



**ÜNİVERSİTE ÖĞRENCİLERİNİN SINAVLARINA
VE BAŞARILARINA YÖNELİK ÖNGÖRÜ
ÇALIŞMASI**

**2023
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

Fatma Tuğçe TÜRK

**Tez Danışmanı
Prof. Dr. İlhami Muharrem ORAK**

**ÜNİVERSİTE ÖĞRENCİLERİNİN SINAVLARINA VE BAŞARILARINA
YÖNELİK ÖNGÖRÜ ÇALIŞMASI**

Fatma Tuğçe TÜRK

**Tez Danışmanı
Prof. Dr. İlhami Muharrem ORAK**

**T.C.
Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Yüksek Lisans Tezi
Olarak Hazırlanmıştır**

**KARABÜK
Aralık 2023**

Fatma Tuğçe TÜRK tarafından hazırlanan “ÜNİVERSİTE ÖĞRENCİLERİNİN SINAVLARINA VE BAŞARILARINA YÖNELİK ÖNGÖRÜ ÇALIŞMASI” başlıklı bu tezin Yüksek Lisans Tezi olarak uygun olduğunu onaylarım.

Prof. Dr. İlhami Muharrem ORAK

İmzası

Tez Danışmanı: Bilgisayar Mühendisliği Anabilim Dalı

.....

Bu çalışma, jürimiz tarafından Oy Birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans Tezi olarak kabul edilmiştir. 29/11/2023

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Doç. Dr. Salih GÖRGÜNOĞLU (KÜ)

.....

Üye : Dr. Öğr. Üyesi Oğuzhan MENEMENCİOĞLU (KBÜ)

.....

Üye : Prof. Dr. İlhami M. ORAK (KBÜ)

.....

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Yüksek Lisans derecesini onamıştır.

Doç. Dr. Zeynep ÖZCAN

.....

Lisansüstü Eğitim Enstitüsü Müdürü

“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Fatma Tuğçe TÜRK

ÖZET

Yüksek Lisans Tezi

ÜNİVERSİTE ÖĞRENCİLERİNİN SINAVLARINA VE BAŞARILARINA YÖNELİK ÖNGÖRÜ ÇALIŞMASI

Fatma Tuğçe TÜRK

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Prof. Dr. İlhami Muharrem ORAK

Kasım 2023, 55 sayfa

Bu çalışmanın amacı; üniversite öğrencilerinin ders notlarına ait verilerle öğrenci başarı analizi yapmak ve başarı sonuçlarına göre çıkarımlarda ve öngörülerde bulunmaktır. Son zamanlarda bilgisayar teknolojilerinin hızla ilerlemesi ve eldeki verilerin artması ile birlikte veri analizi çalışmaları diğer alanlarda olduğu gibi eğitim alanında da büyük önem kazanmıştır. Eldeki verileri anlamlandırmak ve faydalı hale getirmek ilgili alanla ilgili önemli çıkarımların elde edilmesine fırsat sağlayacaktır.

Günümüzde öğrenci verilerinin çok sayıda ve çeşitli olması veri büyüdükçe yapılan analizlerin daha kapsamlı, ayrıntılı ve çok yönlü araştırmayı önemli hale getirmektedir. Bu önemden hareketle bu çalışmada Karabük Üniversitesindeki okuyan öğrencilerin verileri analiz edilerek ders başarılarına etki eden faktörlerin belirlenmesi amaçlanmıştır. Bu çerçevede Bilgisayar Mühendisliği bölümündeki farklı programlarda okuyan öğrencilerin bölüm ve programlarına göre ders

başarılarını incelemek, tahmin etmek ve öğrencilerin dersi öğrenmeye yönelik davranışlarını anlamak için çıkarımlarda bulunulmuştur.

Bu çalışmada Karabük Üniversitesi Bilgisayar Mühendisliği Bölümündeki öğrencilere ait olan veriler üzerinde işlemler yapılmıştır. Elde edilen verilere göre öğrenci notlarının ders, dersin grubu, program, dönem bilgileriyle ilişki durumları analiz edilmiştir. Veri madenciliği yöntemleri ve algoritmaları kullanılarak derslerin birbiriyle ilişkisi, başarı oranlarının tespiti ve analizi yapılmıştır. Yorumlaması ve değerlendirmesi Sonuçlar kısmında belirtilmiştir.

Anahtar Kelimeler: Veri bilimi, analiz, istatistik, akademik başarı, öngörü, veri ön işleme

Bilim Kodu: 92429

ABSTRACT

M. Sc. Thesis

PREDICTION STUDY FOR UNIVERSITY STUDENTS' EXAMS AND SUCCESS

Fatma Tuğçe TÜRK

**Karabük University
Institute of Graduate Programs
Department of Computer Engineering**

Thesis Advisor:

Prof. Dr. İlhami Muharrem ORAK

November 2023, 55 pages

The aim of this study is to analyze student achievement with the data of university students' course grades and to make inferences and predictions based on the achievement results. Recently, with the rapid advancement of computer technologies and the increase in available data, data analysis studies have gained great importance in the field of education as in other fields. Making sense of the available data and making it useful will provide the opportunity to make important inferences about the relevant field.

Today, the large number and variety of student data makes it important to conduct more comprehensive, detailed and multifaceted research as the data grows. Based on this importance, this study aims to analyze the data of the students studying at

Karabük University and to determine the factors affecting their course success. In this framework, inferences were made to examine and predict the course success of students studying in different programs in the Department of Computer Engineering according to their departments and programs and to understand students' behaviors towards learning the course.

In this study, the data belonging to the students in Karabük University Computer Engineering Department were analyzed. According to the data obtained, the relationship of student grades with course, course group, program period information was analyzed. Using data mining methods and algorithms, the relationship between the courses and the success rates were determined and analyzed. Interpretation and evaluation are stated in the results and recommendations section.

Key Word: Data science, data mining, analysis, statistics, academic achievement, forecasting, data pre-processing

Science Code: 92429

TEŐEKKÜR

Bu tez alıŐmasını; planlanmasında ve yürütülmesinde ilgi ve desteęini esirgemeyen engin bilgi ve tecrübelerinden yararlandığım saygı deęer hocam sayın Prof. Dr. İlhami Muharrem ORAK' a sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ.....	xi
ÇİZELGELER DİZİNİ	xii
SİMGELER VE KISALTMALAR DİZİNİ	xiv
BÖLÜM 1	1
GİRİŞ	1
BÖLÜM 2	5
LİTERATÜR ARAŞTIRMASI	5
BÖLÜM 3	10
VERİ DEĞERLENDİRME YÖNTEMLERİ VE ÖNEMİ	10
3.1. VERİ ANALİZİ	10
3.1.1. Veri Analizi Süreci.....	11
3.1.2. Veri Analizi Türleri.....	11
3.2. VERİ MADENCİLİĞİ YÖNTEMLERİ.....	12
3.2.1. Birliktelik Kuralları (Apriori)	12
3.2.2. Karar Ağacı (J48).....	13
3.2.3. Eğitimde Veri Analizinin Önemi	15
BÖLÜM 4	17
VERİ SETİ VE ÖN İŞLEMLER	17

	<u>Sayfa</u>
4.1. VERİ ELDE ETME SÜRECİ	17
4.1.1. Araştırmaya Dahil Edilme Kriterleri.....	17
4.1.2. Araştırmadan Dışlanma Kriterleri.....	18
4.2. VERİ SETİNİN ÖZELLİKLERİ	18
4.3. VERİ ÖN İŞLEMLERİ.....	21
4.3.1. Veri Temizleme.....	21
4.3.2. Veri Birleştirme.....	23
4.3.3. Veri Dönüştürme	25
4.3.4. Veri İndirgeme	26
4.4. ÖN İŞLEMLERDEN GEÇİRİLEN VERİLERİN ANALİZİ.....	27
BÖLÜM 5	29
SONUÇLAR.....	29
5.1. VERİ ANALİZİ SONUÇLARI	29
5.2. VERİ MADENCİLİĞİ İLE ELDE EDİLEN SONUÇLAR	37
5.2.1. Dersi İlk Alışta Geçme-Kalma Durumları	37
5.2.1.1. Veri Kümesi	37
5.2.1.2. Birliktelik Kuralları ile Elde Edilen Sonuçlar	38
5.2.1.3. Karar Ağacı İle Elde Edilen Bulgular	39
5.2.2. Dersi İkinci Alışta Geçme-Kalma Durumları	44
5.2.2.1. Veri Kümesi	44
5.2.2.2. Birliktelik Kuralları İle Elde Edilen Sonuçlar.....	44
5.2.2.3. Karar Ağacı ile Elde Edilen Sonuçlar	45
BÖLÜM 6	48
DEĞERLENDİRMELER VE ÖNERİLER	48
6.1. DEĞERLENDİRMELER	48
6.2. ÖNERİLER.....	50
KAYNAKLAR	52
ÖZGEÇMİŞ	55

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 4.1. Bölümlere göre toplam öğrenci sayısını veren kod	27
Şekil 4.2. Grup sayısı 53 olan derse ait durum bilgisi kodu	28
Şekil 4.3. Grup numarası “53”olan dersin tekrarlama kodu	28
Şekil 4.4. Durum ve tekrar bilgisini veren kod	28
Şekil 5.1. (BLM112) Programlama Dilleri II dersi karar ağacı	41
Şekil 5.2. (BLM224) Elektronik dersi karar ağacı	42
Şekil 5.3. (BLM225) Nesneye Dayalı Programlama dersi karar ağacı	42
Şekil 5.5. (BLM112) Programlama Dilleri II dersi karar ağacı	46
Şekil 5.6. (BLM227) Veri Yapıları karar ağacı	47

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 4.1. Alınan veri setinin içeriği	19
Çizelge 4.2. Alınan veri setine eklenen özellikler	20
Çizelge 4.3. Veri setinin ilk hali	20
Çizelge 4.4. Veri setine eklenen özellikler	20
Çizelge 4.5. Devre analizi dersi için “Şube Kodu”	21
Çizelge 4.6. Negatif not girilen iki farklı dersin gösterimi	22
Çizelge 4.7. Aynı döneme ait Endüstri Stajı dersi	22
Çizelge 4.8. F1 Harf Koduna ait tutarsız veri örneği	23
Çizelge 4.9. ID verilmeden önceki veri seti	24
Çizelge 4.10. ID verildikten sonraki veri seti.	24
Çizelge 4.11. Aynı derslerin Grup sütununda birleştirilmesi	25
Çizelge 4.12. Dönem yılının yıl ve dönem sütununa ayrılması	26
Çizelge 4.13. Veri setindeki ÇAP programındaki öğrencilerin silinmesi	26
Çizelge 5.1. Programa göre öğrenci dağılımı.	30
Çizelge 5.2. Bölüm derslerine göre program bazında yüzdeler başarı oranı.	30
Çizelge 5.3. Programlara göre en az başarılı dersler	31
Çizelge 5.4. Programlara göre en başarılı dersler	32
Çizelge 5.5. Ders gruplarına göre başarı oranları	34
Çizelge 5.6. Programlara göre öğrenci dönem başarı oranı	35
Çizelge 5.7. Ön koşullu derslerin başarı oranları	36
Çizelge 5.8. Yaş grupları	36
Çizelge 5.9. En fazla öğrencisi bulunan yaş grubu	37
Çizelge 5.10. Birliktelik kuralı ile dersi ilk alışıta geçme-kalma tahmini (bütün sınıflar beraber)	39
Çizelge 5.11. Karar Ağacı ile dersi ilk alışıta geçme-kalma tahmini (bütün sınıflar beraber)	40
Çizelge 5.12. Karar ağacı şile 1.sınıf derslerine göre 2.sınıf derslerini sınıflandırma... 41	
Çizelge 5.13. Karar ağacı ile 2.sınıf derslerine göre 3.sınıf derslerini sınıflandırma 43	

Sayfa

Çizelge 5.14. Birliktelik kuralı ile ilk alıřta geemeyen ğrencilerin derslerinin tahmini.....	45
Çizelge 5.15. Karar ađacı ile ilk alıřta dersi geemeyen ğrencilerin ders tahmini (bütün derslere gre)	45
Çizelge 5.16. Karar ađacı ile ilk alıřta dersi geemeyen ğrencilerin ders tahmini (1.sınıf derslerine gre 2.sınıf dersleri)	47

SİMGELER VE KISALTMALAR DİZİNİ

KISALTMALAR

TDK	: Türk Dil Kurumu
ÇAP	: Çift Anadal Programı
UE	: Uzaktan Eğitim
VA	: Veri Analizi

BÖLÜM 1

GİRİŞ

Geçmişten günümüze öğrencilere öğretilen bilgi ve beceriler değişen zamanın koşullarına göre her geçen gün güncellenerek değişmekte ve yenilenmektedir. Değişen konularla birlikte uygulanacak olan öğretme yöntemi de zenginleştirilerek yenilenmektedir (Akın ve Kudak, 2020). Bu koşullar altında üniversitede geçirilen zaman önemli bir değere sahiptir ve derecesini tamamlayan öğrenci için bir getiri sağlamaktadır. Ancak yüksek öğretim sistemine giren genç yetişkinlerin başarıları üzerindeki etkisine ilişkin daha fazla araştırmanın yapılmasıyla daha karmaşık ama ilginç bir durumun ortaya çıktığı görülmüştür (Ndou et al. 2020).

