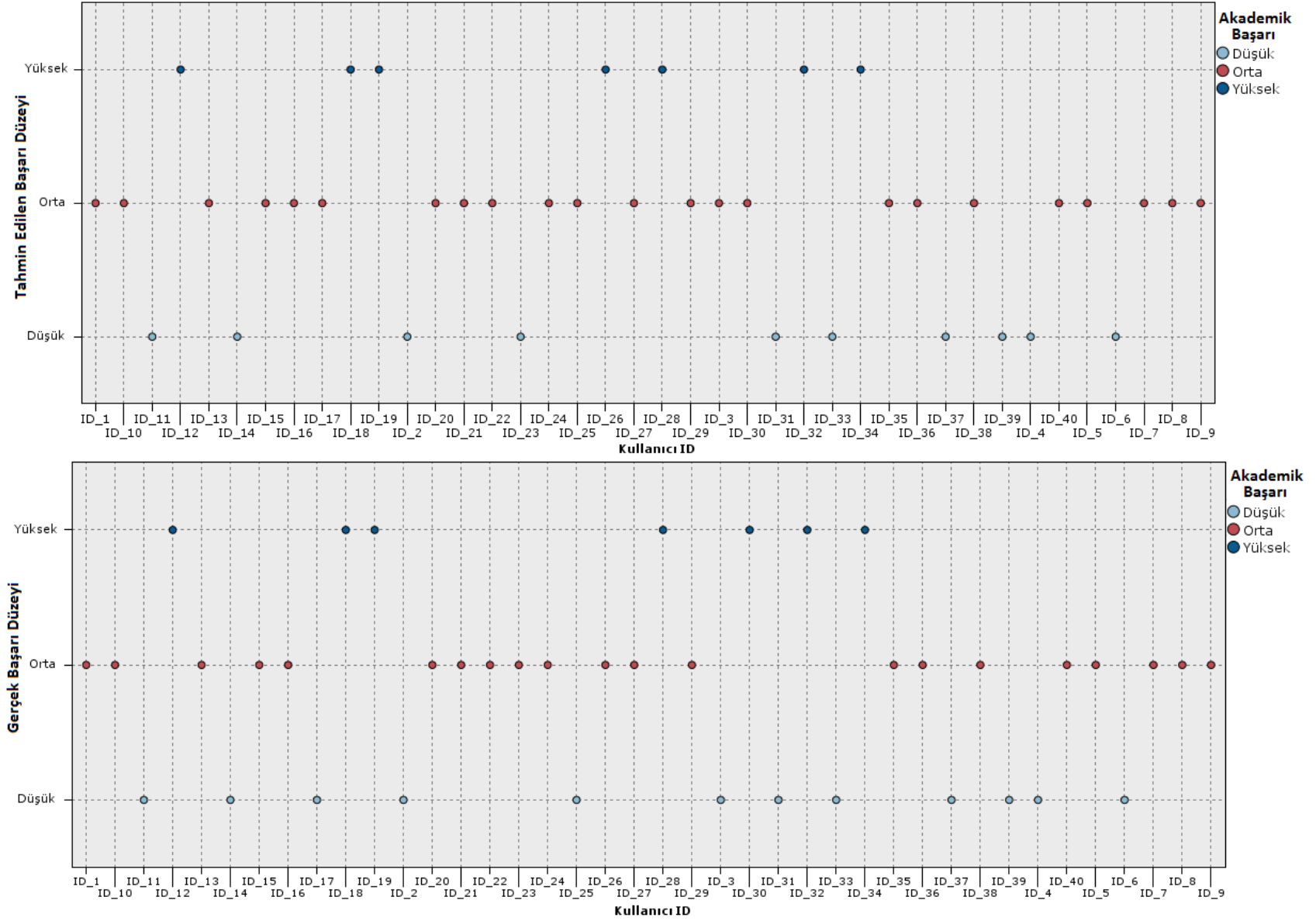


Şekil 4.6. CART Algoritması İle Oluşturulan Karar Ağacı - 2



Şekil 4.6. CART Algoritması İle Oluşturulan Karar Ağacı - 3

Veri yığınınına CART algoritması uygulandıktan sonra öğrencilerin gerçekleştirdikleri eylem ilişkilerine göre akademik başarı düzeyleri belirlenmiştir. Şekil 4.7’de öğrencilerin tahmin edilen ve gerçek akademik başarı düzeyleri ayrıntılı olarak görülmektedir.



Şekil 4.7. Tahmin Edilen ve Gerçek Akademik Başarı Düzeyleri

Şekil 4.7 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.

5. 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 ögelere (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 Öğretim Yönetim Sistemleri (ÖYS) gerçekleştirilen tüm öğrenci hareketlerini kayıt altına aldıklarından eğitsel veri madenciliği çalışmalarına önemli bir veri kaynağı sağlamaktadır.

Bu çalışmada veri madenciliği teknikleri kullanılarak üniversite (lisans) düzeyindeki öğrencilerin ÖYS üzerindeki hareketliliği ile yılsonu akademik başarıları arasındaki ilişkinin belirlenmesi amaçlanmıştır. Bu amaç doğrultusunda aşağıdaki araştırma sorularına 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?

Çalışma kapsamında Öğretim Teknolojileri ve Materyal Tasarımı dersini alan öğrencilerin Moodle 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.

5.1. Birinci Araştırma Sorusuna İlişkin Sonuçlar

Araştırmanın birinci sorusu, öğrencilerin ÖYS üzerindeki hareketliliğinin incelenmesi ve bu hareketlilik sonucu öğrenciler tarafından gerçekleştirilen eylemlerin nasıl bir dağılım gösterdiği üzerinedir. Öğrencilerin ÖYS üzerinde gerçekleştirdikleri eylemlerin dağılımının belirlenmesinde geleneksel istatistiksel yöntemlerden yararlanılmıştır.

Öğrencilerin ÖYS üzerinde gerçekleştirdikleri eylemler incelendiğinde 16 adet farklı eylemin (Tüm gönderilere bak, Ödev yükle, Ödev gönderisine bak, Blog görüntüle, Derse kaydol, Ders geçmişi, Ders kullanıcı raporu, Ders sayfasına bak, Forum kullanıcı raporu, Forum sayfasına bak, Tüm forumları listele, Ders kaynağını görüntüle, Tüm ders kaynaklarını görüntüle, Dosya yükle, Kullanıcı görüntüle, Tüm kullanıcıları listele) gerçekleştirildiği tespit edilmiştir.

ÖYS üzerinde gerçekleştirilen eylemler içerisinde en çok gerçekleştirilen eylemlerin ders kaynağını görüntüle ve ders sayfasına bak eylemleri olduğu görülmüştür. Bu eylemler derse kayıtlı tüm öğrenciler tarafından farklı oranlarda gerçekleştirilmiştir.

5.2. İkinci Araştırma Sorusuna İlişkin Sonuçlar

Araştırmanın ikinci sorusu, öğrencilerin ÖYS üzerinde gerçekleştirdikleri eylemler analiz edildikten sonra bu eylemlerin gerçekleştirilme oranlarına bağlı olarak öğrencilerin akademik başarılarının nasıl ve ne kadar doğruluk oranıyla tahmin edilebileceği üzerinedir. 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. Bu modelin oluşturulmasında kullanılan veri kaynağı 2012-2013 eğitim öğretim yılı bahar dönemi Öğretim Teknolojileri ve Materyal Tasarımı dersini alan öğrencilerin yılsonu akademik başarı notlarıdır. 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 sayılabileceği 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 kümelenmiş, 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 bir öneme sahip işlemdir (Osmanbegović & Suljić, 2012; Romero ve diğerleri, 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 ve diğerleri, 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 Moodle 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 Moodle forum kullanım verilerine ek olarak öğrencilerin Moodle 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 4.5'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.

