

T.C.
KIRIKKALE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DOKTORA TEZİ

Bulut Tabanlı Çevrimiçi Öğrenme Ortamında Etkinlik Öneri Sistemi Tasarımı:
Eğitimsel Veri Madenciliği Uygulaması

Hakan KÖR

Ocak 2017

Bilgisayar Mühendisliği Anabilim Dalında Hakan KÖR tarafından hazırlanan BULUT TABANLI ÇEVİRİMİÇİ ÖĞRENME ORTAMINDA ETKİNLİK ÖNERİ SİSTEMİ TASARIMI: EĞİTİMSEL VERİ MADENCİLİĞİ UYGULAMASI adlı Doktora Tezinin Anabilim Dalı standartlarına uygun olduğunu onaylarım.

Prof. Dr. Erdem Kamil YILDIRIM
Anabilim Dalı Başkanı

Bu tezi okuduğumu ve tezin **Doktora Tezi** olarak bütün gereklilikleri yerine getirdiğini onaylarım.

Prof. Dr. Hasan ERBAY

Danışman

Jüri Üyeleri

Başkan : Prof. Dr. Erdem Kamil YILDIRIM _____

Üye (Danışman) : Prof. Dr. Hasan ERBAY _____

Üye : Prof. Dr. Murat DEMİRBAŞ _____

Üye : Yrd. Doç. Dr. Mustafa COŞAR _____

Üye : Yrd. Doç. Dr. Melih ENGİN _____

...../...../.....

Bu tez ile Kırıkkale Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu Doktora derecesini onaylamıştır.

Prof. Dr. Mustafa YİĞİTOĞLU
Fen Bilimleri Enstitüsü Müdürü

ÖZET

BULUT TABANLI ÇEVİRİMİÇİ ÖĞRENME ORTAMINDA ETKİNLİK ÖNERİ SİSTEMİ TASARIMI: EĞİTİMSEL VERİ MADENCİLİĞİ UYGULAMASI

KÖR, Hakan

Kırıkkale Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı, Doktora Tezi

Danışmanlar: Prof. Dr. Hasan ERBAY

Ocak 2017, 100 sayfa

Bu tez çalışmasında, bulut alt yapısı üzerine kurulmuş çevrimiçi öğrenme platformu için etkinlik öneri modeli oluşturulmuştur. Öncelikle bulut kavramı ve çevrimiçi öğrenme ortamlarında kullanılan bulut yaklaşımları ve bu yaklaşımların sağlayacağı yararlar açıklanmıştır. Sonraki aşamada, öğrenme yönetim sistemi olarak tercih edilen moodle veri tabanından çalışmaya katkı sağlayacak değişkenler süzölmüştür. Başka bir ifadeyle, veri madenciliği işlemlerinde kullanılacak veri ambarı için gerekli veriler moodle öğrenme yönetim sistemi üzerinden elde edilmiştir. Veriler içerisinden tahmin modeli için gerekli değişkenler seçilerek veri madenciliği işlemlerine tabi tutulmuştur.

Veri madenciliği sürecinde geleceğe yönelik tahmin işlemleri için kullanılan karar ağacı sınıflama algoritmaları kullanılmıştır. Yapılan testler sonucu, C5 algoritmasının diğer karar ağacı algoritmalarına göre daha başarılı bir tahmin başarısına sahip olduğu görölmüştür.

C5 algoritmasının işleyiş basamakları model kısmında, oluşturulan karar ağacı yapısına ise bulgular kısmında yer verilmiştir. Ayrıca oluşturulan modelin, web tabanlı uygulama modeli tasarlanarak yayılma kısmına eklenmiştir. Bu model, çevrimiçi öğrenme ortamlarında öğrenen grubu yönlendirici niteliktedir. Bir çok ders materyali ve etkileşim araçları söz

konusu ortamlarda öğrencilerin kullanımına sunulmaktadır. Öğrenmeye istekli ve başarıyı hedefleyen bireyler, bireysel özelliklerine uygun etkinliklerin seçiminde yeterli bilgiye sahip değildir. Oluşturulan model üzerinden, öğrenen bireylere derslere katılmadan önce dönem sonunda elde edebilecekleri tahmini başarı oranlarının verilmesi hedeflenmiştir. Bu tür tahmin algoritmalarının eğitim sürecine sağlayacağı katkılar ve web tabanlı modelin olası yararlarına sonuç kısmında yer verilmiştir.

Veri madenciliği işlemleri dışında, örneklemin demografik yapısını, haftalık başarı, ders materyallerini kullananların başarı oranlarını gösteren temel istatistiksel bilgilere yer verilmiştir.

Sonraki çalışmalarda, araştırmacılara daha başarılı sonuçlar elde etmek için büyük örneklem grubu üzerinde çalışmanın yarar sağlayacağı önerilmiştir. Ayrıca öğrenme grubundaki bireylerin forum, chat, wiki gibi etkileşim araçlarını daha sık kullanmalarına yönelik teşvik edici çözümler getirilmesinin başarıya olumlu katkılar sağlayacağı düşünülmektedir.

Anahtar Kelimeler: Çevrimiçi Öğrenme, Bulut Bilişim, Veri Madenciliği, Eğitimsel Veri Madenciliği, Sınıflama ve Regresyon Karar Ağacı

ABSTRACT

ACTIVITY SUGGESTION SYSTEM DESIGN IN CLOUD BASED ONLINE LEARNING ENVIRONMENT: EDUCATIONAL DATA MINING APPLICATION

KÖR, Hakan

Kırıkkale University

Graduate School of Natural and Applied Sciences

Department of Computer Engineering, Doctorate Thesis

Supervisor: Prof. Hasan ERBAY

January 2017, 100 pages

In this thesis study, an activity suggestion model was formed for a cloud-based online learning platform. The study primarily tried to explain the concept of cloud and cloud approaches used in online learning environments and the possible benefits of these approaches. In the latter stage, variables that would contribute to the study were filtered from the moodle database, which is preferred as a learning management system. In other words, the data acquired for the data warehouse to be used in data mining procedures were acquired via moodle LMS. Among the data, variables were selected for the estimation model and exposed to data mining procedures.

In the data mining process, decision tree classification algorithms were used for prudential estimation procedures. As a result of the tests, it was determined that the algorithm C5 had a more successful estimation than other decision tree algorithms.

While operation steps of the algorithm C5 were involved in the model; the structure of the decision tree was involved in findings. Besides, a web-based application was designed for the model and added to the expansion. This model guides the learner group in online learning environments. Many course materials and interaction devices are presented to the use of students in the

aforementioned environments. Individuals who are eager for both learning and success do not have sufficient knowledge in selecting appropriate activities for their personal features. This model aims to provide approximate success rates to be acquired at the end of the term to learner individuals before participating in courses. Possible contributions of such estimation algorithms to the process of education and possible benefits of the web-based model were involved in the conclusion.

Apart from data mining procedures; basic statistical information showing the demographic structure of the sample, weekly success and success rates of individuals who used course materials were also involved.

It is recommended to conduct future studies on a larger sample group in order for researchers to obtain more successful results. It is also recommended to encourage individuals in the learning group for using interaction devices like forum, chat, wiki more frequently so as to make positive contributions to success.

Key Words: Online Learning, Cloud Computing, Data Mining, Educational Data Mining, Classification and Regression Tree

TEŐEKKÜR

Tez alıőmam boyunca beni motive eden, ynlendiren ve yardımlarını esirgemeyen danıőman hocam Sayın Prof. Dr. Hasan ERBAY'a teőekkr ederim.

alıőmam sırasında akademik bilgilerini payőan ve tezime katkı saėlayan Yrd. Do. Dr. Melih Engin'e, ėr. Gr. mer Faruk Akmeőe'ye, veri madenciliėi kısmında yardım aldıėım Yrd. Do. Dr. mr Demirer'e ve Arő. Gr. Emre Dnder'e, tezin Trke gramer yapısını inceleyen ėr. Gr. Hacı Ahmet Bulut'a, veri ambarı oluőturma iőlemlerine destek olan programcı Onur ztrk'e, doktora alıőmalarım boyunca desteėini esirgemeyen tm arkadaşlarıma ve hocalarıma, ayrıca doktora eėitimim srecinde zamanlarını aldıėım eőim Sebiha Kr'e, kızım Zeynep Kr'e ve oėlum Mustafa Kr'e, beni yetiőtiren ve desteėini hibir zaman esirgemeyen, aileme teőekkr bir bor bilirim.



İÇİNDEKİLER DİZİNİ

Sayfa

ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER DİZİNİ	vi
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	ix
KISALTMALAR DİZİNİ	x
1. GİRİŞ	1
1.1. Araştırma Amaçları	4
2. LİTERATÜR ÖZETİ: BENZER ÇALIŞMALAR	6
3. BULUT TABANLI ÇEVİRİMİÇİ ÖĞRENME YAKLAŞIMLARI	9
3.1. Çevrimiçi Öğrenme	9
3.1.1. Çevrimiçi Öğrenme Ortamları	10
3.1.2. Çevrimiçi Öğrenme Etkinlikleri	11
3.1.3. Bireyselleştirilmiş Çevrimiçi Öğrenme	14
3.2. Bulut Bilişim ve İlişkili Olduğu Kavramlar	15
3.2.1. Bulut Teknolojisinin Gelişimi	16
3.2.2. Bulut Hizmeti Sunum Modelleri	18
3.2.3. Bulut Uygulaması Dağıtım Modelleri	19
3.2.4. Bulut Tabanlı Çevrimiçi Öğrenme Yaklaşımları	19
3.2.5. Alt Yapı Olarak Bulut	21
4. VERİ MADENCİLİĞİ TANIMI, SÜRECİ, TEKNİKLERİ VE İLİŞKİLİ OLDUĞU ALANLAR	23
4.1. Temel Kavramlar	23
4.2. Veri Madenciliği Nedir?	26
4.3. Veri Tabanından Bilgi Keşfi Basamakları	26
4.3.1. Problemin Tanımlanması	28
4.3.2. Veriyi Anlama	28
4.3.3. Veri Hazırlama	29
4.3.3.1. Veri Temizleme	30
4.3.3.2. Veri Birleştirme	30
4.3.3.3. Veri Dönüştürme	31
4.3.3.4. Veri İndirgeme	31
4.3.4. Modelleme	31
4.3.5. Değerlendirme	33
4.3.6. Yayılma (Uygulama)	33
4.4. Veri Madenciliği Yöntemleri	34
4.4.1. Sınıflama Yöntemleri	37
4.4.1.1. Karar Ağaçları	39
4.4.2. Kümeleme Yöntemleri	44
4.5. Veri Madenciliğinin İlişkili Olduğu Alanlar	45
4.6. Eğitimsel Veri Madenciliği Kavram ve Modelleri	47
4.6.1. Eğitimsel Veri Madenciliği	47
5. YÖNTEM	49
5.1. Araştırma Modeli	49
5.2. Veri Toplama Aracı	50

5.3. Araştırma Evreni ve Örneklem Seçimi.....	52
5.4. Veri Ön İşleme Süreci	52
5.4.1. Öğrenci Profil Verilerinin İncelenmesi.....	52
5.4.1.1. Çevrimiçi Öğrenme Ortamı Verilerinin İncelenmesi	53
5.4.2. Verilerin Hazırlanması	54
5.5. Modelin Kurulması.....	54
5.6. Değerlendirme.....	60
5.7. Modelin Yayılımı (Uygulama)	60
6. BULGULAR	62
6.1. Katılımcıları Tanıtıcı Bilgiler.....	62
6.2. Model Sonuçları	63
7. SONUÇ ve ÖNERİLER	69
KAYNAKLAR	73



ŞEKİLLER DİZİNİ

<u>ŞEKİL</u>	<u>Sayfa</u>
1.1. 2004-2014 Yılları Arası Eğitim Teknolojilerinin Gelişimi.....	1
3.1. Kişiselleştirilmiş Öğrenme Modeli	15
3.2. Alt Yapı Olarak Örnek Bulut Şeması.....	22
4.1. Veri Ambarı Mimarisi.....	25
4.2. CRISP-DM İşleyiş Modeli.....	27
4.3. Veri Hazırlama Basamakları	29
4.4. Veri Madenciliği Model Türleri.....	32
4.5. Geçmiş Açıklamada Kullanılan Veri Madenciliği Yöntemleri	35
4.6. Gelecek Tahminlerinde Kullanılan Veri Madenciliği Yöntemleri	36
4.7. Veri Sınıflandırma Kurallarının Oluşturulması	41
4.8. Yeni Verilerin Sınıfını Tespit Etme İşlemi.....	42
4.9. Veri Biliminin İlişkili Olduğu Alanlar	46
5.1. Araştırmanın Kavramsal Modeli	49
5.2. Çevrimiçi Öğrenme Ortamında Etkinlik Takip Raporu.....	51
5.3. Modelin Kurulması	56
5.4. Karar Ağacı Algoritmaları Sınıflandırma Oranları.....	57
5.5. En İyi Sınıflamaya Sahip Karar Ağacı Algoritmaları	57
5.6. Modelin Web Tabanlı Tasarımı	61
6.1. Konuları Takip Eden Öğrencilerin Demografik Özellikleri	63
6.3. Kullanılan Değişkenlerin Öneri Başarısı Üzerine Etkisi.....	64
6.4. Karar Ağacı Sonuçları	66
6.5. Karar Ağacı Sonuçları (Devamı)	67
6.6. Karar Ağacı Sonuçları (Devamı)	68
7.1. Öneri Sisteminin Diğer Kurumlar İçin Yaygınlaştırılması.....	71

ÇİZELGELER DİZİNİ

<u>ÇİZELGE</u>	<u>Sayfa</u>
3.1. Bulut Bilişimin Tarihsel Gelişimi	17
4.1. Karar Ağacı Algoritmaları	43
4.2. Kümele Yöntemleri.....	45
5.1. Demografik Değişkenler.....	53
5.2. ÇÖS'den Elde Edilen Değişkenler	54
5.3. Ses Dosyası Takip Sayısını Çeken Örnek SQL Sorgusu.....	55
5.4. Tüm Etkinliklerin Takip Sayılarını Çeken SQL Sorgusu	55
6.1. Konu Bazında Ortalama Başarılar	64



KISALTMALAR DİZİNİ

EVM	Eğitimsel Veri Madenciliği
HUZEM	Hitit Üniversitesi Uzaktan Eğitim Merkezi
VTBK	Veri Tabanından Bilgi Keşfi
VM	Veri Madenciliği
MUSKUP	Mugla University Knowledge Discovery Unit
GP-ICRM	Genetic Programming Classification Rule Mining
NIST	Ulusal Standartlar ve Teknoloji Enstitüsü
VPN	Virtual Private Network
NASA	National Aeronautics and Space Administration
ÖYS	Öğrenme Yönetim Sistemi
CRISP-DM	Cross Industry Standard Process for Data Mining
PHP	Personal Home Page
MYSQL	My Structured Query Language
ÇÖS	Çevrimiçi Öğrenme Sistemi

1. GİRİŞ

Son yıllarda bilgisayar ve internet teknolojilerindeki gelişmeler sayesinde bilgisayar bilimleri, eğitim-öğretim etkinliklerini de olumlu yönde etkisi altına almıştır. Matematik, istatistik ve bilgisayar bilimleri sıkı ilişkileri olan disiplinler olarak bilinmektedir. Günümüzde, matematik ve istatistik alanında geliştirilen algoritmalar bilgisayarların yüksek hesaplama kabiliyeti ve görselleştirme olanaklarıyla gerçek yaşama entegre edilebilmektedir.



Şekil 1.1. 2004-2014 Yılları Arası Eğitim Teknolojilerinin Gelişimi

Öğretim teknolojilerinin 2004-2014 yılları arasındaki 10 yıllık süreçteki değişimi Şekil 1.1'de verilmektedir (Anonim, 2015). Şekil incelendiğinde, 2004'te ölçeklenebilir vektör grafikler ve öğrenme nesneleri, 2005'te zeki aramalar ve kablosuz iletişim, 2006'da kişisel yayıncılık ile ses ve görüntü

teknolojileri, 2007'de dijital kütüphaneler, 2008'de çevrimiçi videolar, ders kaydetme, bulut mail, 2009'da mobil teknolojiler, 2010'da e-kitap ve açık içerik , 2011'de yine mobil teknolojiler, 2012'de öğrenme analizleri, 2013'de kitlesel çevrimiçi açık dersler ve büyük veri, 2014'de mobil teknolojiler ve sanal desteklerin ön plana çıktığı görülmektedir.

Öte yandan, bilişim teknolojilerindeki gelişim uzaktan eğitimde kullanılan yöntem ve teknikleri de olumlu yönde etkilemiş, tercih edilme oranlarını ciddi oranlarda arttırmıştır. Yapılan akademik çalışmalar uzaktan eğitim öğrencileri ile örgün eğitim öğrencilerinin dönem sonu performanslarının farklı olmadığını, hatta uzaktan eğitim öğrencilerinin başarılarının daha yüksek olduğunu göstermektedir (Kör, 2013; Karaman, 2015).

Uzaktan eğitim 90'lı yılların başlarında internetin uluslararası bir ağ olarak kullanılmasıyla sanal ortama taşınmış ve 2010 yılı sonrasında çok daha hızlı gelişme göstererek sanal sınıf, sanal gerçeklik, oyun ile öğrenme, bulut bilişim ve mobil öğrenme gibi teknolojiler uzaktan eğitimde etkin biçimde kullanılmaya başlanmıştır (Öztürk, 2014). Geleneksel öğretim yöntemlerinin, bilişim teknolojilerindeki gelişmeleri takip ettiği taktirde yükseköğretimde ders içerikleri zenginleşecektir. Ayrıca, öğrenen bireylerin birbiri ile iletişim kurabileceği farklı ortamların (artırılmış gerçeklik, forum, gerçek zamanlı yazışma) oluşturulması sağlanacaktır.

Verilerin katlanarak hızlı bir şekilde arttığı içinde bulunduğumuz bilgi çağında, işlenmemiş verilerin bir değeri bulunmamaktadır. Bilgi ham petrole benzer. Petrolün işlenerek çeşitli alanlarda kullanılması gibi verilerde işlenerek anlamlı ve değerli hale getirilebilir. Bu işlem veri madenciliği olarak anılır, yani, veri madenciliği en temel tanımıyla verilerden değerli bilgi elde etme işlemidir. Ve bir disiplin olarak bilgisayar bilimleri alanında önemli bir konuma sahiptir. Bu bağlamda, kaydedilen eğitim verileri de veri madenciliği işlemleriyle anlamlı hale getirilmektedir. Eğitimsel veri madenciliği (EVM) olarak adlandırılan bu alanla alakalı ilk olarak 2000 yılında workshop düzenlemiştir. Sonrasında, eğitimsel veri madenciliği terimi ilk olarak 2005 yılında kullanılmış, bu alanda ilk uluslararası konferans 2008 yılında yapılmış

ve ilk el kitabı 2010 yılında yayımlanmıştır (Baker, 2014). Günümüzde EVM ayrı bir disiplin olarak değerlendirilmekte, modelleri geliştirilmeye devam etmekte, birçok dergide EVM alanında akademik yayınlar yayımlanmakta, her yıl düzenli olarak konferanslar düzenlenmektedir.

Çoğu üniversitemizde kullanılan mevcut bilişim sistemleri sadece öğrenci kayıtlarıyla ilgili ve temel işlevleri yapmaktadır. Öğrencilerin ders takip durumları, profilleri, öğrencilerin yaptığı etkinliklerle başarıları arasındaki ilişkileri, başarıya veya başarısızlığa etkisi olabilecek değişkenler mevcut öğrenme yönetim sistemlerinde etkin olarak kullanılmamaktadır.

Çevrimiçi öğrenme yönetim sistemlerinde, web 2.0 teknolojisi sayesinde hazırlanmış birçok etkinlik bulunmaktadır. Öğrenciler, bu etkinliklerden hangisinin başarıya daha çok etkili olacağı konusunda kararsız kalmaktadır. Her bir öğrencinin kişisel özelliklerinin farklı olduğu düşünüldüğünde, bu özelliklerin etkinlik seçiminde etkili olduğu düşünülmektedir. Öğrencilerin yaşı, cinsiyeti, mezun olduğu okul, yaşadığı yer gibi değişkenlerin yanı sıra çevrimiçi öğrenmeye karşı sahip oldukları tutumlar etkinliklere katılım oranını etkilemektedir.

Bu çalışmanın amacı, yükseköğretimde öğrenim gören öğrencilerin çevrimiçi etkinlikleri ve profil özellikleri kullanılarak veri madenciliği uygulaması gerçekleştirmektir. Bu uygulama sayesinde benzer özelliklere sahip öğrenciler sınıflanmakta, bu sınıflamaya göre sisteme sonradan dahil olan öğrencilere başarılı olmaları için çevrimiçi olarak yapması gereken etkinliklerin öneri olarak sunulması amaçlanmaktadır. Diğer bir ifadeyle öğrencilerin akademik başarılarını arttırıcı yönde çevrimiçi etkinlik öneri sistemini oluşturmak istenilmektedir. Bu öneri sisteminde altyapı olarak bulut sistemi kullanımının önemi üzerinde durulmuştur.

Tezin evreni, Hitit Üniversitesi Uzaktan Eğitim Merkezi(HUZEM)'nden ders alan öğrenciler oluşturmaktadır. Örneklemini ise 2015-2016 Bahar döneminde HUZEM'den ders alan 881 öğrenci oluşturmaktadır. Örneklemdaki öğrencilerin, çevrimiçi öğrenme ortamlarında gerçekleştirdikleri etkinlik verileri kaydedilmiştir.

Çevrimiçi öğrenme ortamları, çevrimiçi etkinlikler, bulut bilişim kavramları ve bulut bilişimin gelişim evrelerine, veri madenciliği tanımı ve kavramlarına, veri madenciliği kümeleme, sınıflama ve tahmin yöntemlerine, veri madenciliği hazırlık basamaklarına, eğitimsel veri madenciliği tanımı ve bu alanda kullanılan modellere 3. ve 4. bölümlerde yer verilmiştir.

Materyal ve yöntem kısmında, araştırmanın kavramsal modeli şekillerle ifade edilmiştir. Örneklem seçimi, verilerin toplanması, verilerin temizlenmesi, veri madenciliği basamaklarına 5. bölümde yer verilmiştir.

Bulgular kısmında veri madenciliği uygulaması sonucunda elde edilen sonuçlar tablolar halinde verilmiştir.

Sonuç kısmında, eğitimsel veri madenciliği uygulamasıyla elde edilen sınıflama modeli ve web tabanlı tasarım sonuçlarına yer verilmiştir. Tezin amaçlarının gerçekleşme durumlarına yine bu kısımda yer verilmiştir.

1.1. Araştırma Amaçları

Tezin bu kısmında, araştırmanın amaçlarına yer verilmiştir.

- 1.** Çevrimiçi öğrenme ortamlarında öğrenci profilleri ve sistem üzerindeki etkinliklerinden öğrencilerin akademik başarısının tahmin edilmesi.
- 2.** Veri Madenciliği sürecinde farklı sınıflama algoritmalarının tahmin etme başarısı üzerine etkisinin tespit edilmesi.
- 3.** Çevrimiçi öğrenme ortamlarında, öğrencilerin demografik özellikleri ve çevrimiçi ders takip sayılarına göre konu bazında başarı farklılıklarının belirlenmesi.
- 4.** EVM sürecinde kullanılan tüm değişkenlerin öneri başarısı üzerine etkisinin tespit edilmesi.
- 5.** Öneri sisteminde kullanılan değişkenlerden katkısı en fazla olan değişken veya değişkenlerin belirlenmesi.
- 6.** Web tabanlı etkinlik öneri sistemi modelinin tasarlanması.

Tezin amaçları doğrultusunda, oluşturulan veri seti çeşitli istatistiksel ve veri madenciliği yöntemleriyle test edilmiştir.



2. LİTERATÜR ÖZETİ: BENZER ÇALIŞMALAR

Tez çalışmasının bu kısmında literatür incelenerek benzer çalışmalardan elde edilen sonuçlara kısaca yer verilmiştir.

Bilgi miktarı her 3 yılda iki kat artmaktadır (Charles, 2003). Veri miktarı arttıkça bilginin daha hızlı analiz edilmesi ve bilginin işlenerek eş zamanlı anlamlı hale getirilmesi gerekmektedir. Bilginin analiz edilmesinde önceden kullanılan istatistiksel yöntemlerin yetersiz kaldığı görülmektedir (Glymour Vd., 1996). Bu nedenle önceki yöntemlerin yerini veri tabanlarından bilgi keşfi (VTBK) ve veri madenciliği (VM) olarak adlandırılan yeni nesil teknikler almıştır (Munakata, 1999). Bu yeni teknikler, veri analizinde, insanlara akılcı bir şekilde değerli bilgilerin keşfetmesine yardımcı olur (Guruler Vd., 2010).

Literatür incelendiğinde öğrencilerin demografik özellikleri, tutumları veya öğretim sürecinde yaptıkları etkinliklere göre farklı çalışmaların yapıldığı görülmektedir. Guruler ve diğerleri (Guruler Vd., 2010), Muğla üniversitesinde geliştirdikleri MUSKUP(Mugla University Student Knowledge Discovery Unit Program) programı ile öğrencileri demografik özellikleri ve puanlarına göre sınıflandırmışlardır. Çalışmada, Microsoft karar ağacı sınıflandırma tekniği kullanılarak öğrencilerin özellikleriyle başarıları arasındaki ilişkinin keşfedilmesi amaçlanmıştır.

Eğitmenler periyodik test performansına güvenmek yerine öğrencilerin ne bildiklerini ve her öğrenci için en etkili yöntemleri analiz edebilir. Veri analizine odaklanarak, öğretmenler öğrenmeyi çok daha ayrıntılı yöntemlerle sağlayabilirler. Çevrimiçi araçlar, öğrencilerin okumalara ne kadar zaman ayırdıkları, elektronik kaynaklara nereden ulaştıkları ve ne kadar süre kullandıkları gibi çok daha geniş bir öğrenci etkinlikleri değerlendirmesini mümkün kılan anahtar kavramlardır.

Bazı araştırmaların geliştirme ve değerlendirme aşamaları veri madenciliği, veri analizi ve web yönetici sayfalarından elde edilen bilgiler sayesinde

yapılmaktadır. “Büyük veriler”; öğrenci performans ve öğrenme yaklaşımları gibi bilgiler öğretim sürecinin geliştirilmesini mümkün kılar (West, 2012). Öğretmenler belirli zaman aralıklarında öğrencilerin başarılarını ölçmek yerine her öğrenci için en etkili yöntemi analiz edebilir. Öğretmenler, veri analizine odaklanarak öğrenmeyi daha etkili hale getirebilir (Manyika vd., 2011). Çevrimiçi öğrenme araçları öğrencilerin çeşitli etkinliklerini çok daha geniş çerçevede değerlendirebilir. Örneğin, öğrencilerin okumaya ayırdıkları zaman, elektronik kaynaklara erişim biçimi ve anahtar kavramları ne kadar hızlı öğrendikleri gibi (Castro vd., 2007).

