

**ÇEVİRİMİÇİ ÖĞRENME ORTAMINDAKİ ETKİLEŞİM
VERİLERİNE GÖRE ÖĞRENCİLERİN AKADEMİK
PERFORMANSLARININ VERİ MADENCİLİĞİ YAKLAŞIMI İLE
MODELLENMESİ**

**A DATA MINING APPROACH TO STUDENTS' ACADEMIC
PERFORMANCE MODELING IN ONLINE LEARNING
ENVIRONMENT BASED ON THEIR INTERACTION DATA**

Gökhan AKÇAPINAR

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Bilgisayar ve Öğretim Teknolojileri Eğitimi Anabilim Dalı İçin Öngördüğü

Doktora Tezi

olarak hazırlanmıştır.

2014

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

G¼khan AKÇAPINAR'ın hazırladıđı “Çevrimiçi Öğrenme Ortamındaki Etkileşim Verilerine Göre Öğrencilerin Akademik Performanslarının Veri Madenciliđi Yaklaşımı ile Modellenmesi” başlıklı bu çalışma j¼rimiz tarafından **Bilgisayar ve Öğretim Teknolojileri Eđitimi Anabilim Dalı'nda Doktora Tezi** olarak kabul edilmiştir.

Başkan

Prof. Dr. Buket AKKOYUNLU

Üye (Danışman)

Prof. Dr. Arif ALTUN

Üye

Doç. Dr. Erdem KARABULUT

Üye

Doç. Dr. Halil YURDUG¼L

Üye

Yrd. Doç. Dr. Serpil YALÇINALP

ONAY

Bu tez Hacettepe Üniversitesi Lisansüstü Eđitim-Öğretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 22/07/2014 tarihinde uygun gör¼lmüş ve Enstitü Yönetim Kurulunca/...../..... tarihinde kabul edilmiştir.

Prof. Dr. Berrin AKMAN
Eđitim Bilimleri Enstitüsü M¼d¼r¼

İkinci Tez Danışmanı Onay Bildirimi

1- Öğrenci Bilgileri (Student Info)


Adı Soyadı (Name/Lastname)	Gökhan AKÇAPINAR
Anabilim Dalı (Department)	Bilgisayar ve Öğretim Teknolojileri Eğitimi
Bilim Dalı (Division)	Bilgisayar ve Öğretim Teknolojileri Eğitimi
Statüsü (Status)	<input type="checkbox"/> Y. Lisans (Masters - MA, MS) <input checked="" type="checkbox"/> Doktora (Doctorate - Ph.D.) <input type="checkbox"/> Bütünleşik Doktora (Integrated PhD)

2- Tez Bilgileri (Thesis/Dissertation Info)

Tezin Başlığı-1 (Title of thesis/ dissertation-1)	Çevrimiçi Öğrenme Ortamındaki Etkileşim Verilerine Göre Öğrencilerin Akademik Performanslarının Veri Madenciliği Yaklaşımı ile Modellenmesi
Tezin Başlığı-2 (Title of thesis/ dissertation-2)	A Data Mining Approach to Students' Academic Performance Modeling in Online Learning Environment Based on Their Interaction Data

Hacettepe Üniversitesi
Eğitim Bilimleri Enstitüsü Müdürlüğüne
Yukarıda kimlik ve tez bilgileri bulunan öğrencinin ikinci tez danışmanı olarak, tez çalışmasının bilgim dâhilinde tamamlandığını ve sonuçlandırılan çalışmayı onayladığımı belirtirim.

Hacettepe University
Graduate School of Educational Sciences
I, as the co-adviser of the student whose identification and thesis/dissertation information provided above, approve his/her work.

Danışmanın Ünvanı, Adı ve Soyadı (Title, First and Last Name of Co- adviser)	Prof. Dr. Petek AŞKAR
Kurumu (Institution)	TED Üniversitesi
İmza (Signature)	

ÇEVİRİMİÇİ ÖĞRENME ORTAMINDAKİ ETKİLEŞİM VERİLERİNE GÖRE ÖĞRENCİLERİN AKADEMİK PERFORMANSLARININ VERİ MADENCİLİĞİ YAKLAŞIMI İLE MODELLENMESİ

Gökhan AKÇAPINAR

ÖZ

Bu çalışma kapsamında çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak öğrencilerin akademik performanslarının veri madenciliği yöntemleri ile modellenmesi amaçlanmıştır. Öğrenci performansının çevrimiçi ortamdan elde edilen değişkenlere göre modellenmesi, dersi bırakma eğilimi olan öğrencilerin ve dönem sonundaki olası başarısızlıkların erkenden tahmin edilmesi açısından önemlidir. Bu modeller uyarlanabilir öğrenme ortamlarında aktivite düzeylerine göre öğrencilerin otomatik olarak sınıflandırılmasında ya da otomatik uyarlamalar yapılması amacıyla kullanılabilir. Aynı zamanda geleneksel sınıf ortamları ile karşılaştırıldığında öğrenci sayısının çok daha fazla olduğu çevrimiçi öğrenme ortamlarında öğrencilerin akademik performanslarının izlenmesi açısından da önemlidir.

Çalışmaya 2013 - 2014 ders yılı güz döneminde Hacettepe Üniversitesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi (BÖTE) bölümü 2. sınıf öğrencilerinden Bilgisayar Donanımı dersine kayıtlı 76 öğrenci katılmıştır. Öğrenciler 14 hafta süresince yüz yüze derslere ek olarak ders dışı tüm etkinlikleri (yansıma yazma, tartışmalara katılım, kaynak takibi vb.) gerçekleştirmek amacıyla araştırmacılar tarafından tasarlanan ve geliştirilen çevrimiçi öğrenme ortamını kullanmışlardır. Ortamdan toplanan veriler kullanılarak ortamdaki öğrenci davranışlarını yansıtan 28 adet değişken belirlenmiştir. Öğrencilerin akademik performansı ise geçti - kaldı şeklinde kodlanmıştır.

Birinci araştırma sorusu kapsamında farklı sınıflama algoritmaları ve ön işleme yöntemleri karşılaştırılarak öğrencilerin dönem sonu akademik performanslarını en iyi tahmin edecek algoritma ve değişkenlerin belirlenmesi amaçlanmıştır. Bir diğer araştırma problemi kapsamında bu değişkenler ve seçilen algoritma kullanılarak öğrencilerin akademik performanslarının daha önceki haftalarda tahmin edilip edilemeyeceği araştırılmıştır. Son araştırma problemi kapsamında ise çevrimiçi öğrenme ortamında benzer davranış örüntüsü sergileyen öğrenci gruplarının

belirlenmesi ve bu grupların akademik performans ile ilişkisi araştırılmıştır. Verilerin analizinde veri tabanlarındaki gizli örüntü ve bilgileri ortaya çıkartmak amacıyla yaygın olarak kullanılan veri madenciliği yöntemlerinden yararlanılmıştır. Tahmin analizleri sınıflama algoritmaları yardımıyla yapılırken benzer öğrenci gruplarını belirlemek amacıyla kümeleme algoritmalarından yararlanılmıştır. Tahmin analizlerinde sonuçların genelleştirilmesi için ve kümeleme analizlerinde optimum küme sayısını belirlemek için çapraz geçerlilik yöntemi kullanılmıştır.

Araştırma sonuçları öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak dönem sonundaki akademik performanslarının başarılı bir şekilde tahmin edilebileceğini göstermiştir. En yüksek doğru sınıflama oranına veriler eşit genişlik yöntemine göre kesikli hale dönüştürüldüğü durumda ve gini indeksine göre seçilen ilk 10 değişkenin kullanıldığı durumda ulaşılmıştır. Bu durumda CN2 kuralları ve kNN algoritmaları dersten kalan ve geçen öğrencilerin %86'sını doğru olarak sınıflamıştır. Öğrencilerin dönem sonu akademik performanslarının daha önceki haftalardan tahmin edilip edilemeyeceği ile ilgili analizler incelendiğinde ise üçüncü hafta gibi kısa bir sürede bunun %74 oranında doğru olarak tahmin edilebileceği görülmüştür.

Kümeleme analizleri ile ilgili sonuçlar incelendiğinde ise öğrencilerin çevrimiçi öğrenme ortamında gerçekleştirdikleri aktivitelere göre ideal olarak üç farklı kümeye ayrıldığı görülmüştür. Bu kümelerin tanımlanması konusunda ise iki farklı yaklaşım izlenmiştir. İlk olarak elde edilen kümeler her bir kümede yer alan öğrencilerin aktivite düzeylerine göre Aktif Olmayan, Aktif ve Çok Aktif şeklinde adlandırılmıştır. Daha sonra bu kümelerin akademik performans ile ilişkisi incelenmiş ve ortamda daha az aktivite gösteren öğrencilerin derste düşük başarı sergiledikleri (Düşük Öğrenme), orta düzeyde aktivite gösteren öğrencilerin derste orta düzeyde başarı sergiledikleri (Orta Öğrenme) ve yüksek düzeyde aktivite gösteren öğrencilerin derste yüksek başarı (Yüksek Öğrenme) sergiledikleri sonucuna ulaşılmıştır.

Anahtar kelimeler: Akademik performans tahmini, erken uyarı sistemi, öğrenci gruplama, çevrimiçi öğrenme ortamı, eğitsel veri madenciliği, öğrenme analitiği.

Danışman: Prof. Dr. Arif ALTUN, Hacettepe Üniversitesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Anabilim Dalı

II. Danışman: Prof. Dr. Petek AŞKAR, TED Üniversitesi, Eğitim Fakültesi

A DATA MINING APPROACH TO STUDENTS' ACADEMIC PERFORMANCE MODELING IN ONLINE LEARNING ENVIRONMENT BASED ON THEIR INTERACTION DATA

Gökhan AKÇAPINAR

ABSTRACT

The purpose of this study is to model students' academic performance based on their interaction data in an online learning environment with the help of data mining techniques. Academic performance modeling based on variables that has been extracted from online learning environment is important especially to the early prediction of students' drop-out and probable failure in the learning environment. Moreover, these models can be used to classify students' automatically or to make an automatic adaptation based on their activity level in an adaptive learning environment. Teachers can use these models to monitor their students' learning progress in an online learning environment which have more students than traditional classrooms.

The participants of the study consist of 76 university students enrolled in the Computer Hardware course in 2013 – 2014 fall term at Hacettepe University, Department of Computer Education and Instructional Technology (CEIT). In addition to face to face classes during the course of 14 weeks, the students conducted all other activities (writing reflections, participating in discussions, following course resources, etc.) in the online learning environment which has been designed and developed by researchers. Using the data collected from the environment, 28 variables were created reflecting student behavior in the environment. The academic performance of the students were coded as passed or failed.

As a part of the first research question, it was aimed to compare different classification algorithms and pre-processing methods in order to find the best algorithms and the best predictors of students' end of year academic performance. In another research question, it was investigated if it is possible or not to predict students' academic performance in previous weeks with the help of selected predictors and algorithms. In the last research question, we try to find different groups of students who exhibit similar behavior patterns in an online learning environment via clustering; relationships between these clusters and academic

performance were examined as well. In data analysis, data mining methods and techniques that are widely used to extract hidden patterns and knowledge out of the data set were used.

Prediction analyses were performed with the help of classification algorithms, however, clustering algorithms were used to group similar students. In order to generalize the results obtained in prediction analysis and to find the optimal cluster count in clustering analysis, cross validation method was used.

The results of the study showed that students' academic performance at the end of the year could be predicted successfully based on their interaction data in an online learning environment. The highest classification accuracy was achieved when continuous variables were turned into categorical variables via equal width discretization method and the top 10 ranked variables were used according to Gini Index. In this case, CN2 rules and kNN algorithms classify correctly 86% of students who failed and passed the course. When results related to early prediction of students' end of year academic performance are examined, it can be seen that by the third week, 74% of students can be accurately classified.

When examining the results of the cluster analysis, it was observed that students ideally divided into three different clusters according to the activities carried out in the online learning environment. Regarding the identification of these clusters, two different methods were followed. Firstly, the resulting clusters were named according to the students' activity levels: Non-active, Active and Very Active. Then, these clusters' relationships with academic performance were examined. It was found that students with lower activities in the environment had a lower course success (Low Learning); students with moderate activities in the environment had a moderate course success (Medium Learning); and students with higher activities in the environment had a higher course success (High Learning).

Keywords: Academic performance prediction, early warning system, grouping students, online learning environment, educational data mining, learning analytics.

Advisor: Prof. Dr. Arif ALTUN, Hacettepe University, Department of Computer Education and Instructional Technologies

2nd Advisor: Prof. Dr. Petek AŞKAR, TED University, Faculty of Education

ETİK BEYANNAMESİ

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.



İmza
Gökhan AKÇAPINAR

TEŐEKKÜR

Bu zorlu ve uzun süreçte desteęini hiçbir zaman esirgemeyen deęerli hocam ve danıőmanım Prof. Dr. Arif ALTUN'a teőekkür ederim. Hem yüksek lisans hem de doktora sürecinde birlikte alıőma fırsatı bulduęum ok deęerli hocam ve danıőmanım Prof. Dr. Petek AŐKAR'a sonsuz teőekkürlerimi sunarım.

Araőtırma sürecinin baőından sonuna kadar her tez izleme komitesinde getirdikleri önerileri, akademik bilgileri ve destekleri ile tezimi zenginleőtiren deęerli hocalarım Do. Dr. Halil YURDUGÜL ve Do. Dr. Erdem KARABULUT'a ok teőekkür ederim. alıőmamı büyük bir titizlikle inceleyip deęerli katkıları ile tezime destek olan deęerli hocalarım Prof. Dr. Buket AKKOYUNLU ve Yrd. Do. Dr. Serpil YALINALP'e ok teőekkür ederim.

Tezime getirdikleri katkılardan ve yararlı tartıőmalardan dolayı deęerli arkadaşlarım Dr. Erdal COŐGUN ve Dr. Alper BAYAZIT'a teőekkür ederim. Uzun alıőma saatlerinin motivasyon kaynaęı olan deęerli mesai arkadaşlarıma ve bu süreçte emeęi olan tüm arkadaşlarıma teőekkürlerimi sunarım.

İÇİNDEKİLER

ÖZ	iv
ABSTRACT	vi
TEŞEKKÜR.....	ix
İÇİNDEKİLER.....	x
ÇİZELGELER DİZİNİ.....	xii
ŞEKİLLER DİZİNİ.....	xiii
SİMGELER VE KISALTMALAR DİZİNİ	xiv
1. GİRİŞ.....	1
1.1. Problem Durumu	1
1.2. Eğitsel Veri Madenciliği	5
1.2.1. Eğitsel Ortamlar	5
1.2.2. Kullanılan Veri Kaynakları	6
1.2.3. Ön İşleme Süreci	6
1.2.4. Veri Madenciliği Analizleri	7
1.2.5. Sonuçların Yorumlanması.....	7
1.3. Öğrenme Analitiği.....	7
1.3.1. Veri Toplama ve Ön İşleme.....	8
1.3.2. Analitik ve Eylem.....	8
1.3.3. Son İşleme	8
1.4. Araştırmanın Amacı ve Önemi	9
1.5. Araştırma Problemleri ve Alt Problemler	10
1.6. Çalışmanın Sınırlılıkları	11
1.7. İşlevsel Tanımlar	12
2. İLGİLİ ARAŞTIRMALAR	13
2.1. Çevrimiçi Öğrenme Ortamlarında VM ve ÖA Uygulamaları	13
2.1.1. Tahmin Amaçlı Yapılan Çalışmalar	14
2.1.1.1. Akademik Performansın Tahmin Edilmesi ile İlgili Araştırmalar	15
2.1.1.2. Erken Uyarı Sistemleri ile İlgili Araştırmalar	18
2.1.2. Tanımlama Amaçlı Yapılan Çalışmalar.....	19
2.1.2.1. Benzer Öğrenci Gruplarının Belirlenmesi ile İlgili Araştırmalar	20
2.1.3. Genel değerlendirme	23
3. YÖNTEM	24

3.1. Araştırma Grubu.....	24
3.2. Veri Kaynakları.....	24
3.2.1. Çevrimiçi Öğrenme Ortamı	25
3.2.1.1. Çevrimiçi Öğrenme Ortamının Bileşenleri	25
3.2.1.2. Çevrimiçi Öğrenme Ortamının Teknik Özellikleri.....	29
3.2.2. Akademik Performans Puanları	30
3.3. Ham Veri	30
3.4. Araştırmada Kullanılan Değişkenler	32
3.5. Veri Ön İşleme	35
3.6. Veri Analizi	35
3.6.1. Birinci Problem ve Alt Problemlerine İlişkin Veri Analizi.....	36
3.6.2. İkinci Problem ve Alt Problemlerine İlişkin Veri Analizi	41
3.6.3. Üçüncü Problem ve Alt Problemlerine İlişkin Veri Analizi	41
3.7. Araştırmanın İç ve Dış Geçerliliği.....	43
4. BULGULAR ve TARTIŞMA.....	45
4.1. Birinci Araştırma Problemine İlişkin Bulgular.....	45
4.2. İkinci Araştırma Problemine İlişkin Bulgular	53
4.3. Üçüncü Araştırma Problemine İlişkin Bulgular	55
5. SONUÇ ve ÖNERİLER.....	61
5.1. Sonuçlar	61
5.2. Öneriler	67
5.2.1. Araştırmaya Dönük Öneriler	68
5.2.2. Uygulamaya Dönük Öneriler.....	69
KAYNAKÇA.....	71
EKLER DİZİNİ	78
EK 1. Araştırmada Kullanılan Değişkenlere İlişkin Tanımlayıcı İstatistikler	79
EK 2. Orjinallik Raporu.....	80
ÖZGEÇMİŞ	81

ÇİZELGELER DİZİNİ

Çizelge 3.1: Katılımcıların Demografik Bilgileri.....	24
Çizelge 3.2: Başarı Puanları ile İlgili Tanımlayıcı İstatistikler.....	30
Çizelge 3.3: Araştırmada Kullanılan Değişkenler ve Açıklamaları.....	35
Çizelge 3.4: Örnek Çapraz Tablo	39
Çizelge 4.1: Birinci Araştırma Problemi Kapsamında Yapılan Analizler	45
Çizelge 4.2: Verilerin Sürekli Olduğu Durumda Yapılan Analiz Sonuçları	46
Çizelge 4.3: Verilerin Eşit Frekans Yöntemi ile Kesikli Hale Dönüştürüldüğü Durumda Yapılan Analiz Sonuçları.....	46
Çizelge 4.4: Verilerin Eşit Genişlik Yöntemi ile Kesikli Hale Dönüştürüldüğü Durumda Yapılan Analiz Sonuçları.....	47
Çizelge 4.5: Üç Farklı Özellik Seçme Yöntemine Göre Değişkenlerin Önem Puanları	48
Çizelge 4.6: Kazanç Oranı Yöntemine Göre en Önemli 10 Değişken Seçilerek Yapılan Analiz Sonuçları	49
Çizelge 4.7: Gini İndeks Yöntemine Göre en Önemli 10 Değişken Seçilerek Yapılan Analiz Sonuçları	49
Çizelge 4.8: DVM Ağırlığı Yöntemine Göre en Önemli 10 Değişken Seçilerek Yapılan Analiz Sonuçları	49
Çizelge 4.9: Gini İndeksine Göre en Önemli 10 Değişken.....	50
Çizelge 4.10: kNN Algoritmasının Çapraz Tablo Sonuçları*	51
Çizelge 4.11: kNN Algoritmasına Göre Farklı Haftalarda Alınan Veri ile Oluşturulan Sınıflama Modellerine İlişkin Analiz Sonuçları.....	53
Çizelge 4.12: KNN Algoritmasına Göre Farklı Haftalarda Alınan Veri ile Oluşturulan Sınıflama Modellerine İlişkin Çapraz Tablo Değerleri.....	54
Çizelge 4.13: X-Ortalamalar Algoritmasına Göre Küme Ortalamaları	56
Çizelge 4.14: EM Kümeleme Algoritmasına Göre Küme Ortalamaları	57
Çizelge 4.15: EM ve X-Ortalamalar Kümeleme Algoritmalarına Göre Kümelerde Yer Alan Öğrencilerin Ortalama Başarı Puanları	58
Çizelge 4.16: Dersi Geçen ve Dersten Kalan Öğrencilerin Kümelere Göre Dağılımı .	59

ŞEKİLLER DİZİNİ

Şekil 1.1. EVM Süreci.....	5
Şekil 1.2. ÖA Süreci	8
Şekil 3.1. Araştırma Süreç Modeli	24
Şekil 3.2. Çevrimiçi Öğrenme Ortamı, İleti Bölümü	25
Şekil 3.3. Yeni İleti Yazma Sayfası.....	26
Şekil 3.4. Kaynaklar Sayfası.....	27
Şekil 3.5. Bildirimler Sayfası	28
Şekil 3.6. Silverlight Uygulama Yaşam Döngüsü.....	29
Şekil 3.7. Öğrencilerin Günlere Göre Ortama Giriş Sayıları	31
Şekil 3.8. Öğrencilerin Günlere Göre Yazdıkları İleti Sayıları	31
Şekil 3.9. Öğrencilerin Günlere Göre Ders Kaynaklarında Gezinim Sayıları	31
Şekil 3.10. Birinci Araştırma Problemine İlişkin Analizi Süreci.....	37
Şekil 3.11. Örnek ROC Eğrisi.....	41
Şekil 3.12. İkinci Araştırma Problemine İlişkin Analizi Süreci	41
Şekil 3.13. X-Ortalamlar Kümeleme Analiz Süreci	43
Şekil 3.14. EM Kümeleme Analiz Süreci	43
Şekil 4.1. Karar Ağacı Algoritması ile Elde Edilen Örnek Karar Ağacı.....	52
Şekil 4.2. Farklı Haftalara Ait Sınıflama Performansı	54
Şekil 4.3. Farklı Haftalara ait Sınıflama Performansı.....	55
Şekil 4.4. X-Ortalamlar Algoritmasına Göre Normalleştirilmiş Küme Ortalamaları ...	56
Şekil 4.5. EM Algoritmasına Göre Normalleştirilmiş Küme Ortalamaları	58
Şekil 4.6. EM ve X-Ortalamlar Kümeleme Algoritmalarının Ders Başarısına Göre Karşılaştırması	59
Şekil 4.7. EM ve X-Ortalamlar Kümeleme Algoritmalarının Ders Başarısına Göre Karşılaştırması	60

SİMGELER VE KISALTMALAR DİZİNİ

BIC: Bayes Bilgi Kriteri (Bayesian Information Criterion)

DSO: Doğru Sınıflama Oranı

DVM: Destek Vektör Makinaları (Support Vector Machine)

EAKA: ROC Eğrisi Altında Kalan Alan

EM: Maksimum Beklenti (Expectation Maximization)

EVM: Eğitsel Veri Madenciliği

KDD: Veri Tabanlarından Bilgi Keşfi (Knowledge Discovery in Databases)

kNN: k En Yakın Komşuluk (k-Nearest Neighbors)

ÖA: Öğrenme Analitiği

ÖYS: Öğrenme Yönetim Sistemi

ROC: Alıcı İşletim Karakteristiği (Receiver Operating Characteristic)

VM: Veri Madenciliği

ZÖS: Zeki Öğrenme Sistemi

1. GİRİŞ

1.1. Problem Durumu

Geleneksel sınıf ortamları ile karşılaştırıldığında çevrimiçi öğrenme ortamları öğrencilerin öğrenme aktiviteleri ile ilgili önemli miktarda veriyi veri tabanlarında kayıt edebilmektedir (Greller & Drachsler, 2012; Koedinger, Cunningham, Skogsholm, & Leber, 2008). Fare ve klavye hareketleri seviyesine kadar inebilen bu verilerin Eğitsel Veri Madenciliği (EVM) ve Öğrenme Analitiği (ÖA) gibi yeni yaklaşımlarla analiz edilmesi sayesinde eğitim araştırmacılarının üzerinde çalıştığı karmaşık problemlerin çözülmesi hedeflenmektedir. Bu sayede öğrencilerin davranışlarının, tecrübelerinin, bilgi düzeylerinin modellenmesi, benzer öğrenci profillerinin oluşturulması, uyarlanabilir ve kişiselleştirilebilir ortamlarda kullanılacak bilgilerin üretilmesi mümkün olabilir (Bienkowski, Feng, & Means, 2012). C. Romero ve Ventura (2010) eğitsel ortamlarda Veri Madenciliği (VM) ve ÖA uygulamalarını aşağıda verilen 11 kategoride toplamışlardır:

- Verinin analizi ve görselleştirilmesi
- Geri bildirim sağlama
- Öneri sistemleri
- Öğrenci modellemesi
- İstenmeyen öğrenci davranışlarının belirlenmesi
- Öğrencileri gruplama
- Sosyal ağ analizi
- Kavram haritası geliştirme
- Ders malzemelerinin yapılandırılması
- Planlama ve çizelgeleme
- Öğrenci performansının tahmin edilmesi.

Aşağıdaki bölümde her bir kategorinin kapsadığı çalışmaların amacı açıklanmış ve verilen örneklerle EVM ve ÖA uygulamaları hakkında genel bir çerçeve çizilmesi amaçlanmıştır.

Verinin analizi ve görselleştirilmesi: veri analizi ve görselleştirmenin amacı verideki önemli bilgileri vurgulamak ve karar alma süreçlerini desteklemektir. İstatistik ve bilgi görselleştirme bu amaçla en çok kullanılan yöntemlerdir (C. Romero & Ventura, 2010). Eğitsel ortamlarda bu veriler, öğretmen ve yöneticilerin öğrenciler ve dersler hakkında genel bir izlenim edinmesine yardımcı olması açısından oldukça kullanışlıdır. Haftalık ya da aylık kullanım raporları (Monk, 2005); öğrencilerin çalışma süreleri, ortama giriş sayıları, doğru cevaplama oranları (Feng & Heffernan, 2006) bu sayede elde edilebilecek bilgilerden bazılarıdır.

Geri bildirim sağlama: öğrenci davranışları ile ilgili veriler öğrencilere, eğitimcilere ve yöneticilere karar alma süreçlerini destekleyici geri bildirimler sağlamak amacıyla kullanılabilir. Bu veriler aynı zamanda yazılım geliştiricilere (eğitsel) ürünlerini hızlı bir şekilde gözden geçirme ve geliştirme olanağı sağlamak için kullanılabilir. Mevcut verinin görsel olarak sunulması ya da tanımlayıcı istatistikler verilmesinden farklı olarak burada amaç, verideki yeni, keşfedilmemiş bilgileri ortaya çıkartmaktır (C. Romero & Ventura, 2010). Büyük veri tabanlarında değişkenler arası bilinmeyen ilişkileri ortaya çıkartan ve bunları bir kurallar dizisi olarak ortaya koymayı amaçlayan birliktelik kuralları (Zhang & Zhang, 2002) geri bildirim sağlama amacıyla en çok kullanılan VM yöntemidir.

Öneri sistemleri: öneri sistemlerinin amacı, uyarlanabilir öğrenme ortamlarında öğrencilere kişiselleştirilebilir aktiviteler (bir sonraki görev/problem, ziyaret edilecek bağlantı vs.) sunmaktır. Birliktelik kuralları, kümeleme ve sıralı örüntü madenciliği öneri sistemlerinde en çok kullanılan VM yöntemleridir (C. Romero & Ventura, 2010).

Öğrenci modellemesi: öğrenci modellemesinin amacı, öğrencilerin beceri ve açıklayıcı bilgilerini kapsayan bilişsel modellerinin geliştirilmesidir (C. Romero & Ventura, 2010). VM teknikleri ile öğrenci modelinin otomatik olarak oluşturulması amacıyla motivasyon, doyum, öğrenme stili vb. kişisel özellikler ve öğrenme davranışları birlikte dikkate alınır (Frias-Martinez, Chen, & Xiaohui, 2006). Bayes ağları öğrenci modeli oluşturulması amacıyla kullanılan en temel VM yöntemidir. Çevrimiçi öğrenme ortamlarında öğrencilerin öğrenme stillerinin otomatik olarak belirlenmesi (Özpolat & Akar, 2009), öğrencilerin öğrenme ortamındaki gezinim örüntülerinden çalışma belleği kapasitelerinin belirlenmesi (Ting-Wen, El-Bishouty, Graf, & Kinshuk, 2013) vb. çalışmalar buna örnek olarak verilebilir.

İstenmeyen öğrenci davranışlarının belirlenmesi: istenmeyen öğrenci davranışlarını belirlemedeki amaç görev dışı davranışlar sergileme (Ryan S.J.d. Baker, 2007), sistemi kandırmaya yönelik hareket etme (R. S. Baker, Corbett, Koedinger, & Wagner, 2004), dersi bırakma (Kotsiantis, 2009) vb. sorunlar yaşayan öğrencileri otomatik olarak tespit etmek ve bu davranışları düzeltmeleri için bir fırsat vermektir. Sınıflama ve kümeleme bu amaçla en çok kullanılan VM teknikleridir (C. Romero & Ventura, 2010).

Öğrenci gruplama: benzer öğrenci gruplarının belirlenmesindeki amaç kişisel özellikler, öğrenme tercihleri, bireysel farklılıklar, kullanım verileri vb. açılardan benzerlik gösteren öğrencilerin belirlenip sınıflandırılmasıdır (Cristobal Romero & Ventura, 2013). Benzer öğrenci gruplarının belirlenmesi, uyarlanabilir öğrenme ortamlarında otomatik uyarlamalar yapmak açısından ve öğretmenin öğrencilere uygun geri bildirim vermesi açısından önemlidir. Sınıflama ve kümeleme bu amaçla en çok kullanılan VM teknikleridir (C. Romero & Ventura, 2010).

Sosyal ağ analizi: sosyal ağ analizini bireylere ait özellikler (değişkenler) yerine bireyler arası ilişkilerin ortaya çıkartılması ile ilgilidir. Sosyal ağ, bir insan grubu, bir organizasyonda yer alan insanlar ya da birbirine arkadaşlık, işbirliği, bilgi değişimi gibi sosyal ilişkilerle bağlı bireylerin oluşturduğu bir topluluk olarak düşünülebilir (Freeman, 2004). İşbirlikçi filtreleme (collaborative filtering) yöntemi sosyal ağ analizlerinde yaygın olarak kullanılan VM yöntemidir (C. Romero & Ventura, 2010). İşbirlikçi filtreleme çevrimiçi öğrenme ortamlarında öğrencilere bağlama duyarlı (context aware) öğrenme nesnesi önermek amacıyla (Lemire, Boley, McGrath, & Ball, 2005) ya da öğrenciler için bireysel öneri sistemlerinin oluşturulmasına kullanılmaktadır (Drachsler, Hummel, & Koper, 2008).

Kavram haritası geliştirme: kavram haritaları, bilginin hiyerarşik yapısının açıklandığı ve kavramlar arası ilişkilerin ifade edildiği kavramsal grafiklerdir (Novak & Cañas, 2008). VM teknikleri bu haritaların otomatik olarak oluşturulması konusunda eğitimcilere yardımcı olabilmektedir. Metin madenciliği ve birliktelik kuralları bu amaçla en çok kullanılan VM teknikleridir (C. Romero & Ventura, 2010).

Ders malzemelerinin yapılandırılması: ders malzemelerinin yapılandırılmasının amacı eğitimcilere ve geliştiricilere ders malzemelerini ve öğrenme içeriklerini

otomatik olarak yapılandırılması/geliştirilmesi konusunda yardımcı olmaktadır (C. Romero & Ventura, 2010).