5.3. 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 (Larry Johnson ve diğerleri, 2013). Yapılan çalışma sonucunda elde edilen bilgiler ışığında ileriki araştırmalara yön verebilecek araştırmaya ve uygulamaya dönük öneriler aşağıda sıralanmıştır.

5.3.1. Araştırmaya Dönük Öneriler

Çalışma kapsamında Moodle üzerinde öğrenciler tarafından gerçekleştirilen 16 adet farklı eylemden tüm öğrenciler tarafından gerçekleştirilen ve en çok gerçekleştirilme oranına sahip olan 2 eylem girdi değişkeni olarak belirlenmiştir. Yapılacak diğer çalışmalarda farklı eylemler farklı araştırma soruları kullanılarak incelenebilir.

Moodle tarafından tutulan öğrenci kayıtları 8 farklı alandan (ders adı, tarih, saat, IP adresi, kullanıcı adı, eylem, eylem bağlantıları ve bilgi alanı) oluşmaktadır. Çalışma kapsamında öğrenciler ait eylemler kullanılarak akademik başarı düzeyleri arasındaki ilişki belirlenmeye çalışılmıştır. Yapılacak diğer çalışmalarda farklı alanlar ile öğrencilerin akademik başarı düzeyleri arasındaki ilişkiler incelenebilir.

Öğrencilerin Öğretim Teknolojiler ve Materyal Tasarımı dersine ilişkin Moodle üzerinde gerçekleştirdikleri eylemlere bağlı olarak akademik başarı göstergelerinin belirlenmesinde boylamsal bir çalışma gerçekleştirilebilir.

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 Sinir Ağları, Regresyon, Zaman Serisi Analizleri, Genetik Algoritmalar, Kestirim) kullanılarak öğrenci başarıları tahmin edilmeye çalışılabilir.

5.3.2. 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 geri 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 geri dönüt ve yönlendirme işlemleri sağlanabilir.

ÖYS üzerinde dönem içerisindeki öğrenci hareketliliği analiz edilerek öğrenci profilleri belirlenebilir, öğrencilere gerekli geri dönüt ve dönem bitmeden yönlendirme işlemleri sağlanabilir.

Farklı ÖYS sistemlerinden elde edilen veriler kullanılarak öğrenci hareketliliği analiz edilebilir ve öğrenci akademik başarı düzeyleri arasındaki ilişki incelenebilir.

KAYNAKLAR LİSTESİ

- Akçapınar, G. (2014). “Çevrimiçi Öğrenme Ortamındaki Etkileşim Verilerine Göre Öğrencilerin Akademik Performanslarının Veri Madenciliği Yaklaşımı ile Modellenmesi”, Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi. Eğitim Bilimleri Enstitüsü, Ankara.
- Ali, M.M. (2013). “Role of Data Mining in Education Sector”, International Journal of Computer Science and Computing, 2(4), 374-383.
- AlMazroui, Y.A. (2013). “A Survey of Data Mining in The Context of E-Learning”, International Journal of Information Technology & Computer Science, 7(3), 8-18.
- Altınparmak, M., Kurt, D.İ. ve Kapıdere, M. (2011). “E-Öğrenme ve Uzaktan Eğitimde Açık Kaynak Kodlu Öğrenme Yönetim Sistemleri”, Akademik Bilişim’11, İnönü Üniversitesi, Malatya, 343-351.
- 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”, Yüksek Lisans Tezi, Sakarya Üniversitesi Fen Bilimler Enstitüsü, Sakarya.
- Aydın, S. (2007). “Veri Madenciliği ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama”, Doktora Tezi, Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir.
- Aydın, C.Ç. ve Biroğul, S. (2008). “E- Öğrenmede Açık Kaynak Kodlu Öğretim Yönetim Sistemleri ve Moodle”, Bilişim Teknolojileri Dergisi, Cilt: 1, Sayı: 2.
- Aydın, D. (2011). “Malzeme Bilgisi Dersinin Moodle Açık Kaynak Kodlu Öğrenim Yönetim Sisteminde İşlenişinin Öğrenci Üzerindeki Etkinliğinin Değerlendirilmesi”, Yüksek Lisans Tezi, Bahçeşehir Üniversitesi, Fen Bilimleri Enstitüsü, Bilgi Teknolojileri Anabilim Dalı, İstanbul.
- Ayık, Y.Z., Özdemir, A. ve Yavuz, U. (2007). “Lise Türü ve Lise Mezuniyet Başarısının, Kazanılan Fakülte İle İlişkisinin Veri Madenciliği Tekniği ile Analizi”, Sosyal Bilimler Enstitüsü Dergisi, Cilt: 10, Sayı: 2, 441-454.
- Aynekin, G. (2006). “İnternet İçerik Madenciliğinde Yapay Sinir Ağları ve Bir Uygulama”, Yüksek Lisans Tezi, Uludağ Üniversitesi Fen Bilimler Enstitüsü, Bursa.

- Bahçeci, F. (2015). “Öğrenme Yönetim Sistemlerinde Kullanılan Öğrenme Analitikleri Araçlarının İncelenmesi”, Turkish Journal of Educational Studies, Cilt: 2, Sayı :1, 41-58.
- Baker, R., Growda, S.M. & Corbett, A.T. (2011). “*Automatically Detecting A Student’s Preparation For Future Learning: Help Use is Key*”, 4th International Conference on Educational Data Mining, New York.
- Barahate, S.R. (2012). “*Educational Data Mining As A Trend Of Data Mining In Educational System*”, International Conference & Workshop on Recent Trends in Technology, Proceeding Published in International Journal of Computer Applications.
- Beal, C. R., Qu, L., & Lee, H. (2006). “*Classifying Learner Engagement Through Integration of Multiple Data Sources*”, Paper Presented At The Proceedings Of The 21st National Conference on Artificial Intelligence - Volume 1, Boston, Massachusetts.
- 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.
- Bhise, R.B., Thorat, S.S. & Supekar, A.K. (2013). “*Importance of Data Mining in Higher Education System*”, IOSR Journal of Humanities And Social Science, 6(6), 18-21.
- Bırtıl, F.S. (2012). “*Kız Meslek Lisesi Öğrencilerinin Akademik Başarısızlık Nedenlerinin Veri Madenciliği Tekniği İle Analizi*”, Yüksek Lisans Tezi, Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Anabilim Dalı.
- Bienkowski, M., Feng, M. & Means, B. (2012). “*Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics-An Issue Brief*”, <http://tech.ed.gov/learning-analytics/> Erişim: 15.10.2015
- Bilen, O., Hotaman, D., Aşkın, O.E. ve Büyüklü, A.H. (2014). “*LYS Başarılarına Göre Okul Performanslarının Eğitsel Veri Madenciliği Teknikleriyle İncelenmesi: 2011 İstanbul Örneği*”, Eğitim ve Bilim, 39(172), 78-94.
- Bozkır, A.S., Sezer, E. ve Gök, B. (2009). “*Öğrenci Seçme Sınavında (ÖSS) Öğrenci Başarımını Etkileyen Faktörlerin Veri Madenciliği Yöntemleriyle Tespiti*”, 5. Uluslararası İleri Teknolojiler Sempozyumu (IATS’09), 13-15 Mayıs, Karabük Üniversitesi, Karabük, 37-43.
- Bozkır, A. S. (2009). “*OLAP ve Veri Madenciliği Teknolojilerinden Yararlanılarak Web Tabanlı Bir Karar Destek Sisteminin Gerçekleştirilmesi*”, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimler Enstitüsü, Ankara.

- Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.I. (1984). *“Classification and Regression Trees”*, Belmont, California: Wadsworth.
- Bresfelean, P., Bresfelean, M. and Ghisoïu, N. (2008). *“Determining Students’ Academic Failure Profile Founded on Data Mining Methods”*, Proceedings of the ITI 2008 30th International Conference on Information Technology Interfaces, 23-26.
- Büyüköztürk, Ş., Çakmak, K.E., Akgün, Ö.E, Karadeniz, Ş. ve Demirel F. (2012). *“Bilimsel Araştırma Yöntemleri”*, Pegem Akademi, 13. Baskı, s.14, Ankara.
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J. and Zanasi, A. (1998). *“Discovering Data Mining: From Concept To Implementation”*, Prentice Hall, Upper Saddle River, Nj, 517s.
- Cavalcanti, E. R., Pires, C. E. S., Cavalcanti, E. P., & Pires, V. F. (2012). *“Detection And Evaluation Of Cheating On College Exams Using Supervised Classification”*, Informatics in Education, 11(2), 169-190.
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). *“A Reference Model For Learning Analytics”*, International Journal of Technology Enhanced Learning, 4(5), 318-331.
- Cristóbal, R., Sebastián, V., Mykola, P., & Ryan, S. J. D. B. (2010). *Introduction Handbook of Educational Data Mining* (pp. 1-6): CRC Press.
- Czaja, S.J., Charness, N., Fisk, A.D., Nair, S.N., Rogers, W.A. and Sharit, J. (2006). *“Factors Predicting the Use of Technology: Findings From the Center for Research and Education on Aging and Technology Enhancement(CREATE)”*, Psychol Aging, 21(2), 333–352.
- Çalış, A., Kayapınar, S. ve Çetinyokuş, T. (2014). *“Veri Madenciliğinde Karar Ağacı Algoritmaları İle Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama”*, Endüstri Mühendisliği Dergisi, 25(4), 2-19.
- Dougiamas, M. (2014). *“About Moodle”*, http://docs.moodle.org/22/en/About_Moodle, Erişim: 06.10.2014.
- Dunham, M.H. (2003). *“Data Mining Introductory and Advanced Topics”*, Prentice Hall, Pearson Education Inc., New Jersey, s.8.
- Duran, N., Önal, A. ve Kurtuluş, C. (2006). *“E-Öğrenme ve Kurumsal Eğitimde Yeni Yaklaşım Öğrenim Yönetim Sistemleri”*, Akademik Bilişim Konferansı’06, Pamukkale Üniversitesi, Denizli, URL: <http://ab.org.tr/ab06/bildiri/165.pdf>.

- Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). *“Design and Implementation of a Learning Analytics Toolkit for Teachers”*, Educational Technology & Society, 15 (3), 58–76.
- Elias, T. (2011). *“Learning Analytics: Definitions, Processes And Potential”*, URL: <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf> Eriřim: 16.10.2015
- Ergöl, E. (2013). *“Biliřim Teknolojileri Öğretmen Adaylarının Moodle İle Ders İşlenmesi Hakkındaki Görüşleri”*, Yüksek Lisans Tezi, Eğitim Teknolojileri Anabilim Dalı, Isparta.
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R. (1996). *“Advances in Knowledge Discovery and Data Mining”*, USA: MIT Press.
- 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 Education, Lake Buena Vista, Florida, s. 42-47.
- Gaafar, L. and Khamis, M. (2009). *“Applications of Data Mining for Educational Decision Support”*, Proceedings of the 2009 Industrial Engineering Research Conference, 228-233.
- García, E., Romero, C., Ventura, S., & de Castro, C. (2011). *“A Collaborative Educational Association Rule Mining Tool”*, The Internet and Higher Education, 14(2), 77-88.
- Gülen, Ö. ve Özdemir, S. (2013). *“Veri Madencilięi Teknikleri ile Üstün Yetenekli Öğrencilerin İlgi Alanlarının Analizi”*, Journal of Gifted Education Research, 1(3), 215-226.
- Gürçan, F. (2009). *“Web İçerik Madencilięi ve Konu Sınıflandırması”*, Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi Fen Bilimler Enstitüsü, Trabzon.
- Halees, A. (2008). *“Mining Students Data to Analyze Learning Behavior: A Case Study”*, Department Of Computer Science, Islamic University, Paletsine.
- Hämäläinen, W., & Vinni, M. (2010). *“Classifiers for educational data mining Handbook of Educational Data Mining”*, pp. 57-74, CRC Press.
- Hand, D., Mannila, H., and Smyth, P. (2001). *“Principles of Data Mining”*, USA: The MIT Press, London.
- Hark, C. (2013). *“Öğrencilerin Akıllı Tahtaya İliřkin Tutumlarının İncelenmesine Yönelik Bir Veri Madencilięi Uygulaması”*, Yüksek Lisans Tezi, Fırat Üniversitesi Eğitim Bilimleri Enstitüsü, Elazığ.

- Holsheimer, M. and Siebes, A. (1994). “*Data Mining: The Search for Knowledge in Databases*”, CWI Technical Report, Amsterdam, s. 2.
- Hou, S., Hou, R., Shi, X., Wang, J. & Yuan, C. (2014). “*Research on C5.0 Algorithm Improvement and the Test in Lightning Disaster Statistics*”, International Journal of Control and Automation, Vol.7, No.1, pp.181-190.
- Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). “*Developing Early Warning Systems to Predict Students’ Online Learning Performance*”, Computers in Human Behavior, Vol: 36, Pg: 469-478. doi: <http://dx.doi.org/10.1016/j.chb.2014.04.002>
- Hung, S., Yen, D. C. and Wang, H. (2005). “*Applying Data Mining to Telecom Churn Management*”, Expert Systems With Applications, Pp. 1-10.
- Jiang, H. & Cai, J. (2013). “*Research on Application of Mining Association Rules in Education Evaluation*”, International Conference on Advanced Information and Communication Technology for Education, China.
- Johnson, M.W., Eagle, M.J., Joseph, L., & Barnes, T. (2011). “The EDM vis tool. In M. Pechenizkiy et al. (Eds.)”, Proceedings of the 3rd Conference on Educational Data Mining 2011 (pp. 349–350).
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). The 2011 Horizon Report.
- Kalikov, A. (2006). “*Veri Madenciliği ve Bir E-Ticaret Uygulaması*”, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü.
- Karasar, N. (2012). “*Bilimsel Araştırma Yöntemi*”, Nobel Yayın Dağıtım, 24. Baskı, s.76-81, Ankara.
- Kass, G.V. (1980). “*An Exploratory Technique for Investigating Large Quantities of Categorical Data*”, Applied Statistics, Vol. 29, No. 2, pp. 119–127.
- Kaya, H. ve Köymen, K. (2008). “*Veri Madenciliği Kavramı ve Uygulama Alanları*”, Doğu Anadolu Bölgesi Araştırmaları, 159-164.
- Kayrı, M. (2008). “*Elektronik Portfolyo Değerlendirmeleri İçin Veri Madenciliği Yaklaşımı*”, Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, Cilt: 5, No: 1, 98-110.
- Kelly, D., Tangney, B. (2005). “*First Aid For You: Getting to Know Your Learning Style Using Machine Learning*”, 5th IEEE International Conference on Advanced Learning Technologies, Taiwan.
- Koç, M., Karabatak, M. (2012). “*Sosyal Ağların Öğrenciler Üzerindeki Etkisinin Veri Madenciliği Kullanılarak İncelenmesi*”, E-Journal of New World Sciences Academy, 7(1), 155-164.