Bilgisayara dayalı öğrenme modüllerinin geliştirilmesi, öğrencilerin sistematik, gerçek zamanlı olarak değerlendirilmesini sağlar. Veri madenciliği ve veri analiz yazılımı, öğrencilere ve öğretmenlere akademik performansları hakkında anında geri bildirim sağlayabilir. Bu yaklaşım, öğrencilerin ders bırakma nedeni, ekstra yardıma ihtiyaç duyma veya daha zorlayıcı görevler yapma gibi öğrenci çıktılarına öngörmek için altta yatan modelleri analiz edebilir. Belli öğrenciler için en etkili görülen pedagojik yaklaşımları belirleyebilir (Bienkowski vd., 2012).

Ayrıca, bir öğrencinin dönem sonu performansını tahmin etme yeteneği eğitimde artarak önem kazanmıştır (Baker ve Yacef, 2009). Öğrenci performans tahmin modellerini kullanan önceki çalışmalar, istatistiksel modelleme ve veri madenciliği tekniklerine odaklanmıştır (Gunnarsson ve Alterman, 2012). Bu geleneksel modelleme teknikleri kendi içerisinde sınırlamalar içerebilir. EVM öğrenme çıktılarının tahminlerini iyileştirmek için model ve algoritmaların geliştirilmesine odaklanmıştır (Siemens ve Baker, 2012). Mevcut istatistiksel ve EVM yöntemleri tahmin etme başarısını optimize etme üzerine kurulu bir yaklaşımdan yoksundur. Kötü veya yanlış tahminler, regresyon modellerinin temel gerekliliklerinin ihmal edilmesi durumunda ortaya çıkmaktadır (Harrell, 2001).

Wanli ve diğerleri (Wanli vd., 2015), genetik programlama ve yorumlanabilir sınıflandırma kuralı madenciliği (GP-ICRM) modelini kullandıkları çalışmalarında, öğretmenler öğrencilerin belirli bir zaman diliminde yaptıkları

alıřmalarını incelemiřlerdir. Bu alıřmada, her bir ğrenciye daha iyi ğrenme konsantrasyonu saėlamak iin somut ve bireysel neriler sunulmaktadır. Bu sayede ğrencilerin ğrenme srecinde farkındalıkları arttırılmıřtır.

Bireysel farklılıkları, ğrenme tercihleri, kiřisel zellikleri , kullanılan veriler gibi aılardan benzer ğrenen gruplarının tespit edilmesi EVM arařtırmacılarının zerinde alıřtıėı nemli konular arasındadır. Kmeleme ve sınıflama bu amala arařtırmacılar tarafından en ok tercih edilen VM yntemleridir (Peña-Ayala, 2014).



3. BULUT TABANLI ÇEVİRİMİÇİ ÖĞRENME YAKLAŞIMLARI

Bu kısımda, çevrimiçi öğrenme ortamı ve etkinliklerine, bireyselleştirilmiş çevrimiçi öğrenme ortamlarına, bulut bilişim kavramı ve ilişkili olduğu alanlara, bulut bilişimin gelişim sürecine, eğitimde bulut alt yapısı yaklaşımlarına yer verilmiştir.

3.1. Çevrimiçi Öğrenme

İnsanoğlu, tarih boyunca en büyük değişimleri 1900'lü yılların sonundan günümüze kadar geçen süreçte yaşamıştır. Bu dönemde iletişimden sağlığa, ticaretten eğlenceye, eğitimden sanata birçok alanın yer aldığı sanal bir dünya kuruldu. Yaşananlara bilişim devrimi, dijital devrim veya veri devrimi denilse de eskiden yaşanan değişimlere oranla çok daha fazla bir etkiye sahip olmuştur (Gürsakal, 2014). Söz konusu olan bu büyük değişimler araştırmacıları ve uygulamacıları eğitim alanında yenilikler sunmaya zorlamaktadır (Bayır, 2014).

Çevrimiçi öğrenme kavramı, insanların bilgi ve iletişim teknolojileri aracılığıyla birbirleriyle daha kolay iletişim kurmalarını sağlayan internet teknolojisinin uzaktan eğitime yansımaları olarak ortaya çıkmıştır. Çevrimiçi uygulamalar öğrencilere mekân ve zamandan bağımsız öğrenme imkânı sunmakla beraber (Moore ve Kearsley, 1996), uzaktan eğitimin en hızlı gelişen türüdür (Perraton, 2000). Brahmawong çevrimiçi öğrenmeyi, uzaktan eğitimin planlanma, hazırlık, üretim, sunum ve değerlendirilme aşamalarının i nternet tabanlı gerçekleştirildiği bir sistem olarak tanımlanmaktadır (Brahmawong, 2004). Ayrıca öğrencilere, ders materyallerine internet üzerinden erişebilme, eğitmen ve öğrencilerle eş zamanlı – eş zamansız iletişim olanağı sağlamaktadır (Aase, 2000). Çevrimiçi öğrenme, en basit tanımıyla bilgi, beceri ve duyguların öğrenen gruba, metin, resim, grafik gibi

öğeler kullanarak iletilmesidir (Yalın, 2008). Günümüzde ise bu süreç iki yönlü bir iletişime çevrilmiştir. Yani, çevrimiçi öğrenme, planlanma, hazırlık, üretim, sunum ve değerlendirilme aşamaları içermenin yanı sıra, interaktif ortamlar aracılığıyla öğrencilerinde öğrenme ortamına doğrudan katılmalarını sağlamaktadır.

Yükseköğretimde çevrimiçi, karma ve web tabanlı olarak çok sayıda öğretim yöntemi bulunmaktadır. Kurs ve programların sayısı katlanarak artış göstermektedir. Sloan'nın Amerika Birleşik Devletleri çevrimiçi raporuna göre 2009 sonbahar döneminde 5,6 milyon öğrenci en az bir dersini çevrimiçi olarak almıştır. Bu sayı bir önceki seneye göre bir milyondan fazla artış göstermiştir. Söz konusu artış oran olarak % 20'den fazladır. Ayrıca, yükseköğretim öğrenci potansiyelinin %30'unun en az bir dersi çevrimiçi olarak aldığı ifade edilmektedir(Diaz, 2011). E-öğrenmedeki gelişmelerin yanı sıra, yükseköğretimde bulut tabanlı ve web 2.0 teknolojilerinin gelişiminde patlama olduğuna şahit olunmaktadır (Smith, S.D. ve Caruso, J.B., 2010).

3.1.1. Çevrimiçi Öğrenme Ortamları

Çevrimiçi öğrenmede, ders içeriklerine erişim çoğunlukla internet aracılığıyla sağlanmaktadır (Allen ve Seaman, 2016). Fakat eğitim boyutlarını göz ardı ederek ve sadece teknolojiye önem verilerek etkili öğretim üretilmesi zordur (Horzum, 2015). Bunları dikkate alan araştırmacılar çevrimiçi öğrenmeyi bireylerin, belli öğretimsel temeller göz önüne alınarak tasarlanan ve web üzerinden sunulan içeriklerle etkileşim kurarak öğrenmesi temeline dayanan sistem olarak tanımlamaktadır (Pala ve Erdem, 2015).

Çevrimiçi öğrenme ortamları öğrenme ve öğretme sürecine katkı sağlayacak çok sayıda araca sahiptir. Öğitmenler bu araçları kullanarak ödev ve test hazırlayabilirler, ders materyallerini paylaşabilirler ve tartışmalar yürütebilirler (Romero vd., 2010). Öğrenciler ise çevrimiçi ortamlarda blog, viki, forum, mesajlaşma gibi araçları kullanarak işbirlikçi çalışmalar (Moreno vd., 2007)

veya bireysel öğrenme etkinliklerini planlayabilirler. Diğer bir ifadeyle, bireylerin çevrimiçi öğrenme ortamlarında bilgi ve deneyimlerini oluşturmaları, paylaşımları ve katkı sağlamaları için sanal internet ortamları oluşturulmuştur (Cheung vd., 2013). Çevrimiçi öğrenme ortamları bireylerin mevcut bilgiyi aldıkları, ihtiyaçları doğrultusunda düzenlemeler yaptıkları, yeni bilgileri ürettikleri, üretilen bu bilgileri kullanarak problem çözdükleri ve paylaştıkları ortamlardır. Çevrimiçi ortamlar, öğrenen grubun birbirleriyle ve çevreleriyle kurdukları sürekli etkileşim halinde ürettikleri ortak kavramları da kullanarak kendi kişisel anlamlarını oluşturmaları olarak tanımladığımız öğrenme kavramı için oldukça zengin bir potansiyel güçtür (Pala ve Erdem, 2015).

3.1.2. Çevrimiçi Öğrenme Etkinlikleri

Çevrimiçi ortamlarda kayıtlı grup üyeleri yeni bilgi edinmek, daha yararlı kaynaklara ulaşmak, problem çözme becerilerini kolaylaştırmak, bireyler arası iletişimi artırmak için mesleki becerileri desteklemek gibi etkinliklere ilgi duymaktadırlar (Tseng ve Kuo, 2014). Yao ve diğerleri (Yao vd., 2015), sanal topluluklarda bilgi paylaşma, takım öğrenmesi, sosyal sermaye ve dijital bağlılık arasındaki ilişkiyi incelemişlerdir. Çalışmalarında takım öğrenmesinin bilgi paylaşımı ile sosyal sermayenin de hem bilgi paylaşımı hem de takım öğrenmesi ile olumlu yönde ilişkisinin olduğunu bulmuşlardır. Bu bağlamda bilgi paylaşmanın, katılımcıları sanal ortamlara çeken ve bu ortamlarda kalmalarını sağlayan çok önemli bir rolü olduğu söylenebilir. Aynı yaş grubu öğrenenler arasında bilgi paylaşımları ve sosyal etkileşimler çevrimiçi ortamların başarısını etkilemektedir. Çevrimiçi ortamlarda sürekli ve kalıcı bir etkileşimin olmadığı durumlarda bilgi paylaşımının varlığından söz etmek şüphelidir. Bilgi paylaşımı bilgi elde etmek için önemli bir adımdır (Ma ve Yuen, 2011). Çevrimiçi eğitimde oluşan öğrenmelerin nitelikli olması çok önemlidir. Bu nedenle öğrenme faaliyetleri tasarlanırken bu sürece etki edebilecek unsurlar birlikte düşünülmelidir (Çalışkan, 2002). Çevrimiçi ortamlar, öğrenenlerin etkin olarak öğrenme süreci içerisine çekildiği, karar

verme aşamalarında farklı seçeneklerin sunulduğu ve bunun sonucunda farklı etkinliklere katıldığı ortamlar olarak düşünülmelidir. Çevrimiçi eğitimde öğrenen gruba sunulan ders içerikleri, öğrenenleri etkileşimsiz ekran izleyicisi olmaktan kurtaran türden olmalıdır (Çalışkan, 2002). Diğer yandan çevrimiçi alan uzmanları, çevrimiçi öğrenme ortamı oluşturmada ve devamlılığını sağlamada dikkat edilmesi gereken bazı ilkeler üzerinden durmuşlardır. Bu ilkeler, yoğun iletişim kurmak, açık etkileşime girmek, sınıf dışı etkileşim alanları oluşturmak, etkileşimi kolaylaştırıcı araçlar kullanmak, etkileşimleri yönlendirmek, öğrenme sürecinde yapıcı yaklaşımları kullanmak, destek sistemlerini planlamak şeklinde sıralanabilir (Aydın, 2002).

Çevrimiçi öğrenme ortamları, internet tabanlı teknolojilerin her geçen gün yeni araç ve teknikleri kullanıcılara sunması ile öğrenen gruplar öğrenme etkinliklerini eş zamanlı veya eş zamansız olarak takip edebilmektedir. Çevrimiçi öğrenmenin tüm katılımcıların aynı zamanda öğrenme ortamında olması gerektiği ve herhangi bir kurs, ders veya toplantının internet teknolojileri yardımıyla gerçekleştirildiği türüdür. Oturum sırasında gerçekleştirilen etkinlikler kaydedilerek eş zamansız olarak da kullanılabilir.

- **Eş zamanlı Öğrenme ve Kullanılan Araçlar**

Eş zamanlı eğitim, öğretici ile öğrenen grubun çevrimiçi teknolojiler aracılığıyla etkileşim içerisinde buldukları öğretim türüdür. Eş zamanlı öğrenmede iletişim, ses, video, grafikler ve metin yoluyla gerçekleştirilebilirken araç olarak, sanal sınıf uygulamaları, video konferans ve sohbet odaları kullanılmaktadır (Belanger ve Jordan, 2000).

Sanal sınıf: Sınıf ortamında gerçekleştirilen, geleneksel eğitim faaliyetlerinin çevrimiçi olarak internet üzerinden yürütüldüğü ve öğrenen ile öğretici arasında etkileşimin sağlandığı etkinliklerin tümüne sanal sınıf eğitimi denir. Diğer bir ifadeyle sanal sınıf, öğrenen grubun teknolojik alt yapı aracılığıyla öğretim materyalleri ve öğretici ile iletişime geçtiği sanal ortamlardır.

Beyaz Tahta Uygulamaları: Beyaz tahta uygulamasında öğretim elemanı, ekranın bir bölümünü çizim, grafik ve metin için kullanabilir veya öğrenciyi bu çizime davet edebilir. Tüm katılımcılar beyaz tahta üzerindeki nesneyi eş zamanlı olarak görebilirler. Beyaz tahta uygulamaları öğrenenlerin tüm beklentilerini karşılayacak kadar güçlü olması yanında kullanımı ve tasarımı açısından da basit olması önerilmektedir (Horton, 2003).

Sohbet: Çevrimiçi öğrenme ortamında, öğretim elemanı ve öğrenen arasında gerçek zamanlı iletişim sağlamaktadır. Temel amaç öğrencilere bir oturum süresince anında geri bildirim ve sorularına cevap bulma olanağı sağlamaktır.

Sesli Konferans: Katılımcıların birbirleriyle sesli konuşmasını sağlar. Telefon yoluyla gerçekleştirilebileceği gibi internet üzerinden çeşitli yazılımlarla da yapılabilmektedir. Yalnızca sesli iletişim kurulan ortamlarda, tonlamalar anlatılmak istenen duygunun karşı tarafa yansıtılması açısından önemlidir.

Video Konferans: Farklı yerlerde bulunan kişi veya kişilerin teknolojik imkanlar sayesinde bir araya gelerek ses, görüntü ve veri paylaştığı ortamlardır. Video konferans birçok ülkede uzaktan eğitim alanında ve çeşitli iş sektörlerinde kullanılmaktadır.

- **Eş zamansız Öğrenme ve Kullanılan Araçlar**

Eş zamansız öğrenme, öğrenciye farklı zamanlarda kendi planı, programı ve çalışma düzenine uygun herhangi bir öğreticiyle çevrimiçi bağlantı kurmadan istediği zaman aralığında çalışmalarını yürütme olanağı sağlayan öğretim şeklidir.

Forum: Eş zamansız öğrenme aracı olan forum sayesinde öğrenci öğretim elemanı ve diğer sınıf arkadaşlarına soru sorabilir ve bir konu hakkında tartışabilir. Öğretim elemanı ise forum aracılığıyla öğrencilere ders materyali, ödev, sınav vb. etkinlikleri duyurulabilir.

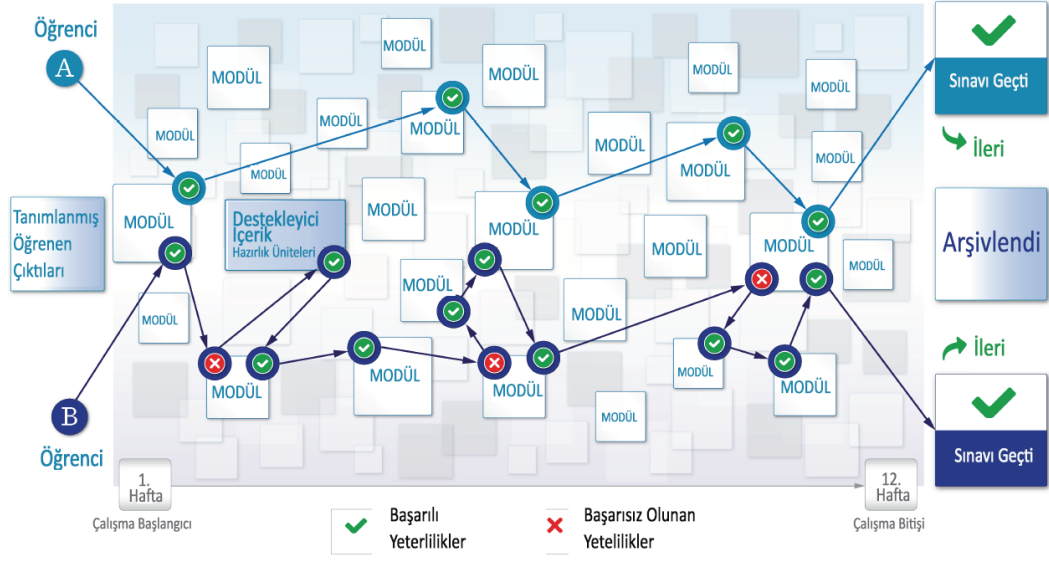
E-posta: Sınıf ortamından uzakta olan öğrenen ve öğreten grubun eş zamansız iletişim olanağı sağlayan bir öğrenme aracıdır. Metin içeren telefon mesajları da bu kategoriye dahil edilmektedir.

3.1.3. Bireyselleştirilmiş Çevrimiçi Öğrenme

Çevrimiçi öğrenme ortamlarında farklı konular hakkında araştırma yapan öğrenenler internet, çoklu ortam ve sanal kütüphanelerde karmaşık bilgi yapılarıyla karşı karşıya kalmaktadır. Söz konusu elektronik ortamlar, çok sayıda bilgiyi farklı erişim seçenekleriyle sunmasına rağmen çoğunlukla bireylerin kendi sorumluluk düzeyleri ve bilgi yönetim yeteneklerine göre yapılandırılmaktadırlar. Fakat bazı durumlarda içeriğin yeteri kadar yapılandırılmadığı ve karmaşık bir yapıya sahip sanal ortamların insan zihninin anlamlandırma, problem çözme ve öğrenme sınırlıklarının üstesinden gelinmesi için bazı bilişsel araçları kullanmaları gerekmektedir (Terga ve Keller, 2005). Çoğu zaman bilişsel araç olarak öğretimi destekleyici yazılımlar düşünülse de konu haritaları şeklinde sunulan bilişsel araçlarda kullanılmaktadır. Örneğin çoğu ders içerikleri çevrimiçi ortamlarda modüler olarak verilmektedir. Bireyler kendi öğrenme hızlarında öğrenme sağlamaktadır. Bireylerin sahip oldukları önceki bilgileri, kavrama ve öğrenme biçimlerinin aynı olmadığı düşünüldüğünde her bireye aynı yöntem ve sırayla ders modüllerinin sunumu uygun olmayacaktır.

Şekil 3.1’de görüldüğü gibi (Ifenthaler, 2014) bireysel öğrenme hızında modüllerin verilmesi daha etkin ve yapılandırılmış bir öğrenme ortamı sağlayacaktır. Bazı durumlarda öğrenme sürecinde sıkıntı yaşayan bireylerin destekleyici modüllere yönlendirildiği ve kimi öğrenenlerin modülleri atlayarak ilerlediği şekilde görülmektedir.

Dinamik Model - Kişiselleştirilmiş ve Adapte Edici Öğrenme



Şekil 3.1. Kişiselleştirilmiş Öğrenme Modeli

3.2. Bulut Bilişim ve İlişkili Olduğu Kavramlar

Ulusal Standartlar ve Teknoloji Enstitüsü (NIST), bulut bilişimi, "yapılandırılması mümkün olan bilgi işlem kaynaklarının (ağlar, sunucular, depolama, uygulamalar ve hizmetler) ortak bir havuzda, kolay ve isteğe bağlı ağ erişimini sağlaması" olarak ifade etmiştir. Bu yetki verme veya erişimi yasaklama işlemleri en az yönetim çabası ile yürütülmektedir(Mell ve Grance, 2012). Bulut bilişim, internetin oluşturulmasından sonra ortaya çıkan ve bilişim dünyasında köklü değişikliklere yol açabilecek bir yenilik olarak ifade edilmektedir (Aksu vd., 2013). Fakat bulut bilişimin tamamen yeni bir kavram olarak düşünülmesi doğru değildir. Esasında 2000'li yılların başından itibaren merkezi bilgisayarlara uzaktan erişim sağlanarak kullanıcılar kendi bilgisayarları üzerinden çalışmalarını yürütmektedir. Bu bağlamda bulut sistemi; geçmişte kullanılmış bir yöntemin günümüzün ekonomik şartlarına göre düzenlenerek ve geliştirilerek, tüm internet kullanıcılarının kullanımına

olanak sağlayacak şekilde yeniden sunulmasıdır (Henkođlu ve Klc, 2013). Diđer bir ifadeyle, servislerin ve uygulamaların internetteki sunucular bnyesinde bulundurulması, internete bađlı herhangi bir cihaz ile uygulama ve servislerin alıřtırılmasıdır(Yksel, 2012). Bulut biliřim, byk bilgi merkezlerindeki (data center) sistem yazılım ve donanımlarına internet zerindeki bir uygulama gibi eriřme hizmetidir. Bu servislerden yazılımın bir servis gibi kullanılmasına "Hizmet Olarak Yazılım" denilmektedir. Bazı firmalar, rnlerini tanımlamak iin "Hizmet Olarak Altyapı" ve "Hizmet Olarak Platform" gibi terimleri kullanmaktadır, ancak bu tanımlamalardan kaınılmalıdır nk kabul edilen tanımlar halen ok eřitlidir. "Dřk dzey" altyapı ve daha st dzey bir "platform" arasındaki izgi net deđildir (Armbrust vd., 2010).

3.2.1. Bulut Teknolojisinin Geliřimi

Bulut biliřim teriminin kkeni aıka belli olmasa da 1950'lere kadar dayanmaktadır. Bilgisayarların ok maliyetli olması nedeniyle řimdilerde zaman paylařımı olarak bilinen kavram o dnemde donanım ve iřlemci paylařımı olarak uygulanmaktaydı.

Zamanla biliřim sektrnde maliyetlerin azalması, bilgisayar ve internet kullanıcı sayısında meydana gelen artıřla birlikte bulut biliřim kavramı da geliřim gstermiřtir. Bulut biliřimin geliřim tablosu izelge 3.1'de grlmektedir.

izelge 3.1 incelendiđinde, fiziksel olarak bilgisayarların paylařımıyla bařlayan bulut biliřim kavramı, gnmzde depolama, elektronik posta hizmetleri, yazılımların ortak kullanılması ve eř zamanlı bilgiye eriřim gibi hizmetler sunarak maliyet, iř gc, gvenlik ve bilginin kolay ynetimi imknlarını kullanıcılarına sunmaktadır (Wikipedia(a), 2016).

Çizelge 3.1. Bulut Bilişimin Tarihsel Gelişimi

Tarih	Gelişim Aşamaları
1950' ler	Şirketlerde ve üniversitelerde kullanılan büyük kapasiteli ana bilgisayarlara zayıf kullanıcı bilgisayarları tarafından ulaşılabilirdi.
1960' lar	John Mc Carthy, bilgi işlem kullanımının gelecekte elektrik ve su hizmeti gibi bir kamu hizmeti olarak sunulacağı öngörüsünü 1960'lı yıllarda ifade etmiştir.
1970' ler	Zaman paylaşımı kavramı ortaya çıkmaya başladı (Strachey, 1959).
1990' lar	Telekomünikasyon şirketleri dengeli bir band genişliği sağlamak için noktadan noktaya (point-to-point) elektrik devreleriyle sanal özel ağlar(VPN) oluşturulmaya başladı.
2000' ler	2002: Amazon web servis, 2006: Amazon s3, ilk gerçek bulut uygulaması 2007-2008: Google ve IBM bulut servis hizmetini oluşturmuştur (Sevli 2011). Temmuz 2010: Rackspace Hosting ve National Aeronautics and Space Administration(NASA) ortaklaşa OpenStack olarak bilinen bir açık kaynak bulut yazılım girişimi başlattı. Mart 2011: IBM firması smart cloud çatısını duyurdu Hazira 2012: Oracle firması Oracle Cloud'u duyurdu

Genel olarak bulut modelleri, 5 temel özellik, 3 hizmet modeli ve 4 dağıtım modelinden oluşmaktadır. Temel özellikler; isteğe bağlı self servisler, geniş ağ erişimi, kaynak havuzu, esneklik ve ölçeklenebilir servislerdir. Hizmet modelleri; yazılım, platform ve alt yapıyı bir hizmet olarak sunmaktadır. Dağıtım modelleri özel bulut, topluluk bulutu, genel bulut ve karma buluttur.

3.2.2. Bulut Hizmeti Sunum Modelleri

Bir bulut sistemi, kullanıcı veya uygulamalarla servis olarak nitelenen çeşitli yollarla etkileşimde bulunabilir. WEB'te bulut bilişimin 4 farklı dağıtım modeli vardır. Bunlar; altyapı, platform, yazılım ve diğer modellerdir (Baun vd, 2011).

Altyapının Servis Olarak Sunulması: Bu model, kullanıcıların veri depolama, bellek ve band genişliği gibi hizmetler için tercih ettiği modeldir. Diğer bir ifadeyle kullanıcılara hiçbir emek harcamadan belli bir maliyet karşılığında donanım altyapısı sunan bulut modelidir.

Platformun Servis Olarak Sunulması: Bu servis modeli altyapı olarak bulut sistemi üzerine kuruludur ve burada hedeflenen son kullanıcı değil geliştiricilerdir. Bu serviste, programlama ve çalıştırma ortamları ile özel programlama dilleri çalıştırılabilir. Bu bulut hizmetini sağlayan servisler, kullanıcılarına web üzerinden ağ desteği, veri tabanı ve işletim sistemi kullanımı için fırsatlar sunar. Platform olarak servis modeli ticari uygulamalarına örnek olarak Google App Engine ve Microsoft Windows Azure gösterilebilir. (Caytiles R.D., 2012; Omotunde vd., 2013).

Yazılımın Servis Olarak Sunulması: Bu bulut yaklaşımında, ağ üzerinden istek yapılan bulut tabanlı yazılım ve uygulamalar son kullanıcılar veya kuruluşların erişimine sunulur. Ödeme yapan kullanıcılar bu hizmetlere erişim hakkına sahiptirler. Bu servisin avantajları şöyle sıranabilir; sadece bir web tarayıcısı aracılığıyla kullanıcılar servis tarafından sunulan yazılımlara erişebilir, bulut içindeki veri merkezleri düşük maliyetle kullanılabilir. Günümüzde yaygın olarak tercih edilen yazılım bulut servisleri şunlardır; Google Docs, Gmail, CRM Online, Salesforce.com, Exchange Online Business Productivity ve Online Suite'dir (Omotunde vd., 2013; Ibikunle, 2011).