Planlama ve çizelgeleme: planlama ve çizelgeleme ile açılacak olan derslerin planlanması, öğrencilerin ders çizelgelerinin oluşturulması, kaynakların uygun şekilde tahsis edilmesi, müfredat oluşturma vb. konularda geleneksel eğitsel süreçlerin zenginleştirilmesi amaçlanmaktadır (C. Romero & Ventura, 2010). Bu amaçla yaygın olarak kullanılan VM yöntemi ise birliktelik kurallarıdır.

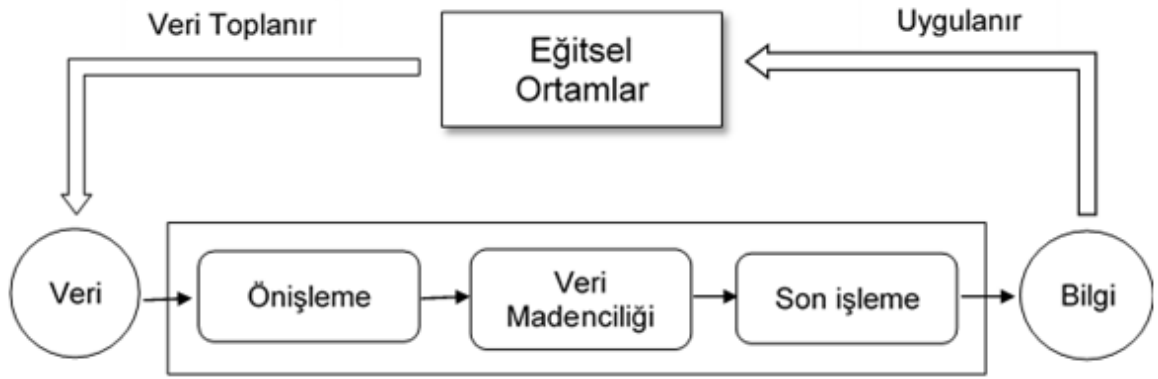
Öğrenci performansının tahmin edilmesi: öğrenme ortamları ile ilgili tahmin çalışmalarının amacı, öğrenciyi tanımlayan ve değeri bilinmeyen bir değişkenin değerinin tahmin edilmesidir (C. Romero & Ventura, 2010). Sınıflama ve regresyon bu amaçla en çok kullanılan VM teknikleridir. Tahmin analizleri ile öğrenci performansı (bilgi, puan vs.) ile ilgili tahmin modelleri oluşturulabilir ve bu tahmin modelleri uyarlanabilir sistemlerde öğrenciyi yönlendirmek veya bir sonraki adımda nasıl bir performans göstereceğini tahmin etmek amacıyla kullanılabilir (Bienkowski ve diğerleri, 2012). Öğrenci performansını tahmin etmeye yönelik kullanılabilir modellerin geliştirilmesi ile birlikte öğretmenler açısından el ile yapılması çok fazla zaman ve emek gerektiren işlerin otomatik olarak yapılması sağlanabilir (Lopez, Luna, Romero, & Ventura, 2012).

EVM ve ÖA araştırmacılarının üzerinde çalıştığı önemli bir konu ise öğrencilerin akademik performanslarını tahmin edecek istatistiksel modellerin oluşturulmasıdır (Chatti, Dyckhoff, Schroeder, & Thüs, 2012; Peña-Ayala, 2014). Bu sayede öğrencilerin bir derste sergileyecekleri performans önceden tahmin edilebilir ve zamanında yapılacak müdahalelerle olası başarısızlıkların önüne geçilebilir (L Johnson, Smith, Willis, Levine, & Haywood, 2006). Öğretim elemanları/öğretmenler, bu modelleri kullanarak öğrencinin gelişimini izleme ve buna uygun müdahale yöntemleri geliştirme konusunda değerli bilgiler üretebilirler.

Akademik performansın modellenmesi temeline dayanan bu çalışma kapsamında ele alınan problemler ve analiz yöntemlerinin daha iyi anlaşılabilmesi amacıyla araştırmanın amacına geçmeden önce ilk olarak EVM ve ÖA alanları ile ilgili bilgilere yer verilmiştir.

1.2. Eğitsel Veri Madenciliği

EVM; öğrenme ortamlarından elde edilen verilerin ortaya çıkartılması için yöntemler geliştirilmesi ve bu yöntemlerin öğrencilerin ve öğrenme ortamlarının daha iyi anlaşılmasında kullanılması ile ilgilenen bir disiplin olarak tanımlanmaktadır (G. Siemens & R.S.J.d. Baker, 2012). García, Romero, Ventura ve de Castro (2011) EVM'yi eğitsel sistemlerden elde edilen ham verinin eğitim yazılımlarının, geliştiricilerin, öğretmenlerin ve araştırmacıların kullanabileceği bilgiye çevirme süreci olarak tanımlamaktadır ve bu süreci Şekil 1.1.'deki gibi özetlemektedir.



Şekil 1.1. EVM Süreci

Aşağıda EVM çalışmalarında izlenen bu sürecin bileşenleri ayrıntılı olarak açıklanmıştır.

1.2.1. Eğitsel Ortamlar

EVM çalışmalarının önemli bir bölümünde veriler, bilgisayar ve web tabanlı eğitim sistemlerinden elde edilmektedir. Peña-Ayala (2014), 2010 – 2013 yılları arası yapılan EVM çalışmalarını incelediği çalışmada incelediği makalelerin yarısına yakınında eğitsel ortam olarak Zeki Öğrenme Sistemlerinin (ZÖS) ve Öğrenme Yönetim Sistemlerinin (ÖYS) kullanıldığını belirtmiştir. Bunun dışında uyarlanabilir ortamlar, çevrimiçi değerlendirme sistemleri (Cristobal Romero & Ventura, 2013), geleneksel sınıf ortamları (Márquez-Vera, Cano, Romero, & Ventura, 2013) ve Web 2.0 araçları (viki, forum, blog vb.) da EVM çalışmalarında verinin elde edildiği ortamlar olarak kullanılmaktadır. Çalışmanın amacına göre bu ortamlar aynı zamanda analiz sonuçlarını uygulamak için de kullanılabilir.

1.2.2. Kullanılan Veri Kaynakları

EVM çalışmalarında veriler, kullanılan ortama (geleneksel, bilgisayar destekli, web destekli vb.) ve araştırma problemine göre değişiklik göstermektedir. Bu veriler log verileri, etkileşim verileri, davranış verileri, gözlem verileri, ölçek verileri ve kontrollü deneylerde ölçülen ölçümlerden elde edilen veriler olabilmektedir (Cristóbal, Sebastián, Mykola, & Ryan, 2010). Bunun dışında göz hareketleri ile ilgili metrikler (Akçapınar, Cosgun, & Altun, 2011; S. Amershi & Conati, 2009), ya da öğrenciler tarafından yazılan mesajlar (Cristóbal Romero, Ventura, & García, 2008) ve açık uçlu sınavlarda öğrencilerin verdiği cevaplar (Cavalcanti, Pires, Cavalcanti, & Pires, 2012) gibi metinler de EVM çalışmalarında veri kaynağı olarak kullanılabilir.

1.2.3. Ön İşleme Süreci

VM çalışmalarında olduğu gibi EVM çalışmalarında da analiz aşamasına geçilmeden önceki en önemli adım verilerin ön işleme sürecidir. Ön işlemenin amacı verinin kalitesini artırmak ve seçilen analiz için ideal değişkenleri elde etmektir (Hämäläinen & Vinni, 2010). Aynı zamanda anlaşılabilir sonuçların üretilmesi açısından da önem arz etmektedir. Ön işleme süreci, verinin türüne ve karmaşıklığına göre değişiklik göstermekle birlikte VM çalışmalarında en çok kullanılan ön işleme adımları aşağıda özetlenmiştir.

- a) **Veri temizleme:** Bu aşamada ilgisiz verilerin silinmesi, kayıp verilerin ya da aykırı gözlemlerin tespit edilmesi gibi işlemler yapılmaktadır.
- b) **Veri dönüştürme:** VM çalışmalarında en sık yapılan veri dönüştürme işlemleri verilerin standartlaştırılması ve sürekli verilerin kesikli verilere dönüştürülmesi işlemidir (Ribeiro, Traina, & Caetano Traina, 2008). Bu dönüştürme işlemi verinin eşit sayıdaki birimlere ayrılması (eşit frekans), eşit uzunlukta belirlenen kesme noktalarına göre belirlenen sayıda aralıklara bölünmesi (eşit genişlik) ya da kullanıcı tarafından belirlenen kesme noktalarına göre (kullanıcı tanımlı) bölünmesi şeklinde yapılabilir (Dougherty, Kohavi, & Sahami, 1995; Han, Kamber, & Pei, 2006).
- c) **Veri birleştirme:** Farklı kaynaklardaki verilerin (ör: öğrencinin demografik bilgileri ve etkileşim verileri) tek bir ortamda birleştirilerek analize hazır hale getirilmesidir.

d) Özellik (nitelik) seçme: İlgisiz ve fazla değişkenleri elimine ederek analizde kullanılacak değişkenlerin seçilmesi sürecidir (Piramuthu, 2004). En önemli ya da en önemsiz değişkenlerin belirlenmesi amacıyla adım adım seçme yöntemi ya da ID3, C4.5, CART gibi karar ağacı algoritmaları kullanılmaktadır (Han ve diğerleri, 2006).

Yukarıda verilen tekniklere ek olarak özellikle log dosyalarının analiz edilmesinde kullanıcıların ve kullanıcılara ait oturumların tanımlanması vb. işlemleri de ön işleme sürecinde yapılmaktadır (Judy, 2010).

1.2.4. Veri Madenciliği Analizleri

EVM sürecinin en önemli bölümü ön işleme sonucu elde edilen veriye VM yöntemlerinin uygulandığı aşamadır. Burada seçilen problem durumuna göre uygun VM yöntemi belirlenir ve analizler yapılır. Yapılan tarama çalışmaları EVM araştırmalarında en çok tercih edilen VM yöntemlerinin sırasıyla: sınıflama; kümeleme; regresyon ve birliktelik kuralları olduğuna işaret etmektedir (Peña-Ayala, 2014; C. Romero & Ventura, 2010). Bu yöntemler EVM çalışmalarında kullanılan VM yöntemlerinin %91'ini oluşturmaktadır (Peña-Ayala, 2014).

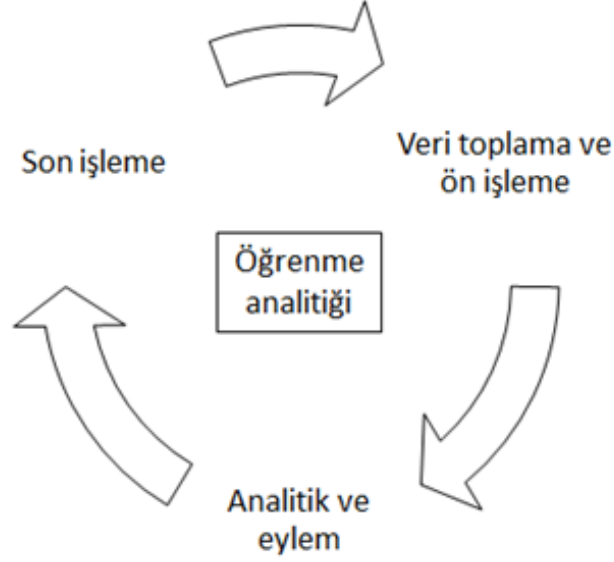
1.2.5. Sonuçların Yorumlanması

EVM sürecinin son aşaması, analiz sonucu elde edilen bilgilerin, model veya örüntülerin yorumlanarak karar verme süreçlerinde ya da eğitsel ortamın iyileştirilmesi amacıyla yorumlanması aşamasıdır (García, Romero, Ventura, & de Castro, 2011; Cristobal Romero & Ventura, 2013).

1.3. Öğrenme Analitiği

Eğitim ortamlarından toplanan verilerin karar verme süreçlerinde kullanımı ile ilgilenen bir diğer araştırma alanı ise ÖA'dır. ÖA; 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 (G. Siemens & R.S.J.d. Baker, 2012). ÖA, EVM'den daha fazla akademik disiplin içermektedir. Bilgisayar bilimleri, istatistik, psikoloji ve öğrenme bilimlerine ek olarak bilgi bilimleri ve sosyolojiden de kavram ve teknikleri kapsamaktadır (Bienkowski ve diğerleri, 2012). ÖA'nın hedefi öğretmen ve okulların eğitsel olanaklarının her bir öğrencinin ihtiyaç ve yeteneğine göre uyarlanmasını olanaklı kılmak olarak tanımlanmaktadır (L. Johnson, Smith, Willis,

Levine, & Haywood, 2011). Şekil 1.2'de görsel olarak sunulan ÖA süreci: (1) veri toplama ve ön işleme, (2) analitik ve eylem, (3) son işleme adımlarından oluşan tekrarlı bir süreçtir (Chatti ve diğerleri, 2012).



Şekil 1.2. ÖA Süreci

Aşağıda bu süreçte yer alan bileşenlerin açıklamalarına yer verilmiştir.

1.3.1. Veri Toplama ve Ön İşleme

EVM sürecinde olduğu gibi ÖA'da da ilk ve en önemli aşama verinin elde edilmesidir. Burada kullanılan veri çok çeşitli eğitsel ortamlardan elde edilebilir. Farklı kaynaklardan toplanan veriler analiz edilmeden önce VM analizlerindeki ön işleme adımları kullanılarak analize hazır hale getirilir.

1.3.2. Analitik ve Eylem

Ön işleme sonucu elde edilen veriye analitik uygulamasının amacı doğrultusunda (daha etkili öğrenme yaşantıları sağlamada yardımcı olacak gizli örüntülerin keşfedilmesi amacıyla) farklı ÖA teknikleri uygulanır (Chatti ve diğerleri, 2012). Analiz sonucu elde edilen bilgilere göre eyleme geçilmesi ise tüm analitik sürecinin temel amacıdır. Bu eylemlerden bazıları; izleme ve analiz, tahmin ve müdahale, mentörlük ve tutorluk, değerlendirme ve geri bildirim, uyarılama, kişiselleştirme ve tavsiye ve yansıma şeklinde sıralanabilir.

1.3.3. Son İşleme

Analitik süreci tekrar eden bir süreç olduğu için analitik uygulamasının iyileştirilmesinin sürekliliğinin sağlanması konusunda son işleme son derece önemli

bir adımdır. Bu adım; farklı kaynaklardan yeni verilerin eklenmesi, veri setinin tekrar gözden geçirilmesi, mevcut değişkenlerin düzenlenmesi, yeni metriklerin eklenmesi ve yeni analitik yönteminin seçilmesi gibi işlemleri kapsamaktadır (Beal, Qu, & Lee, 2006).

EVM ve ÖA iki farklı alan gibi gözükmemektedir; ancak hedefler ve ilgi alanları açısından birçok ortak nokta bulunmaktadır. Bunun dışında kullandıkları teknikler ve yöntemler açısından bir takım farklılıklar bulunmaktadır (Agudo-Peregrina, Iglesias-Pradas, Conde-González, & Hernández-García, 2013; Chatti ve diğerleri, 2012; Cristobal Romero & Ventura, 2013; George Siemens & Ryan S. J. d. Baker, 2012). EVM eğitsel soruları cevaplamak amacıyla VM tekniklerinin eğitsel ortamlardan gelen verilere uygulanması ile ilgilenirken (Chatti ve diğerleri, 2012; Peña-Ayala, 2014; C. Romero & Ventura, 2007; C. Romero & Ventura, 2010) ÖA aynı amaçla, VM tekniklerine ek olarak istatistik, görselleştirme, sosyal ağ analizi, duygu analizi, söylem analizi vb. yöntemleri de kullanmaktadır (Chatti ve diğerleri, 2012).

1.4. Araştırmanın Amacı ve Önemi

Araştırmanın temel amacı, çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak öğrencilerin akademik performanslarının modellenmesidir. Bu doğrultuda;

- Öğrencilerin dönem sonu performanslarının tahmin etmek amacıyla kullanılacak bir sınıflama modelinin oluşturulması,
- Öğrencilerin dönem sonu akademik performanslarının daha erken haftalarda tahmin edilip edilemeyeceğinin araştırılması,
- Benzer davranış örüntüsü sergileyen farklı öğrenci gruplarının tespit edilmesi ve bu grupların akademik performans açısından karşılaştırılması amaçlanmıştır.

EVM ve ÖA çalışmalarındaki en önemli problemlerden bir tanesi de mevcut öğrenme ortamlarından öğrenci aktiviteleri ile ilgili değişkenlerin çıkartılması konusunda karşılaşılan zorluklardır. Aynı zamanda bu ortamlarda hangi tür verilerin toplanacağı, ne kadar verinin ne kadar süre ile toplanacağı, verilerin nasıl depolanacağı, ne tür ön işleme süreçlerinden geçireceği de araştırılması gereken önemli problemlerdir (Bienkowski ve diğerleri, 2012).

C. Romero ve Ventura (2007)'a ise EVM uygulamalarının yaygınlaşması için ihtiyaç duyulan çalışma alanlarını; VM'de uzman olmayan kişilerin kullanabileceği hazır araçlarının geliştirilmesi; VM araçlarının çevrimiçi eğitim ortamlarına entegre edilmesi; verilerin ve modellerin standartlaştırılması, bu sayede geliştirilen araçların farklı ortamlarda tekrar kullanılmasının sağlanması; geleneksel VM algoritmalarının eğitsel ortamlarda kullanılabilir şekilde uyarlanması olarak sıralamaktadır.

Çalışma kapsamında araştırmacılar tarafından yukarıda verilen problemler de dikkate alınarak öğrencilerin öğrenme süreçleri ile ilişkili verilerin toplanabileceği bir çevrimiçi öğrenme ortamı tasarlanmış ve geliştirilmiştir. Çalışmada kullanılan veriler bu ortamdan elde edilmiştir. Araştırma kapsamında geliştirilen çevrimiçi öğrenme ortamı, EVM ve ÖA alanlarında yapılacak çalışmalarında veri toplama amacıyla kullanılabilir bir ortamdır. Aynı zamanda çalışma kapsamında ele alınan araştırma problemleri bu ortamdan toplanan verilerle öğrencilerin akademik performansları arasındaki ilişkiyi ortaya koyması açısından önemlidir.

1.5. Araştırma Problemleri ve Alt Problemler

Araştırmanın temel problemi, çevrimiçi öğrenme ortamından elde edilen verilerin, VM yöntemleri ile analiz edilerek öğrencilerin akademik performanslarının modellenmesinde kullanılıp kullanılmayacağını araştırılmasıdır. Bu temel problem çerçevesinde belirlenen araştırma problemleri aşağıdaki gibidir.

1. Öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak oluşturulan farklı sınıflama modellerinin derse ilişkin akademik performanslarını (geçti - kaldı) tahmin etme başarısı nasıldır?
 - 1.1. Ön işleme sürecinde kullanılan farklı veri dönüştürme tekniklerinin sınıflama performansları üzerine etkisi nasıldır?
 - 1.2. Ön işleme sürecinde kullanılan farklı "özellik seçme" tekniklerinin sınıflama performansları üzerine etkisi nasıldır?
 - 1.3. Daha az sayıda değişken kullanılarak oluşturulan sınıflama modellerinin performansı tüm değişkenler kullanılarak oluşturulan sınıflama modellerine göre nasıldır?
 - 1.4. Öğrenci performansının tahmin edilmesinde hangi değişkenler daha önemlidir?

- 1.5. Performans metrikleri açısından karşılaştırıldığında dersten kalan öğrencilerin tahmin edilmesinde en iyi performans gösteren sınıflama algoritması/algoritmaları ve kullanılan ön işleme teknikleri nelerdir?
2. Birinci araştırma probleminde en iyi performans gösteren sınıflama algoritması ve ön işleme teknikleri kullanılarak dersten başarısız olacak öğrencilerin daha önceden tahmin edilmesi mümkün müdür?
3. Öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak benzer kullanım profiline sahip öğrenciler gruplanabilir mi?
 - 3.1. X-Ortalamalar kümeleme algoritmasına göre öğrenciler kaç farklı gruba ayrılmaktadır ve küme merkezlerine göre bu gruplar nasıl tanımlanabilir?
 - 3.2. EM kümeleme algoritmasına göre öğrenciler kaç farklı gruba ayrılmaktadır ve küme merkezlerine göre bu gruplar nasıl tanımlanabilir?
 - 3.3. Her iki kümeleme algoritmasına göre elde edilen farklı kümeler bu kümelere yer alan öğrencilerin ortalama ders başarıları ile ilişkili olarak tanımlanabilir mi?
 - 3.4. Her iki kümeleme algoritmasına göre dersi geçen ve kalan öğrencilerin kümelere göre dağılımları incelendiğinde hangi algoritma başarılı ve başarısız öğrencileri ayırt etmede daha iyi bir performans göstermiştir?

1.6. Çalışmanın Sınırlılıkları

Araştırma kapsamında oluşturulan modeller öğrencilerin bilgisayar donanımı dersine ilişkin akademik performansları ve bu ders kapsamında çevrimiçi ortamda gerçekleştirdikleri etkinliklerin verileri kullanılarak elde edilmiştir. Bu nedenle; oluşturulan modeller çalışmaya katılan öğrenciler, ele alınan ders ve dersi veren öğretim görevlisi ile sınırlıdır. Çalışmada akademik performans olarak öğrencilerin bilgisayar donanımı dersi sonundaki geçme notları dikkate alınmıştır. Bilgisayar donanımı dersine ilişkin geçerlik ve güvenilirliği test edilmiş standart ölçme araçları olmadığı için buradaki performans, dersi veren öğretim görevlisinin geliştirdiği ölçme araçları ile sınırlıdır.

Uygulama kapsamında öğrencilerden bir döneme ilişkin veri toplandığı için oluşturulan modellerin dış geçerliliği yeni veri setinde test edilememiştir. Bu sınırlılığı aşmak için çalışmadaki tahmin modellerine ilişkin sonuçlar 5k-çapraz-geçerlilik

yöntemi kullanılarak genelleştirilmiştir. Çapraz geçerlilik yöntemi VM çalışmalarında oluşturulan modellerin sonuçlarının genelleştirilmesi (yeni veri setinde test edilmesinin olanaklı olmadığı durumlarda) ve dolaylı olarak örneklem sayısından kaynaklı sınırlılıkların aşılması konusunda başvurulan bir yöntemdir.

1.7. İşlevsel Tanımlar

Çevrimiçi öğrenme ortamı: Dersin çevrimiçi öğrenme etkinliklerinin gerçekleştiği öğrencilerin ders kaynaklarını takip edip, öğrendikleri kavramlarla ilgili yansıma yazdıkları, tartışmalara katıldıkları aynı zamanda öğrenci etkileşimleri ile ilgili verilerin kayıt edilmesine olanak sağlayan, araştırmacılar tarafından tasarlanan ve geliştirilen dersin web ortamı.

Etkileşim (kullanım, davranış) verisi: Çalışma kapsamında öğrencilerin çevrimiçi öğrenme ortamında yaptıkları her türlü etkileşim (okuma, yazma, etiketleme, gezinme, değerlendirme, tartışma vb.) etkileşim verisi olarak adlandırılmıştır.

Ham veri: Veri tabanında yer alan öğrenci etkileşimleri ile ilgili işlenmemiş veriler.

Değişken (metrik): Toplanan ham verilerinin bir takım matematiksel dönüştürmeler sonucu analizlerde kullanılacak verilere dönüştürülmüş hali.

Öğrenci profili: Etkileşim verileri açısından benzer davranış örüntüsü sergileyen öğrencilerin oluşturduğu grupların, ortak özelliklerine göre tanımlanması.

Eğitim veri seti: VM çalışmalarında model oluşturma sürecinde kullanılan veri seti.

Test veri seti: VM çalışmalarında oluşturulan modellerin performansının test edildiği veri seti.

Performans metrikleri: VM çalışmalarında oluşturulan modellerin performansını yorumlamak ve/veya karşılaştırmak amacıyla kullanılan metrikler.

Akademik performans: Öğrencilerin dönem sonunda Bilgisayar Donanımı dersinden kalma veya geçme durumları.

Performansın modellenmesi: Öğrencilerin akademik performanslarının veri madenciliği yöntemleri kullanılarak istatistiksel açıdan modellenmesi.

2. İLGİLİ ARAŞTIRMALAR

Bu bölümde yer verilen çalışmalar ulusal ve uluslararası akademik bilimsel veri tabanları, dergiler, EVM, ÖA ve kullanıcı modelleme konferanslarında yayınlanan bildiriler araştırma problemleri kapsamında taranarak seçilmiştir.

2.1. Çevrimiçi Öğrenme Ortamlarında VM ve ÖA Uygulamaları

EVM çalışmalarının sınıflandırılması konusunda araştırmacılar tarafından farklı yöntemler izlenmiştir (R. Baker & Siemens, in press; R. Baker & Yacef, 2009; R. S. J. d Baker, 2010; Peña-Ayala, 2014; C. Romero & Ventura, 2007; C. Romero & Ventura, 2010; Cristobal Romero & Ventura, 2013; George Siemens & Ryan S. J. d. Baker, 2012). EVM alanında yapılan ilk tarama çalışmasında C. Romero ve Ventura (2007), 1995 – 2005 yılları arasında yayınlanan 81 çalışmayı incelemiş ve EVM alanında yapılan uygulamaları kullanılan VM yöntemlerine göre sınıflandırmışlardır. Buna göre, çalışmaları istatistik-görselleştirme ve web madenciliği olarak iki kategoriye ayırmışlardır. Web madenciliğini de (1) kümeleme, sınıflama ve aykırı değer analizi; (2) biriktelik kuralları ve sıralı örüntü keşfi; (3) metin madenciliği şeklinde alt kategorilere ayırmışlardır.

R. Baker ve Yacef (2009) 45 araştırmaya referans vererek yaptıkları çalışmada EVM araştırmalarını kullanım amaçlarına göre; çevrimiçi öğrenme ortamlarında kullanılan öğrenci modellerinin iyileştirilmesi, alan modellerinin iyileştirilmesi, öğrencilere pedagojik destek sağlanması ve öğrenme ile ilgili deneysel çalışmalar olarak ayırmışlardır.

C. Romero ve Ventura (2010) 2007 yılında yaptıkları tarama çalışmasına 2010 yılına kadar yapılan 225 çalışmayı daha ekleyerek yeni bir analiz yapmışlardır. Bu çalışmada ise EVM uygulamalarını araştırmalarda seçilen görevlere göre 11 kategoriye ayırmışlardır. Bunlar: (1) verinin analizi ve görselleştirilmesi; (2) geri bildirim sağlama; (3) öneri sistemleri; (4) öğrenci performansını tahmin etme; (5) öğrenci modellemesi; (6) istenmeyen öğrenci davranışlarının belirlenmesi; (7) öğrencileri gruplama; (8) sosyal ağ analizi; (9) kavram haritası geliştirme; (10) kurs içeriğinin yapılandırılması; (11) planlama ve zamanlamadır.

En güncel tarama çalışması ise Peña-Ayala (2014) tarafından 2014 yılının başında yayınlanmıştır. Bu çalışmada EVM konusunda 2010 ve 2013 (ilk çeyreği) yılları

arasında yayınlanan dergi makaleleri, kitap bölümleri, konferans ve çalıştay bildirimlerinden oluşan toplam 240 çalışma kullanılan yöntem, teknik, görev, ortam vb. açılardan incelemiştir. Araştırmacı EVM çalışmalarının uygulama alanlarına göre: öğrenci modelleme, öğrenci davranışlarını modelleme, öğrenci performansını modelleme, değerlendirme, destek ve geri bildirim sağlama ve öğretmene yardımcı olma olarak sıralamıştır.

Bu çalışma kapsamında araştırmaların sınıflandırılması konusunda, kullanılan VM yöntemleri dikkate alınmıştır. Çalışmalar, Tahmin ve Tanımlama amaçlı olmak üzere iki ana başlık altında toplanmış olup sunulan çalışmaların diğer araştırmacılar tarafından önerilen kategorilerle de ilişkileri kurulmuştur.

2.1.1. Tahmin Amaçlı Yapılan Çalışmalar

EVM çalışmalarında en çok kullanılan yöntemlerden bir tanesi tahmindir (R. Baker & Yacef, 2009). Tahmin çalışmalarında amaç bir değişkenin değerini (geleneksel istatistik analizindeki bağımlı değişken) başka bir grup değişkeni kullanarak (bağımsız değişkenler) tahmin etmektir. Tahmin modelleri küçük veri setlerinde oluşturulup geçerliliği test edildikten sonra daha büyük ölçekteki veri setlerine uygulanmaktadır (R. Baker & Siemens, in press). Bir tahmin modelinin oluşturulabilmesi için küçük veri setinde, tahmin edilecek değişkenin değerinin bilinmesi gerekmektedir. Bu sayede oluşturulan tahmin modeli bağımlı değişkenin bilinmediği veri setlerinde, bu değişkenin alacağı değer tahmin edilmesi amacıyla kullanılabilir.

EVM çalışmalarında en çok kullanılan tahmin yöntemleri Sınıflama ve Regresyondur. Tahmin çalışmalarında kullanılan yöntem, bağımlı değişkenin türüne göre (kategorik, sürekli, vb.) değişmektedir. Tahmin edilmek istenen değer ikili (ör: 0 - 1) veya kategorik ise sınıflama, sürekli ise regresyon modelleri kullanılmaktadır.

Eğitsel ortamlarda tahmin yöntemleri, öğrenci başarısının tahmin edilmesi (Cristobal Romero, Espejo, Zafra, Romero, & Ventura, 2010), görev dışı (Ryan S.J.d. Baker, 2007) ya da sistemi kandırmaya yönelik (R. S. Baker ve diğerleri, 2004) öğrenci davranışlarının tespit edilmesi, uzaktan eğitimde dersi bırakma eğilimi olan öğrencilerin önceden tahmin edilmesi (Kotsiantis, 2009), öğrencilerin çevrimiçi öğrenme ortamındaki kaybolma düzeylerinin tahmin edilmesi (Akçapınar ve diğerleri, 2011), e-öğrenme ortamlarında öğrencilerin öğrenme stillerinin otomatik

olarak belirlenmesi (Özpolat & Akar, 2009), öğrencilerin öğrenme ortamındaki gezinim örüntülerinden çalışma belleği kapasitelerinin belirlenmesi (Ting-Wen ve diğerleri, 2013), çevrimiçi sınavlarda hile yapan öğrencilerin tespit edilmesi (Burlak, Hernandez, Ochoa, & Munoz, 2006) vb. amaçlarla kullanılmıştır. R. S. J. d Baker (2010) tahmin yönteminin farklı bir kullanım alanına dikkat çekerek, modeli oluşturan değişkenlerden hangilerinin tahmin etmede daha önemli olduğunun belirlenmesi ve doğrudan veri toplamanın mümkün olmadığı ya da veri toplama sürecinin çalışılmak istenen yapıyı değiştirme riski olduğu durumlarda da kullanılabileceğini belirtmiştir.

2.1.1.1. Akademik Performansın Tahmin Edilmesi ile İlgili Araştırmalar

Akademik performans, öğrencilerin bir öğrenme etkinliği sonucu yapılan değerlendirmeden aldıkları puan olarak tanımlanabilir. Bu puan kısa süreli bir öğrenme etkinliğine (ör: bir öğrenme nesnesi) ait olabileceği gibi, bir derse, bir döneme ya da eğitim sürecinin tamamına ilişkin olabilir. Araştırmanın amacına göre sürekli bir değişken olabileceği gibi kategorik (geçti, kaldı, düşük, yüksek, v.b.) olarak da kullanılabilir. Akademik performansı tahmin edilmesi, EVM araştırmacılarının üzerine çalıştığı en eski ve en popüler araştırma konusudur (Cristobal Romero & Ventura, 2013).