- Kotsiantis, S. (2009). "*Educational Data Mining: A Case Study for Predicting Dropout-Prone Students*", International Journal of Knowledge Engineering and Soft Data Paradigms, 1(2), 101-111. doi: 10.1504/ijkesdp.2009.022718
- Kudyaba, S. and Hoptroff, R. (2001). "*Data Mining and Business Inteligence*", USA: Idea Group, s. 8.
- Kumar, S.A. & Vijayalakshmi, M.N. (2013). "*Discerning Learner's Erudition Using Data Mining Techniques*", International Journal on Intergrating Technology in Education, 2(1), 9-14.
- Kurt, Ç. ve Erdem, O.A. (2012). "*Öğrenci Başarısını Etkileyen Faktörlerin Veri Madenciliği Yöntemleriyle İncelenmesi*", Politeknik Dergisi Journal of Polytechnic Cilt: 15 Sayı: 2 s. 111-116.
- Lakhan, S. and Jhunjhunwala, K. (2008). "*Open Source Software In Education*", Educase Quarterly, 2, 32-40.
- Lee, C.S. (2007). "*Diagnostic, Predictive and Compositional Modeling with Data Mining in Integrated Learning Environmentes*", Computers & Education, 49(3), 562-580.
- 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, URL: <http://ceur-ws.org/Vol-894/paper6.pdf>.
- Loh, W. and Shih, Y. (1997). "*Split Selection Methods For Classification Trees*", Statistica Sinica 7, 815-840.
- Lonn, S. and Teasley, D.S. (2009). "*Saving Time Or Innovating Practice: Investigating Perceptions And Uses Of Learning Management Systems*", Computers & Education, 53, 686–694.
- 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, 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.

- Márquez-Vera, C., Cano, A., Romero, C., & Ventura, S. (2013). "Predicting Student Failure At School Using Genetic Programming And Different Data Mining Approaches With High Dimensional And Imbalanced Data", *Applied Intelligence*, 38(3), 315-330.
- Mostow, J. (2004). "*Some Useful Design Tactics for Mining Its Data*", ITS2004 Workshops – Analyzing Student-Tutor Interaction Logs to Improve Educational Outcomes, Maceió, Alagoas, Brazil.
- Onat, A. (2008). "*Veri Madenciliğinin Web Tabanlı Uygulamalarda İnsan Uyumluluklarının Tespiti Üzerine Bir Çalışma*", Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Osmanbegović, E., & Suljić, M. (2012). "*Data Mining Approach for Predicting Student Performance*". *Economic Review*, 10(1).
- Quinlan, J.R. (1986). "*Induction of Decision Trees*", *Machine Learning*, Volume 1, Issue 1, pp 81-106.
- Quinlan, J. R. (1993). "*C4.5: Programs For Machine Learning. Morgan Kaufmann Publishers*", USA.
- Özpolat, E., & Akar, G. B. (2009). "*Automatic Detection of Learning Styles For An E-Learning System*", *Computers & Education*, 53(2), 355-367.
- Peña-Ayala, A. (2014). "*Educational Data Mining: A Survey And A Data Mining-Based Analysis Of Recent Works*", *Expert Systems With Applications*, 41(4, Part 1), 1432-1462. doi: <http://dx.doi.org/10.1016/j.eswa.2013.08.042>
- Piramuthu, S. (1998). "*Evaluating Feature Selection Methods For Learning in Data Mining Applications*", Thirty-First Annual Hawai International Conference on System Sciences, IEEE Computer Society, 6-9 January, Kohala Coast Hawaii USA, 294.
- 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*, Cilt: 1, Sayı: 2, s. 42-58.
- Roiger, R. J. and Geatz, M.W. (2003). "*Data Mining*", USA: Pearson Education, s. 4.
- Romero, C. and Ventura, S. (2007). "*Educational Data Mining: A Survey from 1995 to 2005*", *Expert Systems with Applications*, Cilt: 33, Sayı: 1: 135-146.
- 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.

- Romero, C., & Ventura, S. (2010). “*Educational Data Mining: A Review Of The State Of The Art*”, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40(6), 601-618.
- Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., & Ventura, S. (2010). “*Web Usage Mining For Predicting Final Marks of Students That Use Moodle Courses*”, Computer Applications in Engineering Education, 21 (1), doi: 10.1002/cae.20456
- Romero, C. & Ventura, S. (2013). “*Data Mining In Education*”, Wiley Interdisciplinary Reviews: Data Mining And Knowledge Discovery, 3(1), 12-27.
- 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?*”, 1 st Moodle ResearchConference, 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, Cilt: 15, Sayı: 3, s. 192-202.
- Sever, H. ve Oğuz, B. (2002). “*Veri Tabanında Bilgi Keşfine Formel Bir Yaklaşım*”, Bilgi Dünyası Dergisi, Cilt: 3, Sayı: 2, URL: <http://bd.org.tr/index.php/bd/article/view/244>, Erişim Tarihi: 12.01.2015.
- Sevindik, T., Kayışlı, K. & Ünlükahraman, O. (2012). “*Web Tabanlı Eğitimde Veri Madenciliği*”, Turkish Journal of Computer and Mathematics Education, 3(3), 183-193.
- Sezer, E.A., Bozkır, A.S., Yağız, S. ve Gökçeoğlu, C. (2010). “*Karar Ağacı Derinliğinin CART Algoritmasında Kestirim Kapasitesine Etkisi: Bir Tünel Açma Makinesinin İlerleme Hızı Üzerinde Uygulama*”, Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu, Kayseri.
- Sharma, R. Singh, H. (2013). “*Data Mining in Education Sector*”, International Journal of Electronics & Data Communication, 2(1), 4-8.
- Shearer, C. (2000), “*The Crisp-DM Model: The New Blueprint for Data Mining*” Journal of Data Warehousing, 5(4), 13-23.
- Siemens, G. (2010). “*What Are Learning Analytics?*”, Elearnspace. URL: <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>
- Siemens, G. & Baker, R.S. (2012). “*Learning Analytics And Educational Data Mining: Towards Communication And Collaboration*”, The 2nd International Conference on Learning Analytics And Knowledge, Canada.

- Stallman, R. (2012). “*Why Open Source Misses the Point of Free Software’ Open source is a Development Methodology; Free Software is a Social Movement*”, URL: <https://www.gnu.org/philosophy/open-source-misses-the-point.html>. Erişim Tarihi: 12.01.2015.
- Superby, J.F., Vandamme, J.P. and Meskens, N. (2006). “*Determination Of Factors Influencing The Achievement Of The First-Year University Students Using Data Mining Methods*”, Proceedings Of The 8th International Conference On Intelligent Tutoring Systems, Educational Data Mining Workshop, Jhongali, Taiwan, Pp: 37-44.
- Şeker, Ş.E. (2013). “*İş Zekası ve Veri Madenciliği*”, Cinius Yayınları, İSTANBUL.
- Şengür, D. ve Tekin, A. (2013). “*Öğrencilerin Mezuniyet Notlarının Veri Madenciliği Metotları İle Tahmini*”, Bilişim Teknolojileri Dergisi, Cilt: 6, Sayı: 3.
- Şuşnea, E. (2009). “*Classification Techniques Used in Educational System*”, The 4th International Conference on Virtual Learning, Romania.
- Talavera, L. and Gaudioso, E. (2004). “*Mining Student Data to Characterize Similar Behavior Groups in Unstructured Collaboration Spaces*”, 16th European Conference on Artificial Intelligence (ECAI 2004) - Workshop on Artificial Intelligence, 17–23, Valencia, Spain.
- Taşdemir, M. (2012). “*Öğrenci Başarısına Etki Eden Faktörlerin Regresyon Analizi İle Tespiti*”, Yüksek Lisans Tezi, Dicle Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı.
- Thuarisingham, B.M. (2003). “*Web Data Mining and Applications in Business Intelligence and Counter Terrorism*”, CRC Press LLC, Boca Raton, FL, USA.
- Ting-Wen, C., El-Bishouty, M. M., Graf, S., & Kinshuk. (2013). “*An Approach for Detecting Students' Working Memory Capacity From Their Behavior in Learning Systems*”, Paper Presented at The Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Tiwari, M., Singh, R., Vimal, N. (2013). “*An Empirical Study of Application of DM Techniques for Predicting Student Performance*”, International Journal of Computer Science and Mobile Computing, 2(2), 53-57.
- Tosun, S. (2007). “*Sınıflandırmada Yapay Sinir Ağları ve Karar Ağaçları Karşılaştırması: Öğrenci Başarıları Üzerine Bir Uygulama*”, Yüksek Lisans Tezi, İstanbul Üniversitesi, Fen Bilimleri Enstitüsü.
- Türker, A.Y. (2012). “*Uzaktan Eğitim Öğretim Yönetim Sisteminin Bulanık Çok Kriterli Karar Verme Yöntemleri İle Seçimi*”, Yüksek Lisans Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Kocaeli.