3.2.3. Bulut Uygulaması Dağıtım Modelleri

Özel Bulut: Bu dağıtım modeli kurumların kiraladığı veya kendi oluşturduğu buluttur. Oluşturulan sistem yalnızca kurum tarafından kullanılır.

Kamu Bulut: Kamu hizmetlerini karşılamak için oluşturulmuş bulut dağıtım modelidir. Depolama ve yazılım gibi hizmetleri herkese açık bir şekilde sunar. Diğer dağıtım modellerine göre daha az maliyetlidir. Google AppEngine, Force.com, Amazon EC2 ve S3 kamu bulut örnekleridir.

Karma Bulut: İki veya daha fazla bulut yaklaşımının birleştirilmesiyle oluşturulan bulut dağıtım modelidir. Bulutlar mevcut özelliklerini kaybetmeden yazılımın kullanılması ve verinin taşınmasına izin verecek biçime yapılandırılmış veya özel teknolojik tekniklerle birleştirilmiştir. Planlanan iş yükü noktaları bu model ile kontrol edilebilir.

Topluluk Bulutu: Belirli bir toplulukla bulut alt yapısının paylaşıldığı dağıtım modelidir. Bu model, devlet kurumu veya özel kuruluşlar tarafından paylaşılabilir, böylece ortak hedefleri olan, benzer güvenlik ihtiyacı olan, aynı biçimde yönetilen kurum ve kuruluşlar desteklenir.

3.2.4. Bulut Tabanlı Çevrimiçi Öğrenme Yaklaşımları

Kişilerin veya kurumların günlük yaşam içerisinde vazgeçilmez bir parçası haline gelen bilgisayar ve internet, son zamanlarda bilişim teknolojilerinde meydana gelen hızlı gelişmelerden olumlu yönde etkilenerek farklı alanlarda daha tercih çok edilebilir bir hale gelmiştir. Söz konusu gelişmelerle birlikte etkin bir öğrenme için gelişmiş teknolojinin kullanımı tüm eğitim basamaklarının geleceği haline gelecektir (Al-Zoube vd., 2010). Web' in doğası ve web kaynaklarına erişimde kişisel kullanım sonucunda eğitim sektörü, iş dünyası, istihdam alanları, eğlence sektörü ve sağlık hizmetleri alanları da değişime uğramıştır. Son 15 yıldır, internet sürekli bir değişim

göstererek web 1.0'ın statik yapısı yerini, web 2.0'nin daha dinamik ortamı ve işbirlikçi yapısına bırakmıştır. Son kullanıcılar, ortak yazılım uygulamalarını, bilgi paylaşımlarını ve oluşturdukları yeni servisleri çevrimiçi olarak çalıştırabilmektedir (Safran vd., 2007).

Bulut bilişim, teknolojideki hızlı gelişmelerden en çok etkilenen bilişim alanlarından biridir. Bulut teknolojisi veri sağlamak ve uygulamaları kullanmak için internet ve uzak merkezi sunucuları kullanır. Tüketiciler ve işletmeler, bulut bilişim sayesinde internete bağlı herhangi bir bilgisayar aracılığıyla kişisel dosyalarına erişebilmektedir. Bu teknoloji ile merkezi veri depolama yapılabilmektedir. Böylece daha hızlı bilgi işleme, band genişliği ve depolama sağlanmaktadır (Bora ve Ahmed, 2013).

Özellikle 2000 yılından sonra teknolojide meydana gelen gelişmelerin, eğitim alanında hızla uygulandığı, eğitim-öğretim sürecinde uygulanan mevcut strateji ve yöntemlere yenilikçi uygulamalar kattığı izlenmektedir. Bilgi toplumu olarak öğrenme gereksinimleri göz önüne alındığında, bilgiye zaman ve mekandan bağımsız hızlı erişim sayesinde ihtiyaç anında öğrenmenin önem kazandığı söylenebilir. Ortaya çıkan yer ve zaman bağımsız öğrenme ihtiyacı, "mobil öğrenme" eğitim modelini ortaya çıkarmıştır. Bu yeni öğrenme modeli eğitime çok yeni bir boyut ve ivme kazandırmıştır. Mobil öğrenme, hareket halinde ve ihtiyaç anında bilgiye ulaşmayı kolaylaştırarak, eğitim-öğretim programlarına yenilikçi stratejiler, öğrenme yaklaşımları ve yöntemler getirmektedir (Çakır, 2011).

Benzer ihtiyaçlara sahip eğitim kurumları öğrenim sürecini yönetmek için öğrenme yönetim sistemi yazılımları (ÖYS) kullanmaktadır. Kurumların kullandıkları ÖYS'ler ölçeklenebilirlik, ulaşılabilirlik ve güvenlik gibi yönlerden tehlikeli olabilecek potansiyele sahiptir. Bu noktada bulut bilişim öğrenme sürecinde kullanılan yazılım ve donanım gereksinimlerini karşılayabilecek teknolojik bir yeniliktir (Güldal, 2016).

Çevrimiçi öğrenme ortamlarında, bulut bilişimin sağladığı önemli katkılar arasında yer alan yönetim kolaylığı, servis sağlayıcı etkileşimi ile depolama ortamları, sunucular, servisler ve uygulamalar gibi birçok bilişim kaynağına,

istenildiğinde ve elverişli bir şekilde ağ erişimi sağlayan alt yapıya ihtiyaç duyulmaktadır (Mell ve Grance, 2011; Ercan, 2010). Bu bağlamda adı son zamanlarda sıkça duyulan bulut teknolojileri, geniş bant internet, sanallaştırma gibi diğer yeniliklerin çevrimiçi öğrenme ortamları üzerinde gelecekte daha bariz bir etkiye sahip olacağı düşünülen bilişim dünyasının yeni teknolojik gelişmelerinden biridir (Ercan, 2010).

3.2.5. Alt Yapı Olarak Bulut

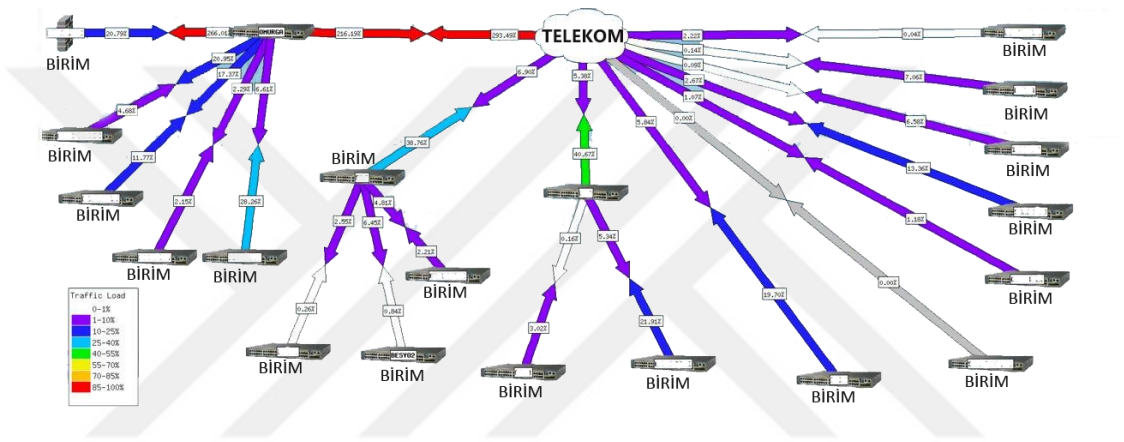
İnterneti sembolize etmek için 1994 yılına kadar bulut sembolü kullanılmıştır (Goyal, 2012). IBM, Google ve birçok üniversite, 2007 yılında bulut bilişim araştırma projeleri üzerinde çalışmış ve 2008 yılında, bilgi teknolojileri servisi kullanıcıları; servis hizmeti alanlar ve servis sağlayanlar olmak üzere ikiye ayrılmıştır. Bu bağlamda, şirketlerin kendi bünyelerindeki yazılım ve donanımları, belli bir takım modellere göre belirtilen servisleri oluşturmaları sonucu bulut bilişim kavramı ortaya çıkmıştır (Sevli, 2011). Gardner kuruluşu, 2010 yılında bilişim şirketleri arasında yapılan bir araştırmaya göre bulut bilişim, öne çıkan ilk üç bilişim teknolojisi arasında yer almaktadır (Koyuncu, 2012).

Özel kurumlar veya devlet kurumların mevcut olan bilişim teknolojileri alt-yapıları bir süre sonra güncelliğini yitirmekte ve yeni alt yapılara ihtiyaç duyulmaktadır. Ancak bulut bilişim tabanlı sistemler kurumların bilişim altyapı maliyetlerini azalttığı gibi daha esnek yazılım ve donanım seçenekleri sunmaktadır (Sarıtaş, 2013). Diğer bir ifadeyle bulut bilişim altyapısı sayesinde eğitime sürekli yatırım yapma gerekliliği ortadan kalkarak daha az maliyetli yazılım lisanslarına ve donanım kaynaklarına daha hızlı erişim sağlanmaktadır (Sevli, 2011).

Eğitim sürecinde iletişim, depolama, planlama, sunum hazırlama gibi işlemlerde ücretsiz çevrimiçi yaklaşımlar büyük avantajlar sağlamaktadır. Bu bulut tabanlı yaklaşımlardan Google Apps bünyesinde çok sayıda uygulama

barındırmakla birlikte Google Drive, Gmail, Takvim, Google E-tablolar, Google Dokümanlar, Google Siteler, Google Slaytlar ve Google Çeviri gibi uygulamalar eğitim amaçlı da kullanılmaktadır (Google, 2015).

Bulut çözümleri, kurumların bulutta daha çok iş yüklemeye yönelik girişimleriyle yeni bir "standart" haline gelmiştir. Bulut tabanlı çözümler, eğitim kurumları için maliyet tasarrufundan, eğitimlerin sunulma süresinin azaltılmasına kadar birçok avantaj sunmaktadır.



Şekil 3.2. Altyapı Olarak Örnek Bulut Şeması

Şekil 3.2' de altyapı olarak bulut tabanlı yaklaşımı ifade eden örnek tasarım görülmektedir. Bu yaklaşımda, bir kurumdaki diğer tüm alt birimler donanımsal ihtiyaçlarını bir ana omurga üzerinden karşılamaktadır. Bu sayede kaynaklar daha etkin kullanılmakta ve tek merkezden yönetilmektedir. Altyapı olarak bulut sistemlerinde depolama ve hız gibi ihtiyaçların güncellendiğinde mevcut omurga üzerine yeni depolama kaynakları veya işlemciler eklenmektedir. Kurulum aşamasında masraflı gibi görünen bulut altyapısına sahip sistemler uzun vadede çok daha ekonomik etkiye sahiptir.

4. VERİ MADENCİLİĞİ TANIMI, SÜRECİ, TEKNİKLERİ VE İLİŞKİLİ OLDUĞU ALANLAR

Bu kısımda, veri madenciliği tanımına, temel kavramlara, veri madenciliği işlem basamaklarına, veri madenciliği yöntemlerine, karar ağacı algoritmalarına, eğitimsel veri madenciliği tanımına ve veri madenciliğinin ilişkili olduğu alanlara yer verilmiştir.

4.1. Temel Kavramlar

Dünya devletlerinin güç politikaları incelendiğinde, birinci sanayi devrimi 18. yüzyılın ikinci yarısından sonra hammadde kaynaklarına bağlı olarak ve büyük mucitlerin etkileri sonucunda ortaya çıkmıştır. İkinci Dünya Savaşı sırasında tank, top, uçak ve roketlerle ilgili radar, atış kontrol sistemleri ve nükleer alanlarla ilgili gelişmelerin bilim ve teknoloji sayesinde geliştirildiği gözlemlenmiştir. Bu bağlamda, dünyanın güçlü devletleri gelişmeleri takip ederek teknolojiye ve bilime daha fazla yatırım yapmaya başlamışlardır. Elektronik, nükleer ve uzay alanlarında belli başlı endüstri ülkeleri arasında devam eden rekabet ve beraberinde geliştirdikleri teknolojilerin üretim sistemlerine de uygulanması ile bilgi çağı olarak adlandırılan yeni bir dönem ortaya çıkmıştır. Bu yeni çağda, gelişmiş sanayi toplumlarını geride bırakan, yoğun bilgi üretimi, bilgisayar teknolojisi ve zekâ ekonomisi değer kazanmıştır. Bilgi, ekonominin en önemli hammaddesi ve başlıca ürünü haline gelmiş, zenginlik oluşturmak için gerek duyulan sermaye varlıklarından bedensel emek, arazi, imalat aletleri ve fabrikalar yerini bilgiye bırakmıştır (Yenilmez, 1993).

Dijital çağ, veriden bilginin üretilmesi, tüketilmesi, uyarlanması, paylaşılması, kaynak ve servislerin dönüşüm işlemlerinde inanılmaz değişimleri beraberinde getirmiştir (Frank, 2015). En basit tanımıyla veri, bilgisayarlar

tarafından işlenebilen herhangi bir durum, sayı veya metin olarak tanımlanabilir (anderson.ucla.edu, 2016). Günlük yaşam içerisinde veri akış halindedir. Bu verilerin elde edildiği yerler; web sunucularına yapılan bağlantılardan, trafik akım algılayıcılarından, uydular, radyolar, bankacılık işlemleri, web sayfası içerikleri, cep telefonları, kredi kartları, televizyon ve bilgisayarlardır. Şehirlerin altyapılarında ise akıllı binalar, trenler, otobüsler, uçaklar, köprüler ve fabrikalardır. Veri akışı öyle hızlıdır ki son iki yılda oluşan veri şimdiye kadar kaydedilmiş toplam insan uygarlığı (insanlık tarihi) verilerini geride bırakmıştır (Shaw, 2014).

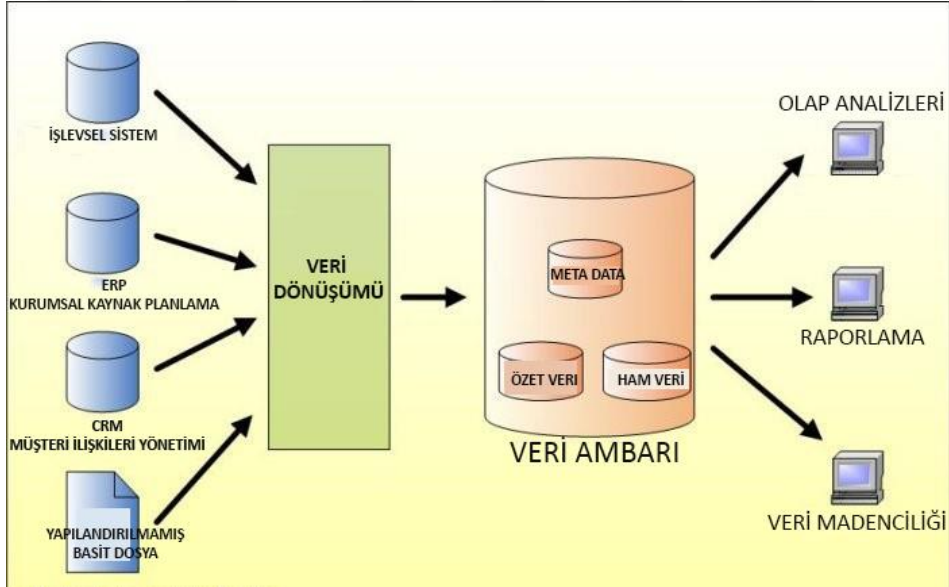
Veri: İşlenmemiş ham ifadelerle veri denilmektedir. Veriler bir konu hakkında temel seviyede sayısal veriler içerir. Örnek vermek gerekirse; bir kurumdaki öğrencilerin bir derste başarı puanı veri olarak kabul edilir.

Enformasyon (Malumat): Anlamlı bir konu etrafında verinin değerli hale getirilmesidir. Diğer bir ifadeyle; belirlenmiş verilerin bir anlam oluşturmak amacıyla işlenmiş haline enformasyon (malumat) adı verilir. Örneğin; başarı puanı bir veridir. Başarı puanları düzenli biçimde ölçülerek bir ilerleme tablosu oluşturulursa, veriyi enformasyona dönüştürülmüş olur. Burada bağlam olarak kabul edilen başarı puanı diğer başka verilerden temizlenmiş, tarih sıralaması yapılmış ve tablo haline dönüştürülmüştür. Elde edilen listeyi bir amaca uygun olarak değerlendirmek mümkündür.

Bilgi: Bilgi, düzenlenmiş verileri yorumlamak için gereksinim duyulan kuralların anlaşılmasıdır. Diğer bir ifadeyle; bilgi, enformasyonlar ile bunlardan yararlanarak neler yapılabileceği hakkında ilişkiyi kavrama yeteneğidir. Veriler; toplanır, sınıflandırılır, tekrar düzenlenir, saklanır, özetlenir, yeniden elde edilir ve iletilir. Son aşamada iletilen bilgidir. Örnek olarak, yukarıda bahsettiğimiz başarı puanını düşünürsek, belirli aralıklarla ölçülen başarı puanlarının geçmiş yıllara göre kıyaslanarak başarının düştüğü veya yükseldiği hakkında yorumlama işlemlerine bilgi denir.

Veri Ambarı: Veri madenciliği işlemleri için gerekli bilgi uygun formatta bir dosya veya veri tabanlarından elde edilir. Fakat işletmelerde kullanılan işlemsel veri tabanları veri madenciliği uygulamalarında kullanılamazlar. Veri

madenciliği uygulamasının yapılabilmesi için bu veri tabanlarında, belirli bir zaman dilimi veya konu odaklı olarak düzenleme, birleştirme ve sabitleme yapılması gerekir (Silahtaroglu, 2016). Söz konusu işlemlerin yapıldığı ilişkiel veri tabanlarına veri ambarı denilmektedir. Bir veri ambarı ortamı, kullanıcılar için veri toplama sürecini yöneten çıkarım, aktarım, dönüştürme ve yükleme çözümü, çevrimiçi analitik işleme altyapısı, kullanıcı analiz araçları ve diğer uygulamaları içerir (Oracle(a), 2016). Şekil 4.1'de çeşitli kaynaklardan toplanan verilerin istenilen biçimde veriye dönüştürülmesi, veri ambarında saklanması ve veri analizi, raporlama gibi işlemlerde kullanılması gösterilmiştir (Inmonna, 2016).



Şekil 4.1. Veri Ambarı Mimarisi

4.2. Veri Madenciliği Nedir?

Veri depolama cihazlarının son yıllarda ucuzlaması ve kolay depolama olanakları nedeniyle saklanan bilgi miktarı katlanarak artış göstermektedir. Sahip olunan büyük veri yığınları bir amaç doğrultusunda işlenmediği sürece bir değer ifade etmemekte ancak çeşitli veri işleme yöntemleriyle anlam kazanmaktadır. Birçok kaynak, verilerden anlamlı bilgi elde etme sürecini veri madenciliği olarak tanımlamaktadır.

- Veri madenciliği, karar verme işlevini daha iyi sağlamak amacıyla önceden bilinmeyen ilişkileri, kümeleri ve veri düzenlerini keşfetmek ve görüntülemek için veri tabanlarını inceleme sürecidir (Benoit, 2002).
- Veri madenciliği, istatistiksel ve matematiksel teknikler kullanılarak veri ambarlarında depolanan büyük miktarlardaki verinin örüntü tanıma teknolojileriyle birlikte incelenmesi yoluyla anlamlı yeni ilişkiler, eğilimler ve örüntüler bulunması sürecidir (Oracle(b), 2016).

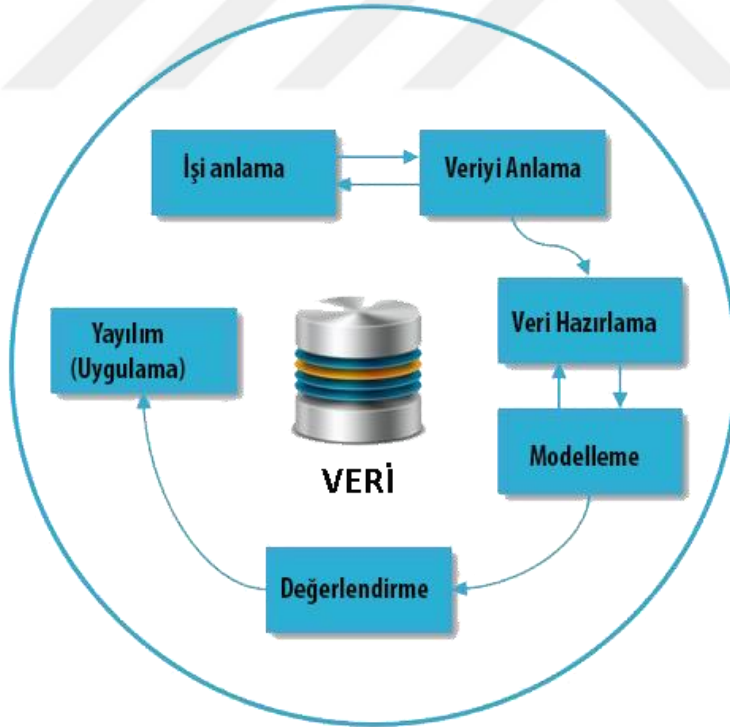
Tanımlar incelediğinde, veri madenciliğinin öne çıkan özellikler şöyle ifade edilebilir;

- Büyük veri depoları üzerinde işlem yapılabilir,
- Karar verme sürecine destek olur,
- Veriler arası ilişkiler ve anlamlar keşfedilir,
- Çeşitli matematiksel algoritmalar ve teknikler kullanılmaktadır.

4.3. Veri Tabanından Bilgi Keşfi Basamakları

Veri madenciliği belli basamakları olan bir süreci kapsamaktadır. Bu basamaklar arasında en çok iş yükünün olduğu kısım verilerin ön hazırlığının yapılmasıdır. Literatürde yer alan bazı çalışmalarda veri tabanından bilgi keşfi kavramı kullanılırken, veri madenciliği ve veri tabanından bilgi keşfi kavramlarını eş anlamlı olarak ifade eden çalışmalarda yer almaktadır.

Veri madenciliğinin kökeni 1980'lere dayanmasına rağmen 90'lı yıllarda hala emekleme aşamasındaydı. Veri madenciliği hala tanımlanmakta ve geliştirilmektedir. Genellikle veri modelleri kümesinde, algoritma analizleri ve özel amaçlı uygulamalarda hafif yığılmalar olmuştur. 1999'da otomobil üreticisi, sigorta sağlayıcı, donanım ve yazılım üreticisi, istatistiksel yazılım üreticisi büyük şirketler veri madenciliğini biçimlendirmek ve bir standart oluşturmak için bir araya gelmiştir. Yapılan çalışmalar sonucu veri madenciliği için çapraz endüstri standardı süreci olan CRISP-DM oluşturulmuştur. Bu standartlaştırma sonucunda veri madenciliği süreci veriden bağımsız işleyen 6 basamaktan oluşmaktadır (North, 2012). Şekil 4.2 bu basamakları göstermektedir. Şekil incelendiğinde, ilk basamakta işin anlaşılması yani veri madenciliği işlemiyle "Ne yapılmak isteniyor?" sorusu cevaplanmaktadır. Sonrasında ise veri madenciliği sürecinin verinin anlaşılması, verinin hazırlanması, modelin oluşturulması, değerlendirme ve son olarak yayılma (uygulama) basamaklarından oluştuğu görülmektedir.



Şekil 4.2. CRISP-DM İşleyiş Modeli

4.3.1. Problemin Tanımlanması

Veri madenciliği süreci problemin belirlenmesi ve gerekli olan ihtiyaçların tespit edilmesiyle başlamaktadır. Bu tespitler genel olarak örüntülere ilişkin sorular ve veri tabanındaki ilişkileri kapsamaktadır. Veri madenciliği çalışmalarında temel bileşen çalışmanın ne için yapıldığının bilinmesidir. Bu da üretilen yeni bir bilgiye duyulan yönetsel ihtiyaçlara ve çalışmanın iş hedeflerinin belirlenmesiyle başlar. Sonraki aşamada ihtiyaçların karşılanmasına için verinin toplanması, işlenmesi, sonuçların raporlanması faaliyetlerine ilişkin süreçle ilgili sorumlulukların tanımlandığı bir plan geliştirilir (Olson ve Shi, 2007). Esasında, neyi bilmek istediğimiz bilinmiyorsa, cevap aranacak sorular tanımlanmadıysa, veri madenciliği çabalarının verimli olması daha az bir olasılıktır (North, 2012).

Tanımlanacak problemlere şu örnekler verilebilir; “Öğrenciler çevrimiçi etkinlikleri yeteri kadar takip ediyor mu?” ve “Öğrencilerin başarılı olabilmeleri için etkinlikleri hangi düzeyde takip etmelidir?”.

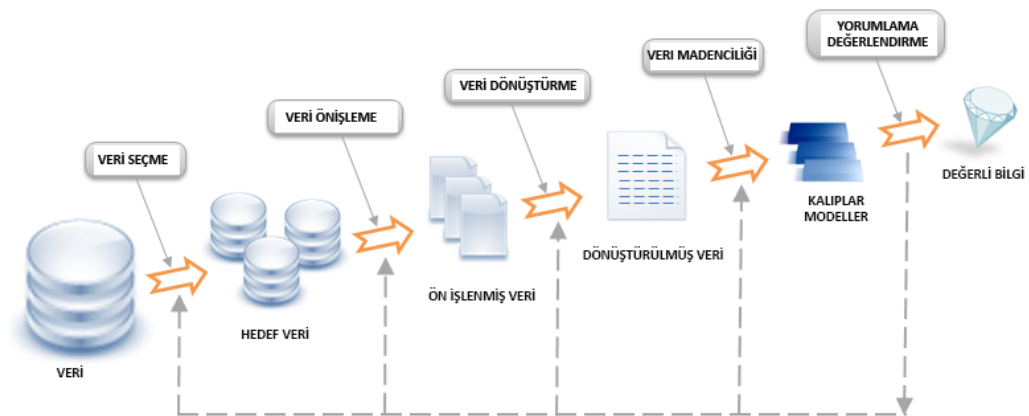
4.3.2. Veriyi Anlama

Veri madenciliği süreci için gerekli olan verilerin standart veri tabanlarından ziyade belirli düzenlemelerle oluşturulmuş merkezi veri ambarlarından elde edilmektedir. Ancak verilerin merkezileştirilmesi yeterli değildir. Bir kurumun verileri toplandıktan sonra bir takım sorulara cevap aranmalıdır. Bunlar; “Veriler nereden geldi?”, “Verileri kim topladı ve standart bir toplama yöntemi var mı?”, “Sütunlar ve veri satırlar ne anlama gelir?”, “Bilinmeyen kısaltmalar var mı?”. Veri madenciliği faaliyetlerinin veri hazırlama aşamasında bir miktar araştırma yapılması gerekebilir. Bazen, belirli verilerin geldiği yerden nasıl toplandığına, nasıl toplandığından, nasıl kodlandığından ve kaydedildiğinden emin olmak için çeşitli birimlerdeki konu uzmanlarıyla görüşmek gerekir. Verilerin doğruluğunu ve güvenilirliğini de doğrulamak çok önemlidir. Eski bir

söz olan "Bu hiç bir şeyden daha iyidir." sözü veri madenciliğinde geçerli değildir. Yanlış veya eksik veriler, bir veri madenciliği faaliyetinde hiçbir şeyden daha kötü olabilir çünkü eksik veya yanlış verilere dayalı kararlar kısmen veya yanlış kararlar alınmasına neden olabilir. Verileri sağlıklı bir şekilde topladıktan, tanımladıktan ve anladıktan sonra, veri madenciliği işlemlerine başlanılabilir.