Lopez ve diğerleri (2012) öğrencilerin Moodle öğrenme yönetim sistemindeki forum kullanım verilerinin ders başarısının önemli bir göstergesi olup olmadığını test etmişlerdir. Bir diğer araştırma problemleri ise sınıf değişkeninin (ders başarısı) bilinmediği durumlarda kümeleme algoritmaları ile aynı sonuca ulaşıp ulaşılmayacağını test edilmesi yönündedir. Araştırmacılar, öğrencilerin forum katılımları ile ilgili olarak 8 değişken belirlemişlerdir ve bu değişkenleri kullanarak öğrencilerin ders başarılarını (geçti – kaldı) tahmin etmeye çalışmışlardır. Bu amaçla birçok sınıflama algoritmasını karşılaştırarak doğru sınıflama oranına göre performanslarını belirlemişlerdir. Araştırma bulgularında 8 değişken kullanarak yapılan sınıflamada en iyi sınıflama oranı %87,7 ile BayesNet algoritması ile elde edilirken yapılan nitelik (özellik) seçme işlemi sonucu belirlenen 6 değişken ile yapılan sınıflamada NaiveBaseySimple algoritması %89,4 doğru sınıflama oranına ulaşılmıştır. Kümeleme algoritmalarında ise EM algoritması ile %89,4 oranında doğru sınıflama sağlanmıştır. Araştırma sonuçları öğrencilerin Moodle ortamındaki ders forumuna katılımlarının ders başarılarının tahmin edilmesinde iyi bir değişken olduğunu göstermiştir. Bir diğer önemli sonuç ise kümeleme analizi ile sınıflama

analizindeki aynı başarının sağlanmasıdır. Bu sonuç başarı puanlarının bilinmediği durumda kümeleme analizleri ile benzer bir sınıflamanın yapılabileceğini göstermesi açısından önemlidir.

Moodle ortamında yapılan diğer bir araştırmada C. Romero, Ventura, Hervás, ve Gonzales (2008) öğrencilerin ders başarısını tahmin etmek amacıyla bir grup sınıflama algoritmasının performansı karşılaştırılmıştır. Araştırmacılar öğrencilerin forum katılım verilerine ek olarak Moodle kullanımlarını yansıttığını düşündükleri 9 farklı değişken seçmişlerdir. Öğrencilerin ders başarısı; geçti, kaldı, iyi, mükemmel şeklinde dördü kategorik bir yapıya dönüştürülerek kullanılmıştır. Algoritma performansları 10k çapraz geçerlilik yöntemi ile genelleştirilerek değerlendirilmiştir. Sınıflama performanslarının karşılaştırılmasında sınıf değişkeninin kategorilere dengeli bir şekilde dağılmadığı durumlarda doğru sınıflama oranının doğrudan kullanılması uygun olmayacağı için bu dengesizliği dikkate alan bir doğru sınıflama oranı hesaplanarak karşılaştırmalar yapılmıştır. Araştırma bulguları incelendiğinde sınıflama algoritmalarının doğru sınıflama oranlarının %60 - %70 arasında değiştiği görülmektedir. Araştırmacılar, doğru sınıflama oranının %70'in üzerine çıkmamasının öğrencilerin tamamının katılmadığı değişkenlerden kaynaklanabileceğini belirtmişlerdir (ör: öğrencilerin %30'u forumlara veya quizlere katılmamış); fakat bu değişkenler öğrencilerin gerçek kullanımını yansıttığı için sınıflama oranlarını olumsuz etkilese de araştırmadan çıkartılmamıştır. Araştırma sonuçları sınıflama sonucu elde edilen kuralların ve bilgilerin öğretmenlerin yeni öğrencileri sınıflamasında ya da karar verme süreçlerinde kullanılabileceğini belirtmişlerdir. Araştırmacılar aynı zamanda öğretmenlerin bu bilgileri sorun yaşayan öğrencilere zamanında müdahale etmek amacıyla kullanabileceklerini belirtmişlerdir.

Osmanbegović ve Suljić (2012) öğrencilerin demografik bilgilerine ek olarak üniversite öncesi başarıları ile ilgili verileri de kullanarak üniversitedeki başarılarını tahmin etmeye yönelik bir sınıflama modeli oluşturmak için üç farklı sınıflama algoritmasının performansını karşılaştırmışlardır. Çalışma 257 üniversite birinci sınıf öğrencisi ile yürütülmüştür. Bir tahmin modeli oluşturmak amacıyla kullanılan veriler dönem başında anket yardımıyla toplanmıştır. Çalışmada kullanılan değişkenler cinsiyet, İnternet erişimleri, aileleri ile ilgili demografik bilgilerden ve lise başarı puanları, üniversiteye giriş puanları ve daha önceden kısmen ya da tam zamanlı

burs alıp almadıkları gibi geçmiş başarıları ile ilişkili 12 değişkenden oluşmaktadır. Başarı değişkeni olarak da öğrencilerin “İşletme Enformatiği” dersindeki geçme notları alınmıştır. Öğrencilerin geçme notları A, B, C, D, E, F şeklinde ifade edilmektedir; fakat araştırmacılar, öğrencilerin bu not gruplarına dengeli bir şekilde dağılmamasının (A = %0,39 B = %7,78 C = %12,06 D = %27,63 E = %28,02 F = %24,12) hata oranını artıracaklarını belirterek başarı değişkenini geçti (%75,88) ve kaldı (%24,12) şeklinde dönüştürerek kullanmışlardır. Sınıflama amacıyla C4.5, Multilayer Perceptron (MLP) ve Naive Bayes (NB) algoritmaları kullanılmıştır ve sonuçların genelleştirilmesi için 3k çapraz geçerlilik yöntemi kullanılmıştır. Kullanılan algoritmaların performansı ise doğru sınıflama oranı ve elde edilen sonuçların kullanım kolaylıklarına göre değerlendirilmiştir. En iyi sınıflama oranı %76,65 ile NB ikinci olarak da %73,93 doğru sınıflama oranı ile C4.5 algoritması ile elde edilmiştir. Araştırmacılar, karar ağacı kullanan algoritmaların (ör: C4.5) sonuçlarının öğrenci ve öğretmenler tarafından anlaşılmasının daha kolay olduğunu belirtmişlerdir. Son olarak kullanılan yöntemin öğrenci performansını artırmada ve doğru zamanda yapılacak müdahalelerle başarısızlık oranlarını azaltmada öğrenci ve öğretmenlere yardımcı olarak kullanılabileceğini belirtmişlerdir.

Öğrenme yönetim sisteminden elde edilen etkileşim verileri kullanılarak öğrenci performansını tahmin etmeye yönelik bir diğer çalışmada Macfadyen ve Dawson (2010) öğrenme yönetim sisteminden elde edilen verileri kullanarak öğrencilerin geçme notlarıyla istatistiksel olarak anlamlı ilişki gösteren 15 adet değişken belirlemişlerdir. Bu değişkenleri kullanarak öğrencilerin akademik performanslarını tahmin etmek amacıyla da lojistik regresyon yöntemini kullanmışlardır. Araştırmacılar oluşturulan sınıflama modelinin dersten kalan öğrencilerin %81'ini doğru olarak sınıfladığını belirtmişlerdir. Tartışma ortamına gönderilen mesaj sayısı, gönderilen mail sayısı ve tamamlanan ödev sayısı ise dönem sonu notları tahmin etmede önemli değişkenler olarak belirtilmiştir. Yapılan bu çalışma öğrenci etkileşimleri ile ilgili öğrenme yönetim sisteminden elde edilen verilerden pedagojik olarak anlamlı bilgiler üretilebileceğini göstermiştir.

Akademik performansı tahmin etmeye yönelik çalışmalar genellikle dönem sonunda toplanan verilerle yapılmaktadır. Buradan elde edilecek sonuçlar şüphesiz ki başarısızlığa neden olan değişkenlerin belirlenmesi açısından önemlidir fakat başarısızlığı önlemeye yönelik çok fazla bilgi sağlayamamaktadırlar. Diğer taraftan

çevrimiçi öğrenme ortamları, öğrenci etkileşimleri ile ilgili olarak sürekli ve zamana bağlı veriler toplanmasına olanak sağlamaktadır. Bu verilerin daha erken zamanlarda analiz edilmesiyle oluşturulacak tahmin modelleri olası başarısızlıkları önlemede etkili olacaktır. Bu tür sistemler alan yazınında erken uyarı sistemleri olarak geçmektedir.

2.1.1.2. Erken Uyarı Sistemleri ile İlgili Araştırmalar

Erken uyarı sistemleri, akademik performansı tahmin etme ile ilgili çalışmaların bir ileri adımı olarak düşünülebilir. Burada amaç akademik performansın çevrimiçi öğrenme ortamlarından gelen değişkenlerle daha erken bir zamanda tahmin edilmesidir. Bu sayede risk altındaki öğrenciler belirlenebilir ve öğretmenlere bu öğrencilere yardımcı olacak iyileştirmeler yapmaları için yardımcı olunabilir (Campbell, DeBlois, & Oblinger, 2007).

Bu konuda yapılan çok fazla çalışma olmamakla birlikte Purdue Üniversitesinde hayata geçirilen Sinyal projesi başarılı bir erken uyarı sistemi örneği olarak literatüre geçmiştir (K. E. Arnold, 2010). Araştırmacılar Moodle öğrenme yönetim sisteminde çalışan Sinyal isimli bir eklenti tasarlanmıştır. Bu eklenti öğretmen tarafından çalıştırıldığı zaman, tahmin algoritmaları yardımıyla öğrenci etkileşimlerini analiz ederek her bir öğrenciye bir risk göstergesi atamaktadır. Trafik ışığı şeklinde yapılandırılan risk göstergesi öğrenciye dersteki durumu ile ilgili olarak kırmızı (dersten başarısız olma ihtimali yüksek); sarı (başarısız olma ihtimali orta düzeyde) veya yeşil (başarılı olma ihtimali yüksek) şeklinde dönüt sağlamaktadır. Üniversitede 2 yıl süreyle yapılan pilot uygulama sonucunda Sinyal sisteminin kullanıldığı derslerde öğrenci başarısında önemli bir artış görülmüştür. Araştırma sonuçlarına göre 220 öğrencinin katıldığı bir dersin ilk haftasında yapılan analizde 45 öğrenci kırmızı (yüksek risk), 32 öğrenci sarı (orta risk) grupta yer alırken ilerleyen haftalarda kırmızı gruptaki öğrencilerin %55'inin sarı gruba (C düzeyi) %24,4'ünün yeşil gruba (A veya B düzeyi) geçtiği, %10,6'sının ise kırmızı grupta kaldığı görülmüştür. Sarı grupta yer alan öğrencilerin ise %31'inin konumunu koruduğu, %69'unun ise yeşil gruba geçtiği görülmüştür.

Hu, Lo, ve Shih (2014) öğrenme yönetim sistemi üzerinden çevrimiçi olarak işlenen bir derste öğrencilerin sistemindeki etkileşimleri ile ilgili verileri veri madenciliği yöntemleri ile analiz ederek öğrencilerin derste sergileyecekleri öğrenme

performansını tahmin edecek bir erken uyarı sistemi geliştirmeyi amaçlamışlardır. Bu amaçla 4, 8 ve 13. haftalara ait verilerle öğrencilerin dersten kalma ve geçme durumlarını tahmin edecek farklı sınıflama modelleri oluşturmuşlardır. Araştırmacılar geliştirdikleri erken uyarı sisteminin, öğrenciler öğrenme ortamını bir süre kullandıktan sonra o anki öğrenme performanslarını etkili olarak tahmin ettiğini belirtmişlerdir. Araştırmacıların ortaya koyduğu deneysel veriler veri madenciliği tekniklerinin erken uyarı sistemi geliştirmede etkili olduğunu göstermiştir. Öğrenme yönetim sisteminden elde edilen zamana bağlı değişkenler ise öğrenci performansını tahmin etmede önemli değişkenler olarak belirtilmiştir.

2.1.2. Tanımlama Amaçlı Yapılan Çalışmalar

VM çalışmalarında en sık kullanılan tanımlayıcı yöntemler kümeleme ve birliktelik kurallarıdır. Kümeleme yönteminde tahmin yöntemlerinden farklı olarak değeri tahmin edilecek bir değişken yoktur. Amaç belirli açılardan benzerlik gösteren elemanları (ör: öğrenciler) gruplayarak kümelere ayırmaktır. Burada yapılan iş temelde veri ile ilgili bir öngörü olmadan veriyi sınıflara ayırmaktır. Bu nedenle kümeleme analizi bazı kaynaklarda danışmansız (unsupervised) sınıflandırma olarak da anılmaktadır (C. Romero & Ventura, 2007). Kümeleme analizlerinde benzer özellikteki elemanları belirlemek için genellikle uzaklık ölçümlerinden yararlanılmaktadır (Cristobal Romero & Ventura, 2013). Verinin kaç farklı kümeye ayrılacağı ile ilgili bir öngörü varsa küme sayısı olarak bu belirtilebilir ya da seçilen algoritma rastgele küme sayıları üreterek verinin ayrılacağı en uygun küme sayısını bulmaya çalışır (R. S. J. d Baker, 2010). Doğru sayıda kümeye ayrılmış bir veride aynı kümede yer alan verilerin birbiri ile daha çok diğer kümelerdeki verilerle daha az benzerlik göstermesi beklenir (R. Baker & Siemens, in press). Eğitsel ortamlarda kümeleme yöntemleri, öğrencilerin (Beal ve diğerleri, 2006) ya da öğrenci davranışlarının gruplanması (Saleema Amershi & Conati, 2007) amacıyla kullanılmaktadır. Alfredo, Félix, ve Àngela (2010) Kümeleme analizinin e-öğrenme bağlamında en çok, öğrenme materyallerinin (birbirlerine benzerliklerine göre) ve öğrencilerin gruplanması (gezinim ve öğrenme davranışlarına göre) amacıyla kullanıldığını belirtmiştir.

Birliktelik kuralları ise büyük veri setlerinde değişkenler arasındaki ilişkileri ortaya çıkartmak için kullanılan bir diğer popüler (tanımlayıcı) VM yöntemidir (Ceglar & Roddick, 2006). Birliktelik kurallarının temeli alışveriş sepeti analizine

dayanmaktadır (Agathe & Kalina, 2010). Amacı veri tabanında bir arada tekrar eden değişkenlerin oluşturduğu örüntüleri ortaya koymaktır. Birliktelik kuralları öğrenme ortamlarında öğrencilerin davranış örüntülerindeki ilişkileri tanımlamak için ya da öğrenme gücünü çeken öğrencilerin veya sıklıkla bir arada yapılan hataların tespit edilmesinde kullanılmıştır (Cristobal Romero & Ventura, 2013). Bunun dışında öğrencilere ihtiyaç duyacakları öğrenme materyallerinin önerilmesi (Jie, 2004); öğretmenlere geri bildirim sağlamak amacıyla öğrencilerin kullanım loglarındaki ilişkilerin ortaya çıkartılması (García ve diğerleri, 2011) gibi amaçlarla da kullanılmaktadır.

2.1.2.1. Benzer Öğrenci Gruplarının Belirlenmesi ile İlgili Araştırmalar

Kişisel özellikleri, öğrenme tercihleri, bireysel farklılıkları, kullanım verileri vb. açılardan benzer öğrenci gruplarının belirlenmesi EVM araştırmacılarının üzerinde çalıştığı bir diğer önemli konudur. Kümeleme ve sınıflama bu amaçla araştırmacılar tarafından en çok tercih edilen VM yöntemleridir (Peña-Ayala, 2014).

Cristóbal Romero ve diğerleri (2008) öğrencilerin Moodle kullanım verilerine göre profillerini belirlemek amacıyla kümeleme analizinden yararlanmışlardır. Öğrenci verilerini k-Ortalamalar (k-Means) kümeleme yöntemi ile analiz etmişlerdir ve küme sayısını ön tanımlı üç olarak tanımlamışlardır. Analiz sonucu elde edilen kümeleri nitelendirmek içinse küme merkezlerini (ortalama / mod) ve standart sapmalarını dikkate almışlardır. Buna göre; Küme 1’de yer alan öğrencilerin hiçbir ödevi yapmadığını, mesaj okuma ve quizlere katılım oranlarının çok düşük olduğunu ve ortamda geçirdikleri sürenin de düşük olduğunu belirtmişlerdir. Küme 2’de yer alan öğrencilerin ise daha fazla mesaj gönderdiklerini, daha fazla ödev yaptıklarını, daha fazla mesaj okuduklarını, quizlerde daha başarılı olduklarını ve ortamda çok daha fazla zaman harcadıklarını belirtmişlerdir. Küme 3’de yer alan öğrencilerin değerlerinin ise Küme 2’de yer alanlardan daha az Küme 1’de yer alan öğrencilerden ise daha fazla olduğu belirtilmiştir. Araştırma örneklemini oluşturan 75 öğrencinin bu üç kümeye dengeli bir şekilde dağıldığı görülmektedir (sırasıyla 24, 22 ve 29). Araştırmacılar öğretmenlerin bu bilgileri öğrencileri çok aktif (küme 2), aktif (küme 3), aktif olmayan (küme 1) şeklinde gruplara ayırmak için ve bu bilgileri de işbirlikli çalışmalarda grupları oluştururken kullanabileceklerini belirtmişlerdir.

Beal ve diğeri (2006), 83 lise öğrencisi ile yaptıkları çalışmada öğrencilerin matematik dersine karşı motivasyonları ile zeki öğrenme sistemindeki kullanım örüntüleri arasındaki ilişkiyi araştırmışlardır. Bu amaçla ilk olarak öğrencilere matematik motivasyon ölçeği uygulamışlar ve öğrencileri ölçeğin alt boyutlarından aldıkları puanlara göre hiyerarşik kümeleme analizi yardımı ile üç gruba ayırmışlardır. Aynı kümede yer alan öğrencilerin ölçeğin alt boyutlarından aldıkları ortalama puanlarına göre grup profili çıkartmışlardır. Buna göre Grup 1 (n=50)'de yer alan öğrencileri matematiğin gelecek yaşamlarında önemli olabileceğini düşünen fakat diğer derslerden daha az seven, matematiğe karşı ortalama motivasyona sahip olarak; Grup 2 (n = 21)'de yer alan öğrencileri, matematiği sevmediğini belirten ve matematik konusunda çok da yeteneklerinin olmadığını düşünen öğrenciler olarak, Grup 3 (n=12)'de yer alan öğrencileri ise yetenekleri konusunda inançları yüksek olan, matematiği diğer derslerden daha çok sevdiğini belirten ve matematik öğrenmenin önemli olduğunu düşünen öğrenciler olarak tanımlamışlardır. Araştırmacılar, öğrencilerin zeki öğrenme sistemindeki kullanım örüntülerini de bağımsız-a, bağımsız-b, tahmin etme, yardım suiistimali ve öğrenme şeklinde beş grupta sınıflamışlardır. Daha sonra belirlenen gruplar ile kullanım stratejilerini eşleştirmişlerdir ve Grup 1'de yer alan öğrencilerin daha çok, problemlerin cevabını tahmin ederek bulmaya çalıştıklarını, Grup 2'de yer alan öğrencilerin problemlerin çözümü ile ilgili öğrenme amaçlı olarak yardım materyallerinden yararlandıklarını Grup 3'de yer alan öğrencilerin ise yardım almadan problemleri doğru olarak çözdüklerini belirlemişlerdir. Bu bulgular öğrenci profil verilerinin pedagojik modellerin oluşturulmasında ve öğrenme stratejilerinin tahmin edilmesindeki potansiyel kullanımına dikkat çekmektedir.

S. Amershi ve Conati (2009) çevrimdışı ve çevrimiçi olmak üzere iki bölüme ayırdıkları çalışmalarında yarı otomatik bir uyarlanabilir öğrenme ortamı tasarlamışlardır. Çevrimdışı bölümünde öğrencilerin öğrenme ortamındaki etkileşim davranışlarını temsil eden değişkenleri kullanarak kümeleme algoritmaları yardımıyla benzer özellikteki öğrencileri gruplamışlardır. Araştırmacılar öğrencilerin etkileşim verilerine göre iki kümeye ayrıldığını belirtmişlerdir ve yaptıkları analizler sonucu bu kümeleri Yüksek Öğrenme ve Düşük Öğrenme şeklinde etiketlemişlerdir. Araştırmanın çevrimiçi bölümünde ise araştırmacılar, kümeleme analizi sonucu elde edilen bu değişkenleri kullanılarak bir tahmin modeli oluşturmuşlardır ve oluşturulan

bu tahmin modelini ortama yeni giren öğrencilerin öğrenme çıktıları (yüksek, düşük) tahmin etmek amacıyla kullanmışlardır. Araştırmacılar geliştirdikleri modeli iki farklı öğrenme uygulamasında yeni öğrencilerle test etmiş ve geliştirilen modelin öğrenci davranışlarını sırasıyla %88,3 ve %86,3 oranlarında doğru olarak sınıfladığını belirtmişlerdir. Araştırmacılar elde ettikleri bulguların öğretmenlerin çevrimiçi ortamların pedagojik faydaları konusundaki farkındalıklarının artırılmasında kullanılabileceğini ve yazılım geliştiricilere bu ortamların iyileştirilmesi konusunda yardımcı olabileceğini belirtmişlerdir.

Xu (2011), kullanıcı grubu öğretmenlerden oluşan bir elektronik kütüphane sitesindeki farklı kullanıcıların çevrimiçi ortamdaki davranışlarını daha iyi anlayabilmek amacıyla kümeleme analizinden yararlanmışır. Elektronik kütüphane, sağladığı çevrimiçi araçlar sayesinde öğretmenlere öğretim materyalleri ve aktiviteler oluşturmak, yeniden düzenlemek veya paylaşmak için bir ortam sağlamaktadır. Kütüphane sitesinin veri tabanı, 2002 – 2010 yılları arasında siteye kayıtlı 5500 öğretmen tarafından oluşturulan 12,000'den fazla proje ve öğretmenler tarafından veri tabanına kayıt edilen 54,000'den fazla çevrimiçi kaynağa bağlantıyı içermektedir. Araştırma kapsamında siteye 2009 yılında kayıt olan 1164 öğretmenden tek seferlik giriş yapanlar ve hiçbir proje oluşturmayanlar çıkartıldıktan sonra kalan 757 öğretmenin verisi kullanılmışır. Araştırmacı analize başlamadan önce ilk olarak bu veriden öğretmenlerin kullanım davranışlarını yansıtacak değişkenleri çıkartmışır. Bu amaçla öğretmenlerin ortamda yaptığı temel aktiviteleri kaynak toplama, proje hazırlama/kullanma ve gezinim olarak üçe ayırmışır. Örtük Sınıf Analizi (Latent Class Analysis) yöntemi ile gerçekleştirdiği kümeleme analizi sonucu ise benzer özellik gösteren öğretmenlerin oluşturduğu yedi farklı kullanıcı tipi tanımlamışır.

Belirli değişkenler açısından benzer davranış örüntüsü sergileyen öğrencilerin/kullanıcıların gruplanarak profillerinin belirlenmesi, uyarlanabilir sistemlere ve öğretmenlere yönelik önemli bilgiler sağlayacaktır. Bu sayede uyarlanabilir öğrenme ortamları oluşturulan öğrenci profiline göre uyarlama yapabilecek öğretmenler ise bu bilgileri grubu daha iyi tanımak ve bireysel geri bildirimler vermek amacıyla kullanabileceklerdir.

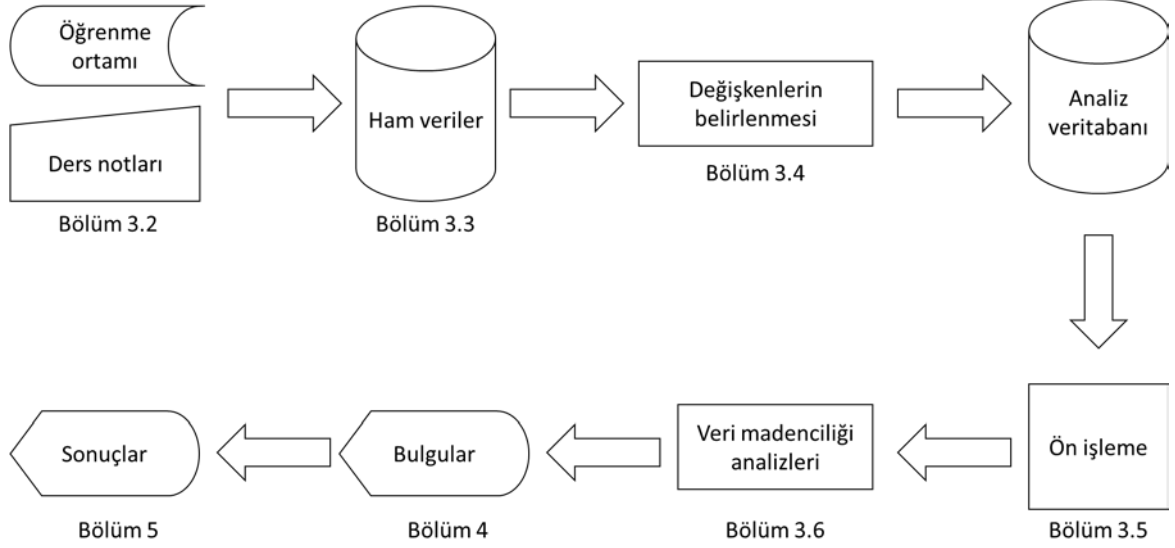
2.1.3. Genel deęerlendirme

Yukarıda verilen alıřmalarda da grldę gibi EVM ve A uygulamaları evrimii ęrenme ortamlarında klasik istatistiksel yntemlerin yetersiz kaldığı ve/veya dřk performans gsterdiği nemli problemlerin zmne ynelik bir alternatif sunmaktadır. ęrenme ortamlarında VM analizlerinin kullanımının yaygınlařmasında arařtırmacıların kolaylıkla eriřebileceęi aık kaynak kodlu VM yazılımlarının geliřtirilmesi ve masast bilgisayarların yksek iřlem kapasitesi gerektiren bu tr analizleri yapabilecek dzeye ulařmasının payı byktr.

Bu alıřma kapsamında da ęrencilerin Bilgisayar Donanım dersinde sergiledikleri akademik performansın evrimii ęrenme ortamındaki veriler iřığında VM yntemleri ile modellenmesi amalanmıřtır. Bu amala ilerleyen blmde detaylı olarak anlatılan ve  ařamadan oluřan bir sre izlenmiřtir. Birinci ařamada farklı sınıflama algoritmaları ve n iřleme yntemleri kullanılarak ęrencilerin dnem sonu performanslarını en iyi tahmin edecek algoritma ve deęiřkenlerin belirlenmesi amalanmıřtır. İkinci ařamada bu deęiřkenler ve seilen algoritma kullanılarak ęrencilerin akademik performanslarının daha nceki haftalarda tahmin edilip edilemeyeceęi arařtırılmıřtır. nc ařamada ise seilen deęiřkenler aısından benzer davranıř rnts sergileyen ęrenci gruplarının belirlenmesi ve bu grupların akademik performanslarının karřılařtırılması amalanmıřtır.

3. YÖNTEM

Araştırmada izlenen süreç veri tabanlarından bilgi keşfi ile ilgili çalışmalarda izlenen süreçlere benzerlik göstermektedir. Şekil 3.1’de özetlenen bu sürecin her bir bileşeni aşağıda detaylı olarak ele alınmıştır.



Şekil 3.1. Araştırma Süreç Modeli

3.1. Araştırma Grubu

Çalışmaya Hacettepe Üniversitesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü 2013 – 2014 Güz Dönemi BTO 213 Bilgisayar Donanımı dersine kayıtlı 76 öğrenci katılmıştır. Katılımcıların demografik özellikleri ile ilgili tanımlayıcı istatistikler Çizelge 3.1’de verilmiştir.

Çizelge 3.1: Katılımcıların Demografik Bilgileri

<i>Cinsiyet</i>	<i>f</i>	<i>%</i>
<i>Erkek</i>	35	46,05
<i>Kadın</i>	41	53,95
<i>Toplam</i>	76	100,00

3.2. Veri Kaynakları

Analizlerde araştırma problemleri doğrultusunda iki farklı veri kaynağı kullanılmıştır. Birinci ve ikinci araştırma problemi için öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri ve dönem sonu başarı puanları kullanılırken üçüncü problemde benzer kullanım davranışı sergileyen öğrencileri gruplamak amacıyla sadece etkileşim verileri kullanılmıştır.

3.2.1. Çevrimiçi Öğrenme Ortamı

Yapılan incelemeler sonucu, mevcut çevrimiçi öğrenme ortamlarından araştırma problemlerine uygun verilerin elde edilemeyeceği tespit edilmiştir. Bu nedenle araştırmacılar tarafından yeni bir öğrenme ortamı tasarlanmıştır. Geliştirilen bu ortam sayesinde öğrencilerin çevrimiçi öğrenme süreçleri ile ilgili ayrıntılı verilerin toplanması mümkün olmuştur. Aşağıda çevrimiçi öğrenme ortamının bileşenleri, öğrencilerin bu ortamda gerçekleştirebileceği roller ve bu ortamdan elde edilen veriler ayrıntılı olarak ele alınmıştır.

3.2.1.1. Çevrimiçi Öğrenme Ortamının Bileşenleri

İleti bölümü: Öğrenciler “İletiler” sayfasını kullanarak kendi yazdıkları iletileri toplu olarak görebilir veya diğer öğrenciler tarafından yazılan iletileri okuyabilir, yorum yazabilir, değerlendirebilir. İletiler, yazılma zamanına göre ters sıralı olarak listelenmektedir (Şekil 3.2.).

The screenshot shows the 'İletiler' (Messages) page in the LCMS system. The page is displayed in a browser window with the URL 'bote2.hacettepe.edu.tr/LCMS/'. The page has a dark blue header with navigation tabs: 'İletiler', 'Tartışma', 'Kaynaklar', 'Bildirimler', 'Duyurular', and 'Kavramlar'. The user's name 'Gökhan Akçapınar' and 'Güvenli çıkış' (Safe exit) are visible in the top right corner.

The main content area is divided into two columns. The left column contains a 'Yeni ileti' (New message) section with a search bar and a list of categories: 02: Bilgisayarın Tarihi, 04: Çalışma Mantığı, 06: Anakart, 07: İşlemci, 08: Bellek, 09: Sabit Disk, 10: Diğer depolama birim., 11: Donanım kartları, 15: Genel. Below the categories is a 'Sadece kendi iletilerimi göster' (Show only my messages) checkbox and a 'Sayfa gezinimi' (Page navigation) section with left and right arrows and a '0' in a red box.