- Ünal, T.A. (2014). “*Büyük Veri Ve Eğitimsel Veri Madenciliğinin Eğitim Alanına Katkılarının İncelenmesi*”, 8th International Computer & Instructional Technologies Symposium, Trakya University Edirne.
- Whitmer, J., Fernandes, K. & Allen, W.R. (2012). “*Analytics in Progress: Technology Use, Student Characteristics, and Student Achievement*”, Educase, URL: <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.
- Yaralıoğlu, K. (2004). “Uygulamada Karar Destek Yöntemleri”, İlkem Ofset, İzmir.
- Yıldız, E., Bahçeci, F. (2014). “*Öğrenme Yönetim Sistemlerinde Kullanılan Öğrenme Analitikleri Araçlarının İncelenmesi*”, 8th International Computer & Instructional Technologies Symposium, Trakya University Edirne.
- Zhang, Y., Oussena, S., Clark T. and Kim, H. (2010). “*Use Data Mining To Improve Student Retention In Higher Education: A Case Study*”, In ICEIS 2010: Proceedings of the 12th International Conference on Enterprise Information Systems, Vol: 1, Databases and Information Systems Integration, pages 190-197. INSTICC, Funchal, Portugal.

EKLER

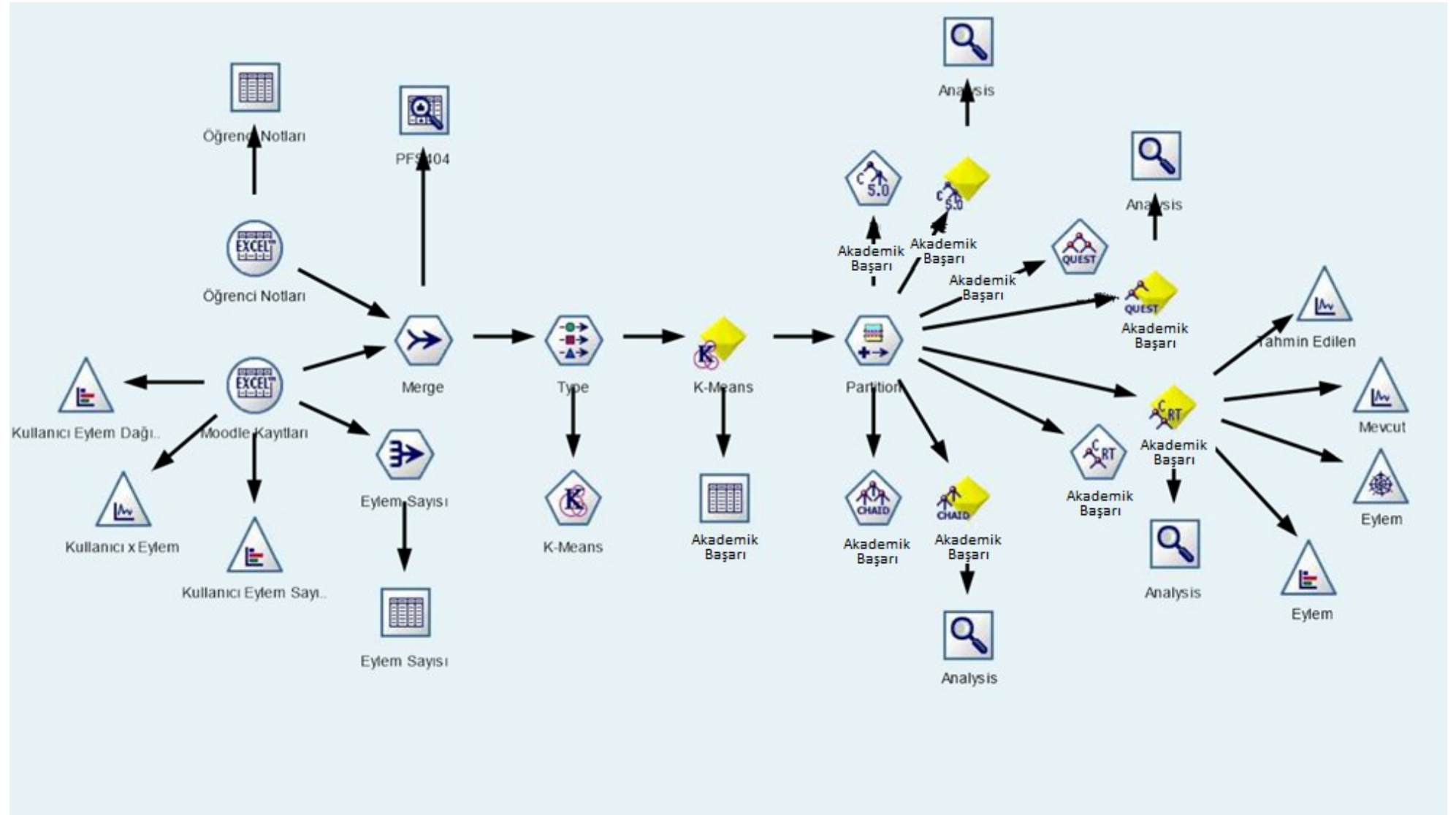
Ek-1. Moodle Eylemleri ve İşlevleri

S.Nu.	Eylem (İngilizce)	Eylem (Türkçe)	İşlevi
1	assignment view all	Tüm gönderilere bak	Öğrencilerin ders öğretmeni tarafından verilen tüm ödev gönderilerini görüntülemesi işlemidir.
2	assign upload	Ödev yükle	Öğrencilerin ödev gönderilerine ait dijital içerikleri (Dosya, Resim, Ses, Video vb.) belirlenen zaman dilimi içerisinde sisteme yüklemesi işlemidir.
3	assign view	Ödev gönderisine bak	Öğrencilerin ders öğretmeni tarafından verilen ödev gönderileri ait ödev gönderi durumunu, not ve son teslim tarihini görüntülemesi işlemidir.
4	blog view	Blog görüntüle	Öğrencilerin blog sayfası açma ve görüntülemesi işlemidir.
5	course enrol	Derse kaydol	Öğrencilerin bir derse katılımcı olarak kaydedilme isteği gönderme ve derse kaydedilmesi işlemidir.
6	course recent	Ders geçmişi	Öğrencilerin bir derse kaydedildikleri tarihten itibaren gerçekleşen ders aktivitelerini, ders kaynak ve aktivite güncellemelerini, forum ve ödev gönderilerini görüntülemesi işlemidir.
7	course user report	Ders kullanıcı raporu	Öğrencilerin kendilerine ait gerçekleştirdikleri (Ders Adı, Tarih, Saat, IP Adresi, Kullanıcı Adı, Eylem, Eylem Bağlantıları, Bilgi Alanı) aktiviteleri görüntülemesi işlemidir.
8	course view	Ders sayfasına bak	Öğrencilerin kayıtlı olduğu dersler içerisinde seçtiği derse ait ders içeriğinin (Kurs Planı, Konular, Dokümanlar vb.) listelendiği ilk giriş sayfasını görüntülemesi işlemidir.
9	forum user report	Forum kullanıcı raporu	Öğrencilerin forum gönderi raporlarını görüntülemesi işlemidir.