4.3.3. Veri Hazırlama

Veri madenciliğinin en önemli basamaklarından biridir. Sonraki aşamalarda model kurulurken karşımıza çıkabilecek herhangi bir sorunda bu basamağa geri dönülüp, tekrar gözden geçirildikten sonra süreç tekrarlanır. En çok çaba ve zaman gerektiren kısım verinin ön işleme aşamasıdır (Sund, 2002). Farklı türde veri girişlerinin olduğu durumlarda kullanımı mümkün olmayan veriler oluşmaktadır. Örneğin bu çalışmada incelenen veri içerisinde "Anne mezuniyet durumu nedir?" sorusuna "ilk okul" ve "ilkokul" gibi aynı soruya verilen farklı cevap biçimleri bulunmakta. Bunun gibi tutarsız ve sonraki aşamada sorun olacak veriler düzeltilerek veri madenciliği için uygun hale getirilmelidir. Kaliteli bir hazırlık aşaması veri madenciliği sürecini olumlu yönde etkilemektedir. Verilerin hazırlanmasında geçen sürenin veri setinin büyüklüğü ile doğru orantılı olduğu düşünülebilir.



Şekil 4.3. Veri Hazırlama Basamakları

Şekil 4.3'te de görüleceği üzere veri hazırlama aşaması veri temizleme, veri birleştirme, veri dönüştürme ve veri indirgeme işlemlerini kapsamaktadır.

4.3.3.1. Veri Temizleme

Veri madenciliğinin bu basamağındaki amaç, kirli veri olarak nitelendirilen (Silahtaroglu, 2016) kayıp ve gürültünün ortadan kaldırılmasıdır. Ayrıca, eksik ve uç değerler içeren verilerin temizlenmesi bu kısımda yapılmaktadır.

Gürültülü, yanlış veya eksik verilerin temizlenmesi için çeşitli yöntemlerin kullanılması gerekmektedir. Veri temizleme amacıyla kümeleme, histogram tekniği, veri dağılımı incelenebilir veya regresyon teknikleri gibi istatistiksel yöntemler kullanılabilir. Eksik veriler, diğer kayıtların aynı tür sütunlarının ortalaması hesaplanarak bulunabilmekte, bulunduğu sütunda en çok tekrar eden değer ile doldurulabilmekte veya olasılık yöntemleri aracılığıyla kayıp değer yerine kullanılacak bir değer belirlenebilmektedir. Verinin boyutu küçültülerek yapılan veri indirgeme, veri madenciliği algoritmalarının zaman ve depolama alanını azaltan bir yöntemdir (Chapman vd., 2000).

4.3.3.2. Veri Birleştirme

Veri madenciliği işlemi için birçok veri tabanından gelen verilerin belirli kurallar ölçüsünde düzenlenmesi veri birleştirmeyi ifade etmektedir. Aslında bu tanım veri ambarını işaret etmektedir. Başka bir ifadeyle, farklı veri tabanlarından veya kaynaklardan gelen farklı türdeki verinin tek türde birleştirilmesi söz konusudur.

4.3.3.3. Veri Dönüştürme

Veri madenciliği uygulamalarında kullanılan model, algoritma ve teknikler bazı türdeki verilerle uyumlu çalışmayabilir. Bazı algoritmalar, sadece sayısal değerlerle çalışırken bazıları ise “0” ve “1” gibi kategorik değerlerle çalışırlar. Bu gibi durumlarda veriler kullanılan algoritmaya uygun hale getirilmelidir (Silahtaroglu, 2016). Veri dönüştürme sayesinde veriler, veri madenciliği işlemlerine uygun hale getirilmektedirler. Çoğunlukla birleştirme, düzeltme, normalleştirme ve genelleştirme olarak adlandırılan dönüştürme türleri kullanılmaktadır (Şentürk, 2006). Örnek ile açıklamak gerekirse; Öğrenci yaşlarının veri madenciliğine dahil edildiği bir işlemde öğrenci yaşlarının 18 ile 60 yaş aralığında olduğunu varsayalım. Bu durumda çok fazla çeşitte öğrenci yaşı verisinin kayıtlı olması beklenir. Bu durum kategorik yapıda analize uygun bir algortmada istenmeyen bir veri sütunu olacaktır. Bu çeşitliliği önlemek için öğrenci yaşları belirli aralıklarda gruplara ayrılarak veri madenciliği uygulamasına devam edilebilir.

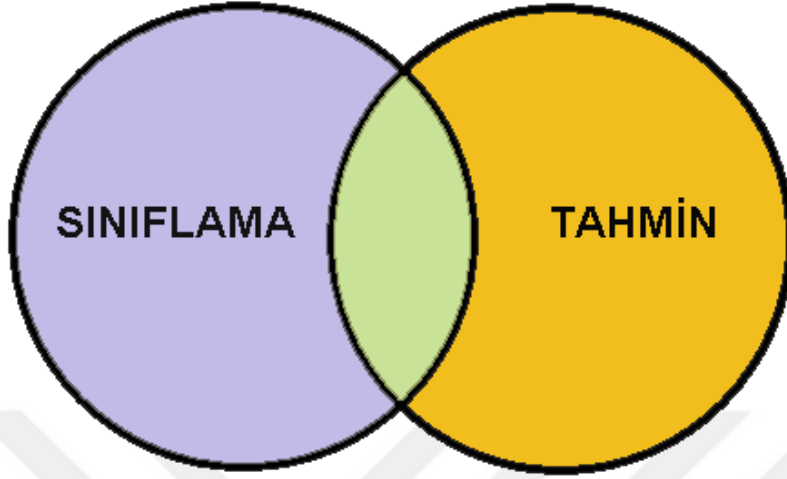
4.3.3.4. Veri İndirgeme

Veri madenciliği işlem süreci bazı durumlarda çok uzun sürebilir. Eğer elde edilen sonucun değişmeyeceğine inanılıyorsa veri setindeki kayıt sayısı veya değişkenlerin sayısı azaltılabilir. Veri indirgeme, veriyi birleştirme veya veri küpü, boyut indirgeme, örnekleme ve genelleme gibi yöntemlerle yapılabilir (Han, 2001).

4.3.4. Modelleme

Veri madenciliği, gerçek dünya gözlemlerinin bilgisayar ortamında temsili bir modelidir. Modeller, verilerinizdeki örüntüleri veya mesajları aramak,

tanımlamak ve görüntülemek için algoritmaların bir uygulamasıdır. Veri madenciliğinde iki temel tür veya model çeşidi vardır: sınıflandırma ve tahmin (North, 2012).



Şekil 4.4. Veri Madenciliği Model Türleri

Şekil 4.4'te görüldüğü gibi, veri madenciliğinin kullandığı model türleri arasında bazı örtüşmeler vardır. Örneğin, karar ağaçları, belirli bir veri kümesinin hangi özelliklerinin belirli bir sonucun en güçlü göstergeleri olduğunu belirlemek için kullanılan bir tahmin modelidir. Sonuç genellikle, bir gözlemin belli bir kategoriye girme ihtimali olarak ifade edilir. Böylece, karar ağaçları doğada öngörücüdür, fakat aynı zamanda verilerimizi sınıflandırmamıza yardımcı olurlar. Modellerin, verilerde tespit edilen kalıplara göre sınıflandırma yapma ve tahmin etmeye yardımcı olabileceği ifade edilebilir. Modeller basit veya karmaşık olabilir. Yalnızca tek bir süreç akışı veya alt işlemler içerebilirler. Modeller veri madenciliğinin hazırlık ve anlaşılması aşamasından, gelişim ve yorumlamaya geçildiği aşamadır.

4.3.5. Deęerlendirme

Model Deęerlendirme, model geliřtirme s¼recinin ayrılmaz bir parçasıdır. Veriyi temsil eden en iyi modeli bulmaya ve seęilen modelin gelecekte ne kadar iyi alıřacaęına karar vermeye yardımcı olur. Veri madencilięinde eęitim iin kullanılan veriler model performansını deęerlendirmede kullanılması uygun deęildir, ¼nk¼ optimum ve uygun modeller kolayca oluřturabilir (Sayad, 2016). Veri madencilięi s¼reci sonunda artık bir sınıflandırma modeli oluřturulsa bile, aklınızda birok soru olabilir. Örneęin, öęrencilerin etkinliklerini öngörerek bir sınıflandırıcı oluřturmak iin önceki etkinlik verilerinin kullanıldığını varsayalım. Sınıflandırıcının gelecek öęrenci davranıřlarını, yani sınıflandırıcının eęitilmedięi gelecekteki öęrenci etkinliklerini ne kadar doęru tahmin edebileceğini bilmek isteriz. Hatta birden fazla sınıflandırıcı oluřturmak iin farklı yöntemler test edilebilir ve bunların doęrulukları karřılařtırmak istenebilir. Ancak sınıflandırma doęruluęu ve bunun tahmin edilmesi, güvenilir bir doęruluk tahmini nasıl elde edilir? Bu gibi sorular sınıflama bařlıęı altında ele alınmaktadır.

4.3.6. Yayılma (Uygulama)

Veri madencilięi iřlemlerinde eęer soruların tanımlanması iyi yapılırsa, veriler hazırlanır ve tanımlanan sorulara cevap bulunursa sonrasında bir model oluřturulduysa artık sonuçlar gerekten kullanma noktasına gelmiř ve kullanılırdır. Bu ařamadaki etkinlikler arasında, modelin ıktılarının kullanıcılar ile buluřturulması, mevcut yönetim veya bilgi sistemleriyle b¼t¼nleřtirilmesi, doęruluk ve performansını artırmak iin modele yeni öęrenmelerin dahil edilmesi, model öęrenme ıktılarının ölç¼lmesi ve görselleřtirilmesi yer alır.

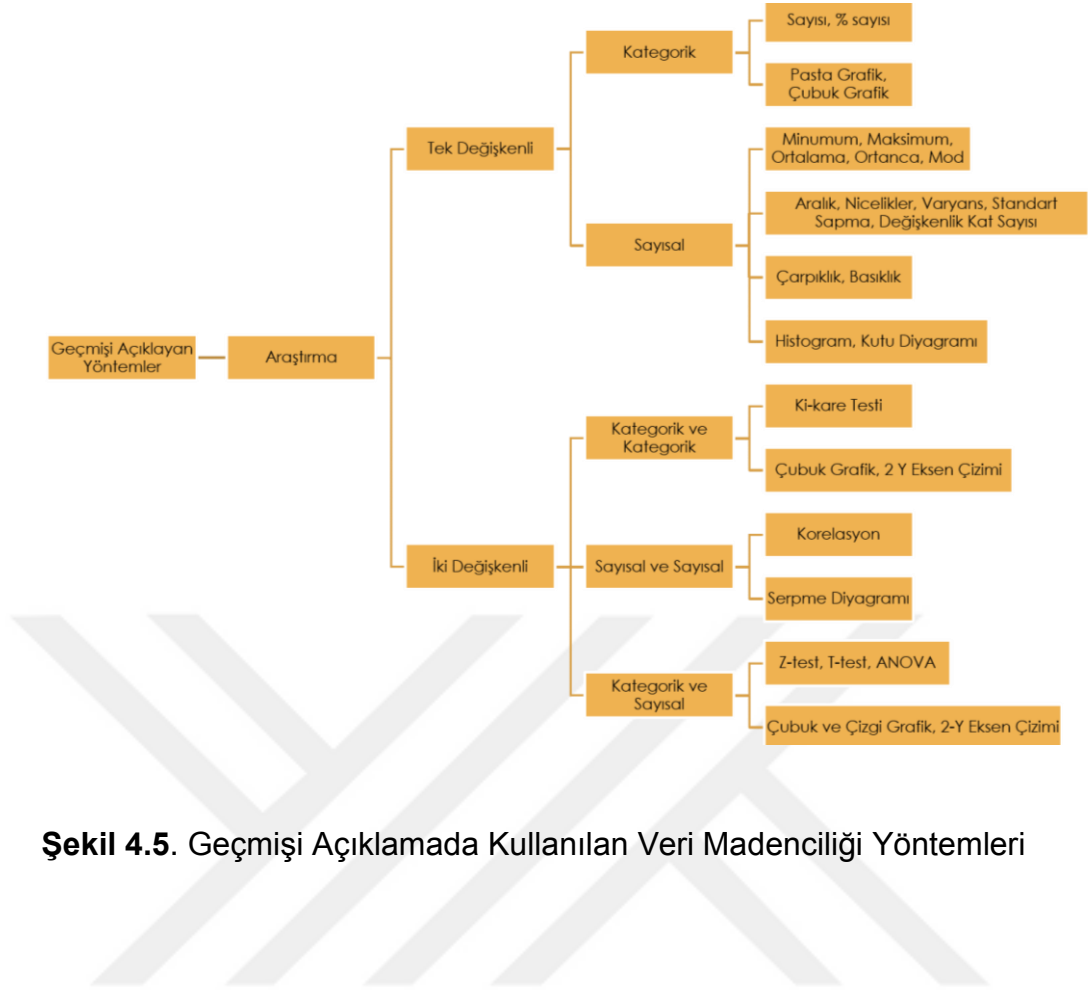
Veri madencilięi iřlemlerinin bilgisayarlar tarafından yapılması iin ok sayıda yazılım oluřturulmuřtur. Bu yazılımlar, kaydedilmiř verilere eřitli modellerin

uygulanmasıyla tablolar halinde ve görselleştirilmiş sonuçlar vermektedir. (North, 2012). Veri madenciliği sürecinde MATLAB, SPSS, rapid miner, R, Python, Weka, Knime gibi bir çok yazılım aracı ve programlama dilleri kullanılmaktadır. Bu tez çalışmasının veri işleme basamaklarında SPSS programı kullanılmıştır.

4.4. Veri Madenciliği Yöntemleri

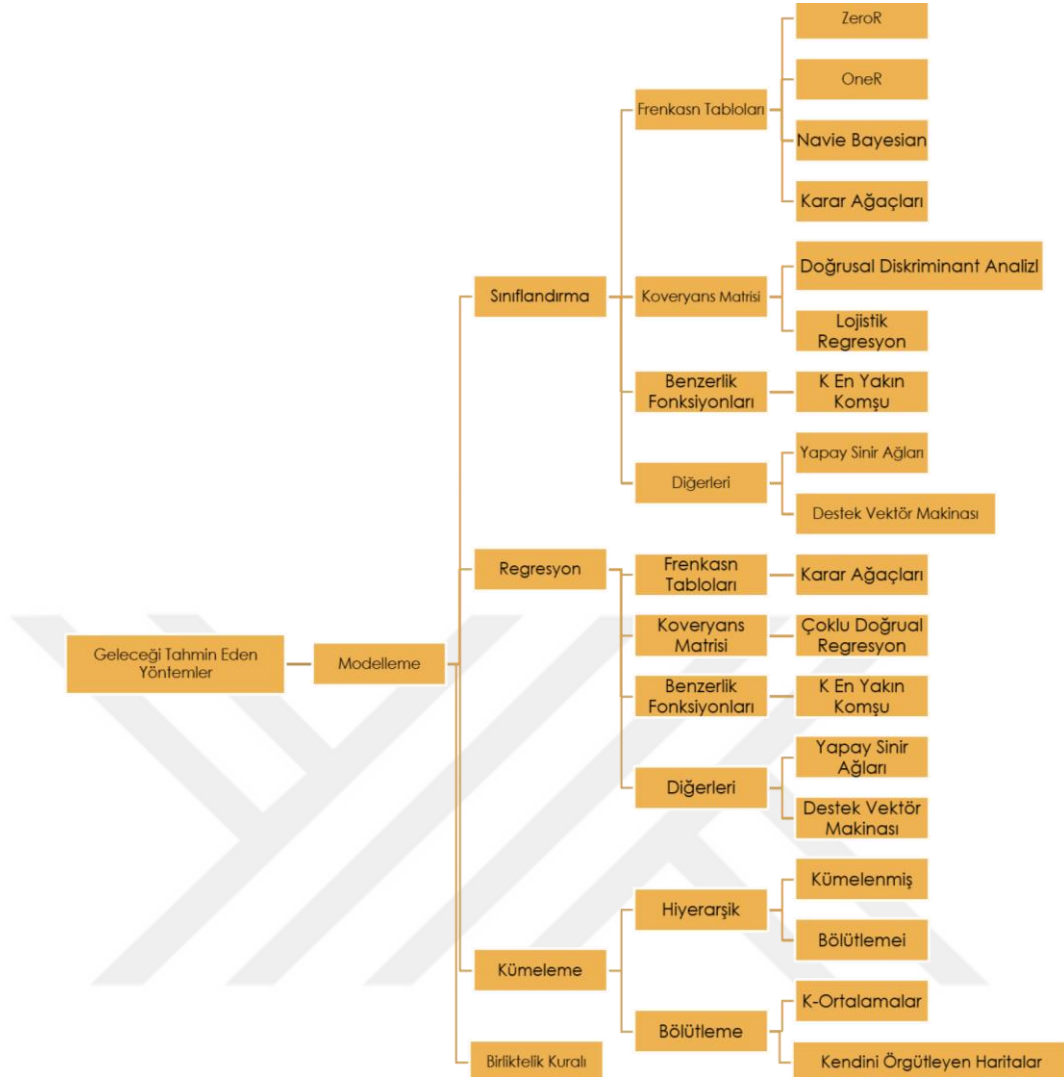
İnternet ortamında bu kadar hızlı artan veri işlenerek bilgi haline getirilmediği takdirde pek bir değeri olmayacaktır. Daha önceden belirttiğimiz gibi işlenmemiş veriler ham petrole benzer (Palmer, 2006; Rotella, 2012). Veri rafine edilmezse, değeri vardır ama kullanılabilir değildir. Ham verilerde benzer şekilde işlendiği zaman değer kazanmaktadır ve çeşitli veri madenciliği yöntemleriyle işlenerek değerli hale getirilebilir.

Veri madenciliği sürecinde, basit bir analizin ötesinde, model ve eğilimlerin keşfi için çok büyük depolama birimleri otomatik taranır. Veri madenciliğinde, verileri bölümlere ayırmak ve gelecekteki olayların olasılığını değerlendirmek için çok yönlü matematiksel algoritmalar kullanılmaktadır. Veri madenciliği veri keşfi olarak da bilinir (Oracle(b),16). Veri madenciliği yöntemleri temelde 3 başlık altında toplanmaktadır. Bunlar; sınıflama, kümeleme ve birliktelik kurallarıdır (Silahtaroglu, 2016).



Şekil 4.5. Geçmiş Açıklamada Kullanılan Veri Madenciliği Yöntemleri

Şekil 4.5'te veri madenciliğinde geçmiş yönelik açıklamaları yapmak için kullanılan yöntemler gösterilmektedir. Şekil incelendiğinde, tek değişkenli ve iki değişkenli olmak üzere 2 ana kategoriye ayrıldığı, tek değişkenli kısmın kategorik ve sayısal olarak 2'ye ayrıldığı ve bu kısımların altında yine alt yöntemler yer almaktadır. Aynı şekilde iki değişkenli yapı kategorik ve kategorik, sayısal ve sayısal, kategorik ve sayısal olmak üzere 3 kısma ayrılmaktadır. Üç ayrı kategoriye ayrılan yapının altında yine çeşitli alt yöntemler görülmektedir.



Şekil 4.6. GelecekTahminlerinde Kullanılan Veri Madenciliği Yöntemleri

Şekil 4.6 incelendiğinde, geleceğe yönelik öngörü elde etmek için sınıflama, regresyon, kümeleme ve birliktelik kuralları gösterilmiştir. Sınıflama yöntemleri önce 4 kategoriye, sonrasında 9 alt kategoriye, regresyon yöntemi önce 4 sonra 5 alt kategoriye, kümeleme yöntemi 2 kategori ve 4 alt kategoriye, son yöntem olarak birliktelik kurallarının tek başlıkta verilerek geleceğe yönelik tahmin yöntemleri özetlenmiştir.

4.4.1. Sınıflama Yöntemleri

Sınıflandırma veri madenciliği yöntemleri arasında en çok bilinenlerden birisidir. Hastalık tanıları, dolandırıcılık tespiti, resim, örüntü tanıma, kalite kontrol çalışmaları ve pazarlama konuları sınıflandırma yönteminin sıkça kullanıldığı alanlardır (Margaret, 2003; Han, 2012). Sınıflandırma, kategorik bir değişkenin (hedef veya sınıf) değerini bir veya daha fazla sayısal veya kategorik değişkeni (öngörücü veya nitelik) temel alan bir model oluşturarak tahmin etmeyi amaçlayan bir veri madenciliği işlevidir. Öğrencilerin yıl sonu değerlendirmesine göre başarılı veya başarısız olarak kategorize etmek için sınıf modeli kullanılabilir. Sınıflandırma yöntemleri 4 ana grup altında toplanır; frekans tablosu, kovaryans matrisi, benzerlik fonksiyonları ve diğer algoritmalar (Sayad, 2016).

Sınıflandırma algoritmalarını, frekans ölçümlerine dayalı temel sınıflandırıcılar bayes sınıflandırıcılar, karar ağacı sınıflandırıcıları ve kural tabanlı sınıflandırıcılar olarak gruplayabiliriz. Sınıflama işlemi ile neler yapılacağı konusunda örnekler çoğaltılabilir. Örneğin bir tıp araştırmacısı kanser tedavisi gören bir hastasına A, B, C gibi tedavi yöntemlerden hangisinin daha uygun olacağını tespit edebilir. Veya bir finans şirketi yöneticisi müşterilerinin riskli grup, güvenli grup olarak ayırarak kredi sağlayacağı ya da kredi vermemesi gereken müşteri potansiyeli belirleyebilir (Han, 2012).

ZeroR: Frekans ölçümlerinin kullanıldığı sınıflandırma yöntemidir. Tüm öngörücüleri göz ardı eden ve hedefe dayalı en basit sınıflandırma yöntemidir. Bu yöntemde, hedef için oluşturulan frekans tablosunda en sık kullanılan değer tespit edilerek sınıflanır.

OneR: İngilizce kelimelerin (One Rule) bir kuralın kısaltması olan OneR, verilerin her öngörüsü için bir kural oluşturan basit ve doğru bir sınıflandırma algoritmasıdır. En küçük toplam hataya sahip olan kuralı "bir kural" olarak seçer. Bir kural oluştururken, her tahmin için hedefe karşı bir frekans tablosu oluşturur. OneR'nin, insanların yorumlamaları kolay olan kurallar üretirken,

modern sınıflandırma algoritmalarından daha az doğru kurallar ürettiği görülmektedir.

Naive Bayesian: Naive Bayesian sınıflandırıcısı, tahmin ediciler arasındaki bağımsızlık varsayımlarıyla Bayes teoremine dayanmaktadır. Naive Bayesian modelini, özellikle çok büyük veri kümeleri için kullanışlı kılan, karmaşık yinelemeli parametre tahminleri olmaksızın oluşturmasıdır. Basitliğine rağmen, Naive Bayesian sınıflandırıcısı genellikle daha gelişmiş sınıflandırma yöntemlerinden şaşırtıcı derecede iyi sonuç verir ve yaygın olarak kullanılır.

Bayes Sınıflandırma, bayes karar teorisinin kullanıldığı istatistiksel işlemler ile tanımlama ve sınıflandırma modellerinde uygulanan temel tekniklerdendir. Modellerin sınıflandırılması olasılıklar biçiminde belirlenir. Sınıflandırma kuralları içerisinde en çok kullanılan “Naive Bayesyen Sınıflandırma” yapay (Neural) sinir ağları veya kara ağaçları gibi diğer sınıflandırma araçlarına göre büyük veri tabanlarında yüksek hız ve tutarlılık sağlamaktadır (Han, 2012).

Naive Bayes Sınıflama Yöntemi verdiği başarılı sonuçlar nedeniyle son yıllarda sıkça kullanılmaktadır. Bu olasılıklı yaklaşımlar verinin nasıl oluşturulduğu hakkında kuvvetli varsayımlar oluşturmakta ve bu varsayımları içeren olasılıksal modeller ortaya koymaktadır. Sonraki aşamalarda üretken model parametrelerini hesaplamak için belirlenmiş örnek derlemeleri kullanılır. Gerçek dünyada bu varsayımlar genelde yanlış olsa da, Naive Bayes sınıflandırıcı, sınıflandırmalar konusunda oldukça başarılı sonuçlar vermektedir (McCallum, 2003).

Algoritmatima Formülü

Bayes Teoremi, $P(c | x)$ 'i $P(c)$, $P(x)$ ve $P(x | c)$ 'den olmak üzere sonraki olasılığı hesaplamının bir yolunu sağlar. Naive Bayes sınıflandırıcısı, bir öngörücünün (x) değerinin verilen bir sınıftaki (c) etkisinin diğer öngörücülerin değerlerinden bağımsız olduğunu varsaymaktadır. Bu varsayıma, sınıfın koşullu bağımsızlığı denir.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$P(c|x)$, öngörücü (nitelik) verilen sınıfın (hedefin) sonraki olasılığıdır.

$P(c)$ sınıfın ön olasılığıdır.

$P(x|c)$, verilen sınıf tahmin edicisinin olasılığıdır.

$P(x)$ tahmini öncül olasılıktır

4.4.1.1. Karar Ağaçları

Makine öğrenmesi araştırmacısı olan J. Ross Quinlan 1970'lerin sonu ve 1980'lerin başında ID3 (Iterative Dichotomiser) olarak bilinen bir karar ağacı algoritması geliştirdi. Daha sonra Quinlan, ID3 algoritmasının yerini alacak olan C4.5'i geliştirdi ve bu algoritma yeni denetlenen öğrenme algoritmalarının karşılaştırıldığı bir kriter haline geldi. 1984 yılında bir grup istatistikçi (L. Breiman, J. Friedman, R. Olshen ve C. Stone), ikili karar ağaçlarının oluşumunu tanımlayan Sınıflama ve Regresyon Ağaçları (CART) adlı kitabı yayınladılar. ID3 ve CART aynı anda birbirinden bağımsız olarak keşfedilmiş, fakat karar ağaçlarını öğrenme verilerinden öğrenmek için benzer bir yaklaşım izlemişlerdir. Bu önemli algoritmalar, karar ağacı kullanımının başlangıç sürecinde büyük bir etki oluşturur.