The right column displays a list of messages. Each message row includes a 'puan' (score), 'cevap' (reply), and 'görüntüleme' (views) column, followed by the message title and content. The messages are sorted by date, with the most recent at the top. The messages are:

- 0 puan, 0 cevap, 10 görüntüleme: RAS, RAS-TO-RAS, CAS. RAS:Aranan bilginin kayıtlı olduğu satıra ulaşırken yaşanan gecikmedir...
- 0 puan, 0 cevap, 4 görüntüleme: HDD, SSD. Teknolojik gelişmelerle birlikte sabit diskler, bütünlük flash belle...
- 0 puan, 0 cevap, 6 görüntüleme: Genişleme yuvaları (External buses). - Anakarta harici donanım elemanlarını bağlamak için kullanılan veriyo... Etiketler: Genişleme Yuvaları (External Buses)
- 1 puan, 1 cevap, 17 görüntüleme: SCSI Kartı. SCSI, Small Computer System Interface'in (Küçük Bilgisayar Sistem Arab... Etiketler: SCSI
- 0 puan, 1 cevap, 4 görüntüleme: TV-Kartı. TV kartı; Bilgisayarlarda TV izlemeye olanak sağlayan bilgisayar parça... Etiketler: TV Kartı
- 0 puan, 1 cevap, 3 görüntüleme: Fax-Modem Kartları. Bilgisayarların telefon kabloları üzerinden lokal yada genel ağlara ba... Etiketler: Fax-Modem Kartı
- 1 puan, 1 cevap, 6 görüntüleme: RJ-45 Konnektörü. Üzerinde 8 adet metal bağlantı noktası bulunan konnektör. Bilgisayarla... Etiketler: RJ-45
- 0 puan, 1 cevap, 6 görüntüleme: Ethernet Kartı. Ethernet kartı, bilgisayar ağlarında bilgisayarlar ağ arasında iletişim... Etiketler: Ethernet Kartı
- 0 puan, 1 cevap, 2 görüntüleme: Mikrofon Girişi. Joystick/MIDI port : Joystick ve MIDI aygıtlarının bağlanmasını sağlar... Etiketler: Mikrofon Girişi

The footer of the page shows 'H . Ü . Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü' and 'v2.1.8'.

Şekil 3.2. Çevrimiçi Öğrenme Ortamı, İleti Bölümü

Öğrenciler yazılan iletiler arasında gezinmek için üç farklı yöntem kullanabilmektedir. Bunlar; sol tarafta bulunan ve haftalara göre işlenen konuları içeren listeden seçim yapmak, arama kutusunu kullanmak ve son olarak “Bildirim” sayfasını kullanarak son yazılan iletiler arasında gezinmek şeklinde sıralanabilir. Sistem öğrencilerin diğer öğrenciler tarafından yazılan ileti ve yorumları “katılıyorum” ya da “katılmıyorum” şeklinde değerlendirme yapmasına da olanak sağlamaktadır.

Öğrenciler aynı zamanda Şekil 3.3.’de örneği görülen “Yeni ileti” ekranını kullanarak, sınıf ortamında öğrendikleri bilgiler ve kavramlar ile ilgili yansımalar yazabilirler. Yeni bir ileti yazmak için öğrenci ilk olarak yansıma yazmak istediği konuyu ön tanımlı listeden seçmelidir. Daha sonra yazacağı içeriğe göre bir başlık seçmelidir.

H.Ü. BÖTE Ders İçerik Yön x
bote2.hacettepe.edu.tr/LCMS/
İletiler Tartışma Kaynaklar Bildirimler Duyurular Kavramlar Gökhan Akçapınar Güvenli çıkış

Yeni İleti

0 puan 0 cevap 10 görüntüleme RAS, RAS-TO-RAS, CAS
RAS:Aranan bilginin kayıtlı olduğu satıra ulaşırken yaşanan gecikmedir...

N_PostCW

Konu: 08: Bellek Başlık: RAM Bellek

İçerik (305 / 500)

Rastgele Erişimli Bellek olarak da bilinen RAM bellek, Von Neumann mimarisindeki merkezi belleğe karşılık gelmektedir. Görevi, sabit disk ile işlemci arasında tampon bölge oluşturmaktır. Yapısı, işlemcinin üzerinde çalışacağı verileri geçici olarak sakladığı elektronik bellek hücrelerinden oluşmaktadır.

Etiketler

Rastgele Erişimli Bellek x RAM x Von Neumann mimarisi x

Kaydet

puan cevap görüntüleme Etiketler: Ethernet Kartı

0 puan 1 cevap 2 görüntüleme Mikروفon Girişi
Joystick/MIDI port : Joystick ve MIDI aygıtlarının bağlanmasını sağlar...

Etiketler: Mikروفon Girişi

H . Ü . Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü v2.1.8

Şekil 3.3. Yeni İleti Yazma Sayfası

Sistemde yazılan içeriği etiketlemek için de bir bölüm bulunmaktadır. Öğrenci yazdığı içeriği etiketlemek için diğer öğrenciler tarafından eklenen etiketleri kullanabileceği gibi yeni etiketler de oluşturabilmektedir. Öğrencilerin yazdıkları

içerikleri etiketlemek için kullanabilecekleri etiket sayısı 10 ile sınırlandırılmıştır. Yansımaların kısa ve öz olması istendiği için içerik alanı 500 karakter ile sınırlandırılmıştır ve öğrencilerin başka bir kaynaktan doğrudan alıntı yapmasının önüne geçebilmek için de bütün sistemde metin girişi yapılan alanlarda kopyala – yapıştır özelliği devre dışı bırakılmıştır.

Ders Kaynakları: Ders kaynakları ile ilgili sayfada haftalık konulara göre derste işlenecek konu ile ilgili sunumlar paylaşılmaktadır. Öğrenciler sanal ortamı kullanarak ders kaynaklarını takip edebilmektedirler (Şekil 3.4.).

The screenshot shows a web browser window with the URL bote2.hacettepe.edu.tr/LCMS/. The page is titled 'Okuma / Yazma' and contains the following text:

Hacettepe Üniversitesi 2013 © 10

Okuma / Yazma

- Plakaların yüzeyi aşağıdaki birimlerden oluşur.
 - Sektör (Sector)
 - Diskten okunabilen ya da yazılabilen en küçük veri birimidir (Genellikle 512, 1024 ya da 2048byte büyüklüğündedir).
 - Küme (Cluster)
 - İşletim sisteminin dosya ya da dizin kaydetmek için kullandığı en küçük birimdir. Birden fazla sektörün bir araya gelmesi ile oluşur (Ör: FAT16 dosya sisteminin küme büyüklüğü 32K'dır). Dosyaların en iyi performans için disk yüzeyinde birbirini takip eden kümelere kayıt edilmesi gerekir. Bir dosyanın diskin farklı bölümlerinde birden fazla kümede yer alması parçalanma (fragmentation) olarak adlandırılır.
 - İz (Track)
 - Sektörlerin bir araya gelerek oluşturduğu halkaya iz denir.
 - Silindir (Cylinder)
 - Farklı plakalardaki izlerin bir araya gelerek oluşturduğu yapıya silindir denir.

The diagrams illustrate the following concepts:

- Sektör (Sector):** A single wedge-shaped unit on a disk platter.
- Küme (Cluster):** A group of sectors on a disk platter.
- Plaka (Platter):** A single circular disk.
- İz (Track):** A circular path on a disk platter.
- Silindir (Cylinder):** A vertical stack of tracks across multiple platters.

At the bottom of the page, it says 'H . Ü . Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü v2.1.8'.

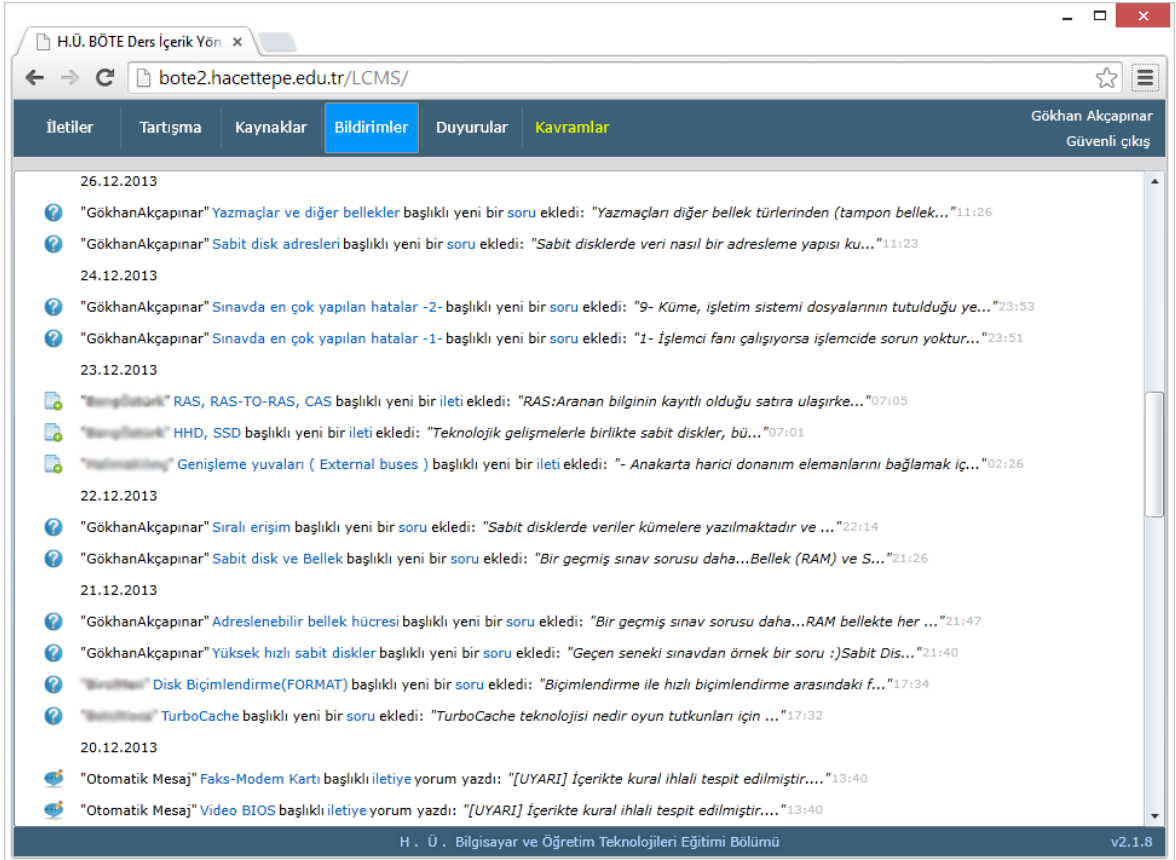
Şekil 3.4. Kaynaklar Sayfası

Tartışma ortamı: Tartışma bölümü öğrencilerin yeni bir tartışma sorusu eklemesine veya mevcut tartışmalara katılabilmesine olanak sağlamaktadır. Tartışma bölümünde öğrenciler veya öğretmenler tarafından gönderilen sorular ortama eklenme tarihine göre ters sıralı olarak listelenir. Öğrenci, tartışma ortamına yeni soru gönderebilir, yazılan sorulara cevap yazabilir, soruları veya cevapları değerlendirebilir. Yeni soru ekleme sayfası “Yeni ileti” sayfasına benzemektedir. Kullanıcı ilk olarak sorunun ilişkili olduğu konuyu seçer, daha sonra sorusunu ekler ve dilerse etiketleri kullanarak soruyu etiketleyebilir. Yazılan sorular öğretmenin

onayından geçtikten sonra tartışma bölümünde yayınlanır. İçerik bölümünde olduğu gibi burada da soru ve cevaplar için 500 karakter sınırı uygulanmıştır.

Oturum ve gezinim: Oturum; öğrencinin çevrimiçi ortama kullanıcı adı ve şifresi ile girmesi ile başlamaktadır ve ortamdaki çıkış yapana kadar geçirdiği süreyi kapsamaktadır. Sistem tarafından öğrencinin ortama her girişi ayrı bir oturum olarak ele alınmaktadır. Oturum sürelerinin güvenilirliğinin artırılması için sayfada belirli bir süre zarfında herhangi bir klavye ya da fare hareketi algılanmadığı durumlarda oturum otomatik olarak sonlandırılmaktadır. Bu süre araştırmacılar tarafından 15 dakika olarak belirlenmiştir. Öğrencinin çevrimiçi ortamda yaptığı her türlü tıklama işlemi ise gezinim verisi olarak kayıt edilmektedir. Gezinim verileri sadece sayfa gezinimlerini değil aynı zamanda sayfa içerisindeki bir takım etkileşim verilerini de kapsamaktadır (içerik ile ilgili yorumların açılması vs.).

Bildirimler: Bildirimler sayfasında ortamdaki son aktiviteler ile ilgili iletiler yer almaktadır. Bildirimler günlük şekilde tarihsel olarak ters sırada ve her aktiviteden 20'şer tane içerecek şekilde kullanıcılara sunulmaktadır (Şekil 3.5.).

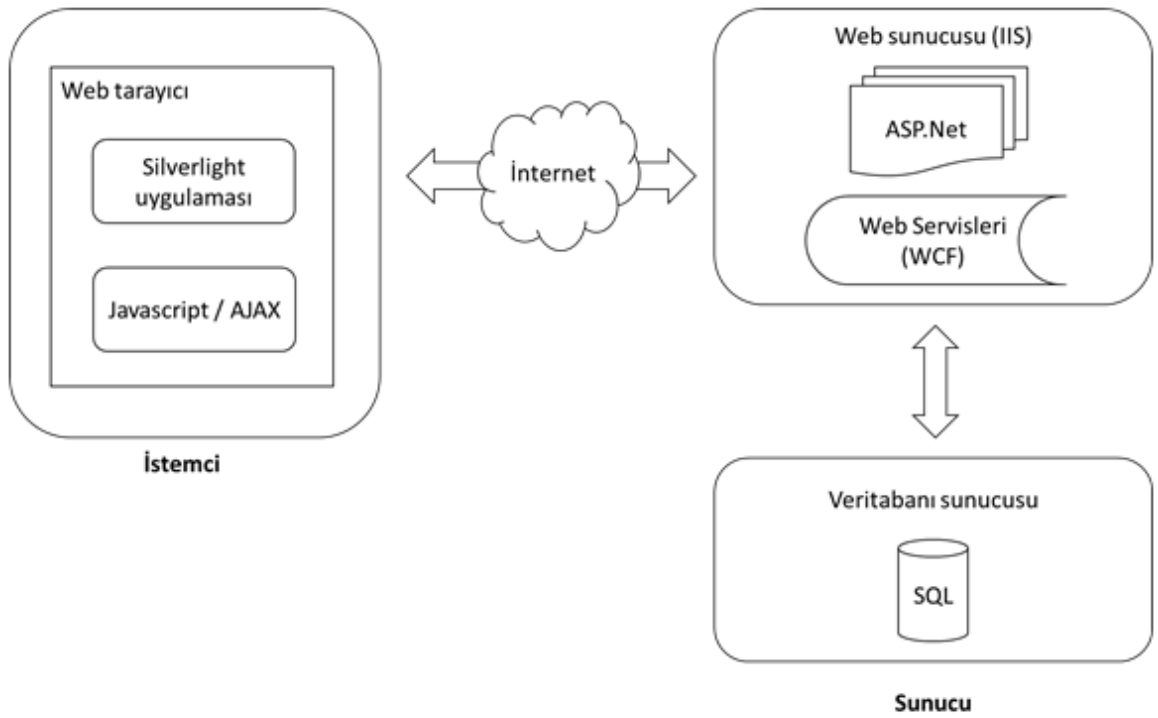


Şekil 3.5. Bildirimler Sayfası

Duyurular: Duyurular bölümünde ders ile ilgili duyurular paylaşılmaktadır, öğretmen dilerse duyuruları e-posta olarak da gönderebilmektedir.

3.2.1.2. Çevrimiçi Öğrenme Ortamının Teknik Özellikleri

Öğrenme ortamı araştırmacılar tarafından C#.Net programlama dili ve Silverlight uygulama geliştirme platformu kullanılarak geliştirilmiştir. Veritabanı yönetim sistemi olarak MS SQL Server kullanılmıştır. Silverlight, Microsoft firması tarafından geliştirilen ve zengin içerikli web uygulamaları geliştirilmesine olanak sağlayan bir uygulama geliştirme platformudur. Sunucu tabanlı programlama dillerinden farklı olarak Silverlight uygulamaları istemci bilgisayarında çalıştırılmaktadır ve sunucu ile web servisleri aracılığı ile haberleşmektedir. Şekil 3.6.'da özetlenen bu süreç sayesinde öğrenci - sistem etkileşimi ile ilgili daha detaylı verilerin (Ör: klavye ve fare hareketleri) toplanması mümkün olabilmektedir.



Şekil 3.6. Silverlight Uygulama Yaşam Döngüsü

Geliştirilen öğrenme ortamı öğrencilerin sürekli olarak erişimine açık olacak şekilde bir sunucuya yüklenmiştir. Bu sunucuda aynı zamanda ortamla ilgili değişkenlerin ve öğrenci etkileşimleri ile ilgili verilerin kayıt edildiği bir de veri tabanı bulunmaktadır.

3.2.2. Akademik Performans Puanları

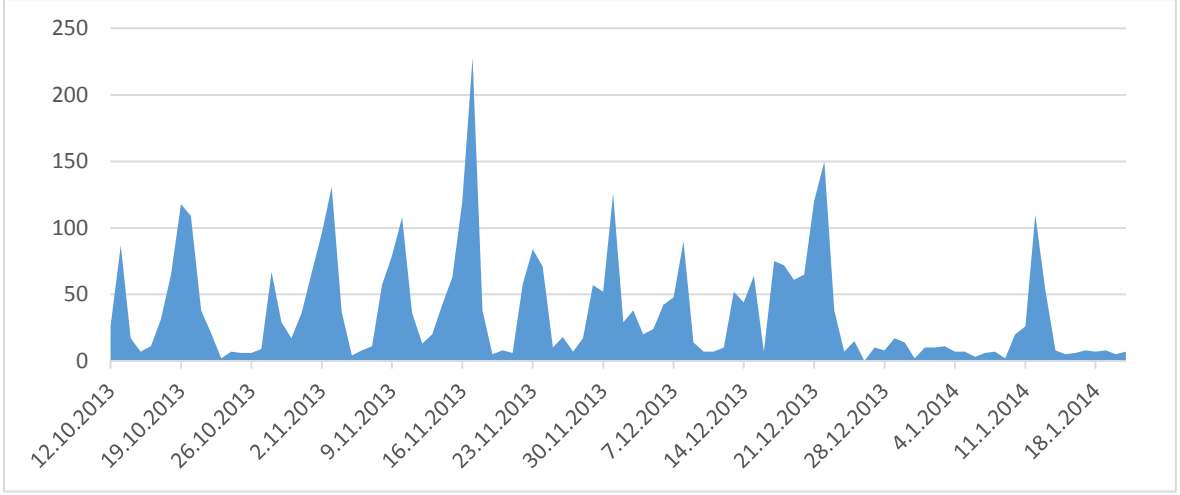
Birinci araştırma sorusu kapsamında çevrimiçi öğrenme ortamındaki etkileşim verileri kullanarak öğrenci performansını tahmin etmeye yönelik bir tahmin modeli oluşturulacaktır. Öğrenci performansının göstergesi olarak öğrencilerin Bilgisayar Donanımı dersine ait dönem sonundaki başarı notları dikkate alınmıştır. Öğrencilere iki ara sınav, bir final olmak üzere toplam üç adet yazılı sınav yapılmıştır. Öğrencilerin performans puanları ara sınavların %25'i ve final sınavının %50'si alınarak hesaplanmıştır. Çizelge 3.2'de bu puanlara ilişkin tanımlayıcı istatistikler verilmiştir. Birinci ve ikinci araştırma problemi kapsamında sınıf değişkeni olarak kullanılan ders başarısı, üniversitenin lisans yönetmeliğinde yer alan göstergeler dikkate alınarak "Kaldı" (n = 27, başarı puanı ≤ 50) ve "Geçti" (n = 49, başarı puanı > 50) şeklinde kesikli hale dönüştürülmüştür.

Çizelge 3.2: Başarı Puanları ile İlgili Tanımlayıcı İstatistikler

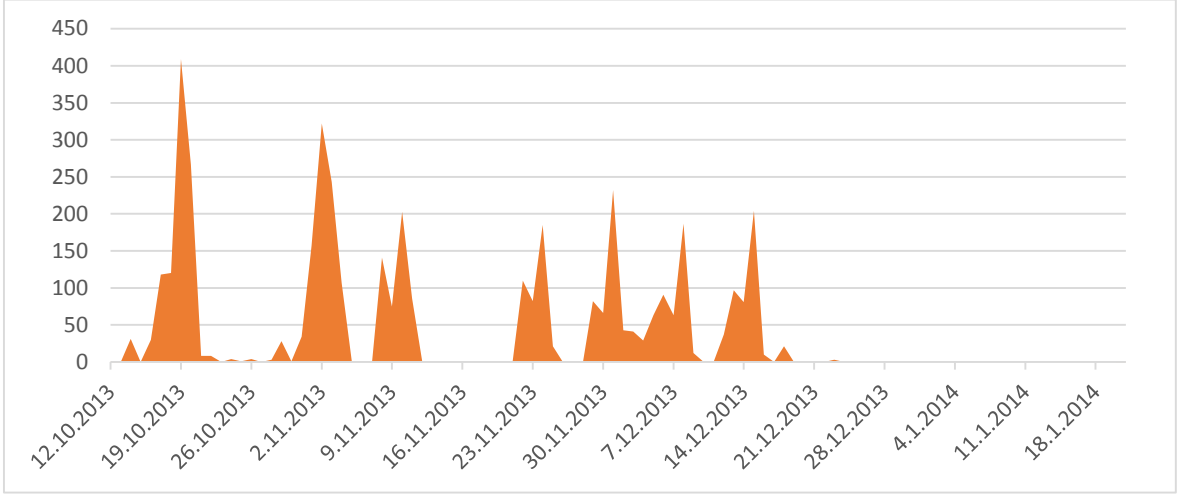
<i>Sınav</i>	<i>N</i>	<i>Ortalama</i>	<i>SS</i>
<i>1. Ara sınav</i>	75	54,20	21,77
<i>2. Ara sınav</i>	73	52,33	17,92
<i>Final sınavı</i>	74	60,59	17,51
<i>Başarı puanı</i>	76	55,49	18,20

3.3. Ham Veri

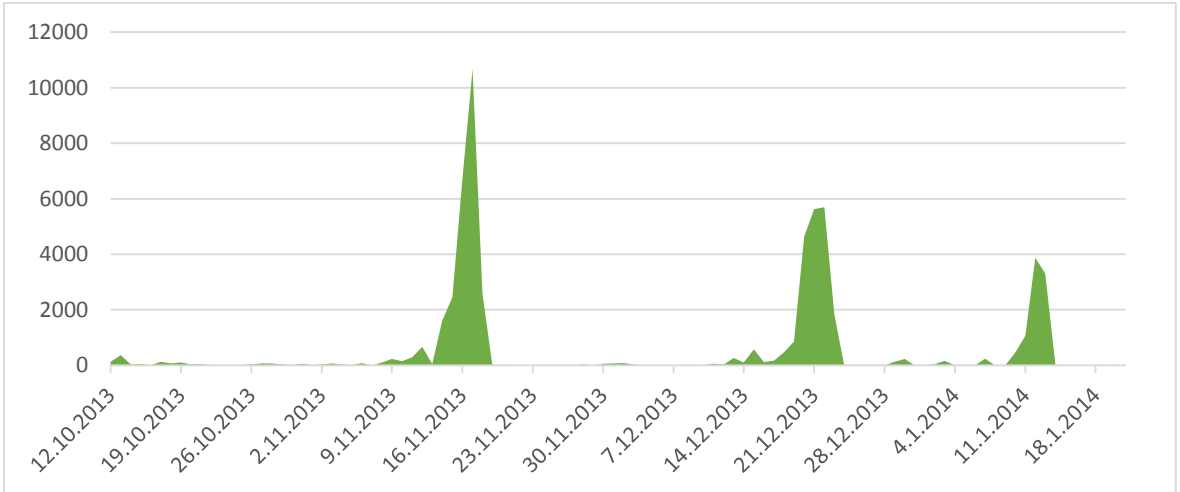
Veri analizine başlamadan önce çevrimiçi öğrenme ortamının veri tabanındaki bilgilerin ve öğrencilerin sınav notlarının bir arada tutulduğu bir analiz veri tabanı oluşturulmuştur. Bu veri tabanı 76 öğrenci tarafından yaklaşık 14 hafta süresince kullanılan öğrenme ortamından elde edilen kullanım verilerini ve buna ek olarak öğrencilerin ders performansını (geçti - kaldı) gösteren bir değişken içermektedir. Bu veriler: 3803 oturum, 119921 gezinim, 4566 yazılı metin, tartışma ortamında 62 adet soru, 297 cevap, toplam 8601 değerlendirme ve yazılan iletileri etiketlemek için kullanılan 3937 etiket verisini içermektedir. Ham verilerden elde edilen bilgilerle öğrencilerin ortamı kullanma alışkanlıklarını gösteren grafikler Şekil 3.7, 3.8 ve 3.9'da verilmiştir. Bu grafikler öğrencilerin 14 haftalık ders süresince ortama ne sıklıkta girdiklerini (Şekil 3.7.), ortamda yazdıkları iletilerin günlere göre dağılımını ve sıklığını (Şekil 3.8.) ve ortamda bulunan ders materyallerini ziyaret etme sıklıklarını göstermektedir (Şekil 3.9.).



Şekil 3.7. Öğrencilerin Günlere Göre Ortama Giriş Sayıları



Şekil 3.8. Öğrencilerin Günlere Göre Yazdıkları İleti Sayıları



Şekil 3.9. Öğrencilerin Günlere Göre Ders Kaynaklarında Gezinim Sayıları

3.4. Arařtırmada Kullanılan Deęiřkenler

Yukarıda ayrıntıları verilen ham verinin doğrudan analizlerde kullanılması mümkün değildir, bu nedenle bir takım dönüşümlerden geçerek analizlerde kullanılacak deęiřkenlerin elde edilmesi gerekmektedir. Analizlerde kullanılacak bu deęiřkenlerin belirlenmesinde öğrencilerin öğrenme ortamda gerçekleřtirdiđi etkileřimler göz önüne alınmıř ve bu etkileřimleri yansıtacak, ölçülebilir metrikler oluřturulmaya çalıřılmıřtır. Veritabanındaki ham veriden deęiřkenleri çıkartma iřlemi geliřtirilen bir araç yardımıyla otomatik olarak yapılmıřtır. Bölüm 3.2.1’de özellikleri ayrıntılı olarak anlatılan sistemdeki öğrenci rolleri ile ilgili deęiřkenler “ileti”, “tartıřma”, “oturum ve gezinim” olmak üzere üç başlıkta toplanmıřtır. Ařađıda bu deęiřkenler ve iliřkili olduđu roller açıklanmıřtır. Bu deęiřkenlerden hangilerinin analizlerde kullanılacađına ise ön iřleme sürecinde karar verilmiřtir.

İletiler ile ilgili deęiřkenler: Öğrenciler ortamda iletiler ile ilgili yapabileceđi etkileřimler: yeni bir ileti oluřturma, yazılan bir iletiyi okuma, iletilere yorum yazma ya da deđerlendirme řeklinde olabilmektedir. Bu etkileřimlerle ilgili oluřturulan deęiřkenler ařađıdaki gibidir.

- t_Ileti: Öğrencinin derste öğrendiđi/öğreneceđi kavramlarla ilgili yazdıđı toplam yansıma sayısı.
- t_BenzersizIleti: Öğrencinin ileti yazdıđı farklı gün sayısı. Bu deęiřken yazılan iletilerin dönem içerisinde düzenli olarak mı yazıldıđını yoksa toplu olarak mı yazıldıđını belirlemek açısından önemlidir.
- t_Etiket: Toplam etiket sayısı öğrencinin kavram listesine yaptıđı katkıyı göstermektedir.
- t_EtiketKullanım: Bu deęiřken, öğrencinin yazdıđı iletelerde kullandıđı toplam etiket sayısını göstermektedir.
- t_IletiDeđerlendirme: Öğrencinin diđer öğrenciler tarafından yazılan iletileri deđerlendirme sayısı.
- t_Yorum: Öğrencinin yazdıđı toplam yorum sayısı.
- t_YorumDeđerlendirme: Öğrencinin diđer öğrenciler tarafından yazılan yorumları deđerlendirme sayısı.

- t_IletiGezinim: Öğrencinin iletiler ile ilgili sayfada gerçekleştirdiği toplam gezinim sayısı.
- t_IletiOkuma: Öğrencinin iletileri (okuma amaçlı) toplam açma sayısı.
- t_YorumGezinim: Öğrencinin iletiler ile ilgili yazılan yorum bölümlerinde gerçekleştirdiği toplam gezinim sayısı.
- o_IletiYazmaSuresi: Öğrencinin ileti yazarken harcadığı ortalama süre.
- o_TusKullanım: Öğrencinin ileti yazarken kullandığı ortalama tuş sayısı.
- o_SilmeTusuKullanım: Öğrencinin ileti yazarken ortalama “Silme” ve “Geri Al” tuşlarını kullanım sayısı.
- o_SayfaDeaktif: Öğrencinin ileti yazarken sayfanın ortalama arka plana gitme sayısı.

Çevrimiçi ödevlerde içeriklerin kaynak belirtilmeden doğrudan alınması (aşırı macılık) önemli bir problemdir. Bu nedenle iletiler ile ilgili değişkenlere yazılan iletilerin orijinalliğini test etmeye yönelik yönelik iki değişkene eklenmiştir.

t_KopyalaYapistir: Öğrencilerin içerikleri yazmadan farklı bir ortamdan kopyala yapıştır yöntemiyle giriş yapması.

t_KopyalletiSayisi¹: Öğrencilerin yazdıkları iletiler arasında benzerlik oranı %70 ve üzeri olan ileti sayısı.

Tartışma ile ilgili değişkenler: Öğrenciler tartışma ortamında yapabileceği etkileşimler: yeni bir soru sorma, yazılan bir soruyu okuma, sorulara cevap yazma ya da değerlendirme şeklinde olabilir. Bu etkileşimlerle ilgili oluşturulan değişkenler aşağıda verilmiştir.

- t_Soru: Öğrencinin tartışma ortamında yazdığı toplam soru sayısı.
- t_Cevap: Öğrencinin tartışma ortamında yazdığı toplam cevap sayısı.
- t_SoruDeğerlendirme: Öğrencinin tartışma bölümünde diğer öğrenciler tarafından yazılan soruları değerlendirme sayısı.

¹ Kopya ileti sayısının hesaplanmasında metin madenciliği analizlerinden yararlanılmıştır; öğrenciler tarafından yazılan iletiler birbirleri ile ve İnternet arama sonuçları ile karşılaştırılarak kopya oranları hesaplanmıştır. Metin madenciliği kullanılarak yapılan benzerlik hesaplamaları sonucu aralarında %70 ve üzeri benzerlik bulunan iletiler benzer ileti olarak ele alınmıştır.

- t_CevapDeğerlendirme: Öğrencinin tartışma bölümünde diğer öğrenciler tarafından yazılan cevapları değerlendirme sayısı.
- t_CevapGezirim: Öğrencinin tartışma ortamında yazılan cevaplar ile ilgili bölümlerde gerçekleştirdiği toplam gezinim sayısı.
- t_TartismaGezirim: Öğrencinin tartışmalar ile ilgili sayfada gerçekleştirdiği toplam gezinim sayısı.

Oturum ve gezinim: Çevrimiçi öğrenme ortamının önemli bir bileşeni de gezinimdir. Öğrenciler kendilerine verilen kullanıcı adı ve şifreler ile ortamda oturum açmaktadır ve yaptıkları her türlü gezinim ve aktivite bu öğrenci ve oturum ile ilişkili olarak kayıt edilmektedir. Oturum ve gezinim davranışları ile ilgili değişkenler aşağıda verilmiştir.