Şüphesiz öğrencilerin okul başarısına etkileyen okul içi ve okul dışı birçok etkenden söz etmek mümkündür (Aslanargun, 2007). Yakın zamana kadar öğrencilerin başarılarını etkileyen derse devam, hazırlık sınavlarına veya düzenli dönem sonu sınavlarına katılarak dersi geçmek, öğrenci sorumluluğu, sınava çalışmak için harcanan zaman vb. konulara odaklanılmıştır (Dragicevic vd. 2014). Bununla birlikte öğrencinin kapasitesi, öğrenciden istenen görevin zorluk derecesi, düşük not alma korkusu, yetersiz hazırlık ve düşük bilgi düzeyi gibi değişkenler de incelenmiştir (Aydın ve Takkaç, 2007). Öğrencilerin başarısı konusunda yapılan kestirim ve karar alma işlemlerinde elde edilen sonuçların doğruluğu üzerinde büyük etkisi bulunduğu için bireysel veya demografik özelliklerin de başarı için önemli olduğu anlaşılmıştır (Aksu ve Doğan, 2018). Bu nedenle artık günümüzde öğrencilerin yaşı, cinsiyeti, medeni durum, ailesinin sosyo-ekonomik durumu, etnik kökeni gibi demografik özellikleri de öğrenci başarısı üzerinde önemli bir etkiye sahip olduğu ileri sürülerek çalışmalar yapılmaya başlanmıştır (Merwe ve Merwe, 2009; Kovacic, 2010; Abbasoğlu, 2020).

Öğrenci başarısı, öğrencilerin kendisi için önemli olduğu kadar yükseköğrenim kurumları için de önemlidir. Bunun nedeni, kalitesi yüksek bir üniversitenin

kriterlerinden birisi, mükemmel akademik başarı siciline dayanmasıdır (Shahiri et al. 2015). Bu nedenle toplumsal kültürü geliştirme ve yeni nesillere aktarma işlevi olan üniversitelerin bu işlevini yerine getirebilmesi için sürekli kendini yenilemesi ve geliştirmesi büyük önem taşımaktadır (Altun vd. 2019). Bu önemden dolayı öğrencilerin başarılarını önceden tahmin etmeleri, hem kurumlara hem de öğretilere öğrencinin performansını izlemeye yardımcı olabilmektedir (Mariame vd. 2021). Bu yolla önleyici tedbirler alınabilmekte ve risk altındaki öğrenciler erken tespit edilerek kurumsal performansın olumsuz etkilenmesi önlenmektedir (Alyahyan ve Düşteğör, 2020).

Öğrencilerin başarısı eğitim kurumları için önemli olmakla birlikte bilgi dünyasındaki hızlı gelişim, artan öğretim standartları, sınıf mevcutlarının artması, teknolojinin eğitimden daha hızlı ilerlemesi gibi nedenlerden dolayı kurumların bir takım yenilikler yapmasını zorunlu kılmıştır (Akçay vd. 2014). Bu yeniliklerden biri teknoloji kullanımınıdır. Bilgi ve iletişim teknolojilerinde yaşanan bu gelişmelerle birlikte üniversiteler teknolojilerden yararlanmaya başlamıştır (Afacan-Adanır, 2021). Teknoloji öğrenme sürecini olumlu yönde etkileyerek güçlü bir öğrenme aracı haline gelmiştir (Mariame et al. 2021). Ancak yeni bilgi ve teknoloji toplum, teknik, ekonomik, sosyolojik, kültürel ve politik düzene ilişkin bir dizi sorunu gündeme getirmiştir (Martinez ve Gomez, 2014). Bu sorunlardan biri yükseköğretim kurumlarındaki eğitim verilerinin önemli oranda büyümesidir. Bu veriler yönetime herhangi bir fayda sağlamadan hızla artmaktadır. Bu zor görevi yönetmek için, iş süreçlerinde üretilen büyük miktarda veriyi işlemek ve bazı yararlı bilgi ve bilgileri çıkarmak için yeni teknikler ve araçlar kullanılmasını gerektirmiştir (Delavari et al. 2008).

Yükseköğretim kurumları için öğrencilerin başarısının tahmin edilmesi çok önemlidir, çünkü öğretim sürecinin kalitesi öğrencilerin ihtiyaçlarını karşılayabilmelerine katkıda bulunmaktadır. Bu anlamda önemli veri ve bilgiler düzenli olarak toplanmakta, ilgili mercilerde değerlendirilmekte ve kalitenin devamlılığı için standartlar belirlenmektedir (Osmanbegovic ve Suljic, 2012). Öğrenci veri tabanlarında bulunan büyük miktarda veri, insanın otomatik analiz tekniklerinin yardımı olmadan en yararlı bilgileri analiz etme ve çıkarma yeteneğini

aşıyor (Algarni, 2016). Dolayısıyla saklanan bu veriler, çok çeşitli özelliklere sahip olmakla birlikte tek başlarına değersiz olup ancak bir amaç doğrultusunda işlendiği zaman bilgiye dönüşüp değer kazanabilmektedir (Akgün ve Bulut-Özek, 2020). Özellikle öğrenci başarısını değerlendirmek ve gerektiğinde akademik müdahaleyi tetiklemek için kullanılabilen çeşitli değerlendirme tekniklerine ihtiyaç vardır (Tinto, 2012).

Günümüzde bilgi yönetimi sürecinde, büyük miktarda veriden değerli ve anlamlı bilgiyi çıkarmak ve keşfetmek için veri madenciliği tekniği kullanılmaktadır (Natek ve Zwilling, 2014). Veri madenciliği; milyarlarca veri kaynağına sahip yazılım sistemlerinin, verilerini ve bilgilerini kıymetli hale getirmek için yapılan çalışmalara verilen isimdir (Kagün ve Maral-Karanfil, 2022). Veri Madenciliği, verileri birçok açıdan veya boyuttan analiz ederek yararlı bilgileri keşfedebilen, bu bilgileri kategorilere ayırabilen ve veri tabanında tanımlanan ilişkileri özetleyebilen güçlü bir yapay zeka aracıdır (Algarni, 2016).

Veri madenciliği, gizli eğilimleri ve kalıpları bulmak için açık bir bilgi tabanı, güçlü analitik yetenekler ve alan uzmanlığının bir karışımını kullanmaktadır. Bununla birlikte çok büyük veya ham veri kümelerinde veri madenciliği, denetimli veya denetimsiz veri madenciliği yöntemleri kullanılarak yapılmaktadır. Çünkü veri madenciliği, farklı verilerle doğrudan etkileşim olmadan gerçekleşemez (Arcinas et al. 2021). Eğitim alanındaki veri madenciliği tekniklerinin uygulanması, eğitim sistemini iyileştirmek için yönetimin karar vermesine yardımcı olacak bazı gizli bilgi ve kalıpları keşfedilmesine yardımcı olmaktadır. Dolayısıyla günümüz eğitim sisteminde, öğrenci özellikleri ve öğrencilerin öğrenme ve başarıları arasındaki etkileşimi göstermede çok önemlidir (Indahyanti et al. 2022).

Bu çalışmada Bilgisayar Mühendisliği Bölümündeki farklı programlarda okuyan öğrencilerin ders notları verileri incelenerek çıkarımlarda bulunulmuştur. Ön işlemler yapılarak hazırlanan veriler üzerinde analiz yapılarak bölüm ve programlarına göre ders başarılarını incelemek, tahmin etmek ve öğrencilerin dersi öğrenmeye yönelik davranışlarını anlamak için değerlendirmeler yapılmıştır. Ayrıca veri madenciliği

algoritmaları ile dersler arasındaki ilişkilerin bulunmasına yönelik çalışmalar yürütülmüştür.

BÖLÜM 2

LİTERATÜR ARAŞTIRMASI

Ayık ve arkadaşları (2007), Atatürk Üniversitesi'nden 1976 yılından itibaren mezun olan ve halen okumakta olan öğrenci bilgilerinin bulunduğu veritabanı üzerinden veri madenciliği teknikleri uygulanarak bir çalışma gerçekleştirmişlerdir. Araştırmaya başlamadan önce öğrencilerin yıllardır birikmiş karmaşık verileri üzerinde uzun bir hazırlık çalışması yapılmıştır. Daha sonra öğrencilerin lise türü ve lise mezuniyet başarısının, kazanılan fakülte üzerindeki etkisine bakılmıştır. Çalışma sonucunda, lise türü ve başarısının arzu edilen bir fakültenin kazanılmasında etkisi olduğu tespit edilmiştir.

Kovacic (2010) öğrencilerin okula devam etme veya okulu bırakmasını etkileyebilecek sosyo-demografik değişkenler (yaş, cinsiyet, etnik köken, eğitim, çalışma durumu ve engellilik) ve çalışma ortamını (kurs programı ve ders bloğu) gibi değişkenler açısından incelemiştir. Ayrıca başarılı ve başarısız öğrencilerin kimler olduğunu çalışmada incelemiştir. Bu çerçevede 2006'dan 2009'a kadar Açık Politeknik okulundaki öğrenci yönetim sisteminde depolanan ve Bilişim Sistemleri dersine kayıtlı 450'den fazla öğrenciyi kapsayan veriler değerlendirilmiştir. Veri madenciliği teknikleri (özellik seçimi ve sınıflandırma ağaçları gibi) ile öğrenci başarısı için en önemli faktörler ve tipik başarılı ve başarısız öğrencilerin profili belirlenmiştir. Çalışma sonucunda başarılı öğrencileri başarısız öğrencilerden ayıran en önemli faktörlerin etnik köken, ders programı ve ders bloğu olduğu tespit edilmiştir.

Alan (2012), Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü öğrencilerinin verilerini kullanarak veri madenciliği yöntemiyle bir çalışma yapmıştır. Çalışmada öğrencilerin okudukları programların, cinsiyetlerinin, Sivas ilinden olup olmamalarının, araştırma görevlisi olup olmadıklarının, ayrıca ders dönemlerinin farklı olup olmamasının başarılarını etkileyip etkilemediği incelenmiştir. Çalışmanın sonucunda doktora programı öğrencilerinin başarı düzeylerinin yüksek lisans

öğrencilerin başarı düzeyinden daha yüksek olduğu fakat cinsiyet, kadro, il ve ders dönemi gibi özelliklerin öğrencilerin başarısını etkilemediğini belirlemiştir.

Martinez ve Gomez (2014) yaptıkları çalışmada 2009-2013 yılları arasında Arjantin'deki iki farklı birimde okuyan öğrencilerin başarılarını etkileyen faktörleri belirlemek için veri madenciliğini kullanmıştır. Araştırma sonucunda yüksek düzeyde başarılı olan öğrencilerin oranının %25,78 olduğunu tespit etmişlerdir. Söz konusu öğrencilerin genel özelliklerine bakıldığında; çoğunun ailesiyle yaşadığı, genellikle belirli verimli saatlerde çalıştıkları, haftada 20 saat çalıştıkları, ebeveynlerinin ilk ve orta öğretim derecesine sahip oldukları, çoğu durumda öğrencilerin amacının, öğrenmeyi öğrenmek veya konuyu tam olarak öğrenmek için çalışmak olduğu, çoğunluğu bekar olduğu ve bunların genelde erkek oldukları tespit edilmiştir.

Baltacı (2018), ilahiyat fakültesine yerleşen öğrencilerin, fakülteye yerleştikten sonraki akademik başarısına etki eden faktörlerin neler olduğu, yerleşme puanlarına ve çeşitli demografik değişkenlerine göre incelemiştir. Veri madenciliği tekniği kullanılarak 2017 yılında bir üniversiteye yerleşen ve halen ilahiyat lisans programlarının ikinci sınıflarında öğrenim gören 3775 öğrenciye ait veri değerlendirilmiştir. Araştırmanın sonucunda, yılsonu not ortalaması ve başarısız olunan ders sayısı arasında negatif yönlü ve yüksek düzeyde bir ilişki belirlenmiştir. Bunun yanında, yılsonu not ortalaması ve yaş arasında yüksek düzeyde negatif ilişki belirlenmiştir. Öğrencinin üniversiteye yerleşme puanı, yılsonu not ortalamasını olumlu yönde etkilerken, başarısız olunan ders sayısını olumsuz etkilemektedir. Üniversite yerleşme puanı, başarısız olunan ders sayısını olumsuz yönde etkilemektedir. Cinsiyet değişkeni yılsonu not ortalamasını olumlu bir şekilde etkilerken, başarısız olunan ders sayısı ile anlamlı bir ilişki içinde değildir. Bunun yanında, üniversiteye yerleşme puanı ile tercih sırası değişkeni arasında da pozitif yönlü oldukça güçlü bir anlamlı ilişki olduğu belirlenmiştir.

Puarungroj ve arkadaşları (2018), öğrencilerin İngilizce yeterliliklerinin, mezun olduktan sonra iş aramak için önemli bir gereklilik olduğunu bunun için dünyanın pek çok üniversitesinde İngilizce dersi verildiğini ancak İngilizce başarısını etkileyen

faktörlerin tam olarak bilinemediğini belirterek bir çalışma gerçekleştirmişlerdir. Bu çerçevede Loei Rajabhat Üniversitesi'nde 2013, 2014 ve 2015 yıllarında okuyan öğrencilerin İngilizce başarılarını ve yordayıcılarını veri madenciliği tekniği ile analiz etmişler. Araştırmada, İngilizce bitirme sınavında öğrenci başarısını tahmin etmek için sınıflandırma modelleri kurmuş ve test edilmiştir. Araştırma sonuçları ayrıca İngilizce seviye belirleme sınavı sonucunun İngilizce mezuniyet sınavı sonucunu tahmin etmede anahtar bir özellik olduğunu göstermiştir.

Altun ve arkadaşları (2019) Akdeniz Üniversitesi Eğitim Fakültesi Sınıf Öğretmenliği Bölümü'nden 2012-2017 yılları arasında mezun olan 578 öğrencinin başarı durumlarını cinsiyetlerine, medeni durumlarına, kayıt yaşı ve 1. sınıf 1. dönem ara sınav puanlarına göre incelemişlerdir. Veri madenciliği yönteminin kullanıldığı çalışmada evli ve erkek olan öğrencilerin başarı düzeylerinin daha düşük olduğunu belirlemişlerdir.