ID3, C4.5, ve CART aç gözlü bir yaklaşımı benimsemiş, yukarıdan aşağıya doğru böl-keşfet biçiminde işleyen algoritmalarıdır. Bu algoritmalar, CHAID algoritması Kaas tarafından 1980 yılında geliştirilmiştir. CHAID algoritması, verileri hızlı ve etkin bir biçimde keşfeden ve istenen sonuca göre kesimler ve profiller oluşturan hızlı, istatistiksel, çok yönlü, ağaç yapılı algoritmadır (IBM(a), 2016). Türkiye'de karar ağaçları algoritmaları kullanarak yürütülen çok sayıda çalışma yapılmıştır. Bunlardan bazıları; karar ağaçları algoritmaları kullanarak hastanelerin durumlarına ilişkin sonuçlar çıkarılmış ve bu sonuçlardan elde edilen verilerle kümeleme analizi yapılarak

gruplandırmıştır (Dolgun, 2008). Başka bir çalışmada, meme kanseri hastalarında yinelemesiz hayatta kalma süresini etkileyen risk etkelerinin tespit edilmesinde karar ağacı yöntemlerinden CRT, CHAID, QUEST, C4.5 ve ID3 ile Kaplan-Meier analizi birlikte kullanılmıştır (Ture, 2009).

Ağaç oluşturulurken, eğitim seti tekrarlanan bir şekilde daha küçük alt gruplara bölünür. Temel karar ağacı algoritması aşağıda gösterilmektedir (Han, 2012).

Yöntem:

- (1) Bir düğüm oluştur N;
- (2) Eğer D'de değişken grupları aynı sınıfta ise, C, ise
- (3) C sınıfı ile etiketli bir yaprak düğüm olarak N' ye dön;
- (4) Eğer Öznitelik listesi boş ise
- (5) D'de çoğunluk sınıfıyla etiketlenmiş bir yaprak düğüm olarak N'ye dön// çoğunluk seçimi
- (6) En iyi bölme kriterini bulmak için uygula Oznitelik_secme_yontemi (D, öznitelik listesi);
- (7) Bölme kriteri ile N düğümünü etiketle;
- (8) Eğer özniteliği bölme ayrık değerli ve çok yönlü bölmelere izin veriyor ise // İkili ağaçlarla sınırlı değil
- (9) Öznitelik listesi <- öznitelik listesi - bölme özniteliği;// Bölme özniteliğini kaldır // bölme özniteliğini kaldır
- (10) bölme kriterinin herbir j sonucu için Dön // her bölme için değişken gruplarını ve gelişen alt ağaçları
- (11) Dj'yi, D'deki j sonucunu tatmin eden veri takımı kümesi olsun; // bir bölme
- (12) Eğer Dj boş ise
- (13) D çoğunluğu sınıfıyla etiketlenmiş bir yaprağı düğüm N'ye bağlayın;
- (14) Değilse oluşturulan karar ağacı (Dj, öznitelik listesi) tarafından döndürülen düğümü düğüm N'ye bağlayın;
- Döngüyü bitir
- (15) N'ye dön;

Yukarıda verilen algortimada, D'deki değişken grupları aynı sınıfta bulunuyorsa, düğüm N sayıda yaprak haline gelir ve bu sınıfla etiketlenir (2. ve 3. adımlar). 4. ve 5. adımların sonlandırma koşullarıdır. Tüm sonlandırma koşulları, algoritmanın en sonunda açıklanmaktadır. Aksi takdirde bu algoritma, bölme kriterlerini belirlemek için öznitelik seçim yöntemini çağırır. Bölme kriteri, D düğümlerini bireysel sınıflara ayırmak veya bölmek için "en iyi" yolu belirleyerek (6. adım), düğüm N'de hangi özniteliğin test edileceğini

söyler. Seçim kriteri ayrıca, testin sonuçlarına göre N düğümünden hangi dalların ayrılacağını söyler. Esasında, bölme, bölme özneliğini gösterir ayrıca bölme noktası veya bölme alt kümesini de belirtebilir. Bölme kriterleri böylece ideal olarak, her daldaki bölmeler mümkün olduğunca "saf" olacak şekilde yapılır. Diğer bir ifadeyle, D'deki değişen grupları bölme kriterinin kendi özel sonuçlarına göre bölünürse, sonuçtaki bölümlerin mümkün olduğunca saf olması beklenir. N düğümü, düğümde bir test görevi gören bölme kriteri ile etiketlenmiştir (7. adım). Bölme kriterinin her bir sonucu için N düğümünden bir dalı ayrılır. D'deki değişkenler buna göre bölümlendirilir (10-11. adım).

Karar ağaçları temelde sınıf etiketli eğitim verilerinden oluşan bir öğrenmedir. Karar ağacı, akış şemasına benzer şekilde dallanan ağaç yapısındadır. Karar ağacında her yapraksız düğüm, bir özellik üzerinde bir testi gösterir, her dal testin bir sonucunu temsil eder, her bir yaprak bir veri sınıfının etiketini gösterir. Ağacın en üstündeki düğüm kök olarak ifade edilir.

Veri sınıflandırma, veri modelinin oluşturulduğu bir öğrenme aşaması ve mevcut verilerin hangi sınıfta olduğunun tespit edildiği aşama olmak üzere iki aşamadan oluşur. Şekil 4.7 ve Şekil 4.8'de veri sınıflandırma basamakları örnek verilerek gösterilmiştir.

Öğrenme Verisi				→	Sınıflama Algoritması		
id	Yaş	Yerleşme Puanı	Durum				
1	18-20	Yüksek	Başarılı				
2	18-20	Düşük	Başarısız				
3	21-23	Yüksek	Başarılı				
4	23+	Düşük	Başarısız				
5	18-20	Yüksek	Başarılı				
6	23+	Yüksek	Başarısız				
...

Sınıflama Kuralları		
Eğer yaş=18-20 ise Başarılı		
Eğer yerleşme puanı=yüksek ise Başarılı		
Eğer yaş=23+ ise Başarısız		

Şekil 4.7. Veri Sınıflandırma Kurallarının Oluşturulması

Test Verisi				→	Sınıflama Kuralları		
id	Yaş	Yerleşme Puanı	Durum				
1	18-20	Yüksek	Başarılı				
2	18-20	Düşük	Başarısız				
3	21-23	Yüksek	Başarılı				
4	23+	Düşük	Başarısız				
5	18-20	Yüksek	Başarılı				
6	23+	Yüksek	Başarısız				
...				
					Yeni Veri		
					Yaş	Yerleşme Puanı	Durum
					18-20	Yüksek	?
					21-23	Düşük	?
					23+	Düşük	?
				

Şekil 4.8. Yeni Verilerin Sınıfını Tespit Etme İşlemi

Şekil 4.7’de verilen örnekte eğitim verisi sınıflandırma algoritmalarıyla işleme tabi tutulduğunda elde edilen sınıfın niteliği durum için karar verir. Öğrenilen model sınıflandırma kuralları biçiminde temsil edilir. Şekil 4.8’de ise sınıflandırma modelinin doğruluğunu karar vermek için test verisi kullanılır. Model doğru kabul edilirse kurallar yeni sınıflandırma işlemlerinde kullanılır.

Burada eğitim verisinin etiketi bilindiğinden bu işlem denetimli öğrenme olarak bilinmektedir. Denetimsiz öğrenmede ise eğitim verilerinin sınıfı veya sınıfların sayısı önceden bilinmemektedir. Sınıflama dışında benzer veri gruplarını tespit etmek için kümeleme işlemi uygulanır. Kısaca, veri sınıflandırma işleminde öncelikle eğitim verisi kullanılır, kurallar öne sürerek model oluşturulur. Ortaya çıkan model sınıflandırma işlemlerinde kullanılır.

Karar ağaçlarında, farklı yöntemleri kullanan öznelik seçme yöntemleri vardır. Bunlar; bilgi kazanımı (information gain), kazanım oranı (gain ratio), gini katsayısı (gini index) ve diğer seçim yöntemleridir.

Karar ağaçları, anlaşılabilirliği ve kolay yorumlanması karar vericiler için aşağıda sıralanan avantajları sağlamaktadır (Chien, 2008).

- Maliyetinin düşük olması,
- Anlaşılmasının, yorumlanmasının kolaylığı ayrıca veri tabanları ile rahatlıkla iletişim kurması,
- Güvenilirliklerinin iyi olması.

Karar ağacı yapısında, karar noktaları kare; şans noktaları daire, karar dalları ve şans dalları çizgilerle sembolize edilir. Son noktalar getirileri veya faydaları gösterir.

Karar ağacı oluşturma basamakları şöyle sıralanır;

- Problemin tanımlanması,
- Karar ağacının oluşturulması / yapılandırılması,
- Olay oluşma olasılıklarının hesaplanması,
- Beklenen kazancın ilgili şans noktası için hesaplanması,
- En yüksek beklenen kazancın (faydanın) ilgili karar noktasına atanması,
- Önerinin sunulması.

Çizelge 4.1. Karar Ağacı Algoritmaları

Karar Ağacı Algoritması	Genel Özellikleri
CRT	Kök veya uç olmayan her bir düğümde iki adet dal bulunur. Gini'ye göre ikili bölme işlemi mevcuttur. Ağacın karmaşıklık ölçüsü budama işlemlerini etkiler. Sınıflandırma ve regresyonu destekleyici bir yapısı vardır. Sürekli hedef değişkenleriyle çalışır. İşleyiş öncesinde verinin hazırlanmasına ihtiyaç duyar.
C4.5 ve C5.0	Düğümlerden ayrılan çoklu dallar ile ağaç oluşturur. Kategori sayısı tahmin edicinin dal sayısına eşittir. Birden fazla karar ağacını tek bir sınıflayıcı ile birleştirir. Ayırma işlemlerinde bilgi kazancı kullanılır. Her yapraktaki hata oranına göre budama işlemi yapılır.
CHAID	Bölme işlemlerinde Ki-kare testleri kullanır. Dal sayısı, iki ile tahmin edicinin en yüksek kategori sayısı arasında değişir.
SLIQ	Hızlı ağaç budama algoritmasına sahip ve hızlı ölçeklenebilir bir sınıflayıcıdır.
SPRINT	Bölme süreci tek bir özelliğin değerine dayanır. Büyük veri setleri için kullanışlıdır. Tüm bellek sınırlamaları üzerinde nitelik listesi veri yapısını kullanarak işleyiş sağlar.

Karar ağaçlarında kullanılan çok sayıda algoritma bulunur. Bunlardan bazıları ID3, C4.5, C5.0, CHAID, CART ve QUEST'tir. Çizelge 4.1'de karar ağacı algoritmaları ve genel özelliklerine yer verilmiştir (Emel, 2005).

4.4.2. Kümeleme Yöntemleri

Kümeleme yöntemi, veri setindeki verileri belirlenmiş benzerlik kriterlerine göre gruplara ayırma işlemidir. Her bir gruba "küme" adı verilir. Kümeleme işleminde küme içindeki elemanların benzerliği fazla, kümeler arası benzerlik ise az olmalıdır.

Kümeleme, tanımlayıcı model veri madenciliği tekniğidir, başka bir ifadeyle gözetimsiz sınıflandırmaya girer. Gözetimsiz sınıflama yönteminde, ilk başta verilen ve henüz sınıflandırılmamış bir küme, veriyi anlamlı alt kümeler oluşturacak şekilde öbeklemektir. Kümeleme işlemi, tamamen gelen verinin niteliklerine göre yapılır.

Kümeleme analizinin kullanılmasında benzer uzaklıklar göz önünde tutularak yararlanılabilecek alternatif ölçüm ve yöntemler bulunmaktadır. Birimler arası mesafeler için standardize euclidyen, euclidyen, kareli euclidyen, minkowski, manhattan mahalanobis veya canberra ölçüleri kullanılabilir. Bu da kümeleme analizinin uygulamada kullanılmasında dikkatli davranmayı zorunlu kılmaktadır. Kümeleme algoritmasıyla veritabanı alt kümelere ayrılır. Her bir kümede yer alan elemanlar içinde buldukları grubu diğer gruplardan ayıran ortak niteliklere sahiptir. Kümeleme yöntemleri çizelge 4.2'de kısaca özetlenmiştir.

Bazı kümeleme algoritmaları, birkaç kümeleme yönteminin birleşmesiyle kullanılabilir, böylece belirli bir algoritmayı yalnızca tek bir kümeleme yöntemi kategorisine özgü olarak sınıflandırmak bazen zordur. Ayrıca, bazı uygulamalarda birkaç kümeleme tekniğinin birleştirilmesini gerektiren kümeleme kriterleri olabilir (Han, 2012).

Çizelge 4.2. Kümele Yöntemleri

Yöntem	Genel Özellikleri
Bölümlenme Yöntemleri	<ul style="list-style-type: none">• Küresel biçimde birbirinden ayrılan kümelerdir,• Uzaklık tabanlı,• Küme merkezini temsil etmek için ortalamayı kullanabilir,• Küçük-orta büyüklükteki veri setleri için etkilidir.
Hiyerarşik Yöntemler	<ul style="list-style-type: none">• Kümelenme hiyerarşik ayrışmadır (çoklu seviyeler),• Hatalı birleştirmeleri veya bölmeleri düzeltemez,• Mikro kümeleme gibi diğer teknikleri veya nesne "bağlantılarını" içerebilir.
Yoğunluk Tabanlı Yöntemler	<ul style="list-style-type: none">• Nesnelerin doğal dağılımları bir yoğunluk fonksiyonu sayesinde tespit edilerek eşik yoğunluğunu aşan bölgeler küme olarak adlandırılır (Akkuş, 2015).
Izgara tabanlı Yöntemler	<ul style="list-style-type: none">• Yoğun çözünürlüklü bir grid veri yapısı kullanır,• Hızlı işlem süresi (genellikle veri nesnelerinin sayısından bağımsız, ancak grid boyutuna bağlıdır).

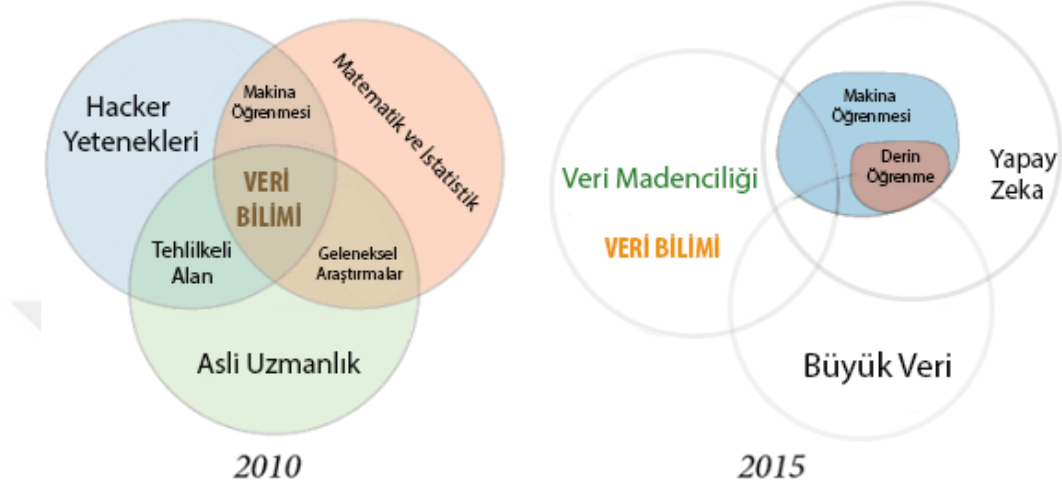
4.5. Veri Madenciliğinin İlişkili Olduğu Alanlar

Veri madenciliği kavramını da kapsayan veri biliminin bağlantılı olduğu disiplinler zamanla değişime uğramıştır. En temel tanımıyla veri bilimi, veriden bilgi veya bilgiler elde etmeye çalıştığınız teknikleri ifade eder.

Şekil 4.9'da veri madenciliği ile ilişkisi olan alanlar görülmektedir. 2010 ve 2015 yılları arasında veri biliminin ilişkili olduğu alanlar farklılık göstermektedir. Şekilde görüldüğü üzere, 2010 yılında matematik ve istatistik, hacker yetenekleri, asli uzmanlıklar gibi alanlarla ilişkili iken 2015 yılında daha çok büyük veri, yapay zeka ve makine öğrenmesi alanlarıyla ilişkili hale gelmiştir(Insidebigdata, 2016).

Veri bilimi alanı ve bu alanda yetişmiş teknik eleman günümüzde çok değerli hale gelmiştir. Bunun en iyi göstergelerinden biri bazı üniversitelerde veri mühendisliği yüksek lisans programları açılmasıdır.

Veri Bilimi Nedir?



Şekil 4.9. Veri Biliminin İlişkili Olduğu Alanlar

Kısaca, veri bilimi günümüzde, makine öğrenmesi, yapay zeka, derin öğrenme ve büyük veri kavramlarıyla sıkı bir ilişki içerisinde.

Makine Öğrenmesi: Genel olarak, makine öğrenmesi, geçmişte yaşanan örneklerle dayalı olarak gelecekte daha iyi sonuçlar elde etmeyi öğrenmek olarak tanımlanabilir (Schapire, 2008). Makine öğrenimi çalışmalarının odaklandığı konu bilgisayarlara veriye dayalı akılcı kararlar aldırma ve karmaşık örüntüleri algılama yeteneği kazandırmaktır. Makine öğrenmesi veri madenciliği, istatistik, örüntü tanıma, olasılık kuramı, gözetimli denetim, yapay zekâ ve kuramsal bilgisayar bilimi gibi alanlarla yakından ilişkilidir (Wikipedia(b), 2016). Makine öğrenmesinde üzerinde durulan nokta kendi kendine işleyen yöntemlerdir. Başka bir deyişle, amaç insan müdahalesi veya yardımı olmaksızın öğrenmeyi otomatik olarak yapan öğrenme algoritmaları tasarlamaktır.

Yapay Zeka: Öncelikle insan olmak üzere doğadaki tüm canlı davranışlarından esinlenerek bu tür davranışları modelleme çalışmalarının genel adıdır. Diğer bir ifadeyle yapay zeka, sezgisel programlama temelinde olan bir yaklaşım olup, insanların yaptıklarını bilgisayarlara yaptırabilme çalışmasıdır. Burada ana hedef, doğadaki varlıkların akıllı davranışlarını yapay olarak oluşturmayı amaçlayan bir kuram üretmektir (Russell, 1995).

Büyük Veri: Kurum ve kuruluşların ciddi büyüklüklerdeki verilerini, doğru tahminler yapabilmek için doğru hızda ve doğru zamanda saklanmasına, işlenmesine ve yönetilmesine olanak sağlayan bir teknoloji olarak tanımlanabilir (İçoğlu, 2015). Son yıllarda internet erişim hız ve kullanıcı sayılarının artmasıyla mevcut veri hacmi katlanarak artmaktadır. İnternet üzerinde veri artışı bir örnekle incelediğinde; 2008 yılı dünya olimpiyatlarında twitter üzerinde 6 milyon kayıtlı kullanıcı 300 bin tweet atarken, 2012 dünya olimpiyatlarında bu sayı 500 milyon kullanıcı ve 400 milyon tweet sayısına ulaşarak 1333 kat artış göstermiştir (IBM(b), 2016).

4.6. Eğitimsel Veri Madenciliği Kavram ve Modelleri

Tezin bu kısmında eğitimsel veri madenciliği tanımına, bu alanda yapılan çalışmalara ve eğitimsel veri madenciliğinde kullanılan algoritmalara yer verilmiştir.

4.6.1. Eğitimsel Veri Madenciliği

Çevrimiçi öğrenme ortamları geleneksel sınıf ortamlarından farklı olarak, öğrencilerin öğrenim etkinliklerini gerçekleştirirken geride bıraktıkları izlerin kayıt altına alınmasına imkân vermektedir. Bu izler; öğrencinin çevrimiçi ortama giriş-çıkış yapma sayısı, ders materyalleri ile etkileşimi, tartışma

forumunda bir soruya cevap yazması olabilir. Günümüzde çevrimiçi ortamlara kayıtlı öğrenci sayısındaki artışlar dikkate alındığında öğrenme süreçleri ile ilgili ciddi miktarda verinin çevrimiçi veri tabanlarına kayıt edildiği bilinmektedir. Fakat, kaydedilen bu verilerin eğitim-öğretimin iyileştirilmesi amacıyla kullanımı basit grafikler ve tanımlayıcı istatistiklerle sınırlı kalmaktadır (Ali, 2013). Ayrıca uzun zamandır farklı alanlarda veri tabanlarındaki gizli örüntü ve eğilimleri tespit etmek amacıyla kullanılan VM yöntemleri, eğitimsel ortamlardan sağlanan bu tür verilerin analizi konusunda önemli bir potansiyele sahiptir. Eğitimsel ortamlarda VM yöntemleri; öğrenci destek ve geri bildirimlerinde, öğrenme kuramlarının değerlendirilmesinde, erken uyarı sistemlerinde, performans kestirimlerinde, öğrenme teknolojilerinde ve gelecekteki öğrenme uygulamalarının geliştirilmesinde kullanılabilir (Greller, 2012).

Verilerdeki gizli yapıların ortaya çıkarılması EVM araştırmacılarının üzerinde çalıştığı önemli bir konudur. Benzer davranış örüntüsü sergileyen öğrenen gruplarının tespiti bunun bir örneğidir ve burada amaç; öğrenme tercihleri, kişisel özellikleri, bireysel farklılıkları gibi açılardan benzer öğrenci gruplarının belirlenmesidir (Romero vd., 2010). Bu sayede öğretim elemanları bu bilgileri kullanarak öğrencilerin gelişim sürecini izleme ve sorun yaşayan öğrencilere uygun müdahale yöntemleri geliştirme konusunda önemli bilgiler üretebilirler. Aynı zamanda bu veriler, uyarlanabilir öğrenme ortamlarında öğrencilerin otomatik olarak sınıflandırması amacıyla veya benzer öğrenci gruplarına yönelik otomatik uyarlamalar yapılması için kullanılabilir.

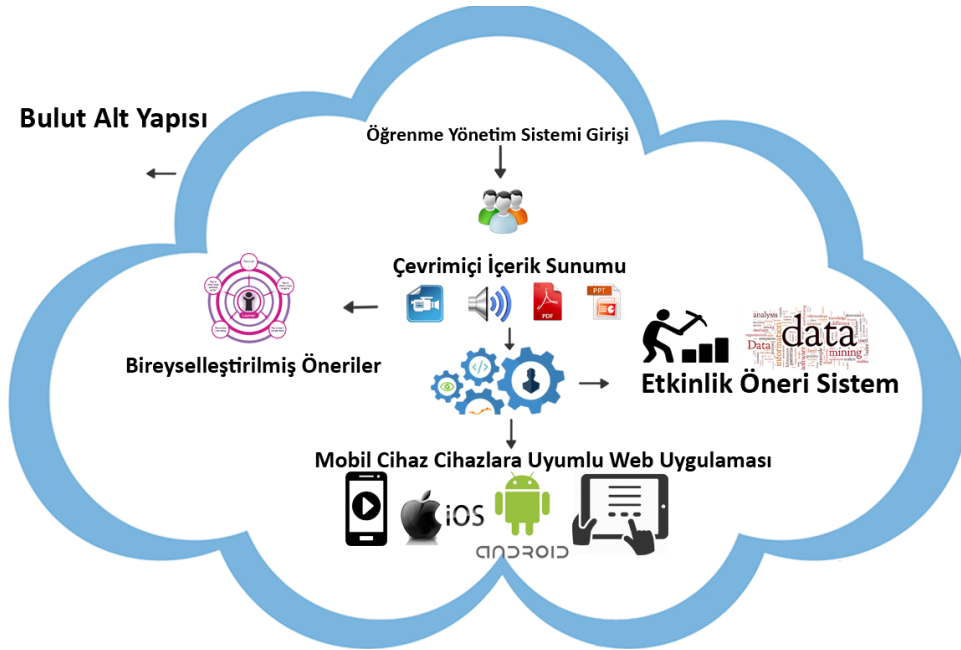
EVM çalışmalarında en sık kullanılan algoritmalar; karar ağacı, en yakın komşu, bayesdir.

5. YÖNTEM

Çalışmanın bu bölümünde araştırma modeli, veri toplama araçları, araştırma evreni, örneklem seçimi ve veri madenciliği basamaklarına yer verilmiştir.

5.1. Araştırma Modeli

Bu tez çalışmasının amaçlarından biri, çevrimiçi öğrenme ortamını kullanan öğrencilere, dönem sonunda elde edebilecekleri akademik başarılarının öngörü olarak sunulmasıdır. Araştırmada temel alınan boyutlar arasında öğrencilerin demografik yapıları, akademik ortalamaları ve gerçekleştirilen çevrimiçi etkinlikler bulunmaktadır. Bu amaca uygun olarak bir araştırma modeli geliştirilmiştir. Araştırmanın kavramsal modeli Şekil 5.1'de gösterilmektedir.



Şekil 5.1. Araştırmanın Kavramsal Modeli

Şekil incelendiğinde altyapı olarak bulut yapısı üzerinde çalışan öğrenme yönetim sistemi görülmektedir. Tasarlanan sisteme yeni bir birim veya üniversite kolaylıkla dahil edilebilmektedir. Eklenen yeni kurumlara işlemci veya depolama yönünden yetersizlik söz konusu olması durumunda donanım altyapısı yatayda genişleme özelliğine sahiptir. Yatayda genişleme ile ana omurga üzerinde yer alan işlemci, bellek(ram) ve depolama birimlerinin sayısı veya kapasitesinin yetersizliği durumunda artırılması kastedilmektedir. Bulut tabanlı bu tür çözümler kuruluş aşamasında yüksek maliyetlidir. Fakat, zamanla yeni server veya donanım malzemeleri almak yerine asıl omurga üzerindeki parçaların niceliklerinin artırılmasıyla bulut yapısı ekonomik yönden daha etkin bir şekilde kullanılmaktadır.