- t_Oturum: Öğrencinin uygulama süresince gerçekleştirdiği toplam oturum sayısı.
- t_BenzersizOturum: Öğrencinin uygulama süresince en az bir kez oturum açtığı farklı gün sayısı.
- t_OturumSuresi: Öğrencinin gerçekleştirdiği oturumlarda toplam ortamda kalma süresi.
- t_Gezinim: Öğrencinin öğrenme ortamında gerçekleştirdiği toplam gezinim sayısı.
- t_KaynakGezinim: Öğrencinin ders kaynakları ile ilgili sayfalarda yaptığı toplam gezinim sayısı.
- t_DuyuruGezinim: Öğrencinin dersle ilgili duyuruların yayınlandığı sayfadaki toplam gezinim sayısı.
- t_BildirimGezinim: Öğrencinin yapılan aktiviteler ilgili bildirimlerin gösterildiği sayfada gerçekleştirdiği toplam gezinim sayısı.

Yukarıda bahsedilen tüm değişkenler Çizelge 3.3.'de özet olarak verilmiştir.

Çizelge 3.3: Araştırmada Kullanılan Değişkenler ve Açıklamaları

No	Değişken	Açıklama
1	t_BenzersizIleti	İleti yazılan farklı gün sayısı
2	t_SoruDeğerlendirme	Tartışma ortamında soru değerlendirme sayısı
3	t_BenzersizOturum	Ortama giriş yapılan farklı gün sayısı
4	t_Gezinim	Ortamda yapılan toplam gezinim sayısı
5	t_EtiketKullanım	Yazılan iletilerde kullanılan etiket sayısı
6	t_Ileti	Öğrenci tarafından yazılan ileti sayısı
7	t_Etiket	Öğrenci tarafından ortama eklenen etiket sayısı
8	t_IletiGezinim	İletilerin bulunduğu sayfada yapılan toplam gezinim
9	t_CevapOkuma	Tartışma ortamında yazılan cevapları açma sayısı
10	t_Cevap	Tartışma ortamında yazılan toplam cevap sayısı
11	t_BildirimGezinim	Kullanıcılar tarafından yapılan aktivitelerin yer aldığı sayfada gezinim sayısı
12	t_Oturum	Gerçekleştirilen toplam oturum sayısı
13	t_CevapDeğerlendirme	Tartışma ortamında değerlendirilen cevap sayısı
14	t_IletiDeğerlendirme	Diğer öğrenciler tarafından yazılan iletileri değerlendirme sayısı
15	t_KopyalaYapıştır	İleti yazarken kopyala – yapıştır kullanım sayısı
16	t_IletiOkuma	Diğerleri tarafından yazılan iletileri açma sayısı
17	t_OturumSuresi	Ortamda kalınan toplam süre
18	o_TusKullanım	İleti yazarken ortalama tuş kullanım sayısı
19	o_SilmeTusuKullanım	İleti yazarken ortalama silme tuşu kullanım sayısı
20	t_KaynakGezinim	Kaynakların yer aldığı sayfada gezinim sayısı
21	t_DuyuruGezinim	Duyuruların yer aldığı sayfada gezinim sayısı
22	t_TartismaGezinim	Tartışma sayfasında gezinim sayısı
23	o_SayfaDeaktif	İleti yazarken ortalama a sayfanın arka planda kalma sayısı
24	t_Yorum	Yazılan iletilere yapılan yorum sayısı
25	t_YorumDeğerlendirme	Yazılan yorumları değerlendirme sayısı
26	o_IletiYazmaSuresi	Ortalama ileti yazma süresi
27	t_KopyalIletiSayisi	Benzerlik oranı yüksek ileti sayısı
28	t_Soru	Öğrencinin tartışma ortamında yazdığı toplam soru sayısı

3.5. Veri Ön İşleme

VM çalışmalarında analizlerden önceki son aşama ön işlemedir. Çalışma kapsamında ele alınan araştırma problemleri birden fazla yöntemi içerdiği için ön işleme süreci her bir yöntem için farklılık göstermektedir. Bu nedenle ön işleme süreci veri analizi kısmında ele alınmıştır. Bu aşamada sadece öğrenciler tarafından katılım oranı çok düşük olduğu için t_Soru değişkeni analizden çıkartılmıştır. Geriye kalan 27 değişken ile ilgili tanımlayıcı istatistikler ise Ek 1.'de verilmiştir.

3.6. Veri Analizi

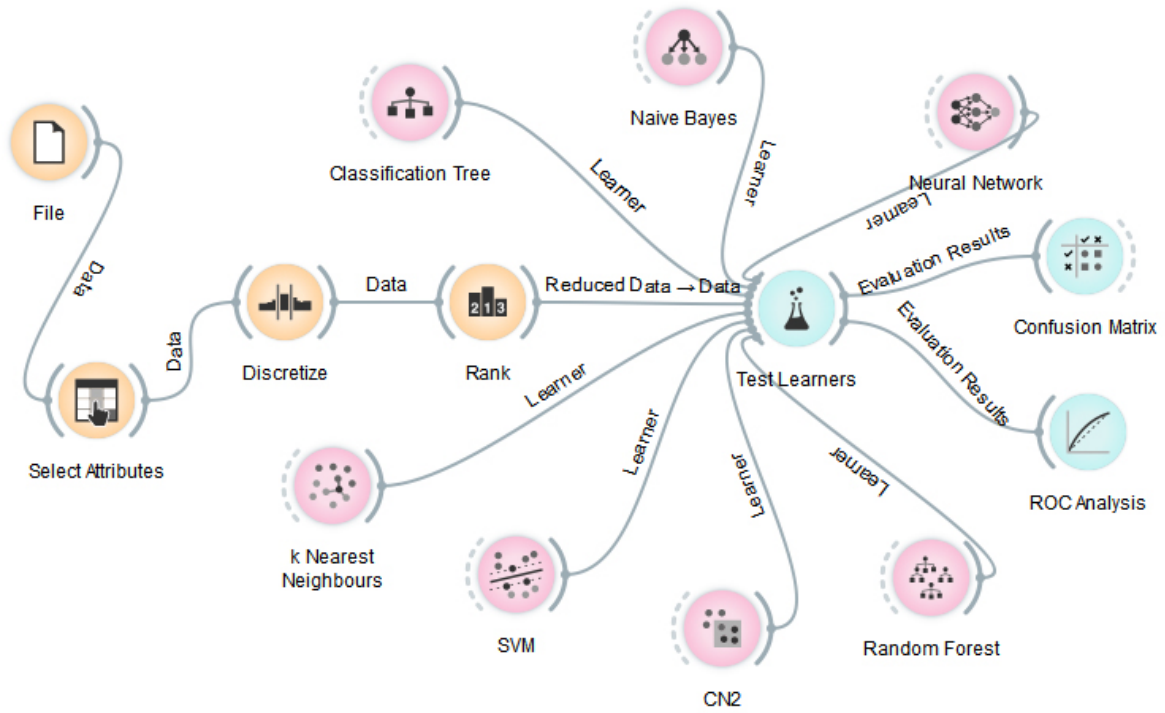
Araştırma problemleri kapsamında farklı veri analizi süreçleri izlendiği için bu bölüm araştırma problemleri doğrultusunda yazılmıştır.

3.6.1. Birinci Problem ve Alt Problemlerine İlişkin Veri Analizi

Birinci araştırma problemi bir tahmin analizidir. Amaç, öğrencilerin çevrimiçi ortamdaki etkileşim verilerini kullanarak dönem sonu akademik performanslarını tahmin edecek bir model oluşturmaktır. Öğrencilerin dönem sonu performansı “Geçti” ve “Kaldı” şeklinde kodlandığı için bu bir sınıflama problemidir. VM çalışmalarında sınıflama amacıyla kullanılacak çok sayıda algoritma bulunmaktadır. Algoritmaların sınıflama performansı ise veri setinin özelliklerine göre farklılıklar gösterebilmektedir. Bu nedenle veri setine uygun sınıflama algoritmasının seçilmesi önemlidir (C Romero, Olmo, & Ventura, 2013). Farklı algoritmalar ile aynı analizi tekrar ederek, performans ölçütlerine göre veri setinde en iyi performans gösteren algoritmanın seçilmesi yaygın olarak kullanılan bir yöntemdir (Akçapınar, Çoşgun, & Altun, 2013; Osmanbegović & Suljić, 2012; Cristobal Romero ve diğerleri, 2010).

Sınıflama analizi sonuçlarının ne amaçla kullanılacağı da algoritma seçiminde dikkat edilmesi gereken bir diğer noktadır. Örneğin “kara kutu” olarak da anılan Naive Bayes algoritması oldukça iyi sınıflama performansı sergileyebilmektedir; fakat elde edilen modelin kullanıcılar tarafından anlaşılması da bir o kadar zordur (Cristobal Romero ve diğerleri, 2010). Bu tür algoritmalar sınıflama performansının modelin yorumlanmasından daha önemli olduğu durumlarda tercih edilmelidir (Dreiseitl & Ohno-Machado, 2002). Diğer taraftan “beyaz kutu” algoritmalar olarak da anılan kural tabanlı ya da ağaç tabanlı algoritmalar ise uzman olmayan kişiler tarafından bile kolaylıkla anlaşılabilir sonuçlar üretmektedir (Osmanbegović & Suljić, 2012).

Bu çalışma kapsamında da benzer bir yaklaşım izlenerek literatürde sınıflama amacıyla sıklıkla kullanılan kural tabanlı, ağaç tabanlı, fonksiyon tabanlı ve çekirdek tabanlı yedi farklı algoritma (naive bayes, random forest, Destek vektör makinaları, karar ağacı, yapay sinir ağları, CN2 kuralları ve k en yakın komşuluk algoritması) seçilmiştir. Sınıflama algoritmalarının performansı ise Doğru Sınıflama Oranı (DSO), Doğruluk, Kesinlik, F-Ölçütü ve ROC altında kalan alan (EAKA) performans ölçütlerine göre karşılaştırılmıştır. Sınıflama analizleri Şekil 3.10.’da verilen sürece göre Orange VM yazılımı (Demšar ve diğerleri, 2013) ile yapılmıştır. Çalışmada kullanılan sınıflama algoritmaları ve performans metrikleri aşağıda açıklanmıştır.



Şekil 3.10. Birinci Araştırma Problemine İlişkin Analizi Süreci

Naive Bayes (NB): NB algoritması, olasılıksal bir sınıflayıcıdır. Eğitim verisini kullanarak bağımlı değişkenlerin koşullu olasılıklarını kestirir ve bunları yeni verilerin sınıflandırmasında kullanır.

Karar Ağacı (Classification Tree): Karar ağaçları VM çalışmalarında tahmin modelleri oluşturmak amacıyla yaygın olarak kullanılan bir yöntemidir. Karar ağacı yöntemini kullanan çok sayıda algoritma bulunmaktadır. C&ART, CHAID, C4.5, ID.3 bunlardan bazılarıdır. Karar ağaçlarında seçilen bilgi ölçütüne göre (gini indeksi, kazanım oranı vb.) veriyi en iyi ayıran değişken belirlenir ve bu doğrultuda ilk düğüm oluşturulur daha sonra bölünme sonucu elde edilen veri için bu işlem tekrarlanır ve veri tekrar alt düğümlere bölünür. Bu işlem her bir düğümde kullanıcı tarafından belirlenen sayıda eleman kalana kadar ya da kullanıcı tarafından belirlenen seviyede düğüm oluştuğu zaman sonlandırılır. Çıktı olarak üretilen karar ağaçlarının yorumlaması kolay olduğu için tercih edilmektedir.

Random Forest (RF): (Breiman), tarafından geliştirilen RF algoritması ağaç tabanlı bir algoritmadır. Karar ağacı algoritmalarının aksine (tek bir ağaç üretmek yerine) bootstrap metodu kullanılarak yüzlerce hatta binlerce karar ağacı üretilmesi esasına dayanır. Değişkenlerin de rastgele seçildiği her bir karar ağacı için veri setinden rastgele bir örneklem seçilir. Seçilen örneklemin 2/3'ü karar ağacı oluşturmak için

kullanılır ve oluşturulan her bir karar ağacının sınıflama performansı veri seti üzerinde test edilir. Algoritma tarafından bu işlem sırasında her bir karar ağacına oy verilir. İşlem bittiğinde en yüksek oyu alan (en düşük hata oranına sahip) ağaç seçilir ve sınıflama o ağaca göre yapılır.

Destek Vektör Makinaları (DVM): DVM (Support Vector Machines) algoritması çekirdek (kernel) tabanlı bir algoritmadır. Doğrusal olarak sınıflanabilen bir veride sınıfları birbirinden ayıran marjini en büyük olan doğrusal fonksiyonu bulmayı amaçlar. Doğrusal olmayan veri setlerinde ise veri daha yüksek boyutlu uzaya aktarılır ve sınıflama işlemi marjini en büyük olan hiper-düzleme göre yapılır (Chang & Lin, 2011).

Yapay Sinir Ağları (Neural Network): Yapay sinir ağları algoritması insan beyninde yer alan nöronların çalışmasından esinlenerek geliştirilmiş bir algoritmadır (Nisbet, Elder, & Miner, 2009). Lojistik ya da doğrusal fonksiyonları kullanarak sınıflama ve regresyon analizlerinde kullanılmaktadır. Elde edilen sonuçların uzman olmayan kişiler tarafından yorumlanması ve anlaşılması zordur (Hämäläinen & Vinni, 2010). Yapısı gereği analiz sürecinin önemli bir bölümü gizli (kara kutu olarak da adlandırılmaktadır) olduğu için (Stolzer, 2009) daha çok ses tanıma, görüntü işleme ve örüntü tanıma gibi alanlarda tercih edilmektedir.

CN2 Kuralları (CN2 Rules): CN2 algoritması kural tabanlı bir algoritmadır. AQ ve ID3 sınıflama algoritmalarının güçlü yönleri temel alınarak oluşturulmuştur (Clark & Niblett, 1989). AQ algoritmasında olduğu gibi sınıflama sonucu EĞER-İSE kuralları üretmektedir ve ID3 algoritmasında olduğu gibi gürültülü (veri girişi veya veri toplanması sırasında oluşan sistem dışı hatalar) verilerde de çalışabilmektedir.

k en yakın komşuluk (k Nearest Neighbors): Algoritmanın temel mantığı sınıflandırılacak verinin en yakınındaki k adet örneğin sınıf bilgisine bakarak yeni veriyi çoğunluğun ait olduğu sınıfa atamaktır (Hand, Smyth, & Mannila, 2001). Algoritmanın sınıflandırma işlemindeki başarısını etkileyen faktörlerden biri k değerinin doğru seçilmiş olmasıdır. Eğer k değeri çok büyük seçilirse farklı sınıflara ait örnekler aynı sınıfa dahil edilebilir. K değerinin çok küçük seçildiği diğer bir durumda ise tam tersi şekilde aynı sınıfta bulunması gereken örnekler farklı sınıflara yerleştirilebilir (Mitchell, 1997).

Sınıflama algoritmalarının performansının karşılaştırılması amacıyla kullanılan performans metrikleri aşağıda açıklanmıştır. Bu metrikler sınıflama modelinin doğru olarak ve hatalı olarak sınıfladığı örneklerin bir arada gösterildiği çapraz tablo üzerinden hesaplanmaktadır. Bu nedenle performans metriklerinin açıklanmasında Çizelge 3.4’de örnek olarak verilen ve öğrencilerin ders başarısı ile ilgili olarak “Geçti”, “Kaldı” şeklinde iki değişkenli bir sınıflama problemine ait örnek bir çapraz tablo kullanılmıştır.

Çizelge 3.4: Örnek Çapraz Tablo

		<i>Gerçek durum</i>	
		<i>Kaldı</i>	<i>Geçti</i>
<i>Tahmin</i>	<i>Kaldı</i>	DP (a)	YN (b)
	<i>Geçti</i>	YP (c)	DN (d)

Burada:

DP (Doğru Pozitif): Model tarafından “Kaldı” olarak sınıflandırılan başarısız öğrenci sayısı.

YN (Yanlış Negatif): Model tarafından “Kaldı” olarak sınıflandırılan başarılı öğrenci sayısı.

DN (Doğru Negatif): Model tarafından “Geçti” olarak sınıflandırılan başarılı öğrenci sayısı.

YP (Yanlış Pozitif): Model tarafından “Geçti” olarak sınıflandırılan başarısız öğrenci sayısı.

Bu çapraz tabloya göre performans metrikleri aşağıdaki gibi hesaplanmaktadır.

Doğru sınıflama oranı (DSO): Modelin doğru olarak sınıflandırdığı elemanların toplam eleman sayısına bölümüdür ve aşağıdaki formülle hesaplanmaktadır. Sınıf değişkeninin dengeli dağılmadığı durumlarda elde edilen sonuçlar yanıltıcı olabilmektedir, bu nedenle modelin performansını ölçmek açısından tek başına yeterli bir ölçüt değildir.

$$DSO = \frac{a + d}{a + b + c + d}$$

Duyarlılık (Sensitivity): Modelin başarısız öğrencileri (Kaldı) doğru olarak sınıflandırma olasılığıdır ve aşağıdaki formüle göre hesaplanmaktadır. Bu metrik aynı zamanda **Anma (Recall)** olarak da kullanılmaktadır.

$$\text{Duyarlılık} = \frac{a}{a + b}$$

Seçicilik (Specificity): Modelin “başarılı” öğrencileri (Geçti) doğru olarak sınıflandırma olasılığıdır ve aşağıdaki formüle göre hesaplanmaktadır.

$$\text{Seçicilik} = \frac{d}{c + d}$$

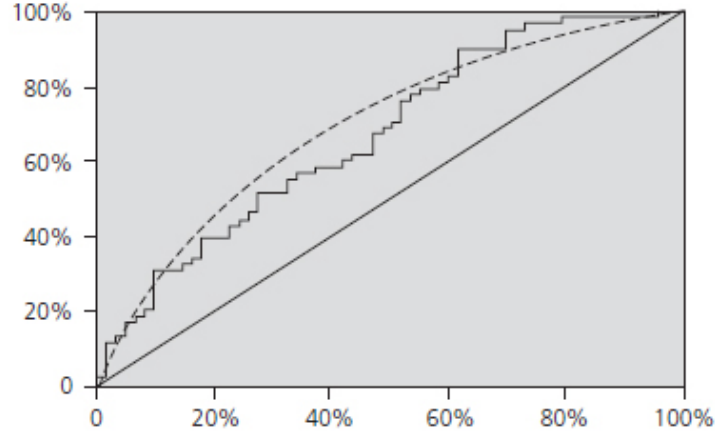
Kesinlik (Precision): Doğru sınıflandırılmış başarısız öğrenci sayısının toplam başarısız öğrenci sayısına oranıdır.

$$\text{Kesinlik} = \frac{a}{a + c}$$

F-Ölçütü (F-measure): Kesinlik ve duyarlılık ölçütlerinin tek başına kullanılması modellerin değerlendirilmesinde yanlış yorumlar yapılmasına neden olabilir, bu nedenle iki ölçütün beraber kullanıldığı F-Ölçütü üretilmiştir. F-Ölçütü, kesinlik ve duyarlılığın harmonik ortalamasıdır.

$$F - \text{Ölçütü} = \frac{2a}{2a + b + c}$$

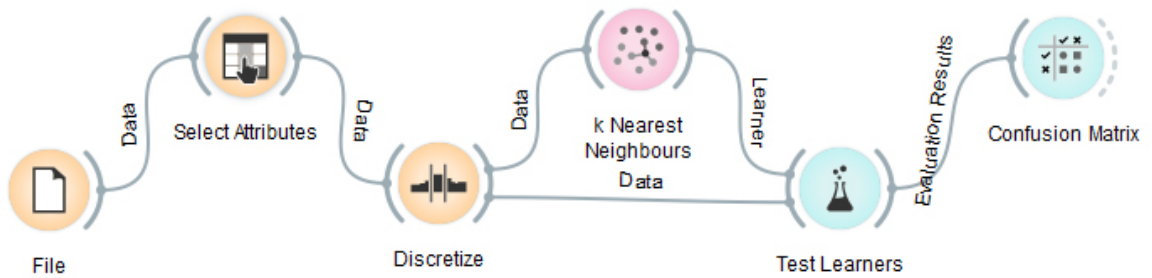
ROC (receiver operating characteristic) altında kalan alan (EAKA): İşaret işlemede bir sezicinin, gürültülü bir kanalda doğru algılama oranının yanlış alarm oranına karşı çizdirilen grafiği ROC eğrisi olarak adlandırılır. Şekil 3.11’de örnek olarak verilen ve farklı sınıflandırıcıları karşılaştırmak için kullanılan ROC eğrilerinde y ekseninde Doğru Pozitif Oranları $a + (a + b)$, x ekseninde ise Yanlış Pozitif Oranları $c + (c + d)$ yer almaktadır. ROC üzerindeki her nokta bir sınıflandırıcının oluşturduğu bir modele karşı düşer. Bu eğrinin altında kalan alan ise sınıflayıcı performansının bir göstergesi olarak alınmaktadır.



Şekil 3.11. Örnek ROC Eğrisi

3.6.2. İkinci Problem ve Alt Problemlerine İlişkin Veri Analizi

İkinci araştırma problemi kapsamında dönem sonunda başarısız olma ihtimali yüksek olan öğrencilerin önceden tahmin edilmesi amacıyla öğrencilerin etkileşim verilerinin yer aldığı analiz veri tabanından 3, 6, 9 ve 12. Haftalara ait veriler alınarak, Şekil 3.12.'de verilen analiz sürecine göre sınıflama modelleri oluşturulmuştur. Sınıflama modelleri oluşturulurken birinci araştırma probleminde en iyi sonuç veren algoritma ve ön işleme yöntemleri dikkate alınmıştır. Elde edilen modellerin performansı ise 14. Haftada alınan veri kullanılarak oluşturulan sınıflama modeli ile karşılaştırılarak yorumlanmıştır. Burada da elde edilen sınıflama modellerinin performanslarının karşılaştırılmasında ve sonuçların genelleştirilmesinde birinci araştırma probleminde anlatılan yöntemler kullanılmıştır.



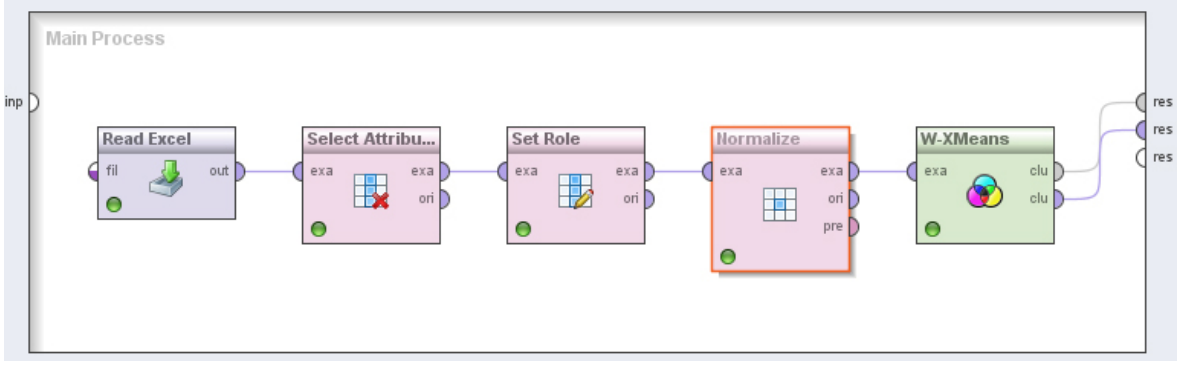
Şekil 3.12. İkinci Araştırma Problemine İlişkin Analizi Süreci

3.6.3. Üçüncü Problem ve Alt Problemlerine İlişkin Veri Analizi

Üçüncü araştırma problemi kapsamında çevrimiçi öğrenme ortamında benzer kullanım davranışları sergileyen farklı öğrenci profillerinin belirlenmesi amacıyla kümeleme analizinden yararlanılmıştır. Kümeleme analizinin amacı küme içi uzaklıkları

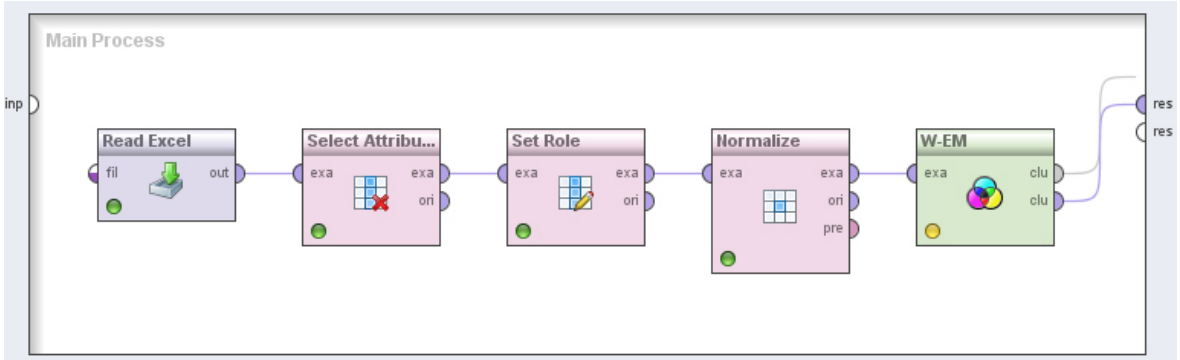
azaltıp, kümeler arası uzaklıkları artırarak, kendi içerisinde benzerlik gösteren elemanlardan oluşan farklı kümeler elde etmektedir (Chien-Sing & Singh, 2004). Kümeleme analizinde verinin kaç farklı kümeye ayrılacağı belirlenmesi konusunda iki farklı yaklaşım bulunmaktadır. Araştırmacılar bunu ön tanımlı olarak belirtilebilir ya da çeşitli algoritmalar kullanılarak verideki optimal küme sayısı bulunabilir (R. J. d. Baker, Corbett, Koedinger, & Roll, 2006; Bienkowski ve diğerleri, 2012). Kümeleme analizi tanımlayıcı bir yöntem olduğu için sınıflama analizinde olduğu gibi farklı algoritmaları karşılaştırarak en iyi performans gösteren algoritmayı seçme gibi bir durum olanaklı olmamaktadır. Bu nedenle, literatürde en çok kullanılan iki kümeleme yöntemi X-Means (X-Ortalamlar) ve EM (Expectation Maximization) kümeleme yöntemleri kullanılarak sonuçları tartışılmıştır. Her iki yöntem için de ideal küme sayısı algoritmalar tarafından belirlenmiştir. Kümeleme analizleri Rapidminer VM yazılımının Weka eklentisinde bulunan W-XMeans ve W-EM algoritmaları kullanılarak Şekil 3.13 ve Şekil 3.14'de verilen analiz süreçlerine göre gerçekleştirilmiştir. Aşağıda bu kümeleme yöntemleri ile ilgili ayrıntılı bilgi verilmiştir.

X-Ortalamlar: X-Ortalamlar kümeleme algoritması, Pelleg ve Moore (2000) tarafından geliştirilmiştir ve K-Ortalamlar kümeleme algoritmasının zayıf yönlerinin geliştirilmesi ile ortaya çıkmış bir kümeleme algoritmasıdır. MacQueen (1967) tarafından bulunan K-Ortalamlar kümeleme algoritmasının amacı n sayıdaki elemanı ön tanımlı olarak belirtilen k tane kümeye ayırmaktır. Bunu yaparken ilk olarak k adet hayali nokta belirlenir. Bu noktalar rastgele olarak seçilir ve küme merkezi olarak adlandırılır. Daha sonra çeşitli uzaklık metrikleri kullanılarak her bir eleman en yakın olduğu küme merkezine atanır. Bu işlem tüm elemanlar bir kümeye atanana kadar devam eder. Yeni oluşan kümelerin merkezi ilk başta rastgele olarak seçilen küme merkezlerinden farklı olacağı için, yeni küme merkezleri hesaplanır ve her bir eleman tekrar bu küme merkezlerine atanır. Bu işlem küme merkezleri sabitlenene kadar devam eder ve bu sayede veri k tane kümeye ayrılmış olur. X-Ortalamlar kümeleme algoritmasının buna getirdiği en önemli katkı k 'nın otomatik olarak belirlenmesidir. Algoritma Bayes Bilgi Kriteri (BIC)'ni kullanarak optimal küme sayısını belirlemektedir. Diğer bir katkı ise büyük veri setlerinde hızlı çalışması konusunda olmuştur.



Şekil 3.13. X-Ortalamalar Kümeleme Analiz Süreci

EM Kümeleme: EM kümeleme algoritmasının adı Expectation Maximization (Maksimum Beklenti) kelimelerinin baş harfinden oluşmaktadır. X-Ortalamalar kümeleme algoritmasından farklı olarak burada bir eleman bir kümeye uzaklık ölçütleri ile kesin bağlı değildir, bunun yerine her bir kümeye bağlı olma olasılığı vardır (Nisbet ve diğerleri, 2009). Bağımsız değişkenler kullanılarak küme sayısı kadar farklı ortalama ve standart sapmada dağılımlar oluşturulur ve elemanlar yakın oldukları dağılıma göre kümelere yerleştirilir (Witten, Frank, & Hall, 2011). Burada küme sayısı kullanıcı tarafından belirleneceği gibi çeşitli yöntemler kullanılarak algoritma tarafından da belirlenebilir. Bu çalışmada kullanılan EM kümeleme analizleri Rapidminer VM yazılımının Weka paketinde bulunan W-EM modülü kullanılarak gerçekleştirilmiştir. Bu modül çapraz geçerlilik yöntemi kullanarak verideki optimal küme sayısının bulunmasına olanak sağlamaktadır.



Şekil 3.14. EM Kümeleme Analiz Süreci

3.7. Araştırmanın İç ve Dış Geçerliliği

Araştırmada kullanılan tüm sınıflama analizlerinde elde edilen sonuçların genelleştirilmesi için 5k Çapraz Geçerlilik yöntemi kullanılmıştır. Bu yöntemde göre veri 5 eşit veri setine bölünür, ilk 4 veri seti ile sınıflama modeli kurulur (eğitim verisi), elde edilen model geriye kalan 1 veri seti üzerinde test edilir (test verisi). Bu işlem

her bir veri seti, test seti olarak kullanılıncaya kadar devam eder. Bu sayede analiz sonucunda 5 farklı sınıflandırma modeline ait performans ölçümleri elde edilir. Daha sonra bu verilerin ortalaması alınarak oluşturulan modelin ortalama performansı elde edilir. Bulgularda verilen sonuçlar bu şekilde elde edilmiş ortalama performans ölçümleridir.

Kümeleme analizlerinde de küme sayısının belirlenmesinde çapraz geçerlilik yöntemi kullanılmıştır. Bu sayede verideki optimal küme sayısı elde edilmiştir. VM çalışmalarında dış geçerlilik, oluşturulan modelin performansının farklı (yeni) bir veri setinde uygulanması ile test edilmektedir. Geliştirilen öğrenme ortamı ilk defa uygulandığı için elde edilen tahmin ve kümeleme modellerinin dış geçerliliği test edilememiştir. İlerleyen çalışmalarda öğrenme ortamının farklı derslerde ve farklı öğrenci gruplarında uygulanması ile bu mümkün olacaktır.

4. BULGULAR ve TARTIŞMA

Bu bölümde, analiz sonucu elde edilen bulgular problem ve alt problemlerin veriliş sırasına göre ele alınmıştır. Araştırmanın temel problemi “çevrimiçi öğrenme ortamından elde edilen verilerin, VM yöntemleri ile analiz edilerek öğrencilerin akademik performanslarının modellenmesinde kullanılıp kullanılmayacağını araştırılması” şeklindedir. Bu doğrultuda belirlenen araştırma problemleri ve alt problemlerine ilişkin bulgular aşağıdaki gibidir.