Aydemir ve arkadaşları (2019), 2018-2019 bahar yarıyılı döneminde alınan Türk Dili II dersi ile ilgili öğrenci başarı durumlarını; öğrencilerin programlarına, ÖSS giriş türüne, ÖSS giriş puanına, ÖSS giriş sıralaması, bir önceki dönem not ortalamasına, ara sınav notuna, sınava çalışma durumuna, mevcut çalışma ile sınavdan kaç puan beklendiğine, sınavın nasıl geçtiğine göre incelemişlerdir. Araştırmada 160 öğrencinin verisi kullanılmıştır. Analizlerde beş farklı ağaç algoritması ile analizi yapılmış ve en iyi başarı oranı RandomForest algoritmasında elde edilmiştir. Gerçek veriler ile tahmin verileri arasındaki korelasyon katsayısı 0.72 bulunmuş ve 0-100 aralığındaki notlar 10.16 ortalama mutlak hata ile tahmin edilmiştir. Çalışma sonucunda öğrencilerin ders geçme notunu önceden tahmin ederek önlemlerin alınabileceği konusunda önerilerde bulunulmuştur.

Abbasoğlu (2020) Yalova ilindeki 4 farklı okuldaki 5, 6, 7 ve 8. Sınıf öğrencilerinden toplam 1395 kişinin E-Okul Yönetim Bilgi Sistemindeki bilgilerini incelemiştir. Çalışmada öğrencilerin demografik özelliklerinin ve sosyoekonomik durumlarının öğrencilerin yılsonu genel başarı ortalamalarını ne düzeyde etki ettiğini belirlemeye çalışmıştır. Verilerin analizi için veri madenciliği yöntemlerini kullanmıştır. Araştırma sonucunda öğrencilerin yaşının, derse devam durumunun,

anne ve babanın ayrı olup olmama durumunun, anne ve baba öğrenim durumunun, gelir düzeyinin, kendisi adına oda olup olmama durumunun ve fen bilimleri alanında kurs görüp görmeme durumunun yılsonu başarı puanlarını etkilediğini tespit etmiştir.

Ndou ve arkadaşları (2020), Güney Afrika yükseköğretim kurumlarındaki kayıtların genişlemesine, mezun olan öğrencilerin yüzdesinde orantılı bir artış eşlik etmediğini, 2000-2017 akademik yılları için 3 yıllık derecelerden ayrılan minimum sürede mezun olan öğrencilerin yüzdesinin düşük kaldığını, %25,7 ile %32,2 arasında değiştiğini ve öğrencilerin %60'ından fazlasının sürekli olarak seçtikleri çalışma alanını kendilerine ayrılan sürede tamamlamadığından, yükseköğretimde mücadele ettiklerini ileri sürerek bir çalışma yapmıştır. Bu bağlamda öğrenci başarısını etkileyen faktörlerin neler olduğunu tespiti için bir veri madenciliği modeli geliştirmişlerdir. Bu doğrultuda bir veri kümesine çeşitli sınıflandırma tekniklerini uygulayarak, bu sınıflandırıcıların öğrenci başarısını artırmak ve olumsuz sonuçlardan kaçınmak amacıyla öğrenci performansını önceden tahmin etmek için kullanılabilceğini göstermişlerdir.

Arcinas ve arkadaşları (2021), yükseköğretim kurumlarının, öğrenci performansını artırmaya yüksek öncelik verdiklerini ancak öğrencilerin başarısına katkı sağlamadan önce öğrencinin mevcut durumunun haritasını çıkarmak gerektiğini ileri sürerek veri madenciliği yöntemini kullanarak bir çalışma gerçekleştirmiştir. Bu amaçla öğrencilerin yükseköğrenim seçimine karar vermek için öğrencilerin davranışlarını, tutumlarını ve performanslarını tahmin etmek için öngörücü araçlar ve prosedürler geliştirilmiştir. Öğrenci performansının erken tahmin edilmesi hedeflenerek, öğrenci başarısını artırmaya dönük adımların neler olabileceği belirlenmiştir. Çalışmada kaliteli bir eğitim standardına ulaşmak için öğrenci performansını tahmin etmek için birkaç girişimde bulunulmuştur; ancak, tahmin doğruluğu tatmin edici olmadığı sonucuna ulaşılmıştır.

Yukarıda verilen örnek çalışmalarda görüldüğü üzere; eğitim sistemlerinde teknoloji kullanımının artmasının, büyük miktarda öğrenci verisinin depolanmasına yol açtığını, bu da öğretme ve öğrenme süreçlerini güçleştirdiği ve çeşitli öngörü sistemlerinin geliştirilmesine neden olduğu görülmektedir. Bu sistemler sayesinde

risk altındaki öğrenciler belirlenebilmekte, farklı öğrenci grupları için öncelikli öğrenme ihtiyaçları belirlenebilmekte, öğrencinin başarısının nasıl tanımlanacağı ve hangi öğrenci özelliklerine odaklanılacağına karar verilebilmekte, mezuniyet oranlarının artırılmasına yardımcı olunmakta, kurumsal performans etkili bir şekilde değerlendirilebilmekte ve okul kaynakları en üst düzeye çıkarılabilmektedir.

BÖLÜM 3

VERİ DEĞERLENDİRME YÖNTEMLERİ VE ÖNEMİ

Verilerin anlamlandırılarak hedeflenen alanla ilgili olarak bir kısım çıkarımlarda bulunmak o alanın dinamiklerinin daha iyi anlaşılmasına ve iyileştirmeler yapılmasına fırsat tanımaktadır. Bu çıkarımların yapılmasına yönelik bu çalışmada farklı açılardan veri setine yaklaşım yapmak amaçlandığından veri analizi kavramı ve veri madenciliği yöntemlerinden faydalanılmıştır. Bu çerçevede ilgili konularla ilgili temel kavramlar üzerinde durulacaktır. Ayrıca buradan elde edilecek sonuçların eğitimdeki öneminden bahsedilecektir.

3.1. VERİ ANALİZİ

Veri analizi, belirli bir amaç doğrultusunda toplanmış veya elde edilmiş verilerin incelenmesi, anlaşılması, yorumlanması ve değerlendirilmesi sürecidir. Bu süreç, verilerin içindeki desenleri, ilişkileri, eğilimleri ve anlamlı bilgileri çıkarmayı amaçlar. Veri analizi, bilimsel araştırmalardan endüstriyel işletmelere kadar birçok alanda kullanılan kritik bir yöntemdir.

Veri analizi, tüm kurumlarda stratejik karar verme süreçlerini etkilemektedir. Uygun yöntemler seçilip doğru yapıldığında hatayı en aza indirmekte veya hataları engelleme imkanına sahiptir. Bu yüzden birçok sektör ve kurum veri analizlerini kullanmaktadır. Bunun sonucunda veri analizi geleceğe yönelik yararlı olacak en doğru kararları almak için başvurulan yöntemlerden biridir. Sonuç olarak veri analizinin önemi gelecekte stratejileri daha doğru belirleyebilmek için büyük önem taşır.

3.1.1. Veri Analizi Süreci

Veri analizi sürecinin amacı işlenmemiş veriden anlam çıkarabilmektedir. Bu sayede en doğru bilgiye ulaşmamızı sağlamaktadır. Veri analizi süreci 4 ya da 5 adımdan oluşmaktadır.

- Problemin tanımlanması
- Probleme yönelik alt yapı oluşturma
- Verinin toplanıp hazırlanması
- Analizin yapılması verinin işlenmesi
- Yorumlama

3.1.2. Veri Analizi Türleri

Veri analizi, Veri analizi, araştırmacıların veriyi anlama, özetleme ve sonuç çıkarma sürecidir. Bu süreç farklı yöntemler ve amaçlar doğrultusunda çeşitli türlerde gerçekleştirilebilir. Araştırmacıların hedeflerine ve sorularına göre farklı analiz yöntemleri seçmeleri gerekebilir. Bu türler, veri analizini farklı bakış açılarından ele alan ve farklı sonuçlar elde etmeyi amaçlayan farklı yaklaşımları temsil eder. Hangi analiz türünün kullanılacağı, veri setinin doğası ve analizin amacına bağlı olarak belirlenir. Veri analizi türlerinden kısaca bahsedilmiştir.

- **Tanımlayıcı Analiz:** Verinin temel özelliklerini ve istatistiklerini anlama
- **Keşifsel Analiz:** Gizli trendleri ve desenleri bulma
- **Açıklayıcı Analiz:** Neden-sonuç ilişkilerini anlama
- **Tahminsel Analiz:** Gelecekteki olayları tahmin etme
- **Zaman Serisi Analizi:** Zaman içindeki verilerin desenlerini inceleme
- **Segmentasyon Analizi:** Veriyi gruplara ayırarak farkları anlama
- **İlişki Analizi:** Değişkenler arasındaki ilişkileri anlama
- **Regresyon Analizi:** Değişkenler arasındaki ilişkiyi inceleme
- **Sentiment Analizi:** Metin verilerinin duygusal tonunu anlama
- **A/B Testi:** Farklı versiyonların performansını karşılaştırma

- **Büyük Veri Analizi:** Büyük ve karmaşık veri setlerini inceleme.

3.2.VERİ MADENCİLİĞİ YÖNTEMLERİ

Bu çalışmada veriler arasındaki doğrudan görülemeyen ilişkilerin olup olmadığını anlamak için Veri Madenciliğinden yararlanılmıştır. Bu amaçla Weka programı kullanılmıştır. Yöntem ve algoritma olarak Birliktelik Kurallarından Apriori algoritması ve Karar Ağacı yönteminde ise J48 algoritması seçilmiştir.

3.2.1. Birliktelik Kuralları (Apriori)

Apriori algoritması, veri madenciliği ve özellikle birliktelik kuralı analizi için kullanılan bir algoritmadır. Bu algoritma, bir veri kümesindeki öğeler arasındaki ilişkileri keşfetmeye yardımcı olur, özellikle hangi öğelerin birlikte sıkça görüldüğünü belirlemek amacıyla kullanılmaktadır.

Birliktelik kuralı analizi sonuçlarını yorumlamak ve anlamak için ölçütler kullanılır. Confidence, lift, leverage ve conviction, kuralların ne kadar anlamlı ve işe yarar olduğunu belirlemeye yardımcı olur.

- **Confidence (Conf):** Confidence, birliktelik kuralının güvenilirliğini ifade eder. Yani bir öğe kümesinin içerdiği öğelerin diğer öğelerle ne sıklıkla birlikte bulunduğunu gösterir. Yüksek bir güven değeri, bu kuralın güvenilir ve sık gözlemlenen bir ilişkiyi ifade ettiğini gösterir.
- **Lift:** Lift, birliktelik kuralının rastgelelikle karşılaştırıldığında ne kadar daha anlamlı olduğunu ölçer. Lift değeri 1'den büyükse, bu kuralın rastgele bir ilişkiyi ifade etmediğini, belirli bir bağımlılığın varlığını işaret eder. Değer ne kadar yüksekse, o kadar anlamlı bir ilişki olduğunu gösterir.
- **Leverage (Lev):** Leverage, birliktelik kuralının gözlenen örneklerle rastgele örnekler arasındaki farkı ölçer. Pozitif bir leverage değeri, bu kuralın gözlenen örneklerle rastgelelik arasında anlamlı bir fark olduğunu gösterir.
- **Conviction (Conv):** Conviction, birliktelik kuralının bağımsızlığını ifade eder. Yani bir kuralın ne kadar bağımsız veya diğer öğelerle ne kadar az ilişkili olduğunu gösterir. Yüksek bir conviction değeri, kuralın

bağımsızlığını kaybettiğini ve diğer öğelerle güçlü bir ilişki içinde olduğunu gösterir.

3.2.2. Karar Ağacı (J48)

Karar Ağacı, veri madenciliği ve veri analizi alanında sıkça kullanılan bir sınıflandırma algoritmasıdır. J48, C4.5 algoritmasının Java programlama dili için geliştirilmiş bir uygulamasıdır. Bu algoritma, veri kümesindeki desenleri tanımlayarak ve sınıflandırarak veri madenciliği ve veri analizi problemlerini çözmek için kullanılır.

Karar ağacı, bir hiyerarşik ağaç yapısı olarak temsil edilir. Bu ağaç, veri kümesindeki özelliklerin ve sınıfların ilişkisini açıklar. Karar ağacı, bir veri örneğini sınıflandırmak için kullanılan bir dizi karar kuralını içerir. Her iç düğüm, bir özellik üzerinde bir karar noktasını temsil eder ve her yaprak düğüm, bir sınıf etiketini veya sonuç sınıfını belirtir.

Karar ağacının belirli bir düğümünün özelliklerini, sınıf tahminlerini ve düğümün içeriğini açıklamak için kullanılır.

- **≤ 0.5 (veya benzeri bir eşitlik/koşul):** Belirli bir düğümdeki bölme koşulunu ifade eder. Karar ağacı, bir özellik değerini belirli bir eşik değeriyle karşılaştırarak örnekleri iki alt düğüme böler.
- **gini:** Düğümün Gini impurity (Gini belirsizliği) değerini gösterir. Gini impurity, düğümdeki örneklerin sınıflarına göre ne kadar homojen veya heterojen olduğunu ölçer. Daha düşük bir Gini impurity değeri, daha homojen bir düğümü ifade eder.
- **samples:** Düğümde toplam kaç örnek bulunduğunu gösterir.
- **value:** Düğümdeki örneklerin her sınıfa ait sayılarını gösterir. İlk değer, bir sınıfa ait olan örnek sayısını temsil eder, ikinci değer ise diğer bir sınıfa ait olan örnek sayısını temsil eder.
- **class:** Düğümdeki örneklerin sonuç olarak hangi sınıfa ait olduğunu belirtir.