Bulut yapısı merkezi bir yönetim sağlamaktadır. Birimlerde kullanılan donanım altyapısının işleyişi hakkında istatistikler tutulmaktadır. Herhangi bir arıza olması durumunda hangi noktada sorunun olduğu anında tespit edilebilmektedir. Ayrıca bulut arayüzü üzerinden yeni bir sunucunun tanımlanması, veri tabanı açılması veya depolama alanlarının artırılması kolay bir şekilde yapılabilmektedir. Bu modelde tasarlanan web uygulaması kısmında mobil cihazlara uyumlu bir temanın tercih edilmesiyle modelin akıllı telefon ve tablet gibi taşınabilir cihazlarla rahatlıkla kullanılabilmesi ve görsel açıdan sorun yaşanılmaması hedeflenmiştir.

5.2. Veri Toplama Aracı

Araştırmada, veri toplama kaynağı olarak moodle çevrimiçi öğrenme yazılımı kullanılmıştır. Bu aşamada çevrimiçi öğrenme ortamlarında gerçekleştirilen öğrenci etkinlikleri ve öğrencilerin profilleri incelenmiştir.

Şekil 5.2' de görüldüğü gibi 2015-2016 Bahar döneminde Atatürk İlkeleri ve İnkılap Tarihi-II dersini çevrimiçi olarak takip eden öğrencilerin, konu bazlı olarak materyal takip sayılarına veri tabanından ulaşılmıştır. Araştırmada, öğrenenler moodle ÖYS platformunu kullanmışlardır. Moodle, programlama

dili olarak PHP ve veri tabanı olarak MYSQL kullanmaktadır. Öğrenenlerin, sisteme login olma sayıları, forumu görüntüleme sayısı, mesaj yazma sayıları, kullanıcı raporlama sayıları, takip ettikleri ders materyal türleri (video, PDF, ses dosyası, sunu) veri tabanı üzerinden çeşitli sorgulamalarla elde edilmiştir.

Zaman	Kullanıcının tam adı	user	Event context	Component	Etkinlik adı	Açıklama	Origin
31 Mar, 00:59		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '13972' viewed the course with id '2'.	web
31 Mar, 00:59		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '13972' viewed the course with id '2'.	web
31 Mar, 00:59		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '13972' viewed the section number '7' of the course with id '2'.	web
31 Mar, 00:58		-	Kaynak: 7. Hafta PDF Dökümanı	Kaynak	Görüntülenen ders modülü	The user with id '13972' viewed the 'resource' activity with course module id '43'.	web
31 Mar, 00:58		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '13972' viewed the section number '7' of the course with id '2'.	web
31 Mar, 00:58		-	Kaynak: 1. Hafta Ders Videosu	Kaynak	Görüntülenen ders modülü	The user with id '10001' viewed the 'resource' activity with course module id '10'.	web
31 Mar, 00:57		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '13972' viewed the course with id '2'.	web
31 Mar, 00:56		-	Kaynak: 1. Hafta Ses Dosyası	Kaynak	Görüntülenen ders modülü	The user with id '10001' viewed the 'resource' activity with course module id '16'.	web
31 Mar, 00:55		-	Kaynak: 1. Hafta Ders Videosu	Kaynak	Görüntülenen ders modülü	The user with id '10001' viewed the 'resource' activity with course module id '10'.	web
31 Mar, 00:55		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '10001' viewed the section number '1' of the course with id '2'.	web
31 Mar, 00:55		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '10001' viewed the course with id '2'.	web
31 Mar, 00:44		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '11908' viewed the course with id '2'.	web
31 Mar, 00:44		-	Kurs: Atatürk İlkeleri ve İnkılap Tarihi-II	Sistem	Ders görüntülendi	The user with id '11908' viewed the section number '2' of the course with id '2'.	web

Şekil 5.2. Çevrimiçi Öğrenme Ortamında Etkinlik Takip Raporu

Şekil 5.2'de çevrimiçi öğrenme ortamını takip eden kullanıcıların istatistiksel verileri görülmektedir. Öğrencilerin çevrimiçi ders ve etkileşim araçlarına erişim sayılarına raporlama işlemleriyle ulaşılmıştır.

5.3. Arařtırma Evreni ve Örneklem Seçimi

Arařtırmanın evrenini, 2015-2016 eğitim-öğretim yılı Bahar döneminde Hitit Üniversitesinde Uzaktan Eğitim Merkezine kayıtlı lisans ve önlisans programlarından birine kayıtlı 3988 öğrenci oluşturmaktadır. Bu çalışma, Hitit Üniversitesi Uzaktan Eğitim Merkezi (HUZEM)'nden izin alınarak yürütülmüştür. HUZEM biriminde araştırma yapılan dönemde uzaktan eğitim yöntemiyle yürütölen Atatürk İlkeleri ve İnkılap Tarihi-II dersini çevrimiçi olarak takip eden 1760 öğrencin çevrimiçi etkinlik verileri kaydedilerek veri seti oluşturulmuştur. Verilerin temizlenmesi işlemi sonrasında çevrimiçi dersleri takip eden 881 öğrenci örneklem olarak seçilmiştir. Örneklem seçiminde amaçlı örnekleme yöntemi kullanılmıştır. Amaçlı örnekleme, evrenin problemin en uygun bir kesiminin gözlem konusu yapılmasıdır (Sencer, 1989). Betimsel arařtırmalarda en az %10 örneklem alınır, küçük evrenlerde ise %20'ye ihtiyaç duyulur (Gay, 1987). Bu durumda %20'den fazla bir oranla örneklem sayısının kabul edilebilir olduđu tespit edilmiştir.

5.4. Veri Ön İşleme Süreci

Tez çalışmasının bu bölümünde, öğrenci profil verileri ve öğrenci etkinlik kayıtlarının temizlenme ve birleştirilmesi basamaklarına yer verilmiştir.

5.4.1. Öğrenci Profil Verilerinin İncelenmesi

Arařtırmaya katılan öğrencilerin demografik özellikleri çizelge 5.1'de görölmektedir. Bu özellikler, öğrencilerin cinsiyeti, yaşadığı şehir, anne mesleđi, baba mesleđi, kardeş sayısı, anne mezuniyet durumu, baba mezuniyet durumu, aylık gelir ve mezun olunan lise türüdür. Bu bilgilerin kaydı çevrimiçi öğrenme sistemi(ÇÖS) veri tabanı üzerinde tutulmaktadır.

Çizelge 5.1. Demografik Değişkenler

Değişken Adı	Açıklama
<i>d1_cinsiyet</i>	Öğrenci cinsiyeti, erkek için 1, kadın için 2 değeri ile kodlanmıştır.
<i>d2_yasadigi_sehir</i>	Yaşadığı şehir, şehirler 1, büyükşehirler 2 ile kodlanmıştır
<i>d3_anne_meslegi</i>	Anne mesleği, ev hanımı 1, çalışan 2 ile kodlanmıştır
<i>d4_baba_meslegi</i>	Baba mesleği, memur 1, serbest çalışan 2 ile kodlanmıştır
<i>d5_kardes_sayisi</i>	: Kardeş sayısı, bir kardeş için 1, iki ve üç kardeş için 2, üçten fazla kardeş için 3 ile kodlanmıştır.
<i>d6_anne_mezuniyet</i>	Annesinin en son mezun olduğu okul, ilköğretim için 1, lise ve üzeri için 2 ile kodlama yapılmıştır.
<i>d7_baba_mezuniyet</i>	Babasının en son mezun olduğu okul, ilköğretim için 1, lise ve üzeri için 2 ile kodlama yapılmıştır.
<i>d8_aylik_geli</i>	Ailenin aylık geliri, 1500 TL'ye kadar 1, 1500-3000 TL arası 2, 3000 TL üzeri 3 ile kodlanmıştır.
<i>d9_mezuniyet_turu</i>	Öğrencinin mezun olduğu lise türü, meslek lisesi için 1, normal liseler için 2 ile kodlama yapılmıştır

5.4.1.1. Çevrimiçi Öğrenme Ortamı Verilerinin İncelenmesi

Araştırma örneklemini temsil eden 3988 öğrenciden 881 öğrencinin ÇÖS'ü takip ettiği tespit edilmiştir. Çizelge 5.2'de ÇÖS üzerinde yer alan, öğrencilerin etkinlikleri sonucu elde edilen değişkenlerin isimleri ve açıklamaları yer almaktadır. Bu bilgiler, öğrencilerin kayıtlı olduğu eğitim birimini, elektronik postası var mı, üniversiteye yerleşim puanı, üniversiteye yerleşim puan türü, pdf takip durumu, ses dosyası takip durumu, video takip durumu, sunum dosyası takip durumu, login sayısı, mesaj sayısı, forum görüntüleme sayısı, raporlama kullanım sayısı, başarı puanı değerlerini içermektedir.

Çizelge 5.2. ÇÖS'den Elde Edilen Değişkenler

Değişken Adı	Açıklama
<i>d10_birimi</i>	Öğrencinin kayıtlı olduğu eğitim birimi, önlisans 1, lisans 2
<i>d11_eposta</i>	Öğrenci e-postaya sahip mi?, evet 1, hayır 2 ile kodlanmıştır.
<i>d12_yerlesim_puani</i>	Öğrencinin üniversiteye yerleşme puanı, ortalamasının altında ise 1, ortalama ve üstü ise 2
<i>d13_yerlesim_puan_turu</i>	Üniversiteye yerleşme puan türü, sayısal 1, sözel 2 ve eşit ağırlık 3 ile kodlanmıştır
<i>d14_pdf_takip_durumu</i>	Öğrencilerin PDF materyalini takip sayısı
<i>d15_ses_takip_durumu</i>	Öğrencilerin ses materyalini takip sayısı
<i>d16_video_takip_durumu</i>	Öğrencilerin video materyalini takip sayısı
<i>d17_sunum_takip_durumu</i>	Öğrencilerin sunu materyalini takip sayısı
<i>d18_login_sayisi</i>	Öğrencilerin login olma sayısı
<i>d19_mesaj_sayisi</i>	Öğrencilerin yazdığı mesaj sayısı
<i>d20_forum_goruntuleme_sayisi</i>	Öğrencilerin forum görüntüleme sayısı
<i>d21_raporlama_sayisi</i>	Öğrencilerin gelişim raporlarını inceleme sayısı
<i>d22_basari_puani</i>	Öğrencilerin sınav başarı puanı

5.4.2. Verilerin Hazırlanması

Bu bölümde, işlenen verilerin elde edilmesi işlemlerine yer verilmiştir. Çalışma evreni içerisinden sadece çevrimiçi öğrenme etkinliklerini takip eden öğrencilerin tüm öğrenciler arasından süzülmesi için MYSQL veri tabanı sorguları kullanılmıştır.

5.5. Modelin Kurulması

Araştırmanın bu bölümünde modelin kurulması sürecine yer verilmiştir. Veri madenciliği sürecinin ilk basamağı olan problem tanımlama işlemi tezin amaçları kısmında tanımlanmıştı. Burada kurulan model, araştırma amaçlarına cevap aramak için oluşturulmuştur. Model kurulmadan önce işlenecek veriler ÖYS üzerinden elde edildi. Öğrencilerin ders materyali takip

sayıları ve forum, mesaj gönderme, wiki kullanım istatistikleri, profil bilgileri moodle ÖYS veri tabanından çeşitli SQL sorguları veri hazırlama basamağında elde edilmiştir. Şekil 5.3'teki sorguya benzer çok sayıda sorgu ile derslerin öğrenciler tarafından takip sayıları tespit edilmiştir.

Çizelge 5.3. Ses Dosyası Takip Sayısını Çeken Örnek SQL Sorgusu

```
Select us.username as okulno, cr.fullname as dersadi, count(*) as sessayisi from
mdl_logstore_standard_log as lg
left join mdl_resource as res on res.id = lg.objectid
left join mdl_course as cr on cr.id = lg.courseid
left join mdl_user as us on us.id = lg.userid
where (action='viewed' ) and lg.objecttable = 'resource'
and res.name like '%Ses%' and cr.fullname like '%Ata%' and us.username is not null
group by us.username,cr.fullname;
```

Şekil 5.4'te ÖYS üzerinde yapılan tüm öğrenci etkinliklerinin çekildiği sorgu görülmektedir.

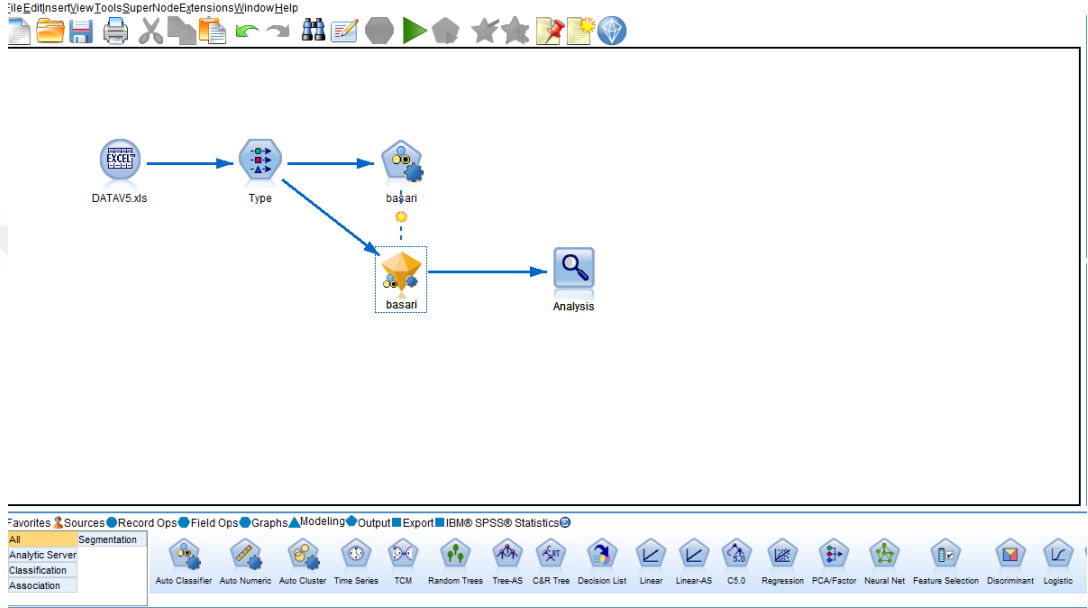
Çizelge 5.4. Tüm Etkinliklerin Takip Sayılarını Çeken SQL Sorgusu

```
select action,target,count(*) as etkinliksayisi from mdl_logstore_standard_log as lg
left join mdl_course as cr on cr.id = lg.courseid
left join mdl_user as us on us.id = lg.userid
where us.username is not null
group by action,target
order by count(*) desc;
```

Elde edilen çevrimiçi öğrenme ortam verilerinin veri madenciliği işlemlerinde karar ağacı algoritmaları kullanılmıştır. Mühendislik çalışma alanlarında, karar ağacı algoritmaları sınıflama yapmak amacıyla yaygın kullanım alanına sahip ve en çok tercih edilen algoritmalar arasındadır. Oluşturulan modelden elde edilen sonuçlara bulgular kısmında yer verilmiştir.

En Uygun Karar Ağacı Algoritmasının Tespiti















Örneklemlerden elde edilen veriler, 30 günlük ücretsiz sürüm SPSS Modeler programı üzerinde işlenmiştir. Veriler programa excel dosyası formatında çağırılmıştır.



Şekil 5.3. Modelin Kurulması

Program içerisinde veri türü ayarları yapıldıktan sonra, otomatik sınıf modelleme algoritmalarıyla işleme tabi tutuldu. Yapılan sınıflandırma işlemi şekil 5.3'te gösterilmiştir.

Şekil 5.4'te verilen sonuçlar incelendiğinde, veri madenciliği sürecinde 12 farklı algoritmanın model oluşturmak için kullanıldığı görülmektedir. Bu sınıflandırma işlemi sonrası başarılı olan algoritmalar üzerinde durulmuştur.

	Model Adı	Model Parametreleri	Model Sayısı
	C5	Default	1
	Logistic regression	Default	1
	Decision List	Default	1
	Bayesian Network	Default	1
	Discriminant	Default	1
	KNN Algorithm	Default	1
	LSVM	Default	1
	Random Trees	Default	1
	SVM	Default	1
	Tree-AS	Default	1
	CHAID	Default	1
	Quest	Default	1
	C&R Tree	Default	1
	Neural Net	Default	1

Şekil 5.4. Karar Ağacı Algoritmaları Sınıflandırma Oranları

Şekil 5.5'e bakıldığında, ortalama başarısı % 70'in üzerinde olan C5, Bayesian Network ve Neural Network algoritmalarından en iyi model çıktısı elde edilmiştir. Bu sonuca göre % 80,136'lık oranla en iyi performansı C5 algoritması göstermiştir.

Grafik	Model	Oluşturma Süresi (Dk)	Maksimum Kâr	Azami Kâr %	Yükseltme (% 30)	Ortalama Başarı	Kullanılan Alanlar	Eğri Altında Kalan Alan
	 C5	<1	417,273	14	1,87	80,136	16	0,731
	 Bayesian Netw...	<1	185,0	20	1,803	73,893	21	0,758
	 Neural Net	<1	135,0	16	1,688	72,304	21	0,72

Şekil 5.5. En İyi Sınıflamaya Sahip Karar Ağacı Algoritmaları

C4.5 ve C5 Algoritmaları

C4.5, Quinlan'ın 1986 yılında geliştirdiği, en yaygın kullanılan karar ağacı algoritması ID3 algoritmasının gelişmiş halidir (Quinlan, 1993). C4.5 algoritmasının ID3' e göre gelişmiş yanları: Karar ağacı oluştururken kayıp veriler dikkate alınmaz. C4.5 algoritması, kayıp verileri mevcut veri ve değişkenler yardımıyla öngörerek kazanım oranı hesaplamasında kullanır (Dunham, 2003).

Özellikle büyük sayıda verilerden oluşan veri setlerinde kullanılan C5.0 algoritması, C4.5'in doğruluk oranını artırmak amacıyla geliştirilmiş bir algoritmadır. Doğruluğu artırmak için boosting algoritmasını da kullanır. Bu algoritmaların diğerlerine göre en büyük farkı normalizasyon yapılmasıdır. ID3 ağaç yapısı üzerinde entropi hesabı yapılır ve bu değere göre karar noktaları belirlenir. C4.5 ağacında ise hesaplanan entropi değerleri oran olarak tutulur. Böylece ağaç üzerinde erişim sıklıklarına göre alt ağaçlar farklı seviyelere taşınabilir. C4.5 ağacının diğer bir farkı da budama işleminin yapılabilmesidir.

C4.5 ve C5.0 algoritmalarının temel özellikleri şöyle sıralanabilir;

- Her adımda bütün özellikler kontrol edilir,
- Her özelliğin normalize edilmiş bilgi kazancı hesaplanır,
- En iyi bilgi kazancını sağlayan özellik karar ağacına eklenir,
- Bütün dallar için bu adımlar tekrar edilir.

ID3 algoritmasında değişkenlerin birçok alt bölümlere ayrılması aşırı öğrenmeye neden olmaktadır. Quinlan, bunu önlemek için C4.5 algoritması için aşağıda verilen kazanım oranını kullanmıştır.

Kazanım Oranı(D,S)= Kazanım (D,S) / Ayırma Bilgisi (D,S)

$$\text{Ayırma Bilgisi (D;S)}= H\left(\frac{|D_1|}{|D|}, \dots, \frac{|D_s|}{|D|}\right)$$

Oluşan Karar Ağacının Kural Yapısı

```
Okul_Turu_Adi <= 1 [ Mode: GEÇER ]
  BirimAdi <= 1 [ Mode: GEÇER ]
    ses <= 1 [ Mode: GEÇER ]
      login <= 2 [ Mode: TEKRAR ]
        pdf <= 1 [ Mode: GEÇER ]
          K_YerlestigiPuanTuru <= 1 [ Mode: TEKRAR ]
            OgrCinsiyeti <= 1 [ Mode: TEKRAR ]
              video <= 1 [ Mode: TEKRAR ] => TEKRAR
              video > 1 [ Mode: GEÇER ] => GEÇER
            OgrCinsiyeti > 1 [ Mode: GEÇER ] => GEÇER
          K_YerlestigiPuanTuru > 1 [ Mode: GEÇER ] => GEÇER
        pdf > 1 [ Mode: TEKRAR ] => TEKRAR
      login > 2 [ Mode: GEÇER ]
        OgrCinsiyeti <= 1 [ Mode: GEÇER ]
          K_Adres_ili <= 1 [ Mode: GEÇER ]
            sunum <= 2 [ Mode: GEÇER ]
              BabaMeslek <= 1 [ Mode: GEÇER ] => GEÇER
              BabaMeslek > 1 [ Mode: GEÇER ]
                pdf <= 1 [ Mode: TEKRAR ]
                  K_YerlestigiPuanTuru <= 1 [ Mode: GEÇER ] => GEÇER
                  K_YerlestigiPuanTuru > 1 [ Mode: TEKRAR ] => TEKRAR
                pdf > 1 [ Mode: GEÇER ] => GEÇER
              sunum > 2 [ Mode: TEKRAR ] => TEKRAR
            K_Adres_ili > 1 [ Mode: GEÇER ] => GEÇER
          OgrCinsiyeti > 1 [ Mode: GEÇER ]
            KardesSayisi <= 2 [ Mode: TEKRAR ]
              AilenizinToplamGeliri <= 1 [ Mode: TEKRAR ]
                pdf <= 1 [ Mode: TEKRAR ] => TEKRAR
                pdf > 1 [ Mode: TEKRAR ]
                  BabaMeslek <= 1 [ Mode: TEKRAR ]
                    pdf <= 2 [ Mode: TEKRAR ] => TEKRAR
                    pdf > 2 [ Mode: GEÇER ] => GEÇER
                  BabaMeslek > 1 [ Mode: GEÇER ]
                    K_Adres_ili <= 1 [ Mode: GEÇER ] => GEÇER
                    K_Adres_ili > 1 [ Mode: TEKRAR ] => TEKRAR
              AilenizinToplamGeliri > 1 [ Mode: GEÇER ]
                video <= 2 [ Mode: GEÇER ] => GEÇER
                video > 2 [ Mode: TEKRAR ] => TEKRAR
            KardesSayisi > 2 [ Mode: GEÇER ]
              pdf <= 1 [ Mode: TEKRAR ] => TEKRAR
              pdf > 1 [ Mode: GEÇER ] => GEÇER
          ses > 1 [ Mode: TEKRAR ]
            KardesSayisi <= 2 [ Mode: TEKRAR ]
              login <= 2 [ Mode: TEKRAR ] => TEKRAR
              login > 2 [ Mode: GEÇER ]
                mesaj <= 1 [ Mode: GEÇER ]
                  AilenizinToplamGeliri <= 1 [ Mode: GEÇER ]
                    BabaMeslek <= 1 [ Mode: GEÇER ] => GEÇER
                    BabaMeslek > 1 [ Mode: TEKRAR ]
                      K_YerlestigiPuanTuru <= 2 [ Mode: TEKRAR ] => TEKRAR
                      K_YerlestigiPuanTuru > 2 [ Mode: GEÇER ] => GEÇER
                  AilenizinToplamGeliri > 1 [ Mode: GEÇER ] => GEÇER
                mesaj > 1 [ Mode: TEKRAR ] => TEKRAR
            KardesSayisi > 2 [ Mode: TEKRAR ] => TEKRAR
  BirimAdi > 1 [ Mode: GEÇER ]
    login <= 2 [ Mode: GEÇER ]
      video <= 1 [ Mode: GEÇER ] => GEÇER
      video > 1 [ Mode: TEKRAR ]
        K_YerlestigiPuanTuru <= 2 [ Mode: GEÇER ]
          OgrCinsiyeti <= 1 [ Mode: TEKRAR ] => TEKRAR
          OgrCinsiyeti > 1 [ Mode: GEÇER ] => GEÇER
        K_YerlestigiPuanTuru > 2 [ Mode: TEKRAR ] => TEKRAR
    login > 2 [ Mode: GEÇER ]
      AnneMeslek <= 1 [ Mode: GEÇER ] => GEÇER
```

```
AnneMeslek > 1 [ Mode: GEÇER ]
video <= 2 [ Mode: GEÇER ]
  BabaEgitimDurumu <= 1 [ Mode: GEÇER ] => GEÇER
  BabaEgitimDurumu > 1 [ Mode: TEKRAR ] => TEKRAR
video > 2 [ Mode: TEKRAR ] => TEKRAR
Okul_Turu_Adi > 1 [ Mode: GEÇER ] => GEÇER
```

5.6. Değerlendirme

Çevrimiçi öğrenme ortamından elde edilmiş veriler temizleme ve birleştirme işlemleri sonrasında elde edilen 22 değişken ile veri madenciliği işlemleri yürütülmüştür. Veri madenciliği işlemleri sonrasında sınıflama algoritmalarından olan C5 algoritmasının yaklaşık % 80 başarı ile doğru öngörüye sahip olduğu görülmüştür. Tahmin başarısı % 70'in üzerinden olduğundan dolayı bu sonuç kabul edilebilir düzeyde bir sınıflama başarısına sahiptir.

5.7. Modelin Yayılımı (Uygulama)

Bu kısımda, oluşturulan modelin web tabanlı olarak uygulanabilirliği ve örnek tasarıma yer verilmiştir.

Şekil 5.6 incelendiğinde, öğrencilerin temel bilgileri, demografik bilgileri ve etkinlik takip sayıları görülmektedir. Oluşturulan modele göre bu tasarımın işleyişindeki yararlar; öğrenciye başarı tahmininin bildirilmesi ve konu bazında başarı sağlamak için öğrencilerin dikkate alması gereken başarı oranlarının verilmesi, öğrencilerin takip etmesi gereken ders materyali ve yarar sağlayacak çevrimiçi etkileşim araçlarının önerilmesi olarak sıralanmaktadır..