Ek-1. Moodle Eylemleri ve İşlevleri (Devam)

S.Nu.	Eylem (İngilizce)	Eylem (Türkçe)	İşlevi
10	forum view forum	Forum sayfasına bak	Öğrencilerin derse ait haber forum sayfasını ya da oluşturulan diğer forum sayfalarını görüntülemesi işlemidir.
11	forum view forums	Tüm forumları listele	Öğrencilerin derse ait oluşturulan tüm forumları görüntülemesi işlemidir.
12	resource view	Ders kaynağını görüntüle	Öğrencilerin derse ait ders içeriğini, ekli dosyaları, derse içeriğiyle ilgili paylaşılan linkleri, ders dokümanlarını, ders sunumları ve ödev açıklamalarını görüntülenmesi işlemidir.
13	resource view all	Tüm ders kaynaklarını görüntüle	Öğrencilerin tüm ders kaynaklarını görüntülemesi işlemidir.
14	upload upload	Dosya yükle	Öğrencilerin dosya yükleme alanına dosya yükleme işlemini gerçekleştirmesi işlemidir.
15	user view	Kullanıcı görüntüle	Öğrencilerin kursa kayıtlı diğer öğrencilerin profil sayfalarını görüntülemesi işlemidir.
16	user view all	Tüm kullanıcıları listele	Öğrencilerin kursa kayıtlı olan tüm öğrencileri listelemesi işlemidir.

Ek-2. SPSS Clementine Kullanılarak Oluşturulan Veri Madenciliği Modeli



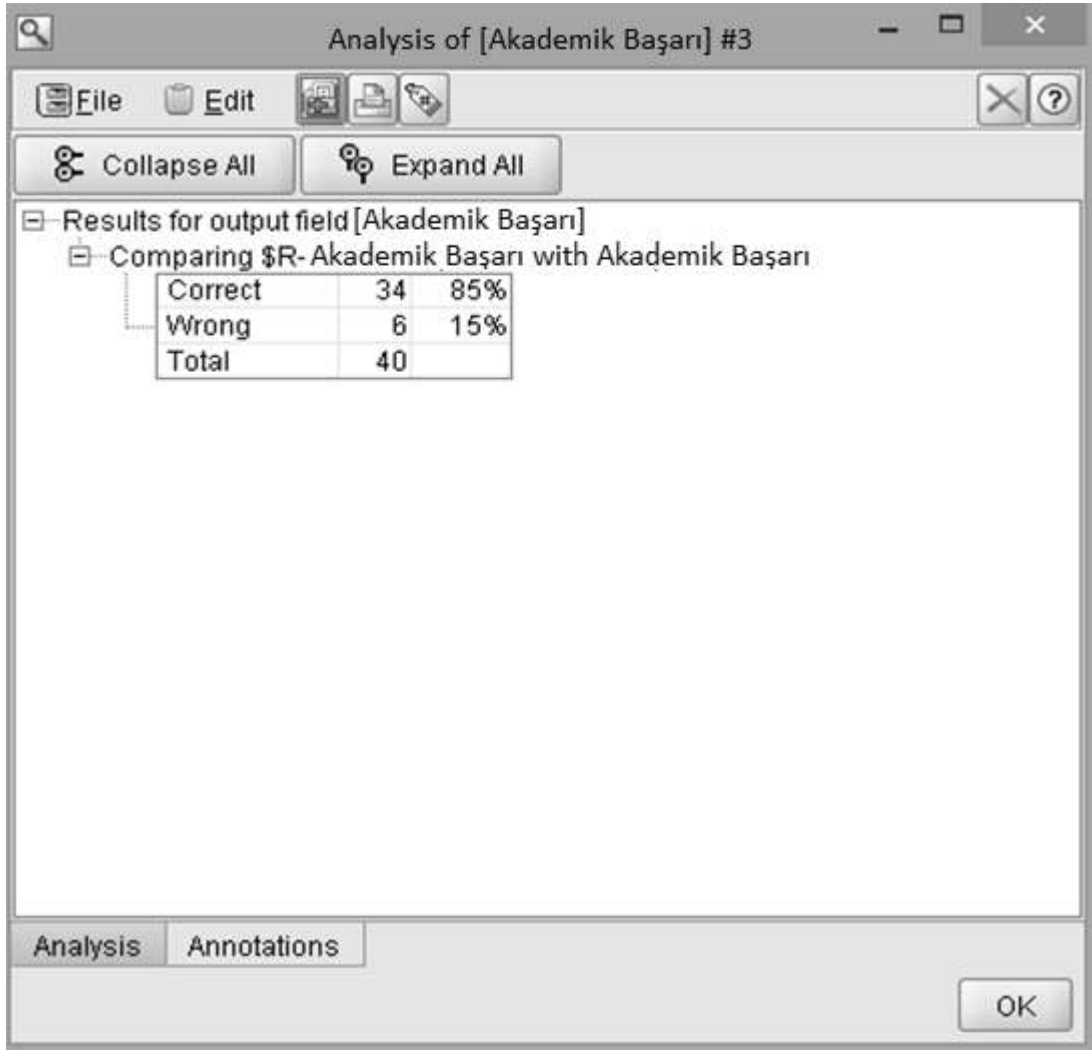
Ek-3. SPSS Clementine’de Kullanılan Düğümler ve İşlevleri

Düğüm	İşlevi
Veri Kaynağı	
 Excel	Excel ortamında kayıtlı olan verilerin SPSS Clementine ortamına aktarılmasını sağlayan düğümdür.
Grafikler	
 Distribution	Bağlı olduğu verilerin dağılım grafiği olarak gösterilmesini sağlayan düğümdür.
 Plot	Bağlı olduğu verilerin saçılım grafiği olarak gösterilmesini sağlayan düğümdür.
 Web	Bağlı olduğu verilerin birbirleriyle olan bağlantı sıklıklarına göre ilişki kuvvetlerinin grafik olarak gösterilmesini sağlayan düğümdür.
Kayıt Seçenekleri	
 Merge	Farklı veri kaynaklarından elde edilen verilerin birleştirilmesi işlemini sağlayan düğümdür.
 Aggregate	Veri kaynaklarından elden edilen verilerin gruplama işlemini gerçekleştirerek verilerin özetlenmesini sağlayan düğümdür.
Çıktı	
 Data Audit	Veriler hakkındaki betimleyici istatistiklerin ve dağılım grafiklerinin gösterilmesini sağlayan düğümdür.
 Analysis	Tahmin edici modellerin geçerliğinin test edilmesini sağlayan düğümdür.
 Table	Bağlandığı noktadaki verileri tabloya dönüştürerek görüntülenmesini sağlayan düğümdür.
Alan Seçenekleri	
 Type	Model düğümlerinden önce kullanılarak veri özelliklerinin tanımlanmasını, model için girdi ve hedef veri alanı seçimlerinin gerçekleştirilmesini sağlayan düğümdür.
 Partition	Model düğümlerinden önce kullanılarak veri kümesini eğitim ve test verisi gibi farklı kısımlara bölerek model geliştirme aşamasında ilgili faaliyetler için kullanılmasını sağlayan düğümdür.