Metrikler, sınıflandırma modelinin performansını anlamak ve karşılaştırmak için kullanılır.

- **Karmaşıklık matrisi (Confusion Matrix):** bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir tablodur. Bu matris, modelin doğru ve yanlış sınıflandırılmış örneklerini gösterir ve aşağıdaki temel bileşenleri içerir:
 - True Positives (TP): Gerçek pozitif örneklerin sayısını temsil eder. Model doğru bir şekilde pozitif olarak tahmin ettiği örneklerdir.
 - True Negatives (TN): Gerçek negatif örneklerin sayısını temsil eder. Model doğru bir şekilde negatif olarak tahmin ettiği örneklerdir.
 - False Positives (FP): Yanlış pozitif örneklerin sayısını temsil eder. Model yanlış bir şekilde pozitif olarak tahmin ettiği örneklerdir.
 - False Negatives (FN): Yanlış negatif örneklerin sayısını temsil eder. Model yanlış bir şekilde negatif olarak tahmin ettiği örneklerdir.
- **Accuracy (Doğruluk):** Modelin doğru tahmin ettiği örneklerin yüzdesini verir. Bu metrik, modelin genel başarısını gösterir. Genellikle tüm tahminlerin ne kadarının doğru olduğunu ölçer ve aşağıdaki formülle hesaplanır:

$$Accuracy = \frac{Doğru\ Tahminler}{Toplam\ Tahminler}$$

- **Precision (Kesinlik):** Pozitif olarak tahmin edilen örneklerin ne kadarının gerçekten pozitif olduğunu ölçer. Precision, modelin pozitif tahminlerinin doğruluk oranını gösterir. Yüksek bir precision, yanlış pozitiflerin az olduğunu gösterir. Yanlış pozitiflerin (yanlış alarm) oranını hesaplamak için kullanılır ve aşağıdaki formülle hesaplanır:

$$Precision = \frac{Gerçek\ Pozitifler}{Doğru\ Pozitifler + Yanlış\ Pozitifler}$$

- **Recall (Duyarlılık):** Gerçek pozitif örneklerin ne kadarının doğru bir şekilde tanımlandığını ölçer. Recall, modelin gerçek pozitifleri ne kadar yakaladığını gösterir. Yüksek bir recall, pozitif örneklerin kaçının yakalandığını gösterir. Kaç pozitif örneğin doğru bir şekilde tespit edildiğini belirler ve aşağıdaki formülle hesaplanır:

$$Recall = \frac{Gerçek\ Pozitifler}{Doğru\ Pozitifler + Yanlış\ Negatifler}$$

- **F1-Score:** Kesinlik (precision) ve duyarlılık (recall) metriklerini bir araya getirir ve modelin performansını genel olarak ölçer. Dengesiz sınıflara sahip veri kümeleri için daha bilgilendirici bir ölçüdür. İdeal bir F1-Score, yüksek bir kesinlik ve yüksek bir duyarlılığı gösterir. Aşağıdaki formülle hesaplanır:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

3.2.3. Eğitimde Veri Analizinin Önemi

Veri analizi, yükseköğretimde büyük bir öneme sahiptir. Yüksek öğretim kurumları, öğrenci başarısı, öğretim kalitesi, araştırma etkinlikleri, yönetim süreçleri ve daha birçok alanda veri üretirler. Yüksek öğretim kurumları, öğrencilerin ders notları, sınav sonuçları, projeler ve diğer öğrenme değerlendirmeleri ile ilgili verileri toplar. Bu verilerin analizi, hangi derslerin öğrenciler için daha zorlayıcı olduğunu, hangi dönemlerde daha düşük performans gösterildiğini ve hangi öğrenci gruplarının ek destek veya rehberliğe ihtiyaç duyduğunu anlamak için kullanılabilir.

Veri analizi, hangi öğretim materyallerinin öğrenci başarısını artırdığını ve öğrenci, ilgisini çektiğini anlamak için kullanılabilir. Bu şekilde öğrenci odaklı öğretim materyalleri geliştirilebilir ve öğrenci merkezli öğrenme yaklaşımları desteklenebilir. Üniversitelerdeki araştırma etkinliklerinin değerlendirilmesi ve yönetimi için veri analizi kullanılır. Akademik yayınlar, proje başvuruları, araştırma fonları ve benzeri konularda veri analizi, araştırma alanındaki eğilimleri ve başarıları anlamak için kullanılır.

Veri analizi, yükseköğretim politikalarının geliştirilmesi ve yönlendirilmesinde yardımcı olabilir. Öğrenci kabul kriterleri, mezuniyet oranları, çeşitlilik ve dahil etme gibi konularda veri analizi, stratejilerin daha etkili bir şekilde tasarlanmasına yardımcı olur. Üniversitelerin yönetim ve idari süreçleri de veri analizi ile izlenebilir. Bütçe yönetimi, kaynak tahsisi, personel performansı gibi alanlarda veri analizi,

kurumsal performansı deęerlendirmek ve gerektięinde iyileřtirmeler yapmak için kullanılabilir.

Yükseköęretimde, cinsiyet, etnik köken, sosyoekonomik durum gibi faktörlerden kaynaklanan eşitsizlikleri tespit etmek ve azaltmak amacıyla veri analizi kullanılabilir. Bu şekilde, farklı öęrenci gruplarının başarılarını eşit hale getirmek için stratejiler geliştirilebilir.

Sonuç olarak, yükseköęretimde veri analizi, kurumların daha iyi kararlar almasına, öęrenci başarısını artırmasına, öęretim ve araştırma kalitesini optimize etmesine ve genel olarak eğitim deneyimini geliřtirmesine yardımcı olur.

BÖLÜM 4

VERİ SETİ VE ÖN İŞLEMLER

Yapılan bu çalışmada Karabük Üniversitesi Mühendislik Fakültesi/Bilgisayar Mühendisliği bölümü için 2009-2014 kayıt tarihleri için farklı programlarda öğrenim görmüş öğrencilerin lisans öğrenim verileri kullanılmıştır. Öğrencilerin programlarına göre sınav ders başarılarının değişik açılardan analiz ve değerlendirmesi yapılmıştır. Verilerin anlamlı hale getirilmesi için veri ön işlemleri ve analizi yapılmıştır. Veri ön işlemleri ve analiz için Microsoft Excel programı, sorgulama yapılması için ise Microsoft SQL Server Management Studio programı kullanılmıştır.

4.1. VERİ ELDE ETME SÜRECİ

Bu çalışma Karabük Üniversitesi Mühendislik Fakültesi/Bilgisayar Mühendisliği bölümü 2009-2018 yılları arasında “%30 İngilizce Birinci Öğretim”, “%30 İngilizce İkinci Öğretim”, %100 İngilizce Birinci Öğretim”, %100 İngilizce İkinci Öğretim” ve “Uzaktan Eğitim Programları” olmak üzere 5 farklı programda öğrenim görmüş öğrencilerin lisans öğrenim verileri kullanılarak yapılmıştır. Daha yakın dönemdeki verilerin alınmaması farklı otomasyon sistemi kullanılması sebebiyle bütüncül veri setinin elde edilememesi sebebiyledir. Araştırmada kullanılan öğrenci verileri, Karabük Üniversitesi öğrenci işleri birimindeki veri tabanından alınmıştır.

4.1.1. Araştırmaya Dahil Edilme Kriterleri

Alınan veri seti Öğrenci Adı Kayıt Tarihi, Doğum Yılı, Program, Ders Hareket ID, Şube Kodu, Kodu, ders ADI, Notu, Dönemi, Sınıfı, Sınav Tipi, Ortalama Notu, Harf Kodu olmak üzere 14 özellikten oluşmaktadır. Araştırmada alınan veri seti

özelliklerinden “Şube Kodu” özelliği herhangi bir anlam içermediğinden dolayı bu alan dahil edilmeyerek 13 özellik kullanılmıştır.

4.1.2. Araştırmadan Dışlanma Kriterleri

Bu çalışmada demografik bilgilerden olan **cinsiyet** ve **doğum yerinin** öğrenci başarısına etkisinin de incelenmesi amaçlanmıştı. Ancak, KVKK kapsamı gereğince bu veriler elde edilemediği için diğer veriler kullanılarak sınırlı verilerle analiz yapılmıştır.

4.2. VERİ SETİNİN ÖZELLİKLERİ

Bu çalışmada Karabük Üniversitesi Mühendislik Fakültesi/Bilgisayar Mühendisliği bölümü 2009-2018 kayıt tarihleri için “%30 İngilizce Birinci Öğretim”, “%30 İngilizce İkinci Öğretim”, %100 İngilizce Birinci Öğretim”, %100 İngilizce İkinci Öğretim” ve “Uzaktan Eğitim Programları” olmak üzere 5 farklı programda öğrenim görmüş öğrencilerin lisans öğrenim verileri kullanılarak yapılmıştır.

Veri setinde bulunan “Kayıt Tarihi” sütunu 2009-2018 yıllarında kayıtlı öğrencileri içermektedir. Alınan veriler kapsamlı olarak incelendiğinde 2014 yılından sonraki verilerde müfredat değişikliğinden dolayı bilgilerin tutarsız olduğu görülmektedir. Bu nedenle çalışmada 2009-2014 yılları arasındaki veriler baz alınmıştır. Alınan Veri setinin boyutu 4.359.191 adet veriden oluşmaktadır. Veri setine ön işlemler uygulanıp anlamlı hale getirildikten sonra 6.189.282 adet veri boyutuna ulaşmıştır.

Veri setinde bulunan programlarda “ÇAP”, “Mühendislik tamamlama” öğrencilerinin bilgileri öğrenci sayılarının ve almış oldukları ders sayılarının az olması sebebiyle çalışmaya dâhil edilmemiştir.

Veri setinde dönem bilgisi içerisinde bulunan “Yaz” dönemi her yıl açılmadığı için açıldığı zaman da öğrenci sayılarının ve açılan derslerin değişkenlik göstermesi sebebiyle çalışmaya dâhil edilmemiştir.

Alınan veri seti KVKK kapsamında olduğu için veri setindeki “Öğrenci Adı” bilgisi sadece ilk iki harf ve yıldızlardan oluşmaktadır. Alınan veri seti Öğrenci Adı Kayıt Tarihi, Doğum Yılı, Program, Ders Hareket ID, Şube Kodu, Kodu, , Ders Adı,, Notu, Dönemi, Sınıfı, Sınav Tipi, Ortalama Notu, Harf Kodu olmak üzere 14 özelliğten oluşmaktadır. Çizelge 4.1’de alınan ilk veri seti içeriği görülmektedir.

Çizelge 4.1. Alınan veri setinin içeriği

Özellikler	Veri tipleri	Örnek
Adı	Öğrenci adı karakter dizisi ve özel karakterden oluşmaktadır.	Bü*****
Kayıt Tarihi	Sayısal değerlerden oluşmaktadır.	2009-2017 arası
Doğum Yılı	Sayısal değerlerden oluşmaktadır.	1956-1999 arası
Program	Karakter dizisi değerlerden oluşmaktadır.	ÇAP, Ue,
Ders Hareket ID	Sayısal değerlerden oluşmaktadır.	55127
Şube Kodu	Karakter dizisi ve sayısal verilerden oluşmaktadır.	1,2,A,B,C,148420
Kodu	Karakter dizisi ve sayısal verilerden oluşmaktadır.	BLM320,SOS381
Adı	Karakter dizisi değerlerden oluşmaktadır.	Matematik, Fizik vb.
Notu	Sayısal değerlerden oluşmaktadır.	0-100 arası
Dönemi	Karakter dizisi ve sayısal verilerden oluşmaktadır.	2012-2013 Güz
Sınıfı	Sayısal değerlerden oluşmaktadır.	1,2,3,4
Sınav Tipi	Karakter dizisi değerlerden oluşmaktadır.	Ara sınav, yarıyıl sonu, bütünleme
Notu	Sayısal değerlerden oluşmaktadır.	0-100 arası
Harf Kodu	Karakter dizisi değerlerden oluşmaktadır.	A1,AA,CC,F1,F2,FF

Veri setinin daha verimli ve anlamlı hale gelebilmesi için sorgulama ve ön işlemler yapılmıştır. Böylelikle veri setindeki öğrenci bilgileri 14 özelliğten 21 özelliğe çıkarılmıştır.