Sınav Giriş Bilgileri		Sınav Sonucu		Sınav Kağıdını Görüntüle		Öneri Sistemi		Oturumu Kapat	
Kimlik Bilgileri				Konu Adı		Ortalama	Sizin		
Öğrenci Numarası	-			1. HAFTA (TÜRK İNKILAP HAREKETLERİ VE ATATÜRK DÖNEMİ DIŞ POLİTİKASI)		49,11	50		
Ad	-			2. HAFTA (HALİFELİĞİN KALDIRILMASI)		57,93	66,66		
Soyad	-			3. HAFTA (ÇOK PARTİLİ REJİM DENEMELERİ)		64,18	50		
Birim	Sungurlu Meslek Yüksekokulu			4. HAFTA (EĞİTİM ALANINDA YAPILAN İNKILAPLAR)		76,92	100		
Bölüm	Bilgisayar Programcılığı			5. HAFTA (TOPLUMSAL ALANDA YAPILAN İNKILAPLAR)		71,17	75		
Demografik Bilgiler						74,11	100		
Cinsiyet	Erkek			6. HAFTA (SOSYAL VE HUKUK ALANINDA YAPILAN İNKILAPLAR)		43,98	50		
Yaş	20			7. HAFTA (ATATÜRK DÖNEMİ TÜRK DIŞ POLİTİKASI)					
Doğum Tarihi	1997								
Anne Mesleği	Ev hanımı								
Baba Mesleği	Memur								
Anne Eğitim Durumu	Lise								
Baba Eğitim Durumu	Lise								
Kardeş Sayısı	2								
Aylık Gelir	Orta								
Ekinlik Takip Bilgileri				Sayı	Ortalama	Sistemin Öngörü Başarısı: % 80,136			
PDF	19	10	ÖNERİLER: Profiliniz ve etkinlikleriniz incelendiğinde, dönem sonunda Atatürk İlkeleri ve İnkılap Tarihi-II dersinden başarılı olacağınız öngörülmektedir. Başarınız artması için 3. ve 5. hafta konularına daha çok önem vermeniz gerekmektedir. Ayrıca forum, chat ve wiki gibi bilgi paylaşım ortamlarını kullanmanız başarınızı olumlu yönde etkileyecektir.						
Ses Dosyası	23	2,36							
Video	12	3,7							
Sunu	19	4,31							
Forum	0	1,5							
Chat-Mesaj	0	1,72							
Login Sayısı	0	4,16							

Şekil 5.6. Modelin Web Tabanlı Tasarımı

6. BULGULAR

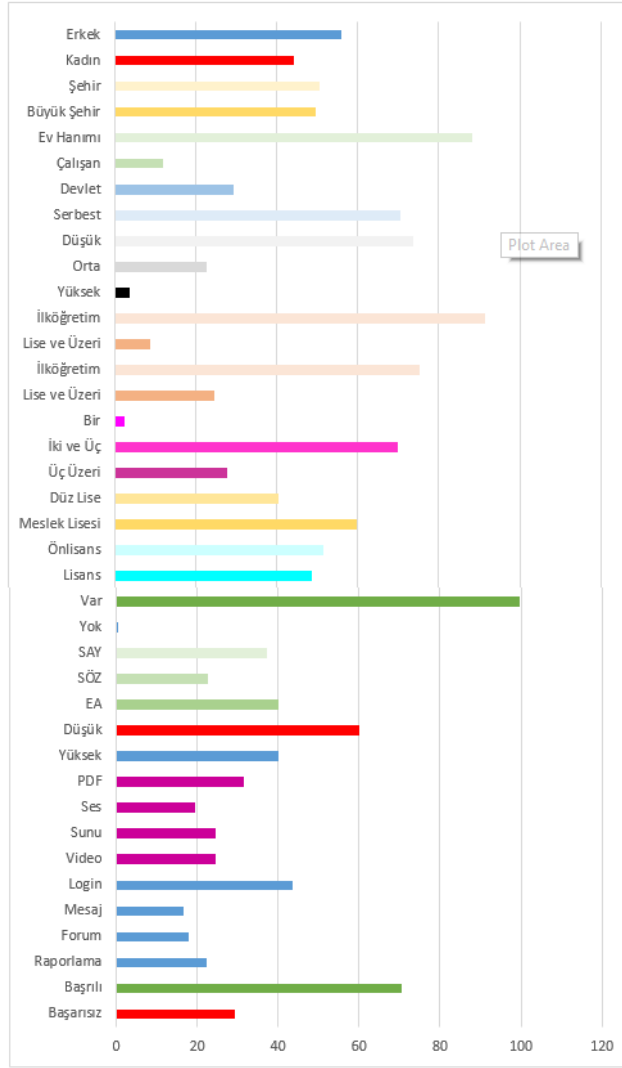
Araştırmanın bu bölümünde elde edilen bulgular çizelgeler halinde verilmiştir. Bulgular bölümünün birinci kısmında öğrencileri tanıtıcı bilgiler; demografik özellikleri, cinsiyet, yaşadığı şehir, gelir düzeyi, mezun olduğu lise türü, anne ve baba mesleği, anne ve babanın mezun olduğu okul bulunmaktadır.

Bulgular bölümünün ikinci kısmında, oluşturulan veri madenciliği modeli sonuçlarına yer verilmiştir.

6.1. Katılımcıları Tanıtıcı Bilgiler

Tezin bu kısmında, çevrimiçi olarak konuları takip eden öğrencilerin demografik bilgilerine ve takip edilen materyal tür ve sayılarına yer verilmiştir. Şekil 6.1'de konuları takip eden öğrenciler incelendiğinde, %55,8'inin erkek, %44,2'sinin kadın, yaşadığı şehir türü değişken oranlarının yakın olduğu, öğrencilerin çoğunluğunun anne mesleği ev hanımı, baba mesleğinin ise serbest meslek, anne eğitim düzeyinin %91,3 ve yine baba eğitim düzeyinin %75,3 oranında ilköğretim, öğrencilerin %59,7'lik oranda meslek lisesi mezunu, %51,4'lik oranda ön lisansa kayıtlı olduğu görülmektedir. Öğrencilerin neredeyse tamamının elektronik posta hesabına sahip olduğu, üniversiteye yerleşim puan türünde sayısal ve eşit ağırlık puan türü oranı birbirine yakın iken sözel puan türüne sahip öğrenci oranı daha azdır. Ayrıca, ders materyalleri takip oranlarına bakıldığında, %31,4'lük oranla PDF türünün ilk sırada olduğu, %29,7'lik oranla sunu ve %24,4'lük oranla video materyal türünün takip edildiği, %19,5 oranla ses materyal türünün ez az takip oranına sahip olduğu görülmektedir.

Değişken Türü		N	%
Cinsiyet	Erkek	492	55,8
	Kadın	389	44,2
Yaşadığı Şehir	Şehir	445	50,5
	Büyük Şehir	436	49,5
Anne Mesleği	Ev Hanımı	777	88,2
	Çalışan	104	11,8
Baba Mesleği	Devlet	259	29,4
	Serbest	622	70,6
Aylık Gelir	Düşük	650	73,8
	Orta	198	22,5
	Yüksek	33	3,7
Anne Eğitimi	İlköğretim	804	91,3
	Lise ve Üzeri	77	8,7
Baba Eğitimi	İlköğretim	663	75,3
	Lise ve Üzeri	218	24,7
Kardeş Sayısı	Bir	21	2,4
	İki ve Üç	615	69,8
	Üç Üzeri	245	27,8
Önceki Mezuniyet	Düz Lise	355	40,3
	Meslek Lisesi	526	59,7
Kayıtlı Olduğu Birim	Önlisans	496	51,4
	Lisans	469	48,6
E-posta	Var	879	99,8
	Yok	2	0,2
Puan Türü	SAY	329	37,3
	SÖZ	200	22,7
Yerleş Puanı	EA	352	40,0
	Düşük	529	60,0
Materyal Türü	Yüksek	352	40,0
	PDF	1707	31,4
Etkileşim	Ses	1062	19,5
	Sunu	1341	24,7
	Video	1329	24,4
	Login	2471	43,5
Sınav Puanı	Mesaj	939	16,5
	Forum	1013	17,8
	Raporlama	1261	22,2
Başarı Ortalaması	Başarılı	622	70,6
	Başarısız	259	29,4



Şekil 6.1. Konuları Takip Eden Öğrencilerin Demografik Özellikleri

6.2. Model Sonuçları

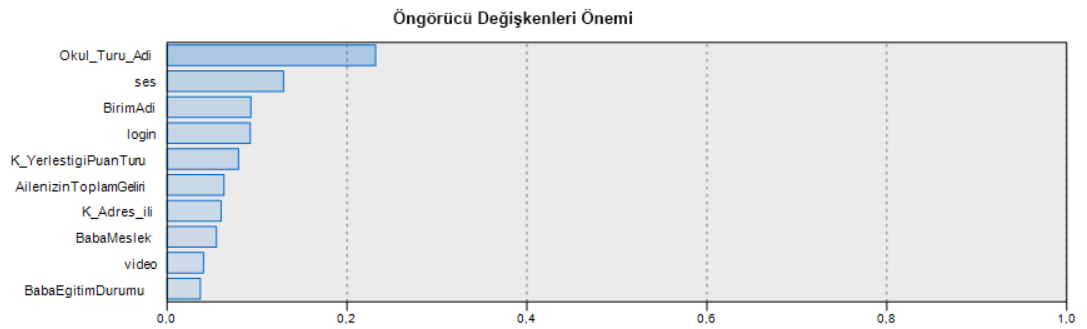
Bu kısımda, kurulan modelin sonucunda oluşan karar ağacı yapısına, ağaç yapısı incelenerek başarılı düğümlere, tahmin etme başarısına, sınıflama başarısına etkisi olan değişkenlere ve haftalık bazda öğrenci başarılarına yer verilmiştir. Şekil 6.2 incelendiğinde, sınıflama uygulaması sonucu oluşan karar ağacı yapısı görülmektedir. Oluşturulmuş karar ağacı ile 1. amaç test edilmiştir ve % 80,136'lık oranla doğru başarı tahmini yapılabilmektedir.

2. amaç sonucu için veri madenciliği sürecinde farklı sınıflama algoritmalarının tahmin etme başarısı üzerinde etkisi incelenmiştir. İlk sırada C5 algoritması yer alırken, ikinci ve üçüncü önemli etkiye bayesian ve neural network algortimalarının sahip olduğu tespit edilmiştir.

3. Amaç sonucu için çevrimiçi öğrenme ortamlarında öğrencilerin konu bazında başarı farklılıkları incelenmiştir. Çizelge 6.1 incelendiğinde, en düşük başarının 7.konuda, en yüksek başarının 4.konuda elde edildiği görülmektedir.

Çizelge 6.1. Konu Bazında Ortalama Başarılar

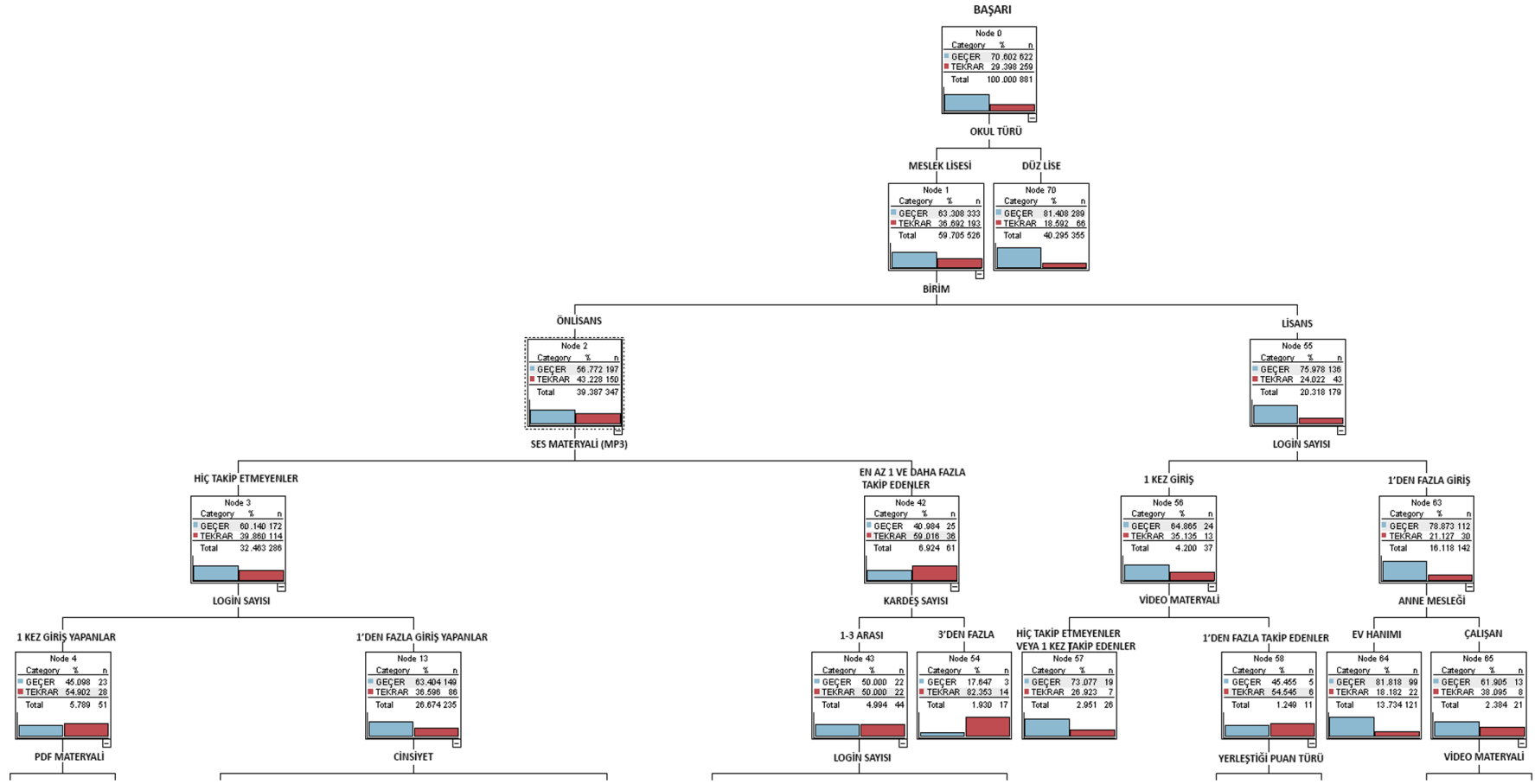
Konu Adı	Soru Sayısı	Ortalama Başarı
1. Konu	2	49,11
2. Konu	3	57,93
3. Konu	2	64,18
4. Konu	3	76,92
5. Konu	4	71,17
6. Konu	2	74,11
7. Konu	4	43,98



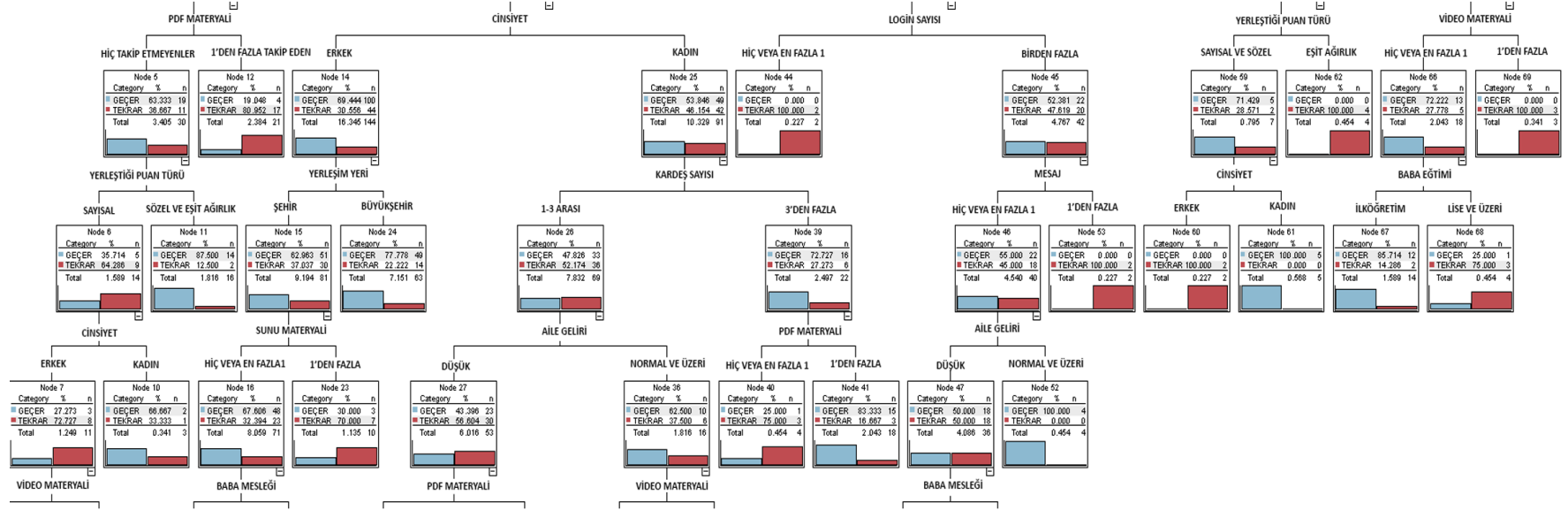
Şekil 6.3. Kullanılan Değişkenlerin Öneri Başarısı Üzerine Etkisi

4. ve 5. Amaç sonucu için, eğitimsel veri madenciliği sürecinde kullanılan değişkenlerin öneri başarısı üzerine etkisi C5 algoritması ile incelenmiş ve sonuca şekil 6.3'de yer verilmiştir. Şekil incelendiğinde, tahmin etme başarısına sırasıyla, lise mezuniyet türü, ses materyal takibi, birim türü, login sayısı, üniversiteye yerleşme puan türü, ailenin aylık geliri, yaşadığı şehir, baba mesleği, video materyal türü, baba eğitim durumunun etkisi olduğu görülmektedir. En fazla etkisi olan değişkenin ise mezun olduğu okul türü olduğu tespit edilmiştir. Şekil 6.4, 6.5 ve 6.6'da veri madenciliği uygulaması sonucu oluşan karar ağacı yapısına yer verilmiştir. Öğrencilerin akademik başarılarının tahmin edilmesinde kök kısmında bulunan değişkenlerin daha etkili olduğu bilinmekte olup şekil 6.4'te yer alan değişkenlerin başarı tahmininde etkisinin daha fazla olduğu söylenebilir. Birbirine bağlı olarak devam eden bu şekillerden üçüncüsünden yani en alt dallara daha yakın olan kısımda tahmin başarısına etkisi daha az olan değişkenler yer almaktadır şekil 6.6'da gösterilmektedir..

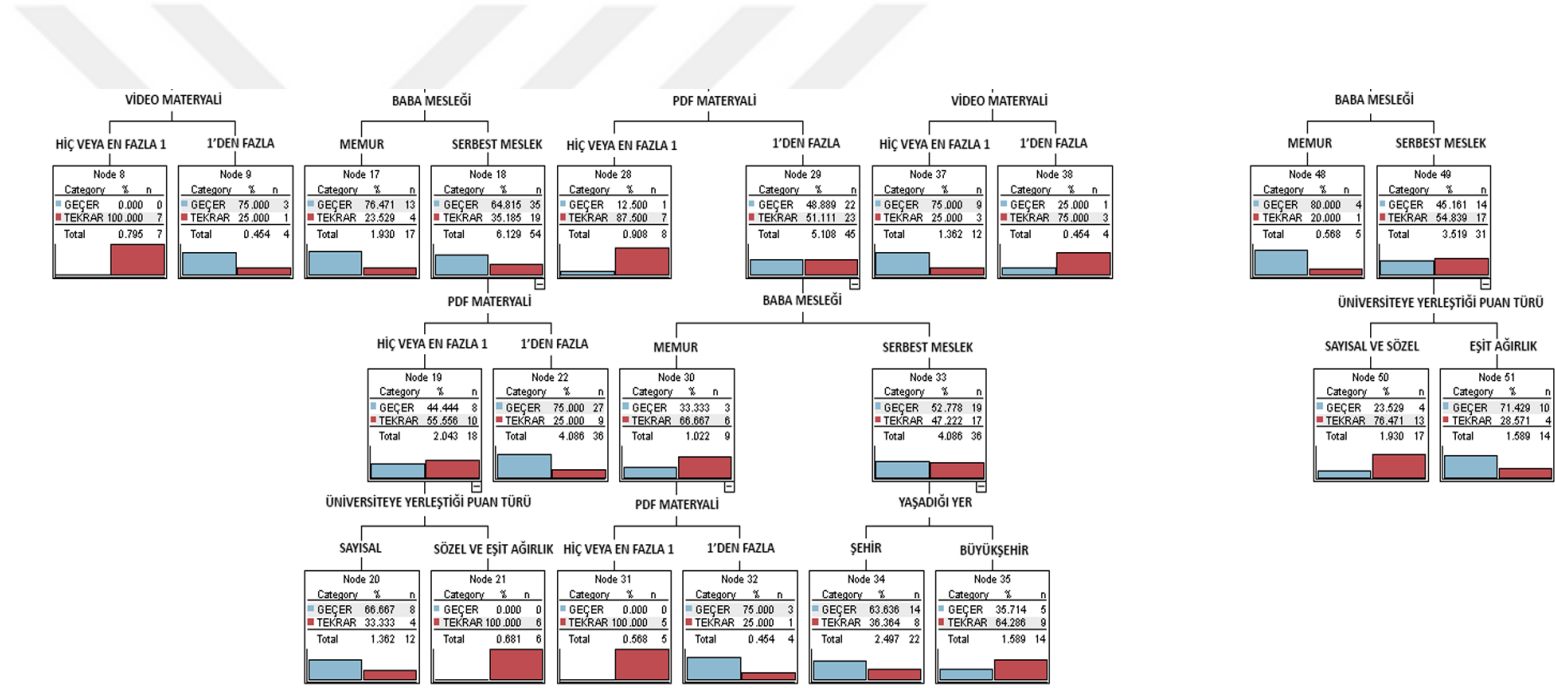
6. amaç sonucu için, oluşturulan öneri modelinin web tabanlı olarak uygulamasının tasarlanması işlemlerine modelin uygulanması kısmında yer verilmiştir. Mevcut moodle ÖYS veri tabanından oluşturulan veri ambarı verileri kullanılarak raporlama ve anlık öneriler sunulmaktadır.



Şekil 6.4. Karar Ağacı Sonuçları



Şekil 6.5. Karar Ağacı Sonuçları (Devamı)



Şekil 6.6. Karar Ağacı Sonuçları (Devamı)

7. SONUÇ ve ÖNERİLER

Çalışmanın bu kısmında elde edilen bulgular doğrultusunda sonuç ve önerilere yer verilmiştir. Seçilen örneklemden elde edilen bulgular yorumlanmış ve öğrencilerin öğrenim gördükleri kurumlara ve öğrencilere öneriler sunulmuştur. Ayrıca örneklem verilerinden genellemeler yapılarak gerekli görülen kurumlar ve ileride bu konuda araştırma yapacak araştırmacılar için çeşitli bilgilere yer verilmiştir.

Literatürde EVM alanında yapılan çalışmalar incelendiğinde, öğrencilerin profilleri ve başarılarına göre(Guruler, 2010) sınıflandırma yapıldığı veya öğrencilerin okumaya ayırdıkları zaman, elektronik kaynaklara nereden bağlandıkları ve anahtar kavramları ne kadar hızlı öğrendikleri (Castro vd., 2007) gibi özelliklere göre sınıflandırıldığı görülmektedir. Bu çalışmada, öğrenci profilleri yanı sıra ders materyallerinin ve çevrimiçi etkinlik takip sayılarının dikkate alınarak sınıflama yapılması ayrıca bu yaklaşımın web tabanlı bir model olarak tasarlanmasının bu alanda yapılan diğer çalışmalara göre yenilikler getirdiği ifade edilebilir.

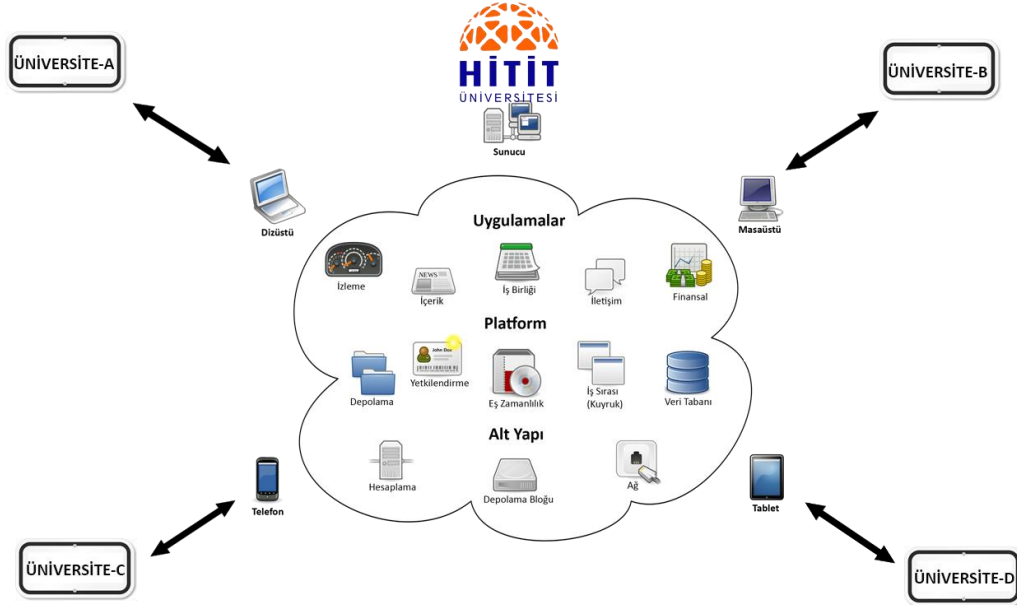
Bulut tabanlı çevrimiçi öğrenme önerisi sisteminin tasarlanması ve uygulanmasına yönelik çalışmanın model kısmında verilen sonuca ve oluşan bulgulara göre;

- Araştırma başlangıcında çalışmaya katkı sağlayacağı düşünülen değişkenlerden, lise mezuniyet türü, ses materyal takibi, birim türü, login sayısı, üniversiteye yerleşme puan türü, ailenin aylık geliri, yaşadığı şehir, baba mesleği, video materyal türü, baba eğitim durumunun katkısı olduğu görülmektedir. Bu sonuca, örneklem büyüklüğü ve diğer etkenlerin neden olduğuna sınırlılıklar kısmında yer verilmiştir. İlerleyen süreçte yeni kayıt olan öğrencilerle örneklem sayısının artacağı, bu bağlamda ÖYS üzerinden daha çok miktarda etkinlik verileri toplanacağı, sistemin öğrenme yeteneğinin ve daha iyi sınıflama sonucu elde edeceği söylenebilir.

- Araştırma sonucu oluşan karar ağacı % 80,136'lık oranda tahmin başarısına sahiptir. Bu sonuç % 70'in üzerinden olduğundan kabul edilebilir düzeydedir.
- Çevrimiçi dersi takip eden öğrencilerin çoğunlukla ders materyallerini takip ettikleri, forum, chat ve wiki gibi etkileşim sağlayacak web 2.0 teknolojilerini kullanmadıkları veya bu araçların kullanım alışkanlığının yeteri kadar olmadığı tespit edilmiştir. Web 2.0 teknolojileri bireyler arası bilgi paylaşımını sağlayan teknolojik ortamlardır. Bu tür ortamları öğrencilerin kullanması hususunda teşvik edici çözümler uygulanmalıdır. Örneğin; wiki kullanarak dersin konusuyla ilgili bilgi girişi yapan, forumda fayda sağlayıcı yeni bir başlık açan veya önceki başlıklara bilgi ekleyen, canlı chat uygulamasını belli sayıda kullanan öğrencilere ödev, proje veya başka bir değerlendirme başlığı altında yıl sonu notunu az da olsa etkileyecek puan verilebilir. Nitekim önceki çalışmalar, öğrencilere ek puan verilmesi durumunda motive edici katkı sağladığı görülmektedir (Ifenthaler, 2014).
- Konu bazda öğrenci başarıları incelendiğinde, önemli derecede farklılıklar olduğu görülmektedir. Konu başarı ortalamalarının her dönem için hesaplanarak kaydedilmesi, öğrencilerin çalışma planı hazırlamasına yardımcı olacaktır. Böylece, ders sorumlusu öğretim elamanı ve kurum yöneticileri başarısızlığın fazla olduğu konu için destekleyici ders materyal, proje veya fazladan canlı ders saatleri ekleyebilir.
- Öğrencilerin yıl sonu başarısına en çok etkisi olan materyalin PDF türü olduğu diğer ders materyallerinin yeteri kadar etkili olmadığı bulgular kısmında görülmektedir. PDF türünün diğer materyallere göre daha fazla tercih edilmesinin nedeni olarak sınırlı kotada internet paketine sahip öğrenciler olduğu düşünülürse, boyutunun az olması en çok tercih edilme nedeni olabilir. Ayrıca video, ses ve sunum materyallerinin masaüstü, tablet ve akıllı telefonlara uygun çözümlükte sunulması kullanıcılara daha tasarruflu internet kullanımı sağlayacaktır.