Ek-3. SPSS Clementine’de Kullanılan Düğümler ve İşlevleri (Devam)

Düğüm	İşlevi
Modeller	
 K-Means	K-Means algoritmasını kullanarak kümeleme işleminin gerçekleştirilmesini sağlayan düğümdür.
 C5.0	C5.0 algoritmasını kullanarak karar ağacı oluşturulmasını sağlayan düğümdür.
 C&RT	CART algoritmasını kullanarak karar ağacı oluşturulmasını sağlayan düğümdür.
 QUEST	QUEST algoritmasını kullanarak karar ağacı oluşturulmasını sağlayan düğümdür.
 CHAID	CHAID algoritmasını kullanarak karar ağacı oluşturulmasını sağlayan düğümdür.
Türetilmiş Modeller	
 K-Means	Türetilmiş K-Means kümeleme algoritmasını temsil eder.
 C5.0	Türetilmiş C5.0 algoritmasını temsil eder.
 QUEST	Türetilmiş QUEST algoritmasını temsil eder.
 C&RT	Türetilmiş CART algoritmasını temsil eder.
 CHAID	Türetilmiş CHAID algoritmasını temsil eder.

Ek-4. CART Algoritması Hata ve Doğruluk Oranı



Analysis of [Akademik Başarı] #3

File Edit

Collapse All Expand All

Results for output field [Akademik Başarı]

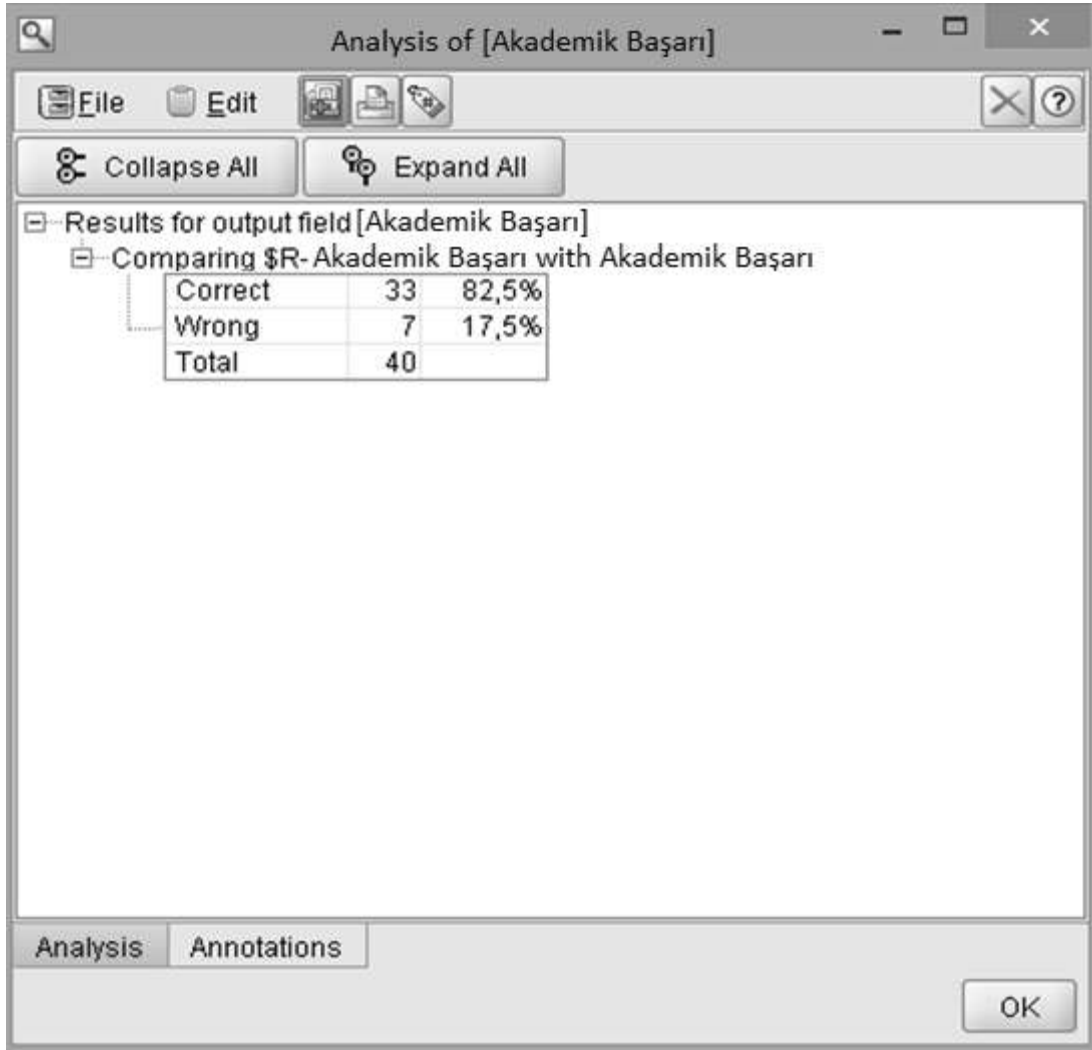
Comparing \$R-Akademik Başarı with Akademik Başarı

Correct	34	85%
Wrong	6	15%
Total	40	

Analysis Annotations

OK

Ek-5. C5.0 Algoritması Hata ve Doğruluk Oranı



Analysis of [Akademik Başarı]

File Edit

Collapse All Expand All

Results for output field [Akademik Başarı]

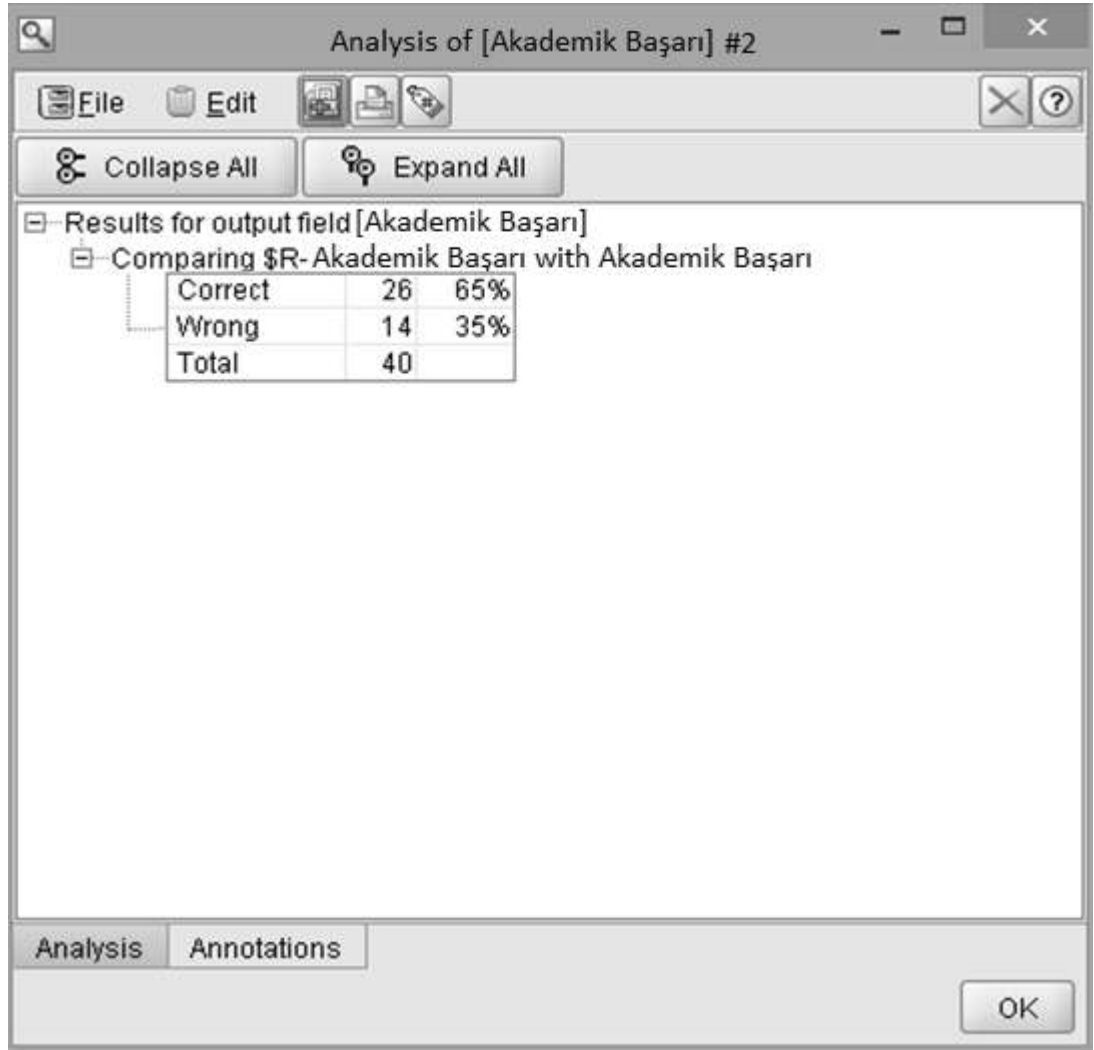
Comparing \$R-Akademik Başarı with Akademik Başarı