Çizelge 4.2. Alınan veri setine eklenen özellikler

Özellikler	Açıklama
ID	Öğrencilere ait eşsiz niteliği belirtir.
Group	Farklı dilde verilen aynı derslerin ortak grubu
Ders_Group	Derslerin Ortak, Teknik, Sosyal sınıflandırma bilgisi
Birinci_Yıl	Dönemin ait olduğu ilk yıl bilgisini içerir
İkinci_Yıl	Dönemin ait olduğu ikinci yıl bilgisini içerir
Durum	“Kaldı” ve “Geçti” bilgisini verir
Tekrar	Dersin kaçınıcı kez alındığı bilgisini içerir

Çizelge 4.3. Veri setinin ilk hali

Adı	Kayıt Tarihi	Doğum Yılı	Program	Ders Hareket ID	Şube Kodu	Kodu	Ders Adı	Notu	Dönemi	Sınıfı	Sınav Tipi	Notu	Harf Kodu
Bü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	89663	1	MSD462	Etik	98	2012-2013 Bahar	4	Ara Sınav	0	A1
Bü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	89663	1	MSD462	Etik	98	2012-2013 Bahar	4	Yarıyıl Sonu Sınavı	0	A1
Me*****	2009	1990	Bilgisayar Mühendisliği (%30 İng.)	101166	1	MSD462	Etik	84	2013-2014 Bahar	4	Ara Sınav	96	A2
Me*****	2009	1990	Bilgisayar Mühendisliği (%30 İng.)	101166	1	MSD462	Etik	84	2013-2014 Bahar	4	Yarıyıl Sonu Sınavı	76	A2

Çizelge 4.4. Veri setine eklenen özellikler

ID	Grup	Ders_Adi	Ders_Group	BirinciYil	İkinciYil	Donem	Harf_Kodu	Durum
1001	21	Etik	Seçmeli	2012	2013	Bahar	A1	GEÇTİ
1001	21	Etik	Seçmeli	2012	2013	Bahar	A1	GEÇTİ
1002	21	Etik	Seçmeli	2013	2014	Bahar	A2	GEÇTİ
1002	21	Etik	Seçmeli	2013	2014	Bahar	A2	GEÇTİ

4.3.VERİ ÖN İŞLEMLERİ

4.3.1. Veri Temizleme

Analizin en kritik ve zaman alıcı sürecidir. Gürültülü, eksik ve tutarsız verilerin silinmesi, düzeltilmesi veya eklenmesi gibi işlemleri gerektirir. Veri setinde bulunan “Şube Kodu” sütunu anlamlı bilgiler sunmadığı için silindi. Tüm dersler için “Şube Kodu” sütununa bakılarak inceleme yapılmıştır. Çizelge 4.5’te “Devre Analizi” dersi için örnek olarak gösterilmiştir.

Çizelge 4.5. Devre analizi dersi için “Şube Kodu”

Şube Kodu	Kodu	Ders Adı	Notu	Dönemi	Sınıfı	Sınav Tipi
B	BLM223	Devre Analizi	60	2014-2015 Güz	2	Ara Sınav
B	BLM223	Devre Analizi	60	2014-2015 Güz	2	Yarıyıl Sonu Sınavı
B	BLM223	Devre Analizi	60	2014-2015 Güz	2	Bütünleme
A	BLM223	Devre Analizi	18	2014-2015Güz	2	Ara Sınav

Veri setinde derslerin ortalama notu gösteren “Notu” sütununda negatif sayı girilen notlar incelenerek düzeltildi. Çünkü ortalama not kısmının eksi değeri sonuçlara etki etmektedir. Negatif sayıların olduğu “Ara sınav, Yarıyıl sonu ve Bütünleme” sınavlarını gösteren “Notu” sütunu tüm dersler için ayrıntılı olarak incelendi. Notu kısmındaki negatif sayılar “0” olarak değiştirildi. Çizelge 4.6’da örnek olarak bu durum gösterilmiştir.

Çizelge 4.6. Negatif not girilen iki farklı dersin gösterimi

Kodu	Ders Adı	Notu	Dönemi	Sınıfı	Sınav Tipi	Notu	Harf Kodu
BLM223	Devre Analizi	0	2014-2015 Güz	2	Ara Sınav	-5	F1
BLM223	Devre Analizi	0	2014-2015 Güz	2	Yarıyıl Sonu Sınavı	-5	F1
BLM112	Programlama Dilleri II	0	2014-2015 Bahar	1	Ara Sınav	-8	F3
BLM112	Programlama Dilleri II	0	2014-2015 Bahar	1	Yarıyıl Sonu Sınavı	-8	F3
BLM112	Programlama Dilleri II	0	2014-2015 Bahar	1	Bütünleme	-8	F3

Ara, yılsonu ve bütünleme notu olmayan dersler ise çalışmamıza katkı sağlamayacağı için silindi. Endüstri Stajı dersi bu duruma örnek olarak Çizelge 4.7’de gösterilmiştir.

Çizelge 4.7. Aynı döneme ait Endüstri Stajı dersi

Adı	Program	Ders Adı	Notu	Dönemi	Sınav Tipi	Notu
Se*****	Bilgisayar Mühendisliği (%30 İng.)	Endüstri Stajı I	0	2013-2014 Güz	Yarıyıl Sonu Sınavı	-3
Sa*****	Bilgisayar Mühendisliği (%30 İng.)	Endüstri Stajı I	0	2013-2014 Güz	Yarıyıl Sonu Sınavı	-3
Yu*****	Bilgisayar Mühendisliği (%30 İng.)	Endüstri Stajı I	0	2014-2015 Yaz	Yarıyıl Sonu Sınavı	-3
Me*****	Bilgisayar Mühendisliği (II) (%30 İng.)	Endüstri Stajı I	0	2014-2015 Güz	Yarıyıl Sonu Sınavı	-4
Me*****	Bilgisayar Mühendisliği (II) (%30 İng.)	Endüstri Stajı I	-4	2015-2016 Güz	Yarıyıl Sonu Sınavı	-4
Me*****	Bilgisayar Mühendisliği (II) (%30 İng.)	Endüstri Stajı I	-4	2015-2016 Güz	Bütünleme	-4
Ab*****	Bilgisayar Mühendisliği (II) (%30 İng.)	Endüstri Stajı I	0	2015-2016 Güz	Yarıyıl Sonu Sınavı	-3

Harf kodu kısmında bulunan “F1” kodu devamsızlıktan kalan öğrencileri belirtir. Veri setinde bulunan derslerde “F1” kodunun ortalama not kısımlarına tutarsız verilerin girilmesi. Ortalama not kısmında negatif olan değerler “0” “verilerek düzeltilmiştir. Aşağıda “F1” kodu ile ilgili örnek ders gösterilmiştir.

Çizelge 4.8. F1 Harf Koduna ait tutarsız veri örneği

Adı	Program	Adı	Dönemi	Sınav Tipi	Notu	Harf Kodu
Er*****	Bilgisayar Mühendisliği - (U.E)	Devre Analizi	2014-2015 Yaz	Ara Sınav	26	F1
Er*****	Bilgisayar Mühendisliği - (U.E)	Devre Analizi	2014-2015 Yaz	Yarıyıl Sonu Sınavı	-5	F1
Em*****	Bilgisayar Mühendisliği - (U.E)	Devre Analizi	2014-2015 Yaz	Ara Sınav	15	F1
Em*****	Bilgisayar Mühendisliği - (U.E)	Devre Analizi	2014-2015 Yaz	Yarıyıl Sonu Sınavı	-5	F1
Mu*****	Bilgisayar Mühendisliği (%30 İng.)	Devre Analizi	2013-2014 Güz	Ara Sınav	41	F1
Mu*****	Bilgisayar Mühendisliği (%30 İng.)	Devre Analizi	2013-2014 Güz	Yarıyıl Sonu Sınavı	-5	F1
İd*****	Bilgisayar Mühendisliği (%30 İng.)	Devre Analizi	2013-2014 Güz	Ara Sınav	42	F1
İd*****	Bilgisayar Mühendisliği (%30 İng.)	Devre Analizi	2013-2014 Güz	Yarıyıl Sonu Sınavı	-5	F1

4.3.2. Veri Birleştirme

Farklı özellikteki verilerin birleştirilip istenilen verilere dönüştürülmesi sürecidir. Bu süreç karmaşıklığı önleyerek istenilen özellikteki verilere ulaşılmasını sağlar. Öğrenci numarasının KVKK kapsamında verilmemiş olması ve Öğrenci Adının anonim hale getirilmiş olmasından dolayı öğrenci kayıtlarının öğrenci bazlı analizinin yapılamayacağından bir eşsiz nitelik oluşturulması ihtiyacına gerek duyulmuştur. Bu duruma çözüm olarak “Adı”, “Kayıt Tarihi”, “Doğum Yılı” ve “Program” bilgileri aynı olan kişilere ID isminde sütun oluşturularak numaralar verildi. Fakat aynı dönem ve aynı sınav bilgileri seçildiğinde 2 farklı not bilgisi olan öğrencilerle karşılaşıldı. Örneğin; Kayıt tarihi, programı, doğum tarihi aynı bilgilere sahip Ab***** öğrencisinin 2014-2015 güz döneminde Fizik dersi ara sınavında iki farklı not çıkmaktadır. Sonrasında bu durumu önlemek için “Arasınav” ve “Yılsonu” bilgileri de eklenerek birleştirme işlemi yapıldı. Bu birleştirme için Excel’deki “DÜŞEYARA” formülü kullanılmıştır.

Çizelge 4.9. ID verilmeden önceki veri seti

Adı	Kayıt Tarihi	Doğum Yılı	Program	Kodu	ADI
Bü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	FOL381	Reading and Speaking at Foreign Language
Bü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	FOL381	Reading and Speaking at Foreign Language
Me*****	2009	1990	Bilgisayar Mühendisliği (%30 İng.)	BLM216	File Organization
Me*****	2009	1990	Bilgisayar Mühendisliği (%30 İng.)	FOL282	Technical Foreign Language II
Se*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	BLM423	Bilgisayar Grafiği
Se*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	BLM499	Endüstri Stajı II
Lü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	FOL381	Reading and Speaking at Foreign Language
Lü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)	FOL381	Reading and Speaking at Foreign Language

Çizelge 4.10. ID verildikten sonraki veri seti.

ID	Adi	Kayıt Tarihi	Doğum Yılı	Program
1001	Bü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)
1001	Bü*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)
1002	Me*****	2009	1990	Bilgisayar Mühendisliği (%30 İng.)
1002	Me*****	2009	1990	Bilgisayar Mühendisliği (%30 İng.)
1010	Me*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)
1010	Me*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)
1013	Me*****	2009	1989	Bilgisayar Mühendisliği (%30 İng.)
1013	Me*****	2009	1989	Bilgisayar Mühendisliği (%30 İng.)
3010	Me*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)
3010	Me*****	2009	1991	Bilgisayar Mühendisliği (%30 İng.)
1072	Me*****	2010	1991	Bilgisayar Mühendisliği (%30 İng.)
1072	Me*****	2010	1991	Bilgisayar Mühendisliği (%30 İng.)

Veriler incelendiğinde Bilgisayar Mühendisliği bölümünün 5 farklı programında aynı derslerin Türkçe ve İngilizce olarak verildiği anlaşıldı. Bu dersler için “Grup” adında sütun oluşturuldu. Böylelikle “Kodu“ sütununda farklı gözükse ama aynı olan dersler birleştirilerek ders başarı analizi için kolaylık sağlandı. Bu sütunda örneğin; “Data Communication Systems” ve “Veri İletişim Sistemleri” isiminde olan 2 ayrı ders “53” numarası altında birleştirilmiştir.

Çizelge 4.11. Aynı derslerin Grup sütununda birleştirilmesi

ID	Adi	Kayıt_Tarihi	Doğum_Yili	Program	Kodu	Grup	Ders_Adi
1036	Ab*****	2009	1989	Bilgisayar Mühendisliği (II) (%30 İng.)	CME325	53	Data Communication Systems
1071	Ab*****	2010	1984	Bilgisayar Mühendisliği (%30 İng.)	CME325	53	Data Communication Systems
1220	Ab*****	2010	1991	Bilgisayar Mühendisliği (U.E)	BLM325	53	Veri İletişim Sistemleri
1276	Ab*****	2010	1992	Bilgisayar Mühendisliği (İngilizce) (II)	CME325	53	Data Communication Systems

“Ders_Group” adında yeni bir sütun açılarak dersler **ortak, sosyal ve teknik** adında üç sınıf olarak sınıflandırılmıştır. Bu birleştirme işlemleri farklı açılardan durum analizleri yapılmasını kolaylaştırmıştır.

Ortak dersler: Tüm mühendisliklerde ve fakültelerde görülen ortak derslerden oluşmaktadır. Örneğin; Türk Dili, Yabancı Dil, Matematik, Diferansiyel Denklemler, Genel Fizik vb.

Sosyal dersler: Bölüm dışı ve ders yükü az olan derslerdir. Örneğin; Değerler Eğitimi, Etik vb.

Teknik dersler: Bölüm dersleridir. Örneğin; Programlama Dilleri, Bilgisayar Ağları, Veri Madenciliği vb.

4.3.3. Veri Dönüştürme

Dönüştürme işlemi, verilerin analiz için istenilen formlara dönüştürülmesidir. Bu işlemler verilerin düzeltilmesi, birleştirilmesi ve normalleştirilmesi gibi işlemleri içerebilir. Veri setinde bulunan Donem_Yili sütunu yıl ve dönem bilgisini içermesinden dolayı “BirinciYıl”, İkinciYıl ve Donem adında yeni veriler oluşturuldu. Böylelikle dönem ve yıllar bazında başarı analiz işlemlerinin yapılması için kolaylık sağlandı.

Çizelge 4.12. Dönem yılının yıl ve dönem sütununa ayrılması

DonemYili	BirinciYil	İkinciYil	Donem
2012-2013 Bahar	2012	2013	Bahar
2012-2013 Bahar	2012	2013	Bahar
2013-2014 Güz	2013	2014	Güz
2013-2014 Güz	2013	2014	Güz

4.3.4. Veri İndirgeme

Gereksiz ve fazla olan bazı değişkenlerin veri setinden çıkartılarak veri boyutunun azaltılması işlemidir. Bu sayede analiz yapılacak olan veri kümesinin en iyi veri sonuçlarına ulaştırılması amaçlanmaktadır. Veri setinde bulunan “Kayıt Tarihi” sütunu 2009-2018 yıllarında kayıtlı öğrencileri içermektedir. Alınan veriler kapsamlı olarak incelendiğinde 2014 yılından sonraki verilerde müfredat değişikliğinden dolayı bilgilerin tutarsız olduğu görüldü. Bu nedenle 2009-2014 yıllarını içeren bilgiler alındı diğer yıllar ise çıkartılarak indirgeme işlemi yapıldı.