4.1. Birinci Araştırma Problemine İlişkin Bulgular

Problem durumu: Öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak oluşturulan farklı sınıflama modellerinin derse ilişkin akademik performanslarını (geçti - kaldı) tahmin etme başarısı nasıldır?

Birinci araştırma problemi kapsamında ayrıntıları Çizelge 4.1’de verilen altı farklı analiz gerçekleştirilmiştir. Birinci analizde veriye herhangi bir işlem yapılmadan sınıflama performansları hesaplanırken, ikinci ve üçüncü analizde farklı veri dönüştürme tekniklerinin sınıflama performansı üzerine etkisine bakılmış, sonraki üç analizde de farklı özellik seçme yöntemlerinin sınıflama performansı üzerine etkileri araştırılmıştır. Tüm analizlerde aynı sınıflama yöntemleri ve performans metrikleri kullanılmıştır.

Çizelge 4.1: Birinci Araştırma Problemi Kapsamında Yapılan Analizler

No	Ön işleme yöntemi		Sınıflama yöntemleri	Performans metrikleri	Sonuçlar
	Veri dönüştürme	Özellik seçme			
1	Yok	Yok	Naive Bayes	DSO Duyarlılık Seçicilik EAKA F-Ölçütü	Çizelge 4.2
2	Eşit frekans	Yok	Random Forest		Çizelge 4.3
3	Eşit genişlik	Yok	Destek Vektör M. Karar Ağacı		Çizelge 4.4
4	Eşit genişlik	Kazanç oranı	Yapay Sinir Ağları		Çizelge 4.5
5	Eşit genişlik	Gini indeksi	CN2 Kuralları		Çizelge 4.6
6	Eşit genişlik	DVM ağırlıkları	k En Yakın Komşuluk		Çizelge 4.7

Birinci analizde değişkenlere herhangi bir müdahale yapılmamıştır ve tüm değişkenler herhangi bir veri dönüştürme işlemi yapılmadan analize katılmıştır. Çizelge 4.2’de verilen analiz sonuçları incelendiğinde en yüksek doğru sınıflama oranına %83 ile yapay sinir ağları algoritması ulaşırken, diğer algoritmaların da %65 ve üzeri performans sergilediği görülmüştür.

Çizelge 4.2: Verilerin Sürekli Olduğu Durumda Yapılan Analiz Sonuçları

Yöntem	DSO	Duyarlılık	Seçicilik	EAKA	F-Ölçütü
<i>Karar ağacı</i>	0,65	0,48	0,73	0,63	0,49
<i>CN2 kuralları</i>	0,66	0,52	0,73	0,67	0,52
<i>Naive Bayes</i>	0,74	0,63	0,80	0,80	0,63
<i>Yapay Sinir Ağları</i>	0,82	0,67	0,90	0,90	0,72
<i>kNN</i>	0,76	0,67	0,82	0,84	0,67
<i>Random Forest</i>	0,75	0,56	0,86	0,85	0,61
<i>DVM</i>	0,80	0,59	0,92	0,85	0,68

VM çalışmalarında ön işleme süreci oldukça önemlidir. Bu aşamada yapılacak müdahalelerle modellerin tahmin performansları artabilmektedir. Bu nedenle birinci analizde bulunan sonuçlardan daha iyi sonuçlar alınıp alınamayacağını test etmek amacıyla iki farklı ön işleme yöntemi test edilmiştir. Birinci alt problem bağlamında (1.1) ön işleme sürecinde kullanılan farklı veri dönüştürme tekniklerinin sınıflama performansları üzerine etkisi olup olmadığı araştırılmıştır. Bu amaçla literatürde sıklıkla kullanılan iki farklı veri dönüştürme tekniği seçilerek elde edilen sınıflama modellerinin performansları karşılaştırılmıştır. Birinci analizde sürekli değişkenler eşit frekans yöntemi kullanılarak üç eşit frekansta parçaya ayrılmıştır ve analiz edilmiştir. İkinci analizde ise değişkenler eşit genişlik yöntemi kullanılarak üç eşit genişlikte parçaya ayrılmıştır ve aynı analiz bu değişkenlerle tekrar edilmiştir. Bu iki analize ilişkin sınıflama analizi sonuçları sırasıyla Çizelge 4.3 ve Çizelge 4.4’de verilmiştir. Veri dönüştürme yöntemlerinin performansları Çizelge 4.2’de sonuçları verilen taban model (herhangi bir müdahale yok) ile karşılaştırıldığında her iki veri dönüştürme işleminde de algoritmaların taban modele göre sınıflama performansını artırdığı görülmektedir (DVM algoritması hariç).

Çizelge 4.3: Verilerin Eşit Frekans Yöntemi ile Kesikli Hale Dönüştürüldüğü Durumda Yapılan Analiz Sonuçları

Yöntem	DSO	Duyarlılık	Seçicilik	EAKA	F-Ölçütü
<i>Karar ağacı</i>	0,75	0,56	0,86	0,73	0,61
<i>CN2 kuralları</i>	0,75	0,52	0,88	0,77	0,60
<i>Naive Bayes</i>	0,74	0,63	0,80	0,74	0,63
<i>Yapay Sinir Ağları</i>	0,83	0,74	0,88	0,89	0,75
<i>kNN</i>	0,73	0,63	0,78	0,79	0,62
<i>Random Forest</i>	0,79	0,63	0,88	0,86	0,68
<i>DVM</i>	0,78	0,37	1,00	0,86	0,54

Çizelge 4.4: Verilerin Eşit Genişlik Yöntemi ile Kesikli Hale Dönüştürüldüğü Durumda Yapılan Analiz Sonuçları

<i>Yöntem</i>	<i>DSO</i>	<i>Duyarlılık</i>	<i>Seçicilik</i>	<i>EAKA</i>	<i>F-Ölçütü</i>
<i>Karar ağacı</i>	0,79	0,81	0,78	0,76	0,73
<i>CN2 kuralları</i>	0,77	0,59	0,86	0,84	0,64
<i>Naive Bayes</i>	0,75	0,63	0,82	0,74	0,64
<i>Yapay Sinir Ağları</i>	0,83	0,78	0,86	0,84	0,76
<i>kNN</i>	0,77	0,67	0,82	0,82	0,67
<i>Random Forest</i>	0,78	0,67	0,84	0,87	0,68
<i>DVM</i>	0,75	0,44	0,92	0,85	0,56

Özellikle ağaç tabanlı algoritmalar (karar ağacı ve CN2 kuralları), verilerin kategorik hale dönüştürülmesi ile doğru sınıflama oranlarını %9 ile %14 arasında artırmıştır. Eşit frekans ve eşit genişlik yöntemine ilişkin sonuçlar karşılaştırıldığında ise eşit genişlik yönteminin özellikle dersten kalan öğrencilerin tahmin edilmesinde (duyarlılık) tüm algoritmalar açısından daha iyi performans sergilediği anlaşılmaktadır.

Bu araştırma problemi kapsamında cevap aranan bir diğer soru ise (1.2) ön işleme sürecinde kullanılan farklı özellik seçme tekniklerinin sınıflama performansları üzerine etkisidir. Özellik seçme işleminin amacı, tahmin analizlerinde hedef değişkenin tahmin edilmesinde daha önemli olan değişkenleri belirlemektir. Bu sayede tahmin işleminin tüm değişkenler yerine daha az sayıdaki değişkenle gerçekleştirilmesi mümkün olabilmektedir. Önemli değişkenleri belirlenmesi konusunda çeşitli algoritmalar kullanılmaktadır ve bunlar kullandıkları yönteme göre her bir değişkene bir puan atamaktadır. Daha sonra değişkenler bu puana göre sıralanarak seçilen belirli sayıda değişkenle analizler yapılmaktadır. Bu çalışmada özellik seçme işlemi Orange VM yazılımında bulunan üç farklı özellik seçme algoritmasına (kazanç oranı, gini indeksi, DVM ağırlığı) göre yapılmıştır. Özellik seçme analizi sonucu her bir değişkenin farklı algoritmalarından aldıkları puanlar Çizelge 4.5.'de sunulmuştur.

Çizelge 4.5: Üç Farklı Özellik Seçme Yöntemine Göre Değişkenlerin Önem Puanları

No	Değişken	Kazanç oranı	Gini indeksi	DVM ağırlığı
1	t_BenzersizIleti	0,196	0,076	0,064
2	t_SoruDegerlendirme	0,188	0,083	0,361
3	t_BenzersizOturum	0,180	0,078	0,190
4	t_Gezinim	0,158	0,042	0,193
5	t_EtiketKullanım	0,133	0,033	0,004
6	t_Ileti	0,128	0,056	0,141
7	t_Etiket	0,122	0,031	0,055
8	t_IletiGezinim	0,117	0,011	0,024
9	t_CevapOkuma	0,117	0,011	0,041
10	t_Cevap	0,116	0,028	0,049
11	t_BildirimGezinim	0,110	0,011	0,002
12	t_Oturum	0,110	0,037	0,066
13	t_CevapDegerlendirme	0,101	0,007	0,013
14	t_IletiDegerlendirme	0,093	0,024	0,007
15	t_KopyalaYapistir	0,084	0,016	0,007
16	t_IletiOkuma	0,078	0,014	0,004
17	t_OturumSuresi	0,078	0,024	0,039
18	o_TusKullanım	0,073	0,016	0,062
19	o_SilmeTusuKullanım	0,071	0,012	0,470
20	t_KaynakGezinim	0,067	0,018	0,037
21	t_DuyuruGezinim	0,056	0,008	0,007
22	t_TartismaGezinim	0,056	0,008	0,064
23	o_SayfaDeaktif	0,037	0,017	0,016
24	t_Yorum	0,036	0,002	0,013
25	t_YorumDegerlendirme	0,036	0,013	0,059
26	o_IletiYazmaSuresi	0,027	0,004	0,059
27	t_KopyalIletiSayisi	0,024	0,006	0,113

Analiz edilen veride hangi özellik seçme yönteminin daha iyi sonuç vereceğini test etmek amacıyla üç farklı analiz yapılmıştır. Birinci analizde kazanç oranı (Gain ratio) yöntemine göre en yüksek puan alan ilk 10 değişken seçilmiş ve tüm algoritmalar için performans metrikleri hesaplanmıştır. İkinci analizde Gini indeksine göre en yüksek puan alan ilk 10 değişken seçilmiş ve sınıflama analizleri tekrarlanmıştır. Son olarak DVM ağırlığı (SVM weight) yöntemine göre en yüksek puan alan 10 değişken seçilmiş ve aynı analizler tekrarlanmıştır. Her üç analize ilişkin sınıflama analizi sonuçları sırasıyla Çizelge 4.6, Çizelge 4.7 ve Çizelge 4.8’de verilmiştir. Tüm analizlerde değişkenler bir önceki deneyde en iyi sonucu veren eşit genişlik yöntemine göre kesikli hale dönüştürülmüş ve o şekilde analiz edilmiştir.

Analiz sonuçları bir önceki araştırma sorusundan bağımsız olarak incelendiğinde (Çizelge 4.6, Çizelge 4.7 ve Çizelge 4.8 karşılaştırıldığında) en yüksek doğru sınıflama oranına CN2 kuralları (%86) ve kNN (%86) algoritmaları ile Gini indeksine göre yapılan özellik seçme işlemi ile ulaşıldığı görülmektedir (Çizelge 4.7).

Çizelge 4.6: Kazanç Oranı Yöntemine Göre en Önemli 10 Değişken Seçilerek Yapılan Analiz Sonuçları

<i>Yöntem</i>	<i>DSO</i>	<i>Duyarlılık</i>	<i>Seçicilik</i>	<i>EAKA</i>	<i>F-Ölçütü</i>
<i>Karar ağacı</i>	0,76	0,78	0,76	0,75	0,70
<i>CN2 kuralları</i>	0,80	0,81	0,80	0,90	0,75
<i>Naive Bayes</i>	0,80	0,81	0,80	0,85	0,75
<i>Yapay Sinir Ağları</i>	0,82	0,74	0,86	0,87	0,74
<i>kNN</i>	0,83	0,78	0,86	0,91	0,76
<i>Random Forest</i>	0,78	0,67	0,84	0,87	0,68
<i>DVM</i>	0,78	0,59	0,88	0,90	0,65

Çizelge 4.7: Gini İndeks Yöntemine Göre en Önemli 10 Değişken Seçilerek Yapılan Analiz Sonuçları

<i>Yöntem</i>	<i>DSO</i>	<i>Duyarlılık</i>	<i>Seçicilik</i>	<i>EAKA</i>	<i>F-Ölçütü</i>
<i>Karar ağacı</i>	0,79	0,81	0,78	0,76	0,73
<i>CN2 kuralları</i>	0,86	0,85	0,86	0,92	0,81
<i>Naive Bayes</i>	0,76	0,74	0,78	0,84	0,69
<i>Yapay Sinir Ağları</i>	0,80	0,74	0,84	0,88	0,73
<i>kNN</i>	0,86	0,89	0,84	0,90	0,81
<i>Random Forest</i>	0,78	0,74	0,80	0,88	0,70
<i>DVM</i>	0,78	0,59	0,88	0,91	0,65

Çizelge 4.8: DVM Ağırlığı Yöntemine Göre en Önemli 10 Değişken Seçilerek Yapılan Analiz Sonuçları

<i>Yöntem</i>	<i>DSO</i>	<i>Duyarlılık</i>	<i>Seçicilik</i>	<i>EAKA</i>	<i>F-Ölçütü</i>
<i>Karar ağacı</i>	0,79	0,81	0,78	0,76	0,73
<i>CN2 kuralları</i>	0,83	0,74	0,88	0,89	0,75
<i>Naive Bayes</i>	0,80	0,74	0,84	0,81	0,73
<i>Yapay Sinir Ağları</i>	0,82	0,70	0,88	0,88	0,73
<i>kNN</i>	0,83	0,74	0,88	0,91	0,75
<i>Random Forest</i>	0,76	0,70	0,80	0,88	0,68
<i>DVM</i>	0,79	0,56	0,92	0,91	0,65

Bir diğer alt problem kapsamında özellik seçme işlemi sonucu en iyi performans gösteren ve daha az sayıda değişken kullanılarak oluşturulan sınıflama modellerinin performansı ile tüm değişkenler kullanılarak oluşturulan sınıflama modellerinin performansları karşılaştırılmıştır (1.3). Bu amaçla bir önceki analizde en yüksek sınıflama oranlarının yer aldığı Çizelge 4.4 ile önemli değişkenler seçilerek yapılan

analiz sonucu en iyi sınıflama oranlarının yer aldığı Çizelge 4.7 karşılaştırıldığında Gini indekse göre seçilen 10 değişken kullanılarak yapılan sınıflama sonuçlarının tüm değişkenlerin kullanıldığı duruma göre daha başarılı olduğu görülmektedir. Sonuçlar doğru sınıflama oranları açısından karşılaştırıldığında kNN ve CN2 kuralları algoritmalarının performanslarını %9 oranında artırdığı görülmektedir. Bunun dışında yapay sinir ağları algoritmasının doğru sınıflama oranında %3'lük bir düşme olurken diğer algoritmaların doğru sınıflama oranlarının değişmediği görülmüştür. Sonuçlar algoritmaların başarılı ve başarısız öğrencileri ayırt etme ölçüsünün bir göstergesi olan EAKA metriği açısından incelendiği zaman, karar ağacı dışında tüm algoritmaların daha az değişken kullanıldığı durumda ayırt ediciliklerini artırdığı görülmektedir.

Birinci araştırma problemi kapsamında cevap aranan bir diğer soru ise (1.4) öğrencilerin Bilgisayar Donanımı dersine ilişkin akademik performanslarının tahmin edilmesinde hangi değişkenlerin daha önemli olduğudur. Farklı özellik seçme yöntemlerinin karşılaştırıldığı bir önceki araştırma probleminde en iyi sınıflama performansının Gini indeksine göre belirlenen değişkenler ile elde edildiği görülmektedir. Çizelge 4.9'da Gini indeksine göre en önemli 10 değişken ve önem derecelerine ilişkin puanlara yer verilmiştir.

Çizelge 4.9: Gini İndeksine Göre en Önemli 10 Değişken

No	Değişken	Gini İndeksi
1	t_SoruDeğerlendirme	0,083
2	t_BenzersizOturum	0,078
3	t_BenzersizIleti	0,076
4	t_Ileti	0,056
5	t_Gezinim	0,042
6	t_Oturum	0,037
7	t_EtiketKullanım	0,033
8	t_Etiket	0,031
9	t_Cevap	0,028
10	t_OturumSuresi	0,024

Çizelge 4.9.'da sunulan bu değişkenler incelendiğinde öğrencilerin ortama giriş yapmalarının (t_Oturum), bu girişleri farklı günlerde yapmalarının (t_BenzersizOturum), ortamda harcadıkları sürenin (t_OturumSuresi), yaptıkları gezinimlerin (t_Gezinim), öğrendikleri kavramlarla ilgili ileti yazmalarının (t_Ileti), yazdıkları iletileri, öğrendikleri kavramları kullanarak etiketlemelerinin (t_EtiketKullanım), iletileri etiketlemek amacıyla kullanılacak yeni etiketler

oluşturmalarının (t_Etiket), yazdıkları iletileri farklı günlere yaymalarının ($t_BenzersizIleti$), tartışma ortamındaki sorulara cevap yazmalarının (t_Cevap) ve yazılan soruları değerlendirmelerinin ($t_SoruDeğerlendirme$) ders başarılarını tahmin etmede önemli değişkenler olduğu görülmektedir.

Birinci araştırma problemi kapsamında cevap aranan son soru ise (1.5) dersten başarısız olan öğrencilerin belirlenmesinde en iyi performans gösteren sınıflama algoritması/algoritmaları ve kullanılan ön işleme tekniklerinin seçilmesidir. Birinci araştırma probleminden elde edilen sonuçlar bu bağlamda incelendiğinde en iyi sınıflama oranına kNN ve CN2 Kuralları algoritmaları ile eşit genişlik yöntemi ve Gini özellik seçme yönteminin kullanıldığı durumda (Çizelge 4.7) ulaşıldığı tespit edilmiştir. Çizelge 4.10'da kNN algoritması ile elde edilen çapraz tablo sonuçları görülmektedir. Buna göre oluşturulan sınıflama modeli dönem sonunda başarısız olan 27 öğrenciden 24'ünü (%89), başarılı olan 49 öğrenciden de 41'ini (%84) doğru olarak tahmin etmiştir. Buna karşılık gerçekte başarılı olan 8 öğrenciyi (%16) başarısız, gerçekte başarısız olan 3 öğrenciyi de (%11) başarılı olarak sınıflamıştır.

Çizelge 4.10: kNN Algoritmasının Çapraz Tablo Sonuçları*

	<i>Kaldı</i>	<i>Geçti</i>	<i>Toplam</i>
<i>Kaldı</i>	24	3	27
<i>Geçti</i>	8	41	49
<i>Toplam</i>	31	45	76

*Sütunlar tahmin, satırlar gerçek sınıfları göstermektedir.

Bir diğer yüksek sınıflama oranına sahip algoritma olan CN2 Kurallar algoritmasının sonucu elde edilen örnek kurallar ise aşağıda verilmiştir.

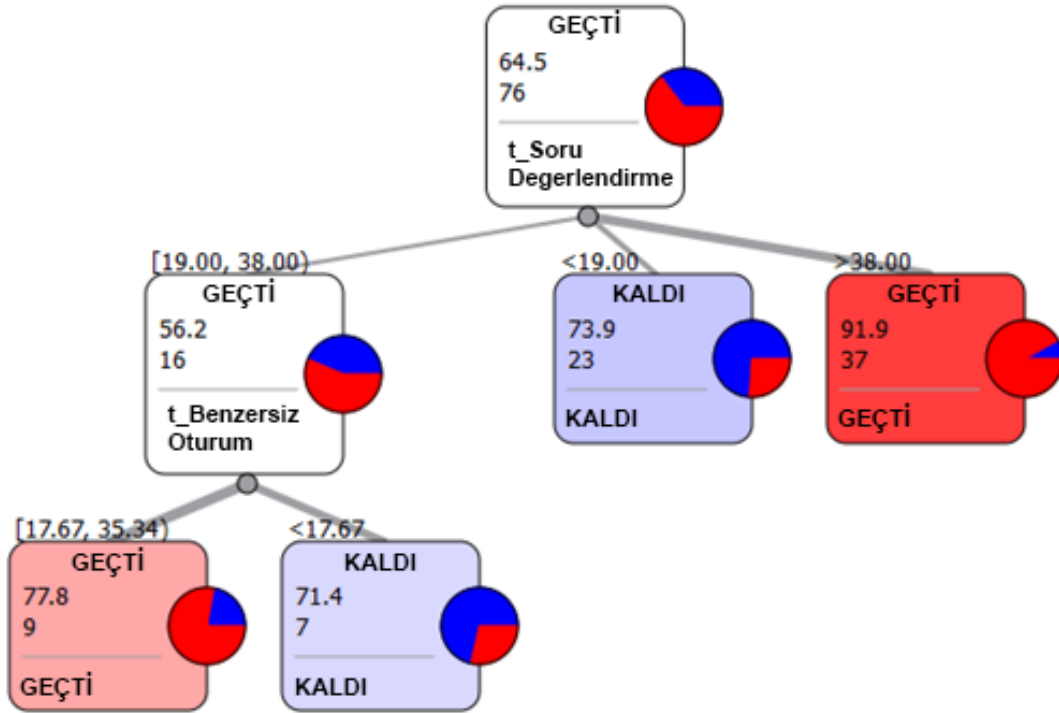
Kural 1: EĞER $t_BenzersizOturum = [17.00, 35.00)$ VE $t_BenzersizIleti = [4.00, 9.00)$ İSE SONUÇ = Geçti

Kural 2: EĞER $t_SoruDeğerlendirme = <19.00$ VE $BenzersizOturum = <17.00$ İSE SONUÇ = Kaldı

Birinci kural ortama giriş yaptığı farklı gün sayısı 17 ile 35 arasında olan ve ileti yazdığı farklı gün sayısı 4 ile 9 arasında olan öğrencilerin başarılı olduğunu göstermektedir. Bu koşulu sağlayan 21 öğrencinin tamamının başarılı olduğu görülmektedir. İki numaralı kural ise değerlendirmede bulunduğu soru sayısı 19'un altında olan ve ortama giriş yaptığı farklı gün sayısı 17'nin altında olan öğrencilerin

başarısız olduğunu söylemektedir. Bu koşulu sağlayan öğrenciler incelendiğinde ise koşulu sağlayan 21 öğrenciden 17'sinin (%81) başarısız olduğu görülmektedir.

Şekil 4.1.'de ise karar ağacı algoritması ile elde edilmiş örnek bir karar ağacına yer verilmiştir. Şekilde verilen karar ağacı incelendiğinde öğrencilerin iki farklı kırılma noktasına göre sınıflara ayrıldığı görülmektedir. Toplam soru değerlendirme sayısına göre yapılan birinci kırılmada ağaç üç dala ayrılmaktadır. Bu dallar incelendiğinde toplam soru değerlendirme sayısı 38'den fazla olan öğrencilerin dönem sonunda dersi geçme olasılıklarının %91.9 olduğu görülmektedir.



Şekil 4.1. Karar Ağacı Algoritması ile Elde Edilen Örnek Karar Ağacı

Toplam soru değerlendirme sayısı 19'un altında olan öğrencilerin ise %73.9 olasılıkla başarısız olduğu anlaşılmaktadır. Toplam soru değerlendirme sayısı 19 ile 38 arasında olan öğrenciler için ise bir diğer değişken olan ortama giriş yapılan benzersiz gün sayısına göre yeni bir bölünme olduğu görülmektedir. Buna göre ortama giriş yaptığı farklı gün sayısı 17'den az olan öğrencilerin %71.4 olasılıkla başarısız olduğu, 17 ile 35 arasında olanların da %77.8 olasılıkla başarılı olduğu anlaşılmaktadır.

4.2. İkinci Araştırma Problemine İlişkin Bulgular

Araştırma problemi 2: Birinci araştırma probleminde en iyi performans gösteren sınıflama algoritması ve ön işleme teknikleri kullanılarak dersten başarısız olacak öğrencilerin daha önceden tahmin edilmesi mümkün müdür?

Dönem sonunda başarısız olma ihtimali yüksek olan öğrencilerin daha erken haftalarda tahmin edilip edilemeyeceğinin araştırılması amacıyla öğrencilerin etkileşim verilerinin yer aldığı analiz veri tabanından 3, 6, 9 ve 12. haftalara ait veriler alınarak, oluşturulan sınıflama modellerinin performansı tüm veri kullanılarak (14. hafta) oluşturulan sınıflama modeli ile karşılaştırılmıştır. Modellerin oluşturulmasında birinci araştırma probleminde elde edilen bulgulardan yararlanılmış ve orada en iyi performans gösteren algoritma ve ön işleme teknikleri kullanılmıştır. Buna göre oluşturulan tüm modellerde birinci araştırma probleminde seçilen on değişken (Çizelge 4.9) ve en yüksek sınıflama performansına sahip olan kNN algoritması kullanılmıştır. Veriler eşit genişlik yöntemi kullanılarak kesikli hale dönüştürülmüştür. Sınıf değişkeni olan akademik performans değişkeni de birinci problemde olduğu gibi Kaldı (n = 22, başarı puanı ≤ 50) ve Geçti (n = 44, başarı puanı > 50) şeklinde kodlanmıştır.

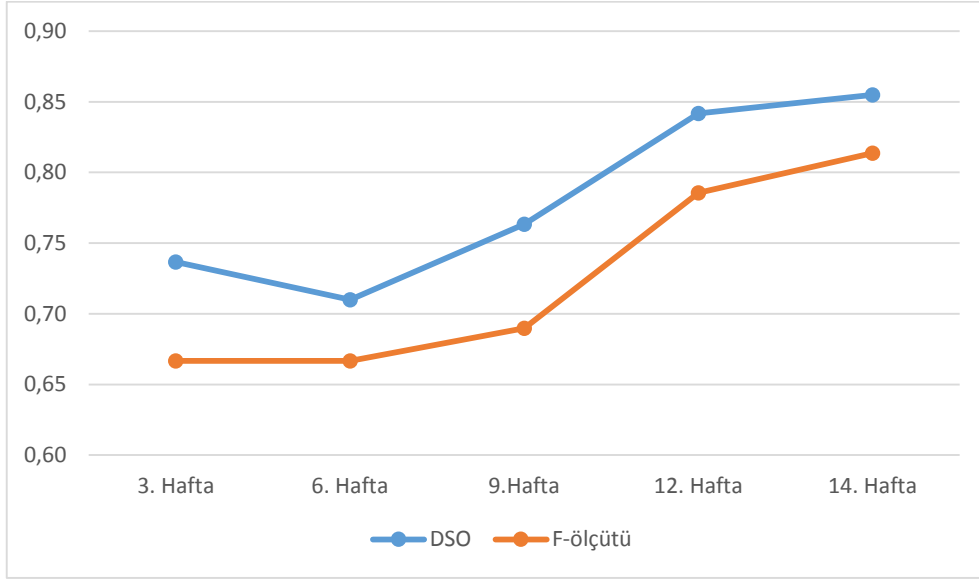
Belirlenen performans metrikleri açısından 3, 6, 9, 12 ve taban modele (14. Hafta) ilişkin sonuçlar Çizelge 4.11’de verilmiştir. Burada verilen sonuçlar 5-k Çapraz geçerlilik yöntemi ile elde edilen genelleştirilmiş sonuçlardır.

Çizelge 4.11: kNN Algoritmasına Göre Farklı Haftalarda Alınan Veri ile Oluşturulan Sınıflama Modellerine İlişkin Analiz Sonuçları

<i>Haftalar</i>	<i>DSO</i>	<i>Duyarlılık</i>	<i>Seçicilik</i>	<i>EAKA</i>	<i>F-Ölçütü</i>
3. Hafta	0,74	0,74	0,73	0,83	0,67
6. Hafta	0,71	0,81	0,65	0,70	0,67
9.Hafta	0,76	0,74	0,78	0,78	0,69
12. Hafta	0,84	0,81	0,86	0,82	0,79
14. Hafta	0,86	0,89	0,84	0,90	0,81

Bu araştırma problemi kapsamında başarısız öğrencilerin tahmin edilmesi amaçlandığı için Çizelge 4.11’de farklı haftalara ait sınıflama modellerinin çapraz geçerlilik işlemi sonucu elde edilen “F-Ölçütü” değerleri ve modellerin genel performansının değerlendirilmesi açısından “DSO” değerleri dikkate alınmıştır.

Şekil 4.2’de görsel olarak sunulan bu iki metriğe ilişkin grafik incelendiğinde 6. Hafta dışında oluşturulan sınıflama modellerinin doğru sınıflama oranının artış eğilimi gösterdiği görülmektedir.



Şekil 4.2. Farklı Haftalara Ait Sınıflama Performansı

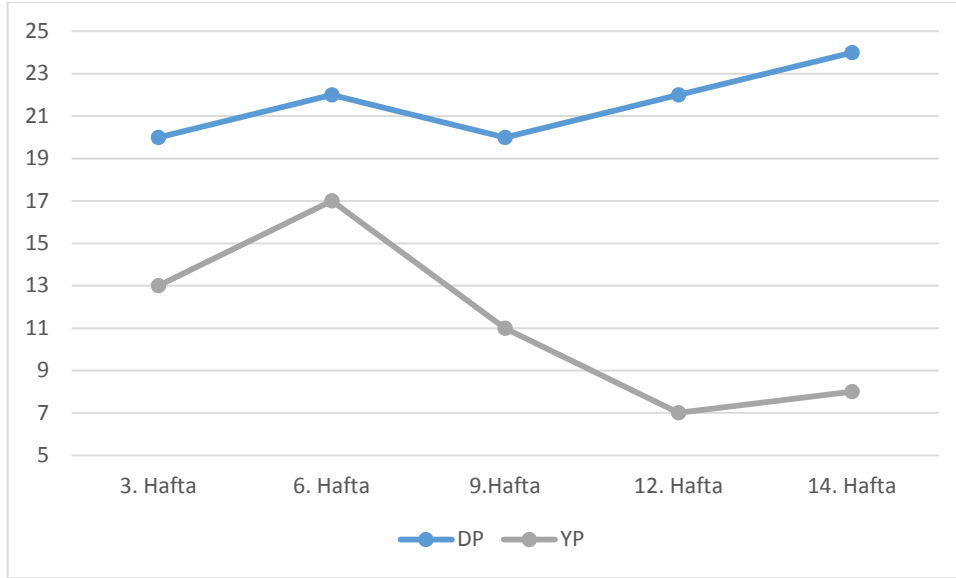
14 Haftalık dönem bittikten sonra alınan veri ile oluşturulan model öğrencilerin %86’sını doğru sınıflarken 3. Haftada alınan veri ile oluşturulan model %74’ünü doğru olarak sınıflamaktadır. Diğer haftalarda da bu oranın %71 ile %84 arasında değiştiği görülmektedir. Aynı şekilde modellerin başarılı ve başarısız öğrencileri ayırt etme gücünün bir göstergesi olan F-Ölçütü değerinin de 6. Hafta dışında bir artış eğilimi gösterdiği anlaşılmaktadır. Elde edilen sınıflama modellerinin haftalık olarak nasıl bir performans sergilediğinin daha iyi anlaşılabilmesi için Çizelge 4.12’de sınıflama modellerine ait çapraz tablolara yer verilmiştir.