Correct	33	82,5%
Wrong	7	17,5%
Total	40	

Analysis Annotations

OK

Ek-6. CHAID Algoritması Hata ve Doğruluk Oranı



Analysis of [Akademik Başarı] #2

File Edit

Collapse All Expand All

Results for output field [Akademik Başarı]

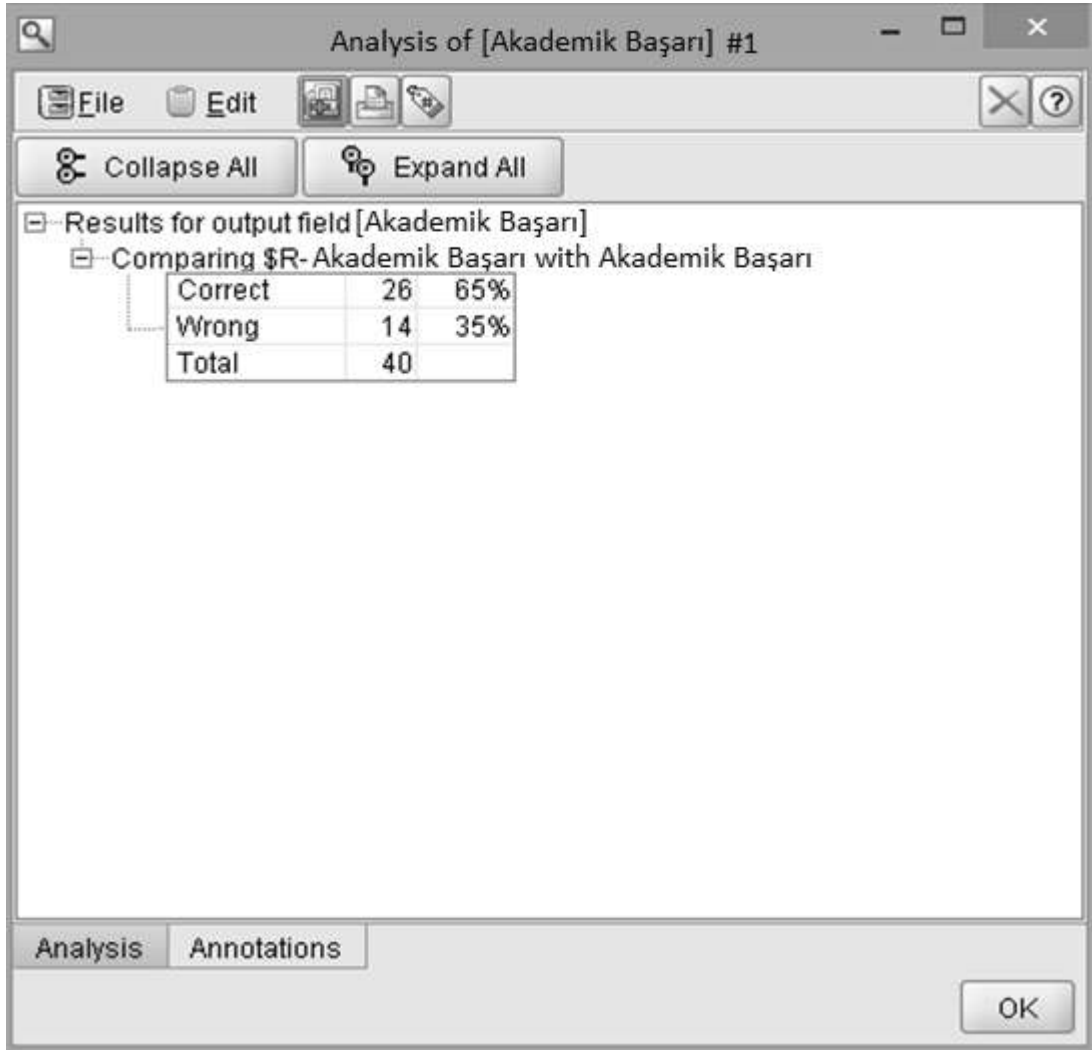
Comparing \$R- Akademik Başarı with Akademik Başarı

Correct	26	65%
Wrong	14	35%
Total	40	

Analysis Annotations

OK

Ek-7. QUEST Algoritması Hata ve Doğruluk Oranı



Analysis of [Akademik Başarı] #1

File Edit

Collapse All Expand All

Results for output field [Akademik Başarı]

Comparing \$R- Akademik Başarı with Akademik Başarı

Correct	26	65%
Wrong	14	35%
Total	40	

Analysis Annotations

OK

Ek-8. İntihal Raporu

The screenshot shows the Turnitin web interface. At the top, there's a navigation bar with the Turnitin logo and a user menu for Halil Ersoy. Below this is a secondary navigation bar with tabs for Assignments, Students, Grade Book, Libraries, Calendar, Discussion, and Preferences. The main content area is titled 'NOW VIEWING: HOME > BILGISAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ : SECTION 1 > ÖZKAN TEZ'. It includes a section 'About this page' explaining the assignment inbox. Below this, the user 'Özkan Tez' is listed with a dropdown for 'INBOX | NOW VIEWING: NEW PAPERS'. A 'Submit File' button is present. To the right, there are links for 'GradeMark Report', 'Edit assignment settings', and 'Email non-submitters'. A table displays the submission details for 'Özkan Özbay' with a 20% similarity score. The footer contains copyright information and links to Usage Policy, Privacy Pledge, Helpdesk, and Research Resources.

Turnitin

Halil Ersoy | User Info | Messages | Instructor | English | Roadmap | Help | Logout

turnitin

Assignments Students Grade Book Libraries Calendar Discussion Preferences

NOW VIEWING: HOME > BILGISAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ : SECTION 1 > ÖZKAN TEZ

About this page

This is your assignment inbox. To view a paper, click the paper's title. To view an Originality Report, click the paper's Originality Report icon in the similarity column. A ghosted icon indicates that the Originality Report has not yet been generated.

Özkan Tez

INBOX | NOW VIEWING: NEW PAPERS

Submit File

GradeMark Report | Edit assignment settings | Email non-submitters

	AUTHOR	TITLE	SIMILARITY	GRADE	RESPONSE	FILE	PAPER ID	DATE
☐	Özkan Özbay	ÖĞRETİM YÖNETİM SİSTEMİ ÜZERİNDE ÜNİVERS...	20%				562164674	19-Aug-2015

Copyright © 1998 – 2015 Turnitin, LLC. All rights reserved.

Usage Policy Privacy Pledge Helpdesk Research Resources