- Bu çalışmada, incelenen ders sözel bir ders olduğu için üniversiteye sayısal puanıyla giriş yapan öğrencilerin başarısının daha düşük olması beklenen bir sonuç olarak görülebilir.

Tez çalışmasında, bulut tabanlı ifadesiyle donanımsal alt yapı kastedilmektedir. Bu çalışmada, söz edilen öneri sisteminin bulut yapısı biçiminde tasarlanması daha kolay yaygınlaştırılmasına olanak sağlayacaktır. Türk Dili, Atatürk İlkeleri ve İnkılap Tarihi ve Yabancı Dil dersleri Yükseköğretim Kurulu(YÖK) tarafından ortak ders olarak nitelendirilmektedir. Bu dersler YÖK kanunu 5/i maddesi gereği tüm üniversitelerde verilmesi zorunlu derslerdir. Web üzerinden başarılı bir şekilde çalışan bir ÖYS üniversitelerdeki öğrencilerin kullanımını sağlayacak yeterli büyüklükte bir bulut alt yapısı üzerine kurulabilir. Bulut sistemleri başlangıç maliyeti yüksek olsa da tüm üniversitelerin zorunlu ortak dersler için ayrı ayrı sunucu kurduğu düşünüldüğünde, ve böyle bir sistemin bulut altyapısıyla sunulması üniversiteler için ciddi ekonomik kazanç sağlayacaktır. Ayrıca bulut tabanlı sistemler yüksek hesaplama kapasitesi ve merkezi yönetime olanak sağlamaktadır. Şekil 7.1 incelendiğinde, önerilen sistemin diğer kurumlar için yaygınlaştırılmasının görsel hali sunulmuştur.



Şekil 7.1. Öneri Sisteminin Diğer Kurumlar İçin Yaygınlaştırılması

Araştırmanın Sınırlılıkları

Bu tez çalışması Hitit Üniversitesi Uzaktan Eğitim Merkezinde çevrimiçi dersleri alan 881 öğrenci ile sınırlıdır. Araştırmada, 7 haftalık ders takip süreci için öğrenen grubu sınıflama işlemi yapılmaktadır. Örneklem olarak alınan kütle sayısının daha fazla olması durumunda sınıflama başarısının daha yüksek olacağı tahmin edilmektedir.

Araştırmada, esasında 7 haftalık süreçte öğrenciler tarafından çevrimiçi olarak takip edilen konulara yönelik konu bazında ayrı ayrı sınıflama ve sonucunda öneriler sunulması öğrenciler için daha yararlı olabilir. Örneklemde yer alan öğrenme grubun beklenenden daha az sayıda etkinlik gerçekleştirmesi nedeniyle 7 haftalık sürecin tamamında yapılan çevrimiçi etkinliklerin tümü baz alınmıştır.

Tez çalışmasında, ÖYS olarak moodle tercih edilmiştir. Moodle öğrenme yönetim sistemi 1998 yılında kullanılmaya başlanmış, dünyada en çok kullanılan ve en çok dil desteğine sahip ÖYS'dir. Fakat moodle'in kendine özgü bir veri tabanı yapısının olması veri elde ederken bazı zorlukları beraberinde getirmiştir. Veri madenciliği işlemleri öncesinde araştırma için anlamlı olacağı tespit edilen değişkenler çeşitli veri tabanı sorguları ile ayrı bir tabloda birleştirilerek elde edilmiştir.

Bu araştırma, sadece Hitit Üniversitesi ile sınırlıdır. Farklı diğer Üniversitelerden daha fazla örneklem sağlanması durumunda daha farklı sonuçlar elde edilebilir.

KAYNAKLAR

Aase S., Higher Learning Goes The Distance, Computer User, 19(10): 16-18, 2000.

Akkuş İ., Veri Madenciliği - Kümeleme ve Kümeleme Yöntemleri, https://www.academia.edu/9860553/Veri_Madencili%C4%9Fi_-_K%C3%BCmeleme_ve_K%C3%BCmeleme_Y%C3%B6ntemleri (Erişim tarihi: 15.06.2016).

Aksu S., Demirel, H., Görgünoğlu, S. Bulut Bilişim Teknolojisi ile Bir Görüntü İşleme Uygulaması. Antalya : Akademik Bilişim 2013, s.1-4, 2013.

Ali L., Asadi, M., Gašević, D., Jovanović, J., Hatala, M., Factors Influencing Beliefs For Adoption Of A Learning Analytic Tool: An Empirical Study , Computers & Education. 62(1): 130-148, 2013.

Allen I. E., Seaman, J. Learning on demand: Online education in the United States, <http://www.onlinelearningsurvey.com/reports/goingthedistance.pdf> (Erişim tarihi: 12.04.2016).

Al-Zoube M., El-Seoud S.A, Wyne M.F., Cloud Computing Based e-learning system. International Journal of Distance Education Technologies. 8(2): 58-71, 2010.

anderson.ucla.edu, Data Mining: What is Data Mining?, <http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm> (Erişim tarihi: 10.03.2016).

Anonim, 10 Years of Educational Technology Trends in Online Learning Infographic, The University of Texas, <http://elearninginfographics.com/educational-technology-trends-in-online-learning-infographic> (Eriřim tarihi, 16 Temmuz 2015).

Armbrust M., Fox., A.,Griffith, R., Joseph, A.,D., Katz, R., Konwinski, A., Lee, G.,Patterson, D., Rabkin, A., Stoica, I., Zaharia, M. A., View of Cloud Computing. ACM New York, NY, USA : Communications of the ACM, 53(4): 50-58, 2010.

Aydın C., H., Çevrimiçi (Online) Öğrenme Toplulukları, Açık ve Uzaktan Eğitim Sempozyumu. Eskişehir: Anadolu Üniversitesi Açıköğretim Fakültesi, 2002.

Baker R. S. J. D., Yacef, K., The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining. Elsevier, 1(1):3-17, 2009.

Baker R., S., Inventado, P.S., Educational Data Mining and Learning Analytics, Learning Analytics: From Research to Practice. Ed: White J.A. Larusson ve White, B. New York : Springer Science+Business Media , 2014.

Baun C., Kunze M., Nimis J., Tai S., Cloud Computing Web-Based Dynamic IT Services, in Cloud Computing. Berlin: Springer, 2011.

Bayır E.A., Çevrimiçi Öğrenme Ortamlarında Sohbet ve E-Posta Kullanımının Öğrencilerin İşlemsel Uzaklık Algılarına Etkisi, Doktora Tezi. Gazi Üniversitesi Fen Bilimleri Enstitüsü, - Ankara, 2014.

Belanger F., Jordan, D. H., Evaluation and implementation of distance learning: Technologies, Tools and Techniques. USA-UK: Idea Group Publishing, 2000.

- Benoit G., Data mining, Annual Review of Information Science and Technology.Elsevier, 36(1): 265-310, 2002.
- Bienkowski M., Feng, M., Means, B., Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief, U.S. Department of Education, Office of Educational Technology, Washington, D.C., 2012.
- Bora U. J., Ahmed M., E-Learning using Cloud Computing. International Journal of Science and Modern Engineering (IJISME). 1(2): 9-13 , 2013.
- Brahmawong C., Guidelines for Internet-based distance education in colleges and universities in Thailand. International Journal of The Computer. 12(2): 7-13, 2004.
- Charles, P., Good, N., Jordan, L.L., Pal, J., How Much Information, <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003> (Erişim Tarihi, 20.04.2016).
- Castro F., Vellido, A., Nebot, A., Mugica, F., Applying Data Mining Techniques to e-Learning Problems. Studies in Computational Intelligence. 62(1):183-221, 2007.
- Caytiles R.D., Lee S., Park B., Cloud Computing: The Next Computing Paradigm. International Journal of Multimedia and Ubiquitous Engineering. South Korea, 7(2): 297-302, 2012.
- Chapman, P., Julian,C., Randy,K., Thomas, K., Thomas, R., Colin, S., Rüdiger, W., CRISP-DM, <https://www.the-modeling-agency.com/crisp-dm.pdf> (Erişim tarihi: 15.05.2016).

- Cheung Christy, M.K., Lee, Matthew, K.O., Lee, Zach, W.Y., Understanding the continuance intention of knowledgesharing in online communities of practice through the post-knowledge-sharing evaluation processes. *Journal of the American Society for Information Science & Technology*. 64(7): 1357-1374, 2013.
- Chien C. F. ve Chen, L. F., Data Mining to Improve Personnel Selection and Enhance Human Capital: A Case Study in High-Technology Industry, *Expert Systems with Applications*. 34(1), 280-290, 2008.
- Çakır H., Mobil Öğrenmeye İlişkin Bir Yazılım Geliştirme ve Değerlendirme. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*. 40(2): 1-9, 2011.
- Çalışkan H., Çevrimiçi Eğitimde Öğrenci Etkileşimi. *Açık ve Uzaktan Eğitim Sempozyumu*. Eskişehir : Anadolu Üniversitesi Açık Öğretim Fakültesi, 2002.
- Diaz, V., Cloud-Based Technologies: Faculty Development, Support, And Implementation, *Journal of Asynchronous Learning Networks*. 15(1): 95-102, 2011.
- Dolgun M.Ö., Türkiye'deki Hastanelerin Veri Madenciliğiyle Gruplandırılması, XI. Biyoistatistik Kongresi, 27–30 Mayıs, Malatya, 2008. Malatya, s. 1-13, 2008.
- Dunham M.H., *Data Mining: Introductory and Advanced Topics*. New Jersey : Pearson Education Inc., 2003.
- Emel G. G., Taşkın, Ç., Veri Madenciliğinde Karar Ağaçları ve Bir Satış Analizi Uygulaması, *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*. 6(2): 221-239, 2005.

- Ercan T., Effective use of cloud computing in educational institutions. *Procedia - Social and Behavioral Sciences*. 12(2): 938–942, 2010.
- Fallon C., Brown, S., *E-learning standards: A guide to purchasing, developing and deploying standards-conformant e-learning* [Kitap]. - [basım yeri bilinmiyor] : USA: St. Lucie Press., 2002.
- Fan J, Han, F., Liu, H., Challenges of Big Data analysis, doi: 10.1093/nsr/nwt032 [Dergi]. - [basım yeri bilinmiyor] : *National Science Review*, 2014. - Cilt 1.
- Frank E.,P., Pharo, N., *Academic Librarians in Data Information Literacy Instruction: A Case Study in Meteorology*. College & Research Libraries (C&RL), Chicago, 2015.
- Gay L. R. ,*Educational research: Competencies for analysis and application* (3rd ed.) [Kitap]. - Columbus : OH: Merrill, 1987.
- Glymour C., Madigan, D., Pregibon, D., & Smyth, P., Statistical inference and data mining. - New York : *Communications of the ACM*, 39(11): 35–41, 1996.
- Google Google Apps for Education, The tools your students want. <http://www.google.com/enterprise/apps/education/products.html> (Erişim tarihi: 15.06.2016).
- Goyal L. C. ve Jatav, P.K., *Cloud Computing: An Overview And Its Impact On Libraries*, *International Journal of Next Generation Computer Applications*. 1(1):9-15, 2012.
- Greller W.,Drachsler, H.,*Translating Learning into Numbers: A Generic Framework for Learning Analytics*, *Educational Technology & Society*, 15(3): 42-57, 2012.

Gunnarsson B. L., Alterman, R., Predicting failure: A case study in coblogging, Computers in Human Behavior. 2nd international conference on learning analytics. Vancouver, 29 Nisan - 2 May, 2012, British Columbia, Canada, s. 263-266, 2012.

Guruler H., Istanbulu, A., Karahasan, M. A., new student performance analysing system using knowledge discovery. Computers & Education. Elsevier, 2010. - Cilt 55(1): 247-254, 2010.

Gülbahar Y., E-öğrenme. - Ankara : Pegem Akademi Yayıncılık, 2009.

Güldal H., Kılıçaslan, Y., Çuhadar, C., Bulut Tabanlı Bir Ders Yönetim Sistemi Yazılımının Geliştirilmesine Dayalı Olarak Öğretim Elemanı ve Öğrencilerin Teknoloji Kabullerinin İncelenmesi. Trakya Üniversitesi Eğitim Fakültesi Dergisi. 6(2): 177-188, 2016.

Gürsakal N., Büyük Veri. - Bursa : Dora Yayıncılık, 2014.

Han J. ve Micheline, K., Data Mining Concepts and Techniques. Morgan Kaufman Publisher, Academic Press, 2001.

Han J., Kamber, M., Pei, J., Data Mining Concepts and Techniques. Elsevier, USA, 2012.

Harrell F. E., Regression Modeling Strategies: With Applications to Linear Models, Springer, 2001.

Henkoğlu, T., Külcü Ö., Bilgi Erişim Platformu Olarak Bulut Bilişim: Riskler ve Hukuksal Koşullar Üzerine Bir İnceleme. Bilgi Dünyası. 4(1): 62–86, 2013.

Horton W, Horton, K., E-learning tools and technologies. Indianapolis : Wiley Corporation, 2003.

Horton W., E-learning by design [Kitap]. - San Francisco : Pfeiffer, John Wiley & Sons, Inc. , 2006.

Horzum M.,B., Interaction, Structure, Social Presence, and Satisfaction in Online Learning, Eurasia Journal of Mathematics, Science & Technology Education, 11(3): 505-512, 2015.

Ibikunle F. Awodele O., Kuyoro S.O., Cloud Computing Security Issues and Challenges. International Journal of Computer Networks (IJCN). 5(3): 225-247, 2011.

IBM(a). Easily identify groups and predict outcomes, <http://www-03.ibm.com/software/products/en/spss-decision-trees> (Eriřim tarihi: 10.07.2016).

IBM(b), Büyük Veri ile Analitik Uygulamalar, ftp://public.dhe.ibm.com/software/pdf/tr/connected2013/01_Buyuk_Veri_ile_Analitik_Uygulamalar.pdf (Eriřim tarihi: 15.08.2017).

Ifenthaler D., From Educational Data Mining to Automated Analysis of Semantics. Potsdam, 2014.

Inmonna, W.H., Data Warehouse, <http://datawarehouse4u.info> (Eriřim Tarihi: 04.01.2016).

Insidebigdata, Learn Data Science: Eight (Easy) Steps, http://insidebigdata.com/2016/10/23/learndatascience_eight_easy_steps (Eriřim tarihi: 28.10.2016)

İçođlu O., Bulut Biliřim ve Byk Veri Arařtırma Laboratuvarı (B3LAB). TUBITAK, Ankara, 2015.

Investopedia What is 'Data Mining' [evrimii] // <http://www.investopedia.com>. - <http://www.investopedia.com>, 2015. - 20 Ocak 2016. - <http://www.investopedia.com/terms/d/datamining.asp>.

Karaman K ve Akgl, İ., İlkokul đrencileri İin Web Tabanlı Deđerler Eđitimi Uygulaması, Uřak niversitesi Sosyal Bilimler Dergisi, 8(3), 87-100, 2015.

Koyuncu, M., Biliřimde Yeni Trend: Bulut Biliřim, <http://www.acikarsiv.atilim.edu.tr/browse/503/17.pdf> (Eriřim tarihi: 20.05.2016).

Kr H., atalođlu, E., Erbay, H., Uzaktan ve rgn Eđitimin đrenci Bařarısı zerine Etkisinin Arařtırılması. Gaziantep University Journal of Social Sciences, 12(3): 267-279, 2013.

Ma W.W.K., Yuen, A.H.K., Understanding online knowledge sharing: An interpersonal relationship perspective. Computers & Education. 56(1): 210–219, 2011.

Manyika J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh,C., Byers, A., Big Data: The Next Frontier for Innovation, Competition, and Productivity, McKinsey Global Institute, 2011.

Margaret D., Data Mining Introductory and Advanced Topics. New Jersey : Pearson Education Inc., 2003.

McCallum A. ve Nigam, K., Comparison of Event Models for Naive Bayes Text, Proceedings of the tenth conference on European chapter of the Association. 5(2): 307-314, 2003.

Mell P., Grance, T., The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology. Nist Special Publication, 145(1): 1-3, 2011.

Moore M. G., Kearsley, G., Distance education: A systems view. Boston : MA: Wadsworth Publishing, 1996.

Moreno L., Gonzalez, C., Castilla, I., Gonzalez, E., Sigut, J., Applying a constructivist and collaborative methodological approach in engineering education. Computers & Education. 49(3): 891-915, 2007.

Munakata T., Knowledge discovery, New York : Communications of the ACM, 42(11): 27-29, 1999.

North M., Data Mining for the Masses.A Global Text Project Book, Rapid Miner, 2012.

Olson D. L., Shi Y., Introduction to Business Data Mining. New York, NY, USA : Mcgraw-Hill College, 2007.

Omotunde A.A., Awodele, O., Kuyoro, S.O. ve Ajaegbu, C., Survey of Cloud Computing Issues at Implementation Level. Journal of Emerging Trends in Computing and Information Sciences, 4(1): 91-96, 2013.

Oracle(a) 2016, Data Warehousing Concepts,
https://docs.oracle.com/cd/B10500_01/server.920/a96520/concept.htm
(Erişim tarihi: 10.03.2016).

Oracle(b) 2016, What Is Data Mining?,

https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm#C HDFGCIJ (Eriřim tarih: 10.04.2016).

Öztürk M., Web Tabanlı Uzaktan Eğitimde Teknolojiye İliřkin Yeni Eğilimler, Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 14(1): 273-288, 2014.

Pala F. K., Erdem, M., Öğretmen Adaylarının Çevrimiçi Tartışma Ortamlarına Yönelik Görüşleri, Turkish Online Journal of Qualitative Inquiry. 6(2): 24-47 , 2015.

Palmer M., Data is the New Oil,

http://ana.blogs.com/maestros/2006/11/data_is_the_new.html (Eriřim tarihi: 10.11.2016).

Peña-Ayala A., Educational data mining: A survey and a data mining-based analysis of recent works. Expert Systems with Applications. 41(4):1432-1462, 2014.

Perraton H. D., Open and distance learning in the developing world, Taylor & Francis Group, Canada, 2000.

Romero, C., Ventura, s., Pechenizkiy, M., Baker, R.S.J.D., Introduction Handbook of Educational Data Mining, Introduction Handbook of Educational Data Mining. CRC Press, 2010.

Quinlan J. R., C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, USA, 1993.

Rotella, P., Is Data The New Oil,

<http://www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/#d82f8cb77a9a> (Eriřim tarihi: 10.12.2015).

Russell S.,J., Norvig, P., Artificial Intelligence A Modern Approach. New Jersey : Alan Apt, 1995.

Safran C., Helic, D., Gütl., C. E-Learning Practices And Web 2.0. International Conference On interactive Computer Aided Learning, Austria , s.1-8,2007.

Sarıtař T. ve Üner N., Eđitimdeki Yenilikçi Teknolojiler: Bulut Teknolojisi,Eđitim ve Öđretim Arařtırmaları Dergisi, 3(2): 192-201, 2013.

Sayad, S., Classification, <http://www.saedsayad.com/classification.htm> (Eriřim tarihi: 12.02.2016)

Sayad S., Model Evaluation, http://www.saedsayad.com/model_evaluation.htm (Eriřim tarihi: 02.02.2016).

Schapiro R., Theoretical Machine Learning,

<https://www.cs.princeton.edu/courses/archive/spring14/cos511> (Eriřim tarihi: 12.03.2016).

Sencer M., Toplum bilimlerinde yöntem, Beta Basım, İstanbul : 1989.

Sevli O., Bulut Biliřimin Eđitim Alanında Uygulanması. Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü Dergisi. Yüksek lisans tezi. Süleyman Demirel Üniversitesi, Isparta,2011.

Shaw, J., Why “Big Data” Is a Big Deal, <http://harvardmagazine.com/2014/03/why-big-data-is-a-big-deal> (Erişim tarihi: 25.12.2015).

Siemens G., Baker, R. S., Learning analytics and educational data mining: towards communication and collaboration, 2nd international conference on learning analytics and knowledge, 29 Nisan-2 Mayıs, 2012, Vancouver, British Columbia, Canada, s. 252-254, 2012.

Silahtaroglu G., Veri Madenciliği Kavram ve Algoritmaları, Papatya Yayıncılık, İstanbul, 2016.

Smith, S.D., Caruso, J.B., The ECAR Study of Undergraduate Students and Information Technology, Colorado: EDUCAUSE Center for Applied Research, 2010.

Strachey, C., Time sharing in large, fast computers. In Proceedings of the IFIP Congress, Paris, s.336-341, 1959.

Sund R., Utilization of Administrative Registers using Statistical Knowledge Discovery. International Workshop on Mining Official Data, National Research and Development Centre for Welfare and Health, Finland, s.1-14, 2002.

Şentürk A., Veri Madenciliği: Kavram ve Teknikler. Ekin Kitabevi, Bursa, 2006.

Tergan S.O., Keller T., Digital concept mapping in learning contexts: Integrating knowledge, arguments and information resources”. Proceedings of 9th International Conference on Information Visualisation, In E. Banissi (Ed.) Los Alamitos, CA : IEEE Computer Society, s.371-377, 2005.

Tseng F. C., Kuo, F.Y. A., study of social participation and knowledge sharing in the teachers' online professional community of practice. Computers & Education. 72(1): 37-47, 2014.

Ture M, Tokatli F, Kurt, İ., Using Kaplan–Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4.5 and ID3) in determining recurrence-free survival of breast cancer patients, Expert Systems with Applications. 36(2): 2017–2026, 2009.

Wanli X., Rui, G.,Eva, P., Sean, G., Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics. Computers in Human Behavior. Elsevier, 47(1): 168-181 ,2015.

West D.M., Big Data for Education: Data Mining, Data Analytics, and Web Dashboards. Washington :<https://www.brookings.edu/wp-content/uploads/2016/06/04-education-technology-west.pdf> (Erişim tarihi: 25.06.2016).

Wikipedia(a) , Cloud computing,
http://en.wikipedia.org/wiki/cloud_computing#history (Erişim tarihi: 20.02.2016).

Wikipedia(b), Makine öğrenimi,
https://tr.wikipedia.org/wiki/Makine_%C3%B6%C4%9Frenimi (Erişim tarihi: 12.12.2015).

Yalın H. İ., İnternet Temelli Eğitim. Ankara : Nobel Yayın Dağıtım, 2008.

Yao C.Y., Tsai, C.C. & Fang, Y.C. Understanding social capital, team learning, members' e-loyalty and knowledge sharing in virtual communities. Total Quality Management & Business Excellence. 26(6):619–631, 2015.

Yenilmez E., H., Bilgi Toplumu Olgusu ve Türkiye Hedef 2000 Yılı.İstanbul : İstanbul Üniversitesi, 1993.

Yüksel H., Bulut bilişim el kitabı,

<http://yukselis.wordpress.com/2012/01/27/bulutbilisim-el-kitabi> (20.05.2016).



ÖZGEÇMİŞ

Adı Soyadı : HAKAN KÖR
Doğum Tarihi : 1981
Yabancı Dil : İngilizce

Eğitim Durumu :
Lisans : Balıkesir Üniversitesi, 2006
Yüksek Lisans : Kırıkale Üniversitesi, 2013

Çalıştığı Kurum/Kurumlar ve Yıl/Yıllar: Milli Eğitim Bakanlığı, 2006-2009
Hitit Üniversitesi, 2009- Devam Ediyor

Yayınları (SCI / SSCI) : Kör, H., Erbay, H., Engin, M., Dündür, E., An Examination of the Correlation Between Science and Technology Attitudes Scale, Frequency of Smartphone Usage Scale and Lif'elong Learning Scale Scores Using the Structural Equation Model, Journal of Baltic Science Education, Vol 16, 2017 (Kabul aldı)

Yayınları (Diğer) : Kör Hakan, Erbay Hasan, Dündür Emre, Demir Emre (2016). Investigating the Relationship between the Effects of Motivation Personal Development and Basic Computer Usage Upon the Student Success in Distance Education Via the Structural Equation Modeling. Participatory Educational Research, 3(2), 51-62. (Yayın No: 3014693)

Kör Hakan, Erbay Hasan, Engin Melih (2016). Technology Leadership of Education Administers and Innovative Technologies in Education A Case Study of Çorum City. Universal Journal of Educational Research, 4(12A), 139-149., Doi: 0.13189/ujer.2016.041318 (Yayın No: 3014747)

Kör Hakan, Erbay Hasan, Demir Emre (2015). Research of Distance Education Students Cloud Perceptions and Use Cases. Participatory

Educational Research (PER), 2015(2), 1-9., Doi: 10.17275/per.15.spi.2.1
(Yayın No: 1707438)

Aksoy Hamit, Kör Hakan, Erbay Hasan, Akmeşe Ömer Faruk (2014). Study on Professional Guidance on Teaching Settings. Procedia - Social and Behavioral Sciences, 141, 1215-1219., Doi: doi:10.1016/j.sbspro.2014.05.208 (Yayın No: 1707596)

Kör Hakan, Aksoy Hamit, Erbay Hasan (2014). Comparison of the Proficiency Level of the Course Materials Animations Videos Simulations E books Used in Distance Education. Procedia - Social and Behavioral Sciences, 141, 854-860., Doi: 10.1016/j.sbspro.2014.05.150 (Yayın No: 1707397)

Kör Hakan, Çataloğlu Erdat, Erbay Hasan (2013). Uzaktan ve Örgün Eğitimin Öğrenci Başarısı ÜzerineEtkisinin Araştırılması. Gaziantep University Journal of Social Sciences (GAUN-JSS), 12(2), 267-279. (Yayın No: 3035530)

Araştırma Alanları : Eğitimsel Veri Madenciliği, Bulut Bilişim, Çevrimiçi Öğrenme Sistemler, Uzaktan Eğitim