Çizelge 4.12: KNN Algoritmasına Göre Farklı Haftalarda Alınan Veri ile Oluşturulan Sınıflama Modellerine İlişkin Çapraz Tablo Değerleri

<i>Haftalar</i>	<i>DP</i>	<i>YP</i>	<i>DN</i>	<i>YN</i>
3. Hafta	20	13	36	7
6. Hafta	22	17	32	5
9. Hafta	20	11	38	7
12. Hafta	22	7	42	5
14. Hafta	24	8	41	3

Şekil 4.3.’de görsel olarak verilen bu tablo incelendiğinde haftalık olarak oluşturulan sınıflama modellerinin dersten başarısız olan 27 öğrenciden 20’sini (%74) 3. Haftadan itibaren doğru olarak sınıfladığı görülmektedir (DP değeri). Son hafta ise 24’ünü (%89) doğru olarak sınıfladığı görülmektedir. Fakat bu değer tek başına

yorumlanması hataya neden olabilmektedir çünkü model her defasında başarısız öğrencileri sınıflarken bir takım başarılı öğrencileri de hatalı olarak başarısız sınıfına atamaktadır (Yanlış Pozitif). Çizelge 4.12’de YP olarak ifade edilen sütunda haftalık olarak bu değerlere yer verilmiştir.



Şekil 4.3. Farklı Haftalara ait Sınıflama Performansı

Şekil 4.3’de görsel olarak sunulan bu değerler incelendiğinde üçüncü hafta oluşturulan sınıflama modelinin normalde dersi geçen 49 öğrenciden 13’ünü (%26) “Kaldı” olarak sınıfladığı son hafta ise bu sayının 8 (%16)’e düştüğü görülmektedir. Bütün haftalar genel olarak incelendiğinde ise 6. Hafta dışında kullanılabilir modeller elde edildiği görülmektedir.

4.3. Üçüncü Araştırma Problemine İlişkin Bulgular

Araştırma problemi 3: Öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak benzer kullanım profiline sahip öğrenciler gruplanabilir mi?

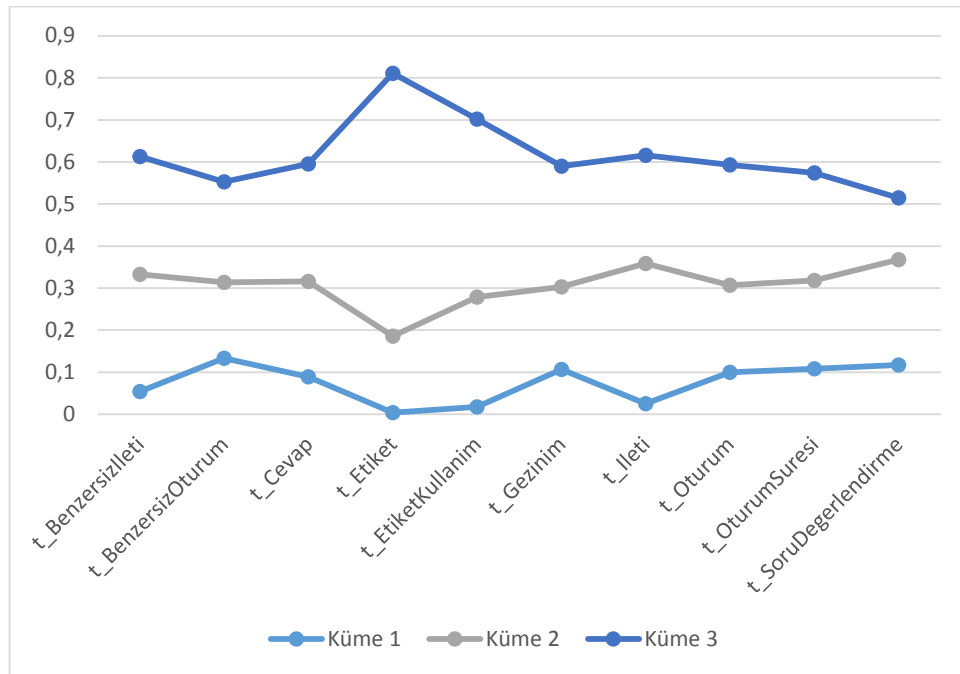
Benzer özellikteki öğrencileri etkileşim verilerine göre gruplamak amacıyla VM yöntemlerinden Kümeleme analizi kullanılmıştır. Kümeleme analizi yaygın olarak kullanılan X-Means (X-Ortalamlar) ve EM (Expectation Maximization) kümeleme algoritmaları kullanılarak gerçekleştirilmiştir. Her iki algortmada da ideal küme sayısının belirlenmesi için çapraz geçerlilik yöntemi kullanılmıştır ve analizler birinci araştırma probleminde seçilen 10 değişken (önemli değişkenler) ile yapılmıştır. Birinci alt problem kapsamında (3.1) X-Ortalamlar kümeleme algoritması ile yapılan analize göre öğrencilerin kaç farklı gruba ayrılacağı ve küme merkezlerine göre bu gruplar nasıl tanımlanabileceği araştırılmıştır. X-Ortalamlar algoritması ile

yapılan analiz sonucu verinin ideal olarak 3 farklı kümeye ayrılacağı tespit edilmiştir. Çizelge 4.13.'de bu üç kümede yer alan öğrencilerin analizde kullanılan değişkenler açısından ortalama değerlerine yer verilmiştir.

Çizelge 4.13: X-Ortalamlar Algoritmasına Göre Küme Ortalamaları

<i>Değişkenler</i>	<i>Küme 1</i>	<i>Küme 2</i>	<i>Küme 3</i>
	n = 21	n = 32	n = 23
<i>t_Benzersizleti</i>	0,74	4,57	8,41
<i>t_BenzersizOturum</i>	8,26	19,52	34,31
<i>t_Cevap</i>	0,96	3,38	6,38
<i>t_Etiket</i>	0,09	4,57	19,97
<i>t_EtiketKullanım</i>	2,52	38,24	96,13
<i>t_Gezinim</i>	463,35	1310,86	2554,25
<i>t_Ileti</i>	3,78	53,29	91,38
<i>t_Oturum</i>	13,74	42,14	81,31
<i>t_OturumSuresi</i>	521,22	1536,71	2777,78
<i>t_SoruDeğerlendirme</i>	10,61	33,33	46,75

Şekil 4.4'de küme ortalamalarının 0 - 1 aralığında normalleştirilmesi ile elde edilen grafik incelendiğinde Küme 1'de yer alan öğrenciler ortama az sayıda giriş yapan, ortamda az zaman geçiren, tartışmalara katılmayan, öğrendikleri kavramlarla ilgili az sayıda ileti yazan öğrenciler olarak tanımlanabilir.



Şekil 4.4. X-Ortalamlar Algoritmasına Göre Normalleştirilmiş Küme Ortalamaları

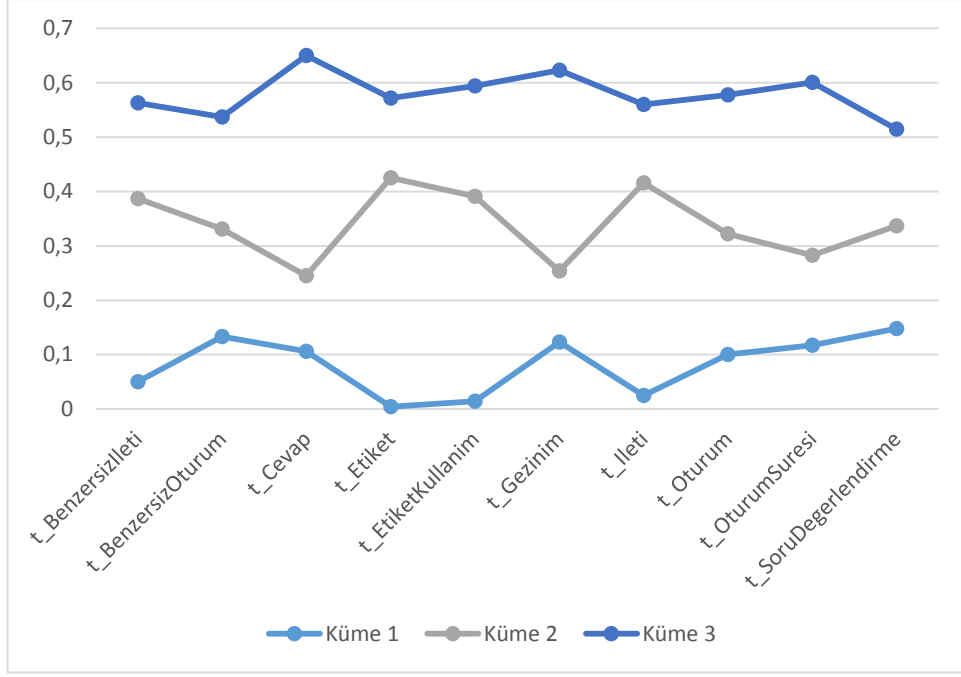
Küme 3 ise tam tersi şekilde ortama düzenli olarak giriş yapan, tartışmalara aktif olarak katılan, ortamda diğerlerine göre daha fazla zaman geçiren öğrenciler olarak tanımlanabilir. Küme 2’de yer alan öğrenciler ise ortamda orta düzeyde aktivite sergileyen, ortalama değerleri Küme 1’de yer alan öğrencilerden daha yüksek, Küme 3’de yer alan öğrencilerden daha düşük olan öğrencilerin yer aldığı küme olarak tanımlanabilir.

Bir diğer alt problem kapsamında (3.2) aynı analiz EM kümeleme algoritması ile tekrar edilmiştir. EM algoritması kullanılarak yapılan kümeleme analizi sonucu da verinin ideal olarak 3 farklı kümeye ayrılacağı tespit edilmiştir. Çizelge 4.14.’de bu kümelerde yer alan öğrencilere ilişkin değişken ortalamaları yer almaktadır. Şekilde 4.5.’de ise normalleştirilmiş küme ortalamalarının grafiği verilmiştir. Değişken ortalamaları incelendiğinde burada da elde edilen kümelerin isimlendirilmesinin X-Ortalamlar algoritması ile aynı şekilde yapılabileceği görülmektedir.

Çizelge 4.14: EM Kümeleme Algoritmasına Göre Küme Ortalamaları

<i>Değişkenler</i>	<i>Küme 1</i>	<i>Küme 2</i>	<i>Küme 3</i>
	n = 17	n = 27	n = 32
<i>t_Benzersizleti</i>	0,82	6,34	9,24
<i>t_BenzersizOturum</i>	9,67	24,13	39,12
<i>t_Cevap</i>	1,44	3,34	8,88
<i>t_Etiket</i>	0,11	13,38	18,00
<i>t_EtiketKullanım</i>	2,41	66,97	101,71
<i>t_Gezinim</i>	671,26	1379,72	3390,94
<i>t_Ileti</i>	4,33	73,13	98,41
<i>t_Oturum</i>	16,59	53,69	96,29
<i>t_OturumSuresi</i>	692,67	1679,88	3570,00
<i>t_SoruDeğerlendirme</i>	15,37	34,94	53,35

Farklı algoritmalarla yapılan kümeleme analizi sonucu elde edilen küme ortalamaları incelendiğinde kümelerde yer alan öğrenci sayılarında ve küme ortalamalarında farklılıklar olmakla birlikte her iki algoritmanın da öğrencileri benzer özellikte üç farklı kümeye ayırdığı görülmektedir.



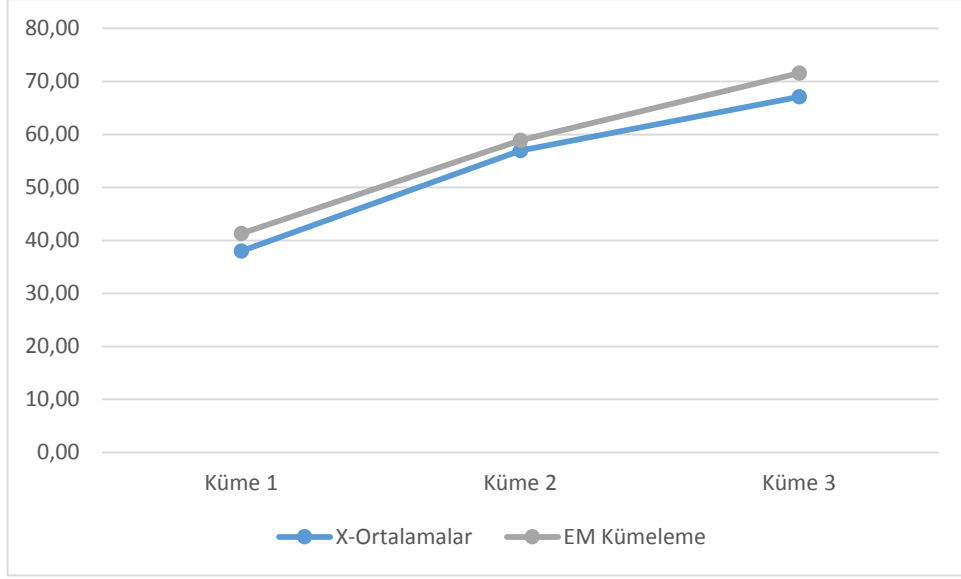
Şekil 4.5. EM Algoritmasına Göre Normalleştirilmiş Küme Ortalamaları

Bir diğer alt problem kapsamında (3.3) her iki kümeleme algoritmasına göre elde edilen farklı kümeler, bu kümelere yer alan öğrencilerin ortalama ders başarıları ile ilişkili olarak tanımlanabilir mi sorusuna cevap aranmıştır. Çizelge 4.15.'de farklı kümeleme algoritmalarına göre kümelere yer alan öğrencilerin ortalama ders başarıları verilmiştir.

Çizelge 4.15: EM ve X-Ortalamalar Kümeleme Algoritmalarına Göre Kümelere Yer Alan Öğrencilerin Ortalama Başarı Puanları

<i>Küme</i>	<i>EM Kümeleme</i>	<i>X-Ortalamalar</i>
<i>Küme 1</i>	41,33	38,00
<i>Küme 2</i>	58,88	56,95
<i>Küme 3</i>	71,59	67,09

Buna göre her iki kümeleme algoritmasının da benzer sonuçlar ürettiği görülmektedir. Şekil 4.6'da da görsel olarak verilen bu ortalamalar incelendiğinde her iki algoritma için de Küme 1'de yer alan öğrencilerin derste düşük başarı gösteren öğrenciler olduğu, Küme 2'de yer alan öğrencilerin ortalama başarıya sahip öğrenciler olduğu ve Küme 3'de yer alan öğrencilerin de diğerlerine göre daha başarılı öğrenciler olduğu görülmektedir.

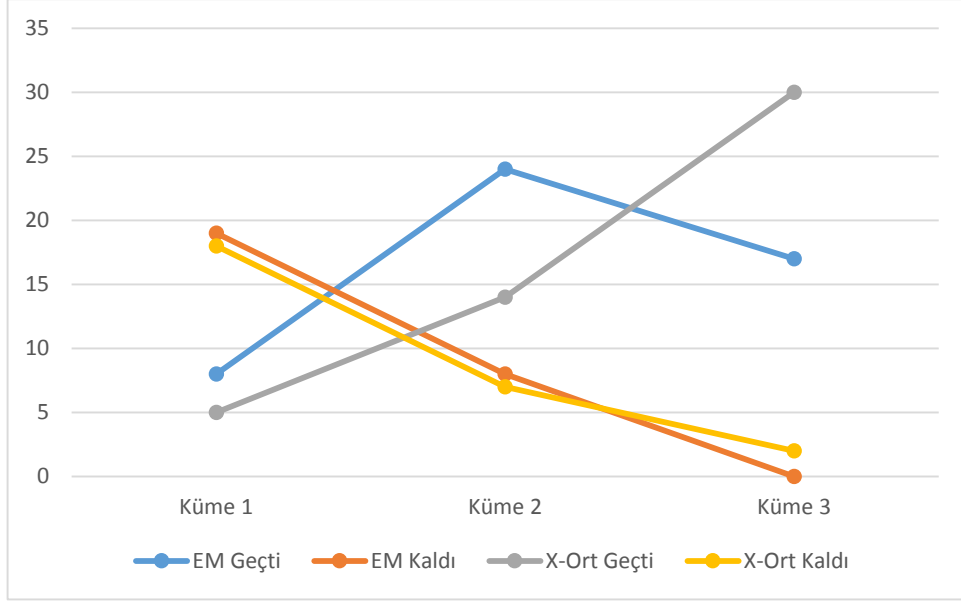


Şekil 4.6. EM ve X-Ortalamlar Kümeleme Algoritmalarının Ders Başarısına Göre Karşılaştırması

Son alt problem kapsamında (3.4) dersi geçen ve kalan öğrencilerin kümelere göre dağılımları incelendiğinde (Çizelge 4.16.) EM Kümeleme algoritmasına göre Küme 1’de yer alan 27 öğrenciden 8’inin dersi geçtiği 19’unun kaldığı, Küme 2’de yer alan 32 öğrenciden 24’ünün dersi geçtiği, 8’inin kaldığı, Küme 3’de yer alan 17 öğrencinin ise tamamının dersi geçtiği görülmektedir. X-Ortalamlar kümeleme algoritmasına göre ise Küme 1’de yer alan 23 öğrenciden 5’inin dersi geçtiği 18’inin kaldığı, Küme 2’de yer alan 21 öğrenciden 14’ünün dersi geçtiği, 7’sinin kaldığı, Küme 3’de yer alan 32 öğrenciden ise 30’unun dersi geçtiği, 2’sinin kaldığı görülmektedir. Şekil 4.7.’de görsel olarak da verilen bu sonuçlar incelendiğinde algoritmaların benzer sonuçlar ürettiği görülmektedir.

Çizelge 4.16: Dersi Geçen ve Dersten Kalan Öğrencilerin Kümelere Göre Dağılımı

<i>Küme</i>	<i>EM Kümeleme</i>			<i>X-Ortalamlar</i>		
	<i>N</i>	<i>Geçti</i>	<i>Kaldı</i>	<i>N</i>	<i>Geçti</i>	<i>Kaldı</i>
<i>Küme 1</i>	27	8	19	23	5	18
<i>Küme 2</i>	32	24	8	21	14	7
<i>Küme 3</i>	17	17	0	32	30	2



Şekil 4.7. EM ve X-Ortalama Kümeleme Algoritmalarının Ders Başarısına Göre Karşılaştırması

İlerleyen bölümde, analizler sonucu elde edilen bulgular ilgili alan yazını doğrultusunda tartışılmış, ileriki araştırmalara ve uygulamalara yönelik önerilere yer verilmiştir.

5. SONUÇ ve ÖNERİLER

5.1. Sonuçlar

Çevrimiçi öğrenme ortamları öğretme ve öğrenmeyi destekleyecek çok sayıda araca sahiptir. Öğretmenler bu araçları kullanarak ders materyallerini paylaşabilirler, ödevler ve testler hazırlayabilirler, tartışmalar yürütebilirler (Cristobal Romero, Espejo, Zafra, Romero, & Ventura, 2013), öğrenciler ise bu ortamlarda viki, blog, mesajlaşma vb. araçları kullanarak işbirlikli çalışmalar yapabilirler (Moreno, Gonzalez, Castilla, Gonzalez, & Sigut, 2007) ya da kendi öğrenme etkinliklerini planlayabilirler. Geleneksel sınıf ortamlarından farklı olarak çevrimiçi öğrenme ortamları, öğrenciler tüm bu etkinlikleri gerçekleştirirken geride bıraktıkları izlerin kayıt edilmesine olanak sağlamaktadır. Bu izler; öğrencinin dersle ilgili duyuruları takip etmesi, tartışma ortamında bir soruya cevap yazması ya da sadece ortama giriş – çıkış yapması olabilir. Yapılan çalışmalar, öğrenme süreçleri ile ilgili olarak veri tabanlarına kayıt edilen bu tür verilerin EVM ve ÖA yaklaşımlarıyla analiz edilmesiyle çok sayıda eğitsel probleme cevap bulunabileceğini göstermektedir. Bu çalışma kapsamında bu problemlerden bir tanesi olan öğrenci performansının modellenmesi ele alınmıştır.

Çalışma kapsamında kullanılan veriler araştırmacılar tarafından tasarlanan ve geliştirilen öğrenme ortamından elde edilmiştir. Öğrencilerin ders dışı etkinlikleri gerçekleştirmek amacıyla kullandıkları bu ortamın veri tabanı 76 öğrencinin Bilgisayar Donanımı dersine ait 14 haftalık kullanım verisini içermektedir. Bu veriler kullanılarak öğrencilerin öğrenme süreçleri ile ilişkili 27 adet değişken üretilmiştir. Bu değişkenler VM teknikleri ile analiz edilerek öğrencilerin akademik performanslarının modellenmesi amaçlanmıştır.

Araştırmanın birinci problemi, öğrencilerin çevrimiçi öğrenme ortamındaki etkileşim verileri kullanılarak bilgisayar donanımı dersine ilişkin dönem sonu akademik performanslarının tahmin edilmesi üzerinedir. Öğrencilerin akademik performansının tahmin edilmesi popüler bir EVM uygulamasıdır (Bousbia & Belamri, 2014; C. Romero & Ventura, 2010). Bugüne kadar bu amaçla birçok farklı ortamda, birçok farklı değişken kullanılarak çeşitli tahmin modelleri oluşturulmuştur. Algoritmaların performansı veri setinin yapısına göre ve belirlenen değişkenlere göre değişiklik gösterebileceği için yeni bir veri setine en iyi sonucu verecek tahmin

algoritmasının analiz yapılmadan seçilmesi kolay bir iş değildir (Akçapınar ve diğerleri, 2013; Osmanbegović & Suljić, 2012; Cristobal Romero ve diğerleri, 2010). Bu nedenle veri setine uygun sınıflama algoritmasının seçilmesi pratikte büyük öneme sahiptir (C Romero ve diğerleri, 2013). Bu amaçla kullanılan en yaygın yöntem ise farklı sınıflama algoritmalarının test edilerek en iyi sınıflama performansını gösteren algoritmanın seçilmesi şeklindedir. Bu çalışmada da benzer bir yaklaşım izlenerek literatürde en çok yer verilen yedi farklı sınıflama algoritmasının öğrenci başarısının tahmin edilmesindeki performansları karşılaştırılmıştır. Aynı zamanda veri ön işleme sürecinde, sınıflama performansı üzerinde etkisi olduğu bilinen veri dönüştürme ve özellik seçme amacıyla kullanılan farklı yöntemlerin de sınıflama performansı üzerindeki etkisi test edilmiştir. Elde edilen farklı sınıflama modellerinin performanslarının karşılaştırılması amacıyla performans metriklerinden yararlanılmıştır. Hedef değişken olan öğrenci performansının göstergesi olarak öğrencilerin Bilgisayar Donanımı dersindeki dönem sonu puanları alınmıştır. Öğrencilerin dönem sonu performansını tahmin etmek amacıyla çevrimiçi öğrenme ortamındaki davranışlarını yansıtan çok sayıda değişken kullanılmıştır. Bu değişkenlerin belirlenmesinde öğrencilerin öğrenme ortamındaki rolleri dikkate alınmıştır. Ders başarısı ise “Geçti”, “Kaldı” şeklinde kodlanarak kullanılmıştır.

Analiz aşamasında ilk olarak veride herhangi bir değişiklik yapılmadan algoritmaların sınıflama performansları test edilmiştir. Daha sonra ön işleme sürecinde veri dönüştürme işleminin sınıflama performansı üzerine etkisi test edilmiştir. Bu amaçla tüm değişkenler iki farklı yöntemle göre (eşit frekans, eşit genişlik) kategorik hale dönüştürülmüştür. Yapılan çalışmalar, kesikli verilerin sınıflama görevlerinde özellikle ağaç tabanlı algoritmalarda sürekli verilere göre daha iyi sonuç verdiği işaret etmektedir (Cristobal Romero ve diğerleri, 2013). Analiz sonuçları veri dönüştürme işlemi yapılmadan elde edilen sonuçlarla karşılaştırıldığında en yüksek doğru sınıflama oranına verilerin eşit genişlik yöntemine göre kesikli hale dönüştürüldüğü durumda ulaşıldığını göstermiştir. Bu nedenle diğer araştırma problemlerinde veriler bu yöntemle göre kesikli hale dönüştürülerek kullanılmıştır.

Tahmin analizlerinin bir diğer kullanım alanı da hedef değişkeni tahmin edecek modeller oluşturmanın yanında bu süreçte diğerlerine göre daha önemli rol oynayan

değişkenlerin belirlenmesidir (R. S. J. d Baker, 2010). Bu sayede elimizdeki tüm değişkenleri kullanmak yerine daha az sayıda değişken ile tahmin modelleri oluşturulabileceği gibi aynı zamanda tahmin edilecek değişken üzerinde hangi değişkenlerin daha önemli olduğunu da belirleme fırsatımız olmaktadır. Bu çalışmada da önemli değişkenleri belirlemek amacıyla üç farklı özellik seçme yöntemine göre en yüksek puan alan 10 değişken seçilmiş ve elde edilen sınıflama modellerinin performansı tüm değişkenlerin kullanıldığı durum (27 değişken) ile karşılaştırılmıştır. Elde edilen sonuçlar Gini indeksi yöntemine göre seçilen 10 değişken ile elde edilen sınıflama performanslarının tüm değişkenler kullanılarak elde edilen sınıflama performansından daha yüksek olduğunu göstermiştir. Çizelge 4.9.'da sunulan bu değişkenler incelendiğinde öğrencilerin ortama giriş yapmalarının (t_Oturum), bu girişleri farklı günlerde yapmalarının, ortamda harcadıkları sürenin (t_OturumSuresi), yaptıkları gezinimlerin (t_Gezinim), öğrendikleri kavramlarla ilgili ileti yazmalarının (t_Ileti), yazdıkları iletileri, öğrendikleri kavramları kullanarak etiketlemelerinin (t_EtiketKullanim), iletileri etiketlemek amacıyla kullanılacak yeni etiketler oluşturmalarının (t_Etiket), yazdıkları iletileri farklı günlere yaymalarının (t_BenzersizIleti), tartışma ortamındaki sorulara cevap yazmalarının (t_Cevap) ve yazılan soruları değerlendirmelerinin (t_SoruDeğerlendirme) ders başarılarını tahmin etmede önemli değişkenler olduğu görülmüştür.

Öğrencilerin ortama katılımlarını gösteren değişkenlerden yazdıkları ileti sayısı (t_Ileti), ortama giriş sayısı (t_Oturum) ve ortamda geçirdikleri toplam süre (t_OturumSuresi) performanslarının tahmin edilmesinde önemli değişkenlerdir. Diğer taraftan öğrencilerin ortama giriş yaptığı farklı gün sayısını gösteren t_BenzersizOturum ve öğrencilerin öğrendikleri kavramlarla ilgili olarak ileti yazdıkları farklı gün sayısını gösteren t_BenzersizIleti değişkenleri de performanslarının tahmin edilmesinde önemli değişkenler olarak ortaya çıkmaktadır. Şekil 3.7, 3.8 ve 3.9'da verilen öğrencilerin ortamı kullanım alışkanlıklarını gösteren grafikler incelendiğinde özellikle sınav haftalarından önceki günlerde ortama girişlerin sıklaştığı görülmektedir. Bu girişler öğrencilerin toplam aktivite sayısını artırmaktadır fakat başarılarına aynı ölçüde yansımayaabilmektedir. Zamana bağlı bu tür değişkenlerin önemli değişkenler arasında yer alması bu açıdan önemlidir.

Etiket kullanımı ile öğrenci performansı arasındaki ilişkisinin incelendiği bir çalışmaya rastlanmamıştır; ancak Lavoué (2011), çevrimiçi öğrenme ortamlarında kullanılan etiketleme özelliğinin yeni kavramların öğrenilmesinde, kavramlar arası ilişkilerin oluşturulmasında ve öğrencilerin kendi iç kavramlarının yansımalarını görmeleri açısından bilişsel ve sosyal öğrenme süreçlerini desteklemede önemli olduğunu belirtmiştir. Bu çalışmada da yazılan içeriklerin etiketlenmesi (t_EtiketKullanım) ve ortama yeni etiketlerin eklenmesi (t_Etiket) öğrenci performansının tahmin edilmesinde önemli değişkenler olarak ortaya çıkmaktadır. Uygulamanın yapıldığı Bilgisayar Donanımı dersinin kavram ağırlıklı bir ders olduğu düşünülürse etiketleme ile ilgili değişkenlerin ön plana çıkması anlam kazanacaktır.

Lopez ve diğerleri (2012), Moodle öğrenme ortamında gerçekleştirdikleri çalışmalarında öğrencilerin tartışma ortamına katılımlarının ders başarılarının iyi birer yordayıcısı olduğu sonucuna ulaşmışlardır. Bu bağlamda belirledikleri önemli değişkenler ise öğrenci tarafından yazılan mesaj sayısı, yazılan cevap sayısı, kullanılan kelime sayısı ve mesajın aldığı ortalama puandır. Macfadyen ve Dawson (2010) ise öğrencilerin Blackboard Vista öğrenme ortamındaki etkileşim verilerini kullanarak ders başarılarını tahmin etmek amacıyla bir Regresyon modeli oluşturmuşlardır. Kalma riski olan öğrencilerin önemli bir çoğunluğunu doğru olarak tahmin eden modeldeki en önemli değişken bu çalışmada da tartışma ortamına gönderilen mesaj sayısı olarak bulunmuştur. Araştırmacılar bu sonucu öğrenmenin sosyal bir süreç olduğunu destekleyen bir bulgu olarak sunmuşlardır. Bu çalışmada da öğrencilerin tartışma ortamına katılımlarının bir göstergesi olan cevap yazma sayıları (t_cevap) ve yazılan soruları değerlendirme sayıları (t_SoruDeğerlendirme) öğrenci performansının tahmin edilmesinde önemli değişkenler olarak ön plana çıkmıştır. Tartışma ortamında soru sorma sayıları ile ilgili değişken ise öğrenciler tarafından yeterli katılım olmadığı için araştırmadan çıkartılmıştır. İleriki araştırmalarda öğrencilerin tartışma ortamını daha aktif olarak kullanmaları sağlanarak bu değişkenin de performans üzerine etkisi araştırılabilir.

Birinci araştırma problemi kapsamında aynı zamanda Bilgisayar Donanımı dersinden kalan öğrencilerin tahmin edilmesinde en iyi performans gösteren sınıflama algoritması/algoritmaları ve kullanılan ön işleme tekniklerinin neler olduğu incelenmiştir. Sınıflama açısından farklı algoritmaların performansları karşılaştırıldığında ise en yüksek doğru sınıflama oranına kNN (%86) ve CN2

Kuralları (%86) ile ulařılırken diđer sınıflama algoritmalarının da %76 ve üzeri performans göstermiřtir. Burada hangi sınıflama algoritmasının seileceđi ise kullanım amacına gre deđiřiklik gsterebilmektedir. Bu nedenle tahmin analizleri ile ilgili bulgular verilirken farklı algoritmaların sonuları dikkate alınmıřtır. rneđin kNN algoritması ile yksek sınıflama oranları elde edilebilmektedir; fakat “kara kutu” olarak da anılan alıřma řekli nedeniyle sınıflama srecinin uzman olmayan kiřiler tarafından anlařılması zor olabilmektedir (Cristobal Romero ve diđerleri, 2010). Diđer taraftan ađa tabanlı algoritmalar (karar ađacı, CN2 kuralları, vb.) karar ađacı ya da eđer-ise kuralları gibi uzman olmayan kiřiler tarafından da kolaylıkla kullanılabilir, anlařılabilir ıktılar retmektedir (Enrique, Romero, Ventura, Carlos de, & Toon, 2010). Burada amacımız sadece đrencileri sınıflamak ise kNN algoritması tercih edilebilir; fakat đrencileri bařarısızlıđa gtren deđiřkenleri belirlemek ve bunlara mdahale etmek istersek CN2 kuralları ya da karar ađacı algoritması ile elde edilen sonuları kullanabiliriz. kNN algoritması ile elde edilen sınıflama modeli, dnem sonunda bařarısız olan 27 đrenciden 24’n (%89), bařarılı olan 49 đrenciden de 41’ini (%84) dođru olarak tahmin etmiřtir. Buna karřılık gerekte bařarılı olan 8 đrenciyi (%16) bařarısız, gerekte bařarısız olan 3 đrenciyi de (%11) bařarılı olarak sınıflamıřtır. İleriki arařtırmalarda bu modeller yeni đrencilerin sınıflandırılmasında kullanılabilir.