Çift Anadal Programındaki (“ÇAP”) öğrencilerinin tüm bilgileri çalışmamıza katkı sağlamadığından dolayı veri setinden silinmiştir. Bu duruma örnek olan örnek veriler Çizelge 4.13’te gösterilmiştir.

Çizelge 4.13. Veri setindeki ÇAP programındaki öğrencilerin silinmesi

Adı	Kayıt Tarihi	Doğum Yılı	Program	ADI
Şe*****	2011	1989	ÇAP - Bilgisayar Mühendisliği (U.E.)	Veritabanı Yönetimi
Şe*****	2011	1989	ÇAP - Bilgisayar Mühendisliği (U.E.)	Mobil Cihazların Programlanması
Şe*****	2011	1989	ÇAP - Bilgisayar Mühendisliği (U.E.)	Yazılım Mühendisliği
Em*****	2011	1991	ÇAP - Bilgisayar Mühendisliği (U.E.)	Dosya Organizasyonu
Mu*****	2012	1992	ÇAP - Bilgisayar Mühendisliği (U.E.)	Algoritmalar
Mu*****	2012	1992	ÇAP - Bilgisayar Mühendisliği (U.E.)	İnternet Tabanlı Programlama
Mu*****	2012	1992	ÇAP - Bilgisayar Mühendisliği (U.E.)	Veri İletişim Sistemleri

4.4. ÖN İŞLEMLERDEN GEÇİRİLEN VERİLERİN ANALİZİ

Verilerin analizi ve sorgulanması için MsSQL programı kullanılmıştır. Alınan veri setindeki öğrenci bilgilerinde öğrenci numarasının KVKK kapsamında verilmemiş olması ve öğrenci Adının anonim hale getirilmiş olması toplam öğrenci sayısının bilinmemesine sebep olmaktadır. Bunun için öncelikle ID numarası verilerek bu durum giderilmiştir. Sonraki aşamada ise MsSQL ile sorgulama yapılarak toplam öğrenci sayısı program bazında bulunmuştur. Toplam öğrenci sayısının bulunması; program, ders, dönem, tekrar alan ve dersi geçme durum analizinin yapılmasını kolaylaştırmıştır. Şekil 4.1’de analiz için bölümlere göre toplam öğrenci sayısını bulan sorgulama işlemi gerçekleştirilmiştir.

```
Ogrenci_Sayilari.sq...d_bell\packard (52)) * X
SELECT COUNT(ID) AS 'Ogrenci_Toplam_Sayi'
FROM Ogrenci_Bolum
SELECT COUNT(*) AS Blg_Muh_Ing
FROM Ogrenci_Bolum
WHERE Program='Bilgisayar Mühendisliği (İngilizce)';
SELECT COUNT(*) AS Blg_Muh_Ing_2
FROM Ogrenci_Bolum
WHERE Program='Bilgisayar Mühendisliği (İngilizce) (II)';
SELECT COUNT(*) AS Blg_Muh_30Ing
FROM Ogrenci_Bolum
WHERE Program='Bilgisayar Mühendisliği (%30 İng.)';
SELECT COUNT(*) AS Blg_Muh_2_30Ing
FROM Ogrenci_Bolum
WHERE Program='Bilgisayar Mühendisliği (II) (%30 İng.)';
SELECT COUNT(*) AS Blg_Muh_UE
FROM Ogrenci_Bolum
WHERE Program='Bilgisayar Mühendisliği - (U.E)';
```

Şekil 4.1. Bölümlere göre toplam öğrenci sayısını veren kod

Çalışmamızda her bir dersi alan öğrenci için dersin başarı yüzdeleri hesaplanmıştır. Bunun için dersi alan öğrenci sayıları bulunduktan sonra durum bilgisi analizi yapılmıştır. Durum bilgisi sütununda “GEÇTİ” ve “KALDI” bilgisi yer almaktadır. Durum bilgisine her bir dersten geçen ve kalan öğrenci sayılarını bulmak için ihtiyaç duyulmaktadır. Bu sebeple ilk öncelikle Excel’de öğrencinin ara sınav, yarı yılsonu, bütünleme notları kullanılarak durum bilgisi belirlenmiştir. Durum bilgisinden sayıların belirlenmesi için MsSQL ‘de işlemler gerçekleştirilmiştir.

```
Gecen_Ogrenci_Bilgi..._bell\packard (52)) X
SELECT DISTINCT ID,Program FROM Ogr_Not_Sis WHERE Grup='53' and Durum='GEÇTİ'
ORDER BY ID
```

Şekil 4.2. Grup sayısı 53 olan derse ait durum bilgisi kodu

Öğrencilerin dersleri kaçınıcı kez aldıkları bilgisi MsSQL’den sorgulanmıştır. Her bir dersin tekrarlama sayısı Excel’de Tekrarlama isimli sütuna eklenmiştir. Buradaki sonuçlar ve Durum sütununda bulunan sonuçlar alınarak dersler, bölüm başarısı üzerine analiz yapılmıştır. Örnek olarak grubu “53” olan dersler için aşağıda MsSQL kodu verilmiştir.

```
Tekrarlama.sql - (...d_bell\packard (53)) X
SELECT ID,Grup,
COUNT(Distinct Donem_Yili) AS CNT
FROM Ogr_Not_Sis
Where Grup='53'
GROUP BY ID,Grup
HAVING COUNT(*) >1;
```

Şekil 4.3. Grup numarası “53”olan dersin tekrarlama kodu

Şekil 4.3’de ise kaçınıcı kez dersi aldığında geçtiği bilgisini veren kod gösterilmiştir.

```
Tekrarla_Gecen_Ogr...._bell\packard (52))* X
SELECT COUNT (ID) As Bir FROM Tekrarlama where Ders_Adi='Circuit Analysis' and Tekrar='1' and Durum='GEÇTİ'
SELECT COUNT (ID) As Iki FROM Tekrarlama where Ders_Adi='Circuit Analysis' and Tekrar='2' and Durum='GEÇTİ'
SELECT COUNT (ID) As Uc FROM Tekrarlama where Ders_Adi='Circuit Analysis' and Tekrar='3' and Durum='GEÇTİ'
SELECT COUNT (ID) As Dort FROM Tekrarlama where Ders_Adi='Circuit Analysis' and Tekrar='4' and Durum='GEÇTİ'
SELECT COUNT (ID) As Bes FROM Tekrarlama where Ders_Adi='Circuit Analysis' and Tekrar='5' and Durum='GEÇTİ'
SELECT COUNT (ID) As Alti FROM Tekrarlama where Ders_Adi='Circuit Analysis' and Tekrar='6' and Durum='GEÇTİ'
```

Şekil 4.4. Durum ve tekrar bilgisini veren kod

Yapılan bu işlemlere göre Excel’ de başarı analizleri hesaplanmıştır. Çıkan sonuçlar “Bulgular” kısmında anlatılmıştır. Bulgular bulunduktan sonra yorum ve öneriler “Sonuçlar ve Öneriler” kısmında anlatılmıştır.

BÖLÜM 5

SONUÇLAR

Veri seti üzerinde öncelikle temel çıkarımlarda bulunacak analiz işlemleri yapıldı. Bu çalışmalarda programlardaki başarı durumları, derslerin başarı oranları gibi bulgular elde edildi. Derslerin birbirlerine etkisi konusunda ise veri madenciliği yöntemlerinden yararlanıldı.

5.1.VERİ ANALİZİ SONUÇLARI

Bu çalışmada Karabük Üniversitesi Bilgisayar Mühendisliği Bölümüne ait olan veriler incelenmiş, öğrenci notlarının;

- Ders (Başarılı, Başarısız)
- Dersin grubu(Teknik, Seçmeli, Sosyal)
- Program (Uzaktan Eğitim,%30 İngilizce I. öğretim, %30 İngilizce II. öğretim, %100 İngilizce I. öğretim, %100 İngilizce II. öğretim)
- Dönem (Bahar, Güz)
- Yaş
- Birbirini etkileyen dersler (Ön koşullu dersler)

bilgileriyle ilişkili olup olmadığı araştırılıp analizi ve yorumlaması yapılmıştır.

Çalışmanın ilk safhasında veri setindeki bilgiler ön işlemlerden geçirilerek hazır hale getirilmiştir. Programa göre öğrenci sayıları bulunup dağılımı Çizelge 5.1’de gösterilmiştir.

Çizelge 5.1. Programa göre öğrenci dağılımı.

Program Adı	Öğrenci Sayısı
Bilgisayar Mühendisliği (U.E)	408
Bilgisayar Mühendisliği (%30 İng.)	391
Bilgisayar Mühendisliği (II) (%30 İng.)	367
Bilgisayar Mühendisliği (İngilizce)	351
Bilgisayar Mühendisliği (İngilizce) (II)	311
Genel Toplam	1828

Çizelge 5.1'e göre en fazla öğrenci sayısının Uzaktan Eğitim programı en az öğrenci sayısının %100 İngilizce ikinci öğretim programında olduğu görülmektedir.

Yapılan çalışmadaki veriler Kayıt Tarihi 2009-2014 öğrencileri için sınırlandırılmıştır. Bu verilerin kullanılmasının sebebi ise müfredat değişikliğinden dolayı harf kodu ve ders adı sütunundaki tutarsız verilerin olmasıdır. Karabük üniversitesinin verilerin alındığı dönemde uygulanan not sistematğine göre; Bir dersin başarılı sayılabilmesi için öğrencilerin A1, A2, B1, B2, C notlarından birinin olması gerekir. Bu çalışmada Programlar arası başarı karşılaştırması yapılabilmesi için her bir programdaki öğrencilerin her bir dersten aldıkları A1, A2, B1, B2 ve C notları için Excel'de ayrı bir "DURUM" sütunu oluşturularak derslerin geçme ve kalma bilgisi bulunmuştur. Aynı programa ait dersi alıp geçen toplam öğrenci sayıları bulunmuştur.

Programa göre alınan tüm derslerin yüzdeleri toplanıp karşılaştırma yapılmıştır. Başarı oranlarının dağılımı Çizelge 5.2'de görülmektedir.

Çizelge 5.2. Bölüm derslerine göre program bazında yüzdellik başarı oranı.

Program Adı	Öğrenci Sayısı	Genel Başarı (%)
Bilgisayar Mühendisliği - (U.E)	408	%73
Bilgisayar Mühendisliği (%30 İng.)	391	%85
Bilgisayar Mühendisliği (II) (%30 İng.)	367	%86
Bilgisayar Mühendisliği (İngilizce)	351	%86
Bilgisayar Mühendisliği (İngilizce) (II)	311	%85
Genel Toplam	1828	

Bu sonuçlar veri setindeki 2009-2014 kayıt tarihli ve dersi tekrarlama bilgileri baz alınarak yapılmıştır. Sonuçlara göre başarı oranlarına baktığımızda, örgün eğitim programlarının yaklaşık olarak başarı oranlarının aynı, en az başarılı olan program ise Bilgisayar Mühendisliği (U.E) olduğu görülmüştür.

Aynı şekilde programların her bir ders için başarı oranları bulunmuştur. Programların en başarısız ve en başarılı olduğu dersler seçilerek Çizelge 5.3 ve Çizelge 5.4'te gösterilmiştir.

Çizelge 5.3. Programlara göre en az başarılı dersler

Program Adı	En az başarılı dersler	Başarı oranı
Bilgisayar Mühendisliği - (U.E)	Oyun Programlama	%32
	Programlama Dilleri I	%46
	Programlama Dilleri II	%47
	Mantık Devreleri	%49
Bilgisayar Mühendisliği (%30 İng.)	Paralel Programlama Esasları	%46
	Oyun Programlama	%66
	Programlama Dilleri II	%70
	Data Communication Systems	%71
Bilgisayar Mühendisliği (II) (%30 İng.)	Paralel Programlama Esasları	%52
	Oyun Programlama	%56
	Programlama Dilleri II	%67
	Data Structures/DataCommunication Systems	%73
Bilgisayar Mühendisliği (İngilizce)	Principles of Parallel Programming	%54
	Game Programming	%58
	Data Communication Systems	%69
	Electronics	%70
Bilgisayar Mühendisliği (İngilizce) (II)	Principles of Parallel Programming	%53
	Programming Languages II	%63
	Data Communication Systems	%64
	Data Structures	%68

Çizelge 5.3'te çıkan sonuçlara göre; ortak olarak “Paralel programlama esasları”, “Programlama Dilleri II” ve “Oyun programlama” derslerinin dört farklı bölümde de en az başarılı olduğu görülmüştür.

Tüm programlarda en başarılı 5 derse ait bilgiler Çizelge 5.4’de verilmiştir. Bu çizelgede en başarılı olan derslerin “Biyoenformatiğe Giriş” ve “Etik” olduğu görülmüştür.

Çizelge 5.4’e göre; Uzaktan Eğitim programı hariç diğer 4 programda Atatürk İlke İnkılap Tarihleri 1, Türk Dili 2, Etik ve Biyoenformatiğe Giriş derslerinin başarılı olduğu görülmüştür.

Çizelge 5.4. Programlara göre en başarılı dersler

Program Adı	En Başarılı Dersler	Başarı Oranı
Bilgisayar Mühendisliği - (U.E)	Biyoenformatiğe Giriş	%99
	Etik	%97
	Veri Madenciliğine Giriş	%99
	Değerler Eğitimi	%96
	Yazılım Mühendisliği	%97
Bilgisayar Mühendisliği (%30 İng.)	Atatürk İlkeleri ve İnkılap Tarihi I-II	% 100
	Biyoenformatiğe Giriş	% 100
	Türk Dili II	% 100
	Ekonomiye Giriş	%99
	Etik	%99
Bilgisayar Mühendisliği (II) (%30 İng.)	Atatürk İlkeleri ve İnkılap Tarihi I-II	% 100
	Biyoenformatiğe Giriş	% 100
	Etik	% 100
	Türk Dili II	% 100
	Türk Dili I	%99
Bilgisayar Mühendisliği (İngilizce)	Atatürk s Principles and History of Revolutions II	%99
	Turkish Language I-II	%99
	Introduction to Bioinformatics	%99
	Ethics	%98
	Atatürk s Principles and History of Revolutions I	%98
Bilgisayar Mühendisliği (İngilizce) (II)	Turkish Language I	% 100
	Introduction to Bioinformatics	%99
	Atatürk s Principles and History of Revolutions I	%99
	Turkish Language II	%99
	Ethics	%99

Ders gruplarının bölüm bazında değerlendirildiđi sonuçlar izelge 5.5'te sunulmuştur. izelge 5.5'e göre; tüm programlarda Güz ve Bahar dönemine ait Sosyal grubuna ait derslerin daha başarılı olduđu görölmüştür.

Ders gruplarının dönem içindeki başarı oranlarına bakılarak izelge 5.6'daki sonuçlar oluşturulmuştur. Dönelere bakıldığında Bilgisayar Mühendisliđi (İngilizce) programı hariç diđer programlarda öğrenim gören öğrencilerin derslerdeki başarısının en yüksek olduđu dönemin “Bahar” dönemi olduđu sonucu ortaya çıkmıştır.

Çizelge 5.5. Ders gruplarına göre başarı oranları

Program Adı	Dönem	Ders Grup	Başarı Oranı
Bilgisayar Mühendisliği - (U.E)	Bahar Dönemi	Teknik	%71
		Ortak	%61
		Sosyal	%86
	Güz Dönemi	Teknik	%67
		Ortak	%63
		Sosyal	%96
Bilgisayar Mühendisliği (%30 İng.)	Bahar Dönemi	Teknik	%88
		Ortak	%91
		Sosyal	%99
	Güz Dönemi	Teknik	%79
		Ortak	%91
		Sosyal	%97
Bilgisayar Mühendisliği (II) (%30 İng.)	Bahar Dönemi	Teknik	%85
		Ortak	%91
		Sosyal	%98
	Güz Dönemi	Teknik	%83
		Ortak	%90
		Sosyal	%98
Bilgisayar Mühendisliği (İngilizce)	Bahar Dönemi	Teknik	%83
		Ortak	%90
		Sosyal	%99
	Güz Dönemi	Teknik	%87
		Ortak	%89
		Sosyal	%97
Bilgisayar Mühendisliği (İngilizce) (II)	Bahar Dönemi	Teknik	%80
		Ortak	%90
		Sosyal	%99
	Güz Dönemi	Teknik	%85
		Ortak	%85
		Sosyal	%98

Çizelge 5.6. Programlara göre öğrenci dönem başarı oranı

Program Adı	Dönem	Başarı Oranı
Bilgisayar Mühendisliği - (U.E)	Bahar Dönemi	%71
	Güz Dönemi	%67
	Ortalama Dönem	%69
Bilgisayar Mühendisliği (%30 İng.)	Bahar Dönemi	%91
	Güz Dönemi	%83
	Ortalama Dönem	%87
Bilgisayar Mühendisliği (II) (%30 İng.)	Bahar Dönemi	%89
	Güz Dönemi	%86
	Ortalama Dönem	%87
Bilgisayar Mühendisliği (İngilizce)	Bahar Dönemi	%88
	Güz Dönemi	%90
	Ortalama Dönem	%88
Bilgisayar Mühendisliği (İngilizce) (II)	Bahar Dönemi	%87
	Güz Dönemi	%87
	Ortalama Dönem	%87

Bilgisayar Mühendisliği bölümünde önemli sayılan bazı kritik dersler vardır. Sisteme göre bu dersler birbirinin devamı sayılmaktadır. Önceki dönemlerde alınan dersin geçilmemesi durumunda devamı niteliğinde olan sonraki dönemlerde açılan ders verilmemektedir. Bu derslere ön koşullu dersler denir. Ön koşullu derslerde ilk dersi geçen öğrenciler devamı olan ikinci dersi alabilmektedir. Çizelge 5.7’de birinci dersi ilk defa alıp geçen öğrencilerin sayısı ile aynı öğrencilerin ikinci dersi de ilk defada alıp geçme sayıları ve başarı yüzdeleri verilmiştir.

Sonuçlara göre Bilgisayar Mühendisliği (U.E.) programının ön koşullu derslerde diğer programlara göre daha başarısız olduğu görülmektedir. Devre Analizi ile Elektronik derslerindeki geçme durumu en düşük durumda olurken (%35 civarında) Matematik 1 ve 2 ile Mikroişlemciler ve Bilgisayar Mimarisi derslerindeki geçme durumu %72 civarında olmuştur.

Çizelge 5.7. Ön koşullu derslerin başarı oranları

Bölüm	Geçen Öğrenci Sayısı / Başarı Yüzdesi											
	Programlama Dil.1	Programlama Dil.2	Ortalama	Matematik 1	Matematik.2	Ortalama	Devre Analizi	Elektronik	Ortalama	Bilgisayar Mimarisi	Mikroişlemciler	Ortalama
Bilgisayar Mühendisliği (%30 İng.)	98	49	%50	172	127	%74	183	64	%35	167	129	%77
Bilgisayar Mühendisliği (II) (%30 İng.)	90	59	%66	171	100	%58	131	44	%34	159	109	%69
Bilgisayar Mühendisliği (İngilizce)	110	52	%47	197	143	%73	158	56	%35	157	113	%72
Bilgisayar Mühendisliği (İngilizce) (II)	69	43	%62	135	103	%76	126	48	%38	141	98	%70
Bilgisayar Mühendisliği (U.E)	136	55	%40	115	30	%26	78	25	%32	51	24	%47

Öğrencilerin yaşlarının ders başarılarına etkisi incelenmiştir. Bunun için Öğrencilerin “Kayıt Tarihi” ve “Doğum Tarihi” verileri baz alınarak yaşları hesaplanmıştır. Çizelge 5.8’ de “Kayıt Tarihi” baz alınarak yaş gruplarına göre öğrenci sayıları verilmiştir.

Çizelge 5.8. Yaş grupları

Bölüm	Yaş Grubu			
	17-22	23-27	28-32	33-47
	Toplam	Toplam	Toplam	Toplam
Bilgisayar Mühendisliği (%30 İng.)	360	19	10	1
Bilgisayar Mühendisliği (II) (%30 İng.)	335	19	11	3
Bilgisayar Mühendisliği (İngilizce)	345	7	0	0
Bilgisayar Mühendisliği (İngilizce) (II)	302	9	0	0
Bilgisayar Mühendisliği (U.E)	210	94	58	45

Sonuçlara göre tüm programlarda en çok öğrenci 17-22 yaş grubunda bulunmaktadır. Uzaktan Eğitim programında 33-47 yaş grubunda diğer programlara göre daha fazla öğrenci olduğu görülmektedir. %100 İngilizce programlarında ise 28-47 yaş aralığında öğrencilerin olmadığı görülmektedir.

Programlarda en fazla öğrenci sayısının olduğu yaşların 18 ve 19 olduğu görülmüştür. Çizelge 5.9’da gösterilmiştir.

Çizelge 5.9. En fazla öğrencisi bulunan yaş grubu

Bölüm	Yaş	
	18	19
	Toplam	Toplam
Bilgisayar Mühendisliği (%30 İng.)	184	114
Bilgisayar Mühendisliği (II) (%30 İng.)	147	134
Bilgisayar Mühendisliği (İngilizce)	159	105
Bilgisayar Mühendisliği (İngilizce) (II)	123	121
Bilgisayar Mühendisliği (U.E)	58	58

5.2. VERİ MADENCİLİĞİ İLE ELDE EDİLEN SONUÇLAR

Veri madenciliği uygulaması veri ön işleme sonucu elde edilen veri setine uygulanmıştır. Bu uygulamalarla derslerin başarı durumunun diğer derslerle ilişkilerini ortaya koymak amaçlanmıştır. Çalışmada farklı yaklaşımlarla analiz yapılmıştır. Bu çerçevede derslerin yılına bakılmaksızın değerlendirme yapıldığı gibi önceki dönemlerde alınan dersler ile sonraki dönemlerde alınan dersler arasındaki ilişkiye de bakıldı. 4. Sınıf derslerinin seçmeli olması sebebiyle değerlendirme dışında tutulmuştur.

5.2.1. Dersi İlk Alışta Geçme-Kalma Durumları

5.2.1.1. Veri Kümesi

Veri setinde en çok ortak alınan 19 ders belirlenmiş ve bu derslerin hepsini alan 706 öğrencinin dersleri ilk alışta geçme/kalma durumları üzerinde veri analizleri gerçekleştirilmiştir.

Dersler:

1. BLM114_Olasilik_ve_Istatistik
2. BLM111_Programlama_Dilleri_I
3. BLM112_Programlama_Dilleri_II
4. BLM222_Algoritmalar
5. BLM209_Ayrik_Matematik
6. BLM223_Devre_Analizi
7. BLM224_Elektronik
8. BLM228_Internet_Tabanli_Programlama
9. BLM221_Mantik_Devreleri
10. BLM225_Nesneye_Dayali_Programlama
11. BLM227_Veri_Yapilari
12. BLM226_Veritabani_Yonetimi
13. BLM113_Bilgisayar_Muhendisligine_Giris
14. BLM321_Mikroislemciler
15. BLM323_Sayisal_Analiz
16. BLM325_Veri_İletisim_Sistemleri
17. BLM322_Otomata_Teorisi
18. BLM320_Bilgisayar_Mimarisi
19. BLM324_Isletim_Sistemleri

5.2.1.2. Birliktelik Kuralları ile Elde Edilen Sonuçlar

Belirlenen 19 dersin aralarındaki ilişkilere yönelik birliktelik kuralları ile bakılarak en yüksek başarı seviyesi elde edilen durumlar Çizelge 5.10 da verilmiştir. Bu çizelgede dersten kalanların ve geçenlerin tahmini hedeflenmiştir.

Çizelge 5.10. Birlikte lık kuralı ile dersi ilk alıřta geme-kalma tahmini (bütün sınıflar beraber)

Tahmin Yöntemi	Tahmin Edenler	Öğrenci Sayısı	Tahmin Edilen	Öğrenci Sayısı	Conf	Lift	Lev	Conv
Kalan	BLM111_Programlama_Dilleri_I	332	BLM112_Programlama_Dilleri_II	314	0,95	1,26	0,09 [63]	4,31
	BLM225_Nesneye_Dayali_Programlama							
Kalan	BLM111_Programlama_Dilleri_I	473	BLM112_Programlama_Dilleri_II	435	0,92	1,22	0,11	2,99
Geen	BLM114_Olasilik_ve_Istatistik	315	BLM113_Bilgisayar_Muhendisligi ne_Giris	288	0,91	1,15	0,05 [38]	2,33
	BLM323_Sayisal_Analiz							
	BLM320_Bilgisayar_Mimarisi							
Geen	BLM222_Algoritmalar	315	BLM323_Sayisal_Analiz	286	0,91	1,21	0,07 [50]	2,65
	BLM324_Isletim_Sistemleri							
Geen	BLM223_Devre_Analizi	376	BLM224_Elektronik	341	0,91	1,15	0,06 [43]	2,17

Programlama Dilleri I ve Nesneye Dayalı Programlama derslerinden 332 öğrenci kalmıřtır. Bu kalan öğrencilerden 314’ü de Programlama Dilleri II dersinden kalmıřtır Böylelikle güven seviyesi (confidence) 0,95 çıkmıřtır. Programlama Dilleri I 473 öğrenci kalmıř. Bu kalan öğrencilerden 435’i de Programlama Dilleri II dersinden kalmıř Böylelikle güven seviyesi (confidence) 0,92 çıkmıřtır. Olasılık ve İstatistik, Sayısal Analiz ve Bilgisayar Mimarisi derslerini geen öğrencilerin %91’i Bilgisayar Mühendisliğine Giriř dersini gemiř. Algoritmalar ve İşletim Sistemleri derslerini geen öğrencilerin %91’i Sayısal Analiz dersini gemiřtir. Devre Analizi dersinden kalan 376 öğrencinin %91’i Elektronik dersinden de kalmıřtır.

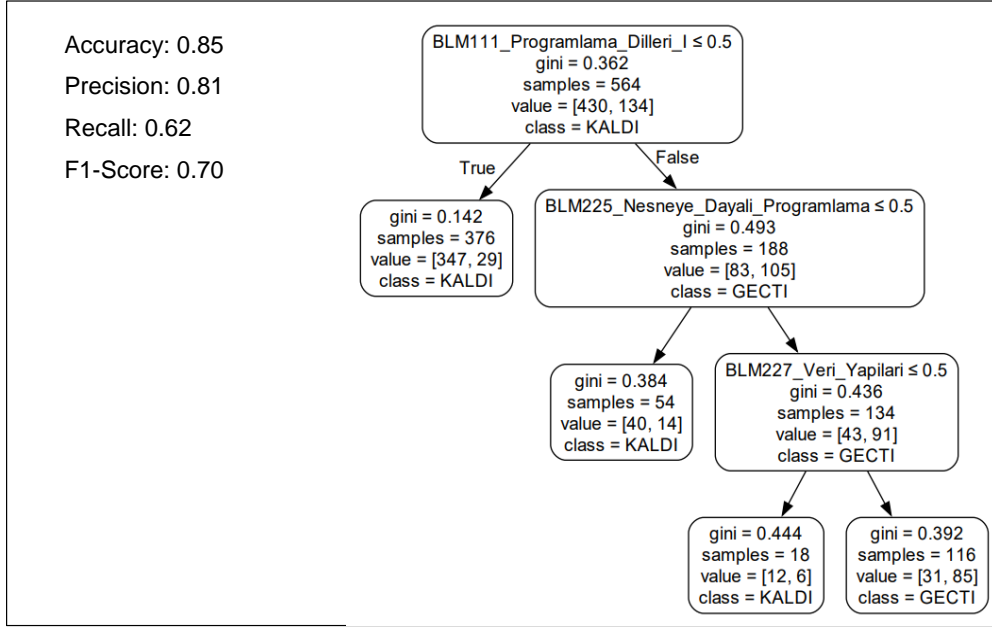
5.2.1.3. Karar Ağacı İle Elde Edilen Bulgular

Öğrencilerin en çok ortak aldıkları 19 ders ilk 3 sınıfı kapsamaktadır. Karar ağaçlarından başarı oranı en yüksek ağacı bulabilmek için tüm sınıfları karıřık ve seilen dersin sınıfının alt sınıflara göre bakarak birok varyasyonda test edilmiřtir. Bu dersler ierisinde en yüksek başarı oranına sahip ağaç seilip, sadece o ağaçla

ilgili detaylar yer verilmiştir. Elde edilen sonuçlar Çizelge 6.1 ve Çizelge 6.3’de verilmiştir.

Çizelge 5.11. Karar Ağacı ile dersi ilk alıştırma geçme-kalma tahmini (bütün sınıflar beraber)

Ders	Accuracy	Depth
BLM112_Programlama_Dilleri_II	0.85	3
BLM111_Programlama_Dilleri_I	0.83	2
BLM113_Bilgisayar_Muhendisligine_Giris	0.82	3
BLM323_Sayisal_Analiz	0.81	3
BLM224_Elektronik	0.77	3
BLM325_Veri_İletisim_Sistemleri	0.77	3
BLM114_Olasilik_ve_Istatistik	0.75	2
BLM223_Devre_Analizi	0.72	3
BLM225_Nesneye_Dayali_Programlama	0.71	2
BLM227_Veri_Yapilari	0.71	2
BLM222_Algoritmalar	0.70	3
BLM324_Isletim_Sistemleri	0.70	3
BLM221_Mantik_Devreleri	0.68	3
BLM320_Bilgisayar_Mimarisi	0.68	2
BLM322_Otomata_Teorisi	0.68	2
BLM226_Veritabani_Yonetimi	0.66	2
BLM228_Internet_Tabanli_Programlama	0.60	3
BLM321_Mikroislemciler	0.58	2
BLM209_Ayrik_Matematik	0.55	2

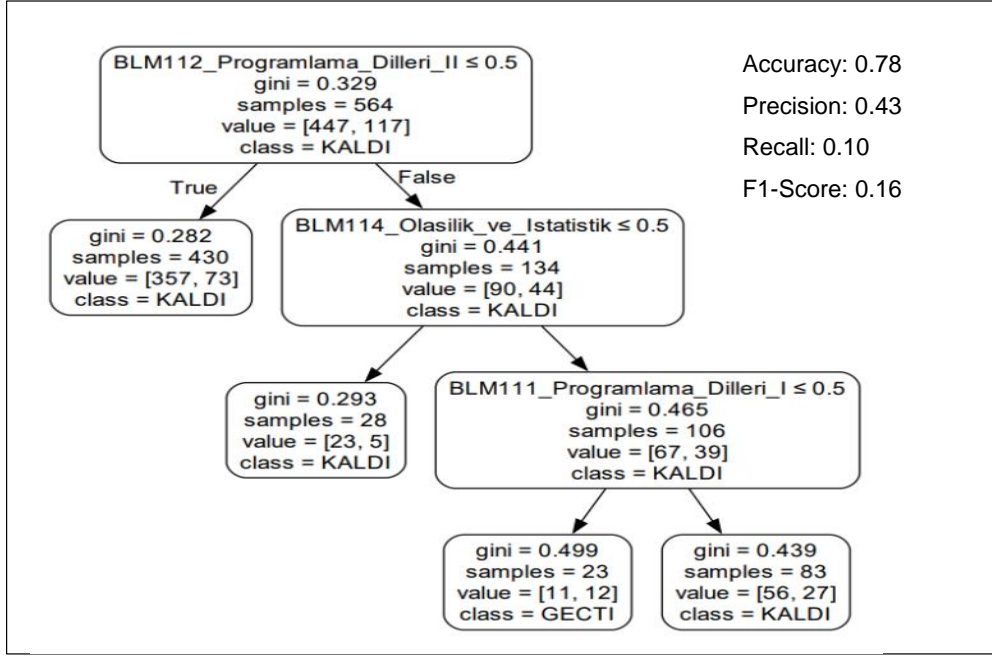


Şekil 5.1. (BLM112) Programlama Dilleri II dersi karar ağacı

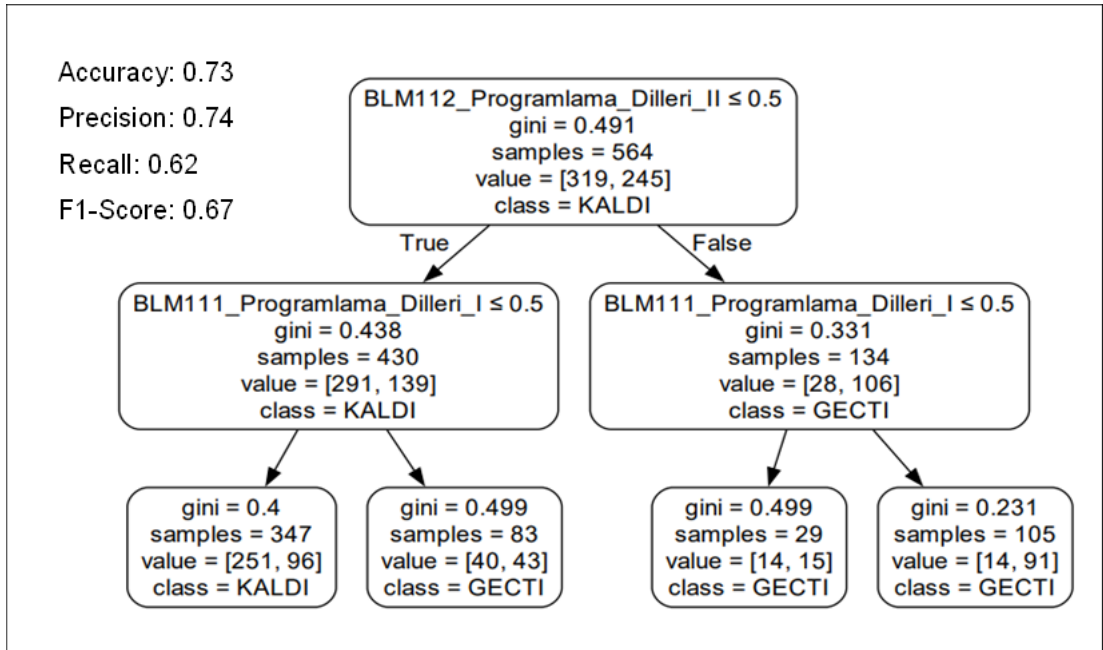
Bu uygulamada sınıf farkı gözeticmeden ilişki tahmini öngörülmesi yapılmıştır. Buna göre en yüksek tahmin başarısı %85 ile Programlama Dilleri-II dersinde sağlanmıştır. Bu dersin başarı durumu Programlama Dilleri-I, Nesneye Dayalı Programlama ve Veri Yapıları ile ilişkili olduğu gözükmemektedir.

Çizelge 5.12. Karar ağacı şile 1.sınıf derslerine göre 2.sınıf derslerini sınıflandırma

Ders	Accuracy	Depth
BLM224_Elektronik	0.78	3
BLM225_Nesneye_Dayali_Programlama	0.73	2
BLM227_Veri_Yapilari	0.73	2
BLM223_Devre_Analizi	0.64	3
BLM209_Ayrik_Matematik	0.63	3
BLM222_Algoritmalar	0.63	3
BLM228_Internet_Tabanli_Programlama	0.63	3
BLM226_Veritabani_Yonetimi	0.59	2
BLM221_Mantik_Devreleri	0.54	3



Şekil 5.2. (BLM224) Elektronik dersi karar ağacı



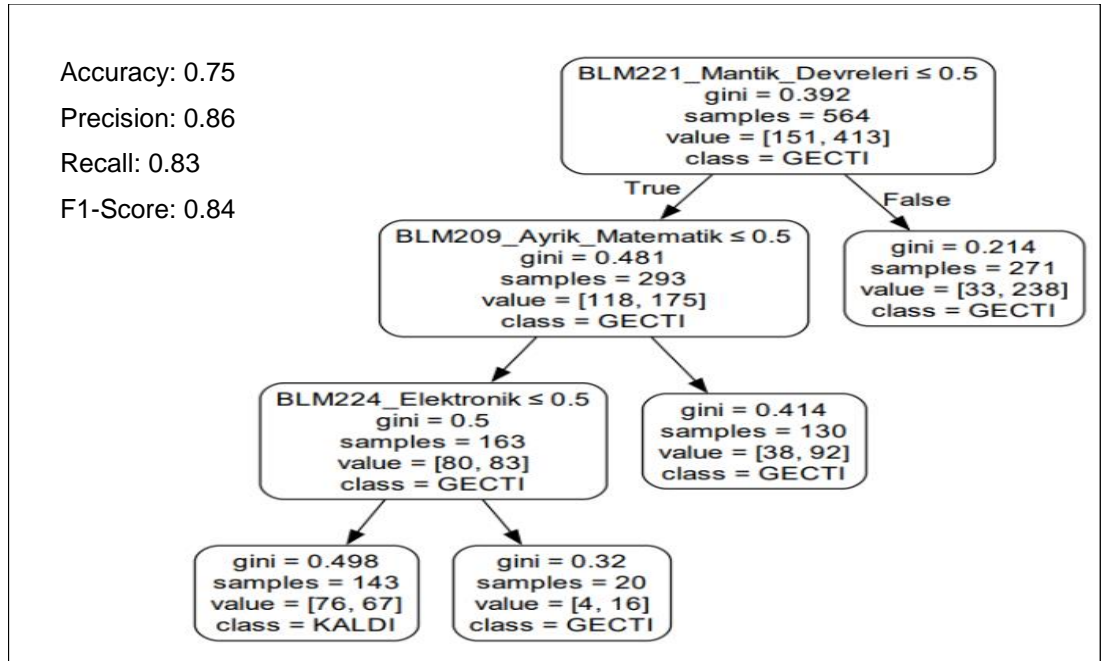
Şekil 5.3. (BLM225) Nesneye Dayalı Programlama dersi karar ağacı

2. Sınıf derslerinin 1. Sınıf derslerine bağlı olarak tahmininde en yüksek başarı Elektronik dersinde gerçekleşmiştir. Bu dersin başarısına etki etmiş olarak gözükten Programlama Dilleri I ve II ile Olasılık dersinin etkisini yorumlamak kolay gözükmemektedir. Ancak 2. Sırada %73 başarı ile yer alan Nesneye Dayalı

Programlama dersinin sadece Programlama Dilleri I ve II ile ilişkili olması tutarlı gözükmemektedir.

Çizelge 5.13. Karar ağacı ile 2.sınıf derslerine göre 3.sınıf derslerini sınıflandırma

Ders	Accuracy	Depth
BLM323_Sayısal_Analiz	0.75	3
BLM325_Veri_İletişim_Sistemleri	0.72	3
BLM324_Isletim_Sistemleri	0.71	3
BLM320_Bilgisayar_Mimarisi	0.68	2
BLM322_Otomata_Teorisi	0.64	3
BLM321_Mikroislemciler	0.54	3



Şekil 5.4. (BLM323) Sayısal Analiz dersi karar ağacı

3. Sınıf derslerinin başarısının 2. Sınıf derslerine bağlı tanımlanmasında en yüksek değer %75 ile Sayısal Analiz dersinde karşılaşılmıştır. Bu ders Mantık Devreleri, Ayrık Matematik ve Elektronik dersleri başarıları ile ilişkilendirilmiştir. Bu sonuçta kısmi ilişki tutarlı gözükmemekte ancak Elektronik dersi ile ilgili ilişki çok anlamlı gözükmemektedir.

5.2.2. Dersi İkinci Alışta Geçme-Kalma Durumları

5.2.2.1. Veri Kümesi

İlk alışta geçemeyip ikincide en çok ortak alınan 10 ders belirlenmiş ve bu derslerin hepsini alan 124 öğrencinin dersleri ikinci alışta geçme/kalma durumları üzerinde veri analizleri gerçekleştirilmiştir.

Dersler:

- 1) BLM111_Programlama_Dilleri_I
- 2) BLM112_Programlama_Dilleri_II
- 3) BLM222_Algoritmalar
- 4) BLM209_Ayrik_Matematik
- 5) BLM223_Devre_Analizi
- 6) BLM224_Elektronik
- 7) BLM221_Mantik_Devreleri
- 8) BLM225_Nesneye_Dayali_Programlama
- 9) BLM227_Veri_Yapilari
- 10) BLM226_Veritabani_Yonetimi

5.2.2.2. Birliktelik Kuralları İle Elde Edilen Sonuçlar

Dersleri ilk alışta geçemeyen öğrenciler üzerinde yapılan birliktelik analizi sonuçları Çizelge 5.14'de verilmiştir. Buna göre Veri Yapıları dersinden kalanların tamamı Programlama Dilleri II dersinden