İkinci arařtırma problemi kapsamında ise đrencilerin dnem sonu performansının daha nceden tahmin edilip edilemeyeceđi arařtırılmıřtır ve bu amala 3, 6, 9, 12. haftalarda alınan verilerin đrenci bařarısını tahmin etmedeki performansı incelenmiřtir. Analiz sonuları 3. Haftada oluřturulan modelin dnem sonunda dersten bařarısız olan 27 đrenciden 20’sini (%74) dođru olarak sınıfladıđını gstermektedir. Aynı model dersi geen 49 đrenciden 13’n (%26) de kaldı olarak sınıflamıřtır. İlerleyen haftalarda alınan verilerle oluřturulan modellerde ise bu oran bařarısız đrencilerin tahmin edilmesi ynnde artarken, bařarılı đrencilerin bařarısız olarak sınıflanma oranında da dřme gzlenmiřtir. Son hafta elde edilen en iyi sınıflama modelinde ise bu oran 24’e (%89) 8 (%16) olarak elde edilmiřtir. Burada elde edilen sonular đrencilerin akademik performansının nceden tahmin edilebileceđini gstermektedir.

Eđitsel veri madenciliđi ve đrenme analitiđi alıřmalarının en nemli hedeflerinden bir tanesi de dřk performanslı ya da problem yařayan đrencileri belirlemek ve

zamanında yapılacak müdahalelerle buna neden olan davranışlarını deęiştirme olanaęı sağlamaktır (Kimberly E. Arnold & Pistilli, 2012). Davranışların deęişmesinde geri bildirimler son derece önemli deęişkenlerdir ve eğitim teknolojilerindeki gelişmelerle birlikte öğrencilere anlık ve daha sık dönüt verilmesi mümkün olmaktadır (Bienkowski ve dięerleri, 2012; Tanes, Arnold, King, & Remnet, 2011). Purdue üniversitesinde uygulanan Sinyal projesi, öğrencilerin anlayacağı şekilde verilen bireysel geri bildirimlerin öğrenciler tarafından olumlu şekilde algılandığını göstermektedir (Bienkowski ve dięerleri, 2012). Araştırmacılar yaptıkları çalışmada VM analizleri ile elde edilen bilgileri öğrencilere geri bildirim vermek amacıyla kullanmış ve bunun öğrenci başarısını artırmada etkili olduğunu göstermişlerdir (K. E. Arnold, 2010). Bu nedenle öğrenci performansının erken haftalarda tahmin edilmesi, verilecek geri bildirimlerle olası başarısızlıkları önlemek amacıyla kullanılabilir.

İlk iki araştırma problemi kapsamında oluşturulan tahmin modelleri, kullanılan deęişkenlerle öğrencilerin akademik performansları arasındaki ilişkilerin ortaya konması açısından önemlidir. Bu sayede oluşturulacak tahmin modelleri, öğrencilerin otomatik olarak sınıflandırılması amacıyla kullanılabilir; ancak başarılı ve başarısız öğrencilerin tanımlanması konusunda çok fazla bilgi sağlamayabilir. Buradan yola çıkarak planlanan üçüncü araştırma problemi kapsamında öğrencilerin çevrimiçi öğrenme ortamındaki 14 haftalık kullanım verileri kümeleme yöntemi ile analiz edilmiş ve benzer kullanım örüntüsü gösteren öğrenciler tespit edilerek bu grupların akademik performansları analiz edilmiştir.

Bu amaçla literatürde sıklıkla kullanılan iki farklı kümeleme algoritması seçilmiş (X-Ortalamalar, EM Kümeleme) ve analiz sonuçları karşılaştırılmıştır. Her iki algoritmada da küme sayısı çapraz geçerlilik yöntemi kullanılarak algoritmalar tarafından belirlenmiştir ve sonuç olarak verinin ideal olarak üç farklı kümeye ayrılabilceęi görülmüştür. Bu kümeler, deęişken ortalamalarına göre Çok Aktif Öğrenciler (Küme 3), Aktif Öğrenciler (Küme 2) ve Aktif Olmayan Öğrenciler (Küme 1) olarak tanımlanabilir (Cristóbal Romero ve dięerleri, 2008). Bir dięer alt problem kapsamında elde edilen kümeler ders başarısı açısından da incelenmiş ve ortamda düşük aktivite gösteren öğrencilerin ders başarısının düşük olduęu, orta düzeyde aktivite gösteren öğrencilerin ders başarısının orta düzeyde olduęu ve yüksek aktivite gösteren öğrencilerin ders başarısının da yüksek olduęu görülmüştür.

Buradan yola çıkarak elde edilen kümeler ders başarısı ile ilişkili olarak Kardan ve Conati (2011)'nin sınıflamasına göre Yüksek Öğrenme (Küme 3), Orta Öğrenme (Küme 2) ve Düşük Öğrenme (Küme 1) olarak isimlendirilebilir. Öğrencilerin çevrimiçi öğrenme ortamındaki aktivite düzeyi ile ders başarısı arasındaki pozitif ilişki beklenen bir durumdur; fakat bu durumun modellenmesi önemlidir. Burada elde edilen sonuçlar ileriki çalışmalarda ortama yeni gelen öğrencilerin sınıflandırılması için ve başarısız olma ihtimali yüksek olan öğrencilere olası başarıları ile ilgili dönüt vermek amacıyla kullanılabilir. Aynı zamanda bu küme değişkenleri S. Amershi ve Conati (2009)'nin yaptığı gibi uyarlanabilir ortamlarla entegre çalışan sınıflama modellerinde girdi olarak kullanılabilir. Çevrimiçi öğrenme ortamlarında benzer öğrencileri tanımlamak için kullanılabilir (Valdiviezo, Reátegui, & Sarango, 2013); öğretmenler bu bilgileri yeni gelen öğrencileri sınıflandırmak için (Lopez ve diğerleri, 2012) ya da işbirlikli etkinliklerde çalışma gruplarını oluşturmak için kullanabilirler (Cristóbal Romero ve diğerleri, 2008).

Burada elde edilen küme bilgileri uyarlanabilir ve kişiselleştirilebilir öğrenme ortamları açısından da önemlidir. Çevrimiçi öğrenme ortamlarındaki öğrenci aktivitelerini temel alarak öğrenci profillerinin oluşturulması konusunda sınıflama ve kümeleme gibi VM tekniklerinden sıklıkla yararlanılmaktadır (Bienkowski ve diğerleri, 2012). Otomatik sınıflama işlemi, uyarlanabilir öğrenme ortamlarının ayrılmaz bir parçasıdır. Sistem, görev seçme, gezinim, içerik vb. herhangi bir uyarlama ya da müdahale yapmadan önce öğrencinin mevcut durumunu sınıflandırmalıdır (Hämäläinen & Vinni, 2010). Kümeleme analizi ise sınıflama modellerinde kullanılacak değişkenlerin üretilmesi açısından önemlidir. Bu sayede kümeleme analizi ile tanımlanan gruplar yeni öğrencilerin sınıflandırılması amacıyla kullanılabilir ve gerçek zamanlı uyarlamalar mümkün olabilir (Bouchet, Harley, Trevors, & Azevedo, 2013).

5.2. Öneriler

Veri depolama ve analiz teknolojilerindeki gelişmelerle birlikte her alanda olduğu gibi eğitim alanında da öğrenci davranışlarıyla ilgili çok daha çeşitli verilerin depolanıp anlık olarak analiz edilmesi mümkün olacaktır. Özellikle öğrenci sayısının her geçen gün arttığı çevrimiçi öğrenme platformlarında bu verilerden etkili olarak yararlanma yollarının araştırılması son derece önemli bir konudur. Bu nedenle yeni Medya Konsorsiyumu ve EDUCAUSE tarafından 2013 yılında hazırlanan Horizon Raporu

(Larry Johnson ve diğeri, 2013), çevrimiçi öğrenci aktiviteleri ile ilgili verilerin analizinden elde edilebilecek bilgilerin önemine vurgu yaparken ÖA ve EVM'yi yakın gelecekte eğitim araştırmalarına yön verecek teknolojiler arasında listelemektedirler. Yapılan çalışmadan elde edilen bulgular ışığında ileriki araştırmalara yön verebilecek araştırmaya ve uygulamaya dönük öneriler aşağıda listelenmiştir.

5.2.1. Araştırmaya Dönük Öneriler

Çalışma kapsamında sunulan performans metriklerinin tamamı 5k-çapraz geçerlilik yöntemi kullanılarak hesaplanan değerlerdir. Bu araştırmanın iç geçerliliği açısından önemlidir; fakat elde edilen modellerin dış geçerliliğinin sağlanabilmesi için aynı şartlarda farklı öğrencilerden toplanan yeni veriler üzerinde test edilmesi gerekmektedir.

Araştırma kapsamında belirlenen 28 değişkenden 10 tanesi etkin olarak kullanılmıştır. İleriki çalışmalarda diğer değişkenler de farklı araştırma soruları kapsamında incelenebilir.

Öğrenme ortamının veri tabanında öğrencilerin yaptıkları gezinimlerle ilgili tarihsel kayıtlar tutulmuştur fakat bunlar analizlerde toplam sayılar olarak ele alınmıştır. İleriki çalışmalarda öğrencilerin gezinim örüntüleri ile akademik performansları arasındaki ilişkilere bakılabilir.

Öğrencilerin yazdıkları iletilerdeki kopyala-yapıştır kullanımları ve yazdıkları iletiler arasında kopya oranı yüksek olanlar iki ayrı değişken olarak kayıt edilmiştir; fakat bu çalışma kapsamında bu değişkenler akademik performansın tahmin edilmesi konusunda önemli bulunmamıştır. Akademik performans üzerine olumsuz etkisi olduğu bilinen bu değişkenler ileriki çalışmalarda daha detaylı olarak incelenebilir.

Birliktelik kuralları değişkenler arası ilişkileri ortaya çıkartmak amacıyla yaygın olarak kullanılan bir VM yöntemidir. Çalışma kapsamında kullanılan değişkenler arası ilişkiler bu yöntemle analiz edilerek öğrencilere ve öğretmenlere yararlı bilgiler sağlayacak kurallar oluşturulabilir.

14 hafta süresince öğrencilerle sınıf ortamında da yüz yüze ders yapılmıştır; ancak akademik performanslarının tahmin edilmesinde sadece çevrimiçi ortamdaki değişkenler dikkate alınmıştır. İleriki çalışmalarda sadece çevrimiçi ortamın

kullanıldığı bir derste ya da sınıf ortamından da bir takım metriklerin de katıldığı bir uygulama yapılabilir.

Bilgisayar donanımı dersine ilişkin standart bir ölçme aracı bulunmamaktadır. Donanım dersine ilişkin başarı göstergelerinin belirlenmesi ve standart ölçme araçlarının geliştirilmesi öğrencilerin akademik performanslarını tahmin edecek genelleştirilmiş modellerin oluşturulması ve bu sayede boylamsal çalışmaların yapılabilmesi açısından önemlidir.

Araştırmada kullanılan değişkenler birbirinden bağımsız olarak analize katılmıştır. Bu değişkenlerin temel bileşenler analizi ya da kümeleme analizi gibi yöntemlerle gruplanması ile elde edilecek yeni değişkenler farklı derslerdeki akademik performansın tahmin edilmesi konusunda daha standart modeller üretilmesi amacıyla kullanılabilir.

5.2.2. Uygulamaya Dönük Öneriler

Öğrenci performansının 3. Hafta gibi kısa bir sürede elde edilen verilerle %74 oranında tahmin edilebilmesi olası başarısızlıkların önüne geçilmesi açısından son derece önemlidir. Bu bilgiler öğrencilerle etkili bir dönüt sistemi içerisinde paylaşılarak başarısızlık oranlarının düşmesi amacıyla kullanılabilir.

Öğrenme ortamına eklenecek eklentilerle çevrimiçi öğrenme ortamından toplanan verilerin anlık olarak görselleştirilmesi ve izlenmesi mümkün olabilir. İleriki çalışmalarda bu veriler öğretim tasarımcılarına, sisteme yapacakları müdahalelere karar verme ve etkilerini izleme sürecinde yardımcı olabilir. Öğretmenler bu bilgileri kullanarak öğrencilerin durumunu anlık olarak izleyebilir ve müdahale edebilirler. Öğrenciler de bu bilgileri kendi öğrenme süreçlerini izlemek amacıyla kullanabilirler.

Burada elde edilen küme bilgileri ile oluşturulan sınıflama modelleri ileriki araştırmalarda ortama yeni giren öğrencilerin sınıflandırılmasında kullanılabilir. Öğrencilerin ortamdaki aktiflik düzeyine göre sınıflandırılması ve bunun da ders başarıları ile ilişkili olması öğretmenlere öğrencinin durumu ile ilgili önemli bir veri sağlayacaktır.

Kümeleme analizi sonucu elde edilen kümeler ileriki çalışmalarda ortama yeni gelen öğrencilerin sınıflandırılması amacıyla sınıf değişkeni olarak kullanılabilir. Bu küme verileri aynı zamanda uyarlanabilir öğrenme ortamlarında kişiselleştirme amaçlı kullanılabilir.

Öğrenci performansından bağımsız olarak öğrencilerin etkileşim verilerinin kümeleme yöntemi ile analiz edilmesi sonucu elde edilen bilgiler kurumsal izleme ve ortamın değerlendirilmesi amacıyla kullanılabilir.

KAYNAKÇA

- Agathe, M., & Kalina, Y. (2010). Measuring correlation of strong symmetric association rules in Educational Data *Handbook of Educational Data Mining* (pp. 245-256): CRC Press.
- Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2013). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior* (0). doi: <http://dx.doi.org/10.1016/j.chb.2013.05.031>
- Akçapınar, G., Cosgun, E., & Altun, A. (2011). *Prediction of perceived disorientation in online learning environment with random forest regression*. Paper presented at the Proceedings of the 4th International Conference on Educational Data Mining., July 6-8, Eindhoven, Netherlands.
- Akçapınar, G., Çoşgun, E., & Altun, A. (2013, October 17th - 18th). *Mining wiki usage data for predicting final grades of students*. Paper presented at the International Academic Conference on Education, Teaching and E-learning (IAC-ETeL 2013), Prague, Czech Republic.
- Alfredo, V., Félix, C., & Àngela, N. (2010). Clustering educational data *Handbook of Educational Data Mining* (pp. 75-92): CRC Press.
- Amershi, S., & Conati, C. (2007). *Unsupervised and supervised machine learning in user modeling for intelligent learning environments*. Paper presented at the Proceedings of the 12th international conference on Intelligent user interfaces, Honolulu, Hawaii, USA.
- 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.
- Arnold, K. E. (2010). Signals: Applying academic analytics. *EDUCAUSE Quarterly*, 33(1). <http://www.educause.edu/ero/article/signals-applying-academic-analytics>
- Arnold, K. E., & Pistilli, M. D. (2012). *Course signals at Purdue: using learning analytics to increase student success*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.
- Baker, R., & Siemens, G. (in press). Educational data mining and learning analytics. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences: 2nd Edition*.
- Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *J. Educ. DataMining*, 1(1), 3-17.
- Baker, R. J. d., Corbett, A., Koedinger, K., & Roll, I. (2006). Generalizing detection of gaming the system across a tutoring curriculum. In M. Ikeda, K. Ashley & T.-W. Chan (Eds.), *Intelligent Tutoring Systems* (Vol. 4053, pp. 402-411): Springer Berlin Heidelberg.

- Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). *Off-task behavior in the cognitive tutor classroom: when students "game the system"*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria.
- Baker, R. S. J. d. (2007). *Modeling and understanding students' off-task behavior in intelligent tutoring systems*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA.
- Baker, R. S. J. d. (2010). Data mining *International Encyclopedia of Education (Third Edition)* (pp. 112-118). Oxford: Elsevier.
- 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.
- Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Washington, D.C.
- Bouchet, F., Harley, J., Trevors, G., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining, 5*(2).
- Bousbia, N., & Belamri, I. (2014). Which contribution does EDM provide to computer-based learning environments? In A. Peña-Ayala (Ed.), *Educational Data Mining* (Vol. 524, pp. 3-28): Springer International Publishing.
- Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. doi: 10.1023/a:1010933404324
- Burlak, G. N., Hernandez, A., Ochoa, A., & Munoz, J. (2006, Sept. 2006). *The use of data mining to determine cheating in online student assessment*. Paper presented at the Electronics, Robotics and Automotive Mechanics Conference, 2006.
- Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review, 42*(4), 40.
- 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.
- Ceglar, A., & Roddick, J. F. (2006). Association mining. *ACM Computing Surveys (CSUR), 38*(2), 5. doi: 10.1145/1132956.1132958
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol., 2*(3), 1-27. doi: 10.1145/1961189.1961199
- 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. doi: 10.1504/IJTEL.2012.051815
- Chien-Sing, L., & Singh, Y. P. (2004, 30 Aug.-1 Sept. 2004). *Student modeling using principal component analysis of SOM clusters*. Paper presented at the Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on.

- Clark, P., & Niblett, T. (1989). The CN2 Induction algorithm. *Machine Learning*, 3(4), 261-283. doi: 10.1023/A:1022641700528
- Cristóbal, R., Sebastián, V., Mykola, P., & Ryan, S. J. D. B. (2010). Introduction *Handbook of Educational Data Mining* (pp. 1-6): CRC Press.
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., . . . Zupan, B. (2013). Orange: Data mining toolbox in python. *J. Mach. Learn. Res.*, 14(1), 2349-2353.
- Dougherty, J., Kohavi, R., & Sahami, M. (1995). *Supervised and unsupervised discretization of continuous features*. Paper presented at the ICML.
- Drachsler, H., Hummel, H. G. K., & Koper, R. (2008). Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model. *International Journal of Learning Technology*, 3(4), 404. doi: 10.1504/ijlt.2008.019376
- Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, 35(5-6), 352-359. doi: [http://dx.doi.org/10.1016/S1532-0464\(03\)00034-0](http://dx.doi.org/10.1016/S1532-0464(03)00034-0)
- Enrique, G., Romero, C., Ventura, S., Carlos de, C., & Toon, C. (2010). Association rule mining in learning management systems *Handbook of Educational Data Mining* (pp. 93-106): CRC Press.
- Feng, M., & Heffernan, N. T. (2006). Informing teachers live about student learning: Reporting in the assistent system. *Technology Instruction Cognition and Learning*, 3(1/2), 63.
- Freeman, L. C. (2004). *The development of social network analysis: A study in the sociology of science*: Empirical Press.
- Frias-martinez, E., Chen, S. Y., & Xiaohui, L. (2006). Survey of data mining approaches to user modeling for adaptive hypermedia. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 36(6), 734-749. doi: 10.1109/TSMCC.2006.879391
- 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. doi: 10.1016/j.iheduc.2010.07.006
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15 %6(3), 42-57.
- Hämäläinen, W., & Vinni, M. (2010). Classifiers for educational data mining *Handbook of Educational Data Mining* (pp. 57-74): CRC Press.
- Han, J., Kamber, M., & Pei, J. (2006). *Data mining: Concepts and techniques, Second Edition (The Morgan Kaufmann Series in Data Management Systems)*: Morgan Kaufmann.
- Hand, D. J., Smyth, P., & Mannila, H. (2001). *Principles of data mining*: MIT Press.

- Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, 36(0), 469-478. doi: <http://dx.doi.org/10.1016/j.chb.2014.04.002>
- Jie, L. (2004). *A personalized e-learning material recommender system*.
- Johnson, L., Adams, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2013). The NMC horizon report: 2013 Higher Education Edition.
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2006). The 2011 Horizon report. Austin, Texas: The New Media Consortium, 2011.
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). The 2011 horizon report.
- Judy, S. (2010). Basics of statistical analysis of interactions data from web-based learning environments *Handbook of Educational Data Mining* (pp. 27-42): CRC Press.
- Kardan, S., & Conati, C. (2011, 2011). *A framework for capturing distinguishing user interaction behaviours in novel interfaces*. Paper presented at the The 4th International Conference on Educational Data Mining (EDM 2011).
- Koedinger, K., Cunningham, K., Skogsholm, A., & Leber, B. (2008). *An open repository and analysis tools for fine-grained, longitudinal learner data*. Paper presented at the Educational Data Mining 2008: 1st International Conference on Educational Data Mining, Proceedings.
- 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
- Lavoué, É. (2011). Social tagging to enhance collaborative learning. In H. Leung, E. Popescu, Y. Cao, R. H. Lau & W. Nejdl (Eds.), *Advances in Web-Based Learning - ICWL 2011* (Vol. 7048, pp. 92-101): Springer Berlin Heidelberg.
- Lemire, D., Boley, H., McGrath, S., & Ball, M. (2005). Collaborative filtering and inference rules for context-aware learning object recommendation. *Interactive Technology and Smart Education*, 2(3), 179-188.
- 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. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, 54(2), 588-599. doi: <http://dx.doi.org/10.1016/j.compedu.2009.09.008>
- MacQueen, J. (1967, 1967). *Some methods for classification and analysis of multivariate observations*. Paper presented at the Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, Berkeley, Calif.

- 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. doi: 10.1007/s10489-012-0374-8
- Mitchell, T. M. (1997). *Machine Learning*: McGraw-Hill, Inc.
- Monk, D. (2005). Using data mining for e-learning decision making. *The Electronic Journal of e-Learning*, 3(1), 41-54.
- Moreno, L., Gonzalez, C., Castilla, I., Gonzalez, E., & Sigut, J. (2007). Applying a constructivist and collaborative methodological approach in engineering education. *Computers & Education*, 49(3), 891-915. doi: <http://dx.doi.org/10.1016/j.compedu.2005.12.004>
- Nisbet, R., Elder, J., & Miner, G. (2009). Chapter 7 - Basic algorithms for data mining: A brief overview. In R. Nisbet, J. Elder & G. Miner (Eds.), *Handbook of Statistical Analysis and Data Mining Applications* (pp. 121-150). Boston: Academic Press.
- Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them *Technical Report IHMC CmapTools 2006-01 Rev 01-2008* (Vol. 284): Florida Institute for Human and Machine Cognition.
- Osmanbegović, E., & Suljić, M. (2012). Data mining approach for predicting student performance. *Economic Review*, 10(1).
- Özpolat, E., & Akar, G. B. (2009). Automatic detection of learning styles for an e-learning system. *Computers & Education*, 53(2), 355-367. doi: <http://dx.doi.org/10.1016/j.compedu.2009.02.018>
- Pelleg, D., & Moore, A. W. (2000). *X-means: extending k-means with efficient estimation of the number of clusters*. Paper presented at the Proceedings of the Seventeenth International Conference on Machine Learning.
- 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. (2004). Evaluating feature selection methods for learning in data mining applications. *European Journal of Operational Research*, 156(2), 483-494. doi: [http://dx.doi.org/10.1016/S0377-2217\(02\)00911-6](http://dx.doi.org/10.1016/S0377-2217(02)00911-6)
- Ribeiro, M. X., Traina, A. J. M., & Caetano Traina, J. (2008). *A new algorithm for data discretization and feature selection*. Paper presented at the Proceedings of the 2008 ACM symposium on Applied computing, Fortaleza, Ceara, Brazil.
- 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*, n/a-n/a. doi: 10.1002/cae.20456
- Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., & Ventura, S. (2013). Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education*, 21(1), 135-146. doi: 10.1002/cae.20456

- Romero, C., Olmo, J., & Ventura, S. (2013, July 6 - 9). *A meta-learning approach for recommending a subset of white-box classification algorithms for Moodle datasets*. Paper presented at the 6th International Conference on Educational Data Mining (EDM 2013), Memphis, Tennessee, USA.
- Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Syst. Appl.*, 33(1), 135-146. doi: 10.1016/j.eswa.2006.04.005
- 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. doi: 10.1109/TSMCC.2010.2053532
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27. doi: 10.1002/widm.1075
- Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384. doi: <http://dx.doi.org/10.1016/j.compedu.2007.05.016>
- 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.
- Siemens, G., & Baker, R. S. J. d. (2012). *Learning analytics and educational data mining: towards communication and collaboration*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.
- Siemens, G., & Baker, R. S. J. d. (2012). *Learning analytics and educational data mining: Towards communication and collaboration*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.
- Stolzer, A. (2009). Tutorial B - Data mining for aviation safety: Using data mining recipe "automatized data mining" from STATISTICA. In R. Nisbet, J. Elder & G. Miner (Eds.), *Handbook of Statistical Analysis and Data Mining Applications* (pp. 377-389). Boston: Academic Press.
- Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education*, 57(4), 2414-2422. doi: <http://dx.doi.org/10.1016/j.compedu.2011.05.016>
- Ting-Wen, C., El-Bishouty, M. M., Graf, S., & Kinshuk. (2013, 15-18 July 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.
- Valdiviezo, P., Reátegui, R., & Sarango, M. (2013). *Student behavior patterns in a virtual learning environment*. Paper presented at the Eleventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI'2013), August 14 - 16, 2013 Cancun, Mexico.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*: Morgan Kaufmann Publishers Inc.

- Xu, B. (2011). *Understanding teacher users of a digital library service: A clustering approach*. (3449487 Ph.D.), Utah State University, Ann Arbor. ProQuest Dissertations & Theses Full Text database.
- Zhang, C., & Zhang, S. (2002). *Association rule mining: Models and algorithms*: Springer-Verlag.

EKLER DİZİNİ

EK 1. Arařtırmada Kullanılan Deęiřkenlere İliřkin Tanımlayıcı İstatistikler

No	Deęiřken	Ortalama	SS	Minimum	Maksimum	Ortanca
1	t_Oturum	50,04	35,08	0	160,00	43,50
2	t_BenzersizOturum	22,34	13,18	0	53,00	20,00
3	t_OturumSuresi	1751,95	1302,89	0	5314,00	1352,00
4	t_Ileti	54,34	47,01	0	155,00	53,50
5	t_BenzersizIleti	5,03	3,77	0	14,00	5,50
6	t_KopyaIletiSayisi	8,89	12,04	0	51,00	3,00
7	t_KopyalaYapistir	11,92	22,69	0	83,00	0,00
8	t_Etiket	9,70	13,58	0	52,00	2,00
9	t_EtiketKullanım	51,80	58,81	0	248,00	35,50
10	o_IletiYazmaSuresi	79,12	69,67	0	433,00	73,00
11	o_TusKullanım	146,07	119,30	0	644,00	138,00
12	o_SilmeTusuKullanım	9,91	9,57	0	52,00	9,00
13	o_SayfaDeaktif	3,53	2,69	0	10,00	4,00
14	t_IletiGezinim	180,59	190,89	0	1222,00	154,50
15	t_IletiOkuma	117,24	121,05	0	653,00	88,50
16	t_Yorum	1,83	4,54	0	29,00	0,00
17	t_IletiDegerlendirme	61,97	62,74	0	294,00	38,50
18	t_YorumDegerlendirme	0,70	1,55	0	9,00	0,00
19	t_TartismaGezinim	146,28	171,67	0	1006,00	105,00
20	t_CevapGezinim	101,20	133,13	0	857,00	66,50
21	t_Cevap	3,91	4,92	0	20,00	2,00
22	t_SoruDegerlendirme	32,11	18,90	0	57,00	36,50
23	t_CevapDegerlendirme	18,39	27,56	0	187,00	8,00
24	t_Gezinim	1577,91	1272,73	0	6029,00	1205,00
25	t_KaynakGezinim	751,74	650,48	0	2436,00	589,00
26	t_BildirimGezinim	51,92	88,56	0	456,00	21,00
27	t_DuyuruGezinim	49,86	49,09	0	286,00	37,00

EK 2. Orjinallik Raporu



Search Trash

My Folders

- My Folders
- My Documents**
- Trash

My Documents Documents Settings page 1 of 1

Title	Report	Author	Processed	Actions
<input type="checkbox"/> ÇEVİRİMİÇİ ÖĞRENME ORTAMINDAKİ ETKİLEŞİM VERİLERİNE GÖRE ÖĞRENCİLERİN AKADEMİK PERFORMANSLARININ VERİ MADENCİLİĞİ YAKLAŞIMI İLE MODELLENMESİ 1 part - 21,119 words	<input type="checkbox"/> 1%	GÖKHAN AKÇAPINAR	September 9, 2014 10:12:04 AM EEST	

page 1 of 1

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı	Gökhan AKÇAPINAR
Doğum Yeri	Ankara
Doğum Tarihi	1983

Eğitim Durumu

Lise	Yenimahalle Teknik Lisesi, ANKARA	2000
Lisans	Hacettepe Üniversitesi – Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü	2005
Yüksek Lisans	Hacettepe Üniversitesi – Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü	2009
Yabancı Dil	İngilizce: Okuma (Çokiyi), Yazma (İyi), Konuşma (İyi)	

İş Deneyimi

Çalıştığı Kurumlar	Araştırma Görevlisi Hacettepe Üniversitesi – Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü	2007 -
	Yazılım Geliştirme Uzmanı RDC İlaç Araştırma ve Geliştirme San. A.Ş.	2005 - 2007

Akademik Çalışmalar

Yayınlar

Akçapınar, G., Altun, A. ve Menteş, T. (2012). Hipermetinsel ortamlarda önbilgi düzeylerinin gezinim profilleri üzerine etkisi. *Eğitim ve Bilim*, 37(163), 143-156.

Akçapınar, G., Altun, A., & Coşgun, E. (2014). *Investigating students' interaction profile in an online learning environment with clustering*. Paper presented at the 14th IEEE International Conference on Advanced Learning Technologies (ICALT2014), 7–9 July Athens, Greece.

Akçapınar, G. ve Altun, A. (2014). *Çevrimiçi öğrenme ortamlarında veri madenciliği uygulama süreçlerine genel bir bakış*. Paper presented at the International Conference on New Trends in Educational Tehnology (INTET2014), 14-15 Nisan Famagusta, Kıbrıs.

Akçapınar, G., Coşgun, E., & Altun, A. (2013). *Mining wiki usage data for predicting final grades of students* Paper presented at the International Academic Conference on Education, Teaching and E-learning (IAC-ETeL 2013), Prague, Czech Republic.

Akçapınar, G., Coşgun, E., & Altun, A. (2011). *Prediction of perceived disorientation in online learning environment with random forest regression*. Paper presented at the Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011), July 6-8, Eindhoven, Netherlands.

Akçapınar, G., & Altun, A. (2010). *The effect of prior knowledge on learners' navigation structure*. Paper presented at the IADIS International Conference Cognition and Exploratory Learning in Digital Age (CELDA 2010), Timisoara, Romania.

Akçapınar, G., & Aşkar, P. (2009). *Measuring author contributions to the mediawiki*. Paper presented at the IADIS International Conference WWW/Internet 2009, Rome, Italy.

İletişim

e-Posta Adresi	gokhana@hacettepe.edu.tr
-----------------------	--

Jüri Tarihi	22/07/014
--------------------	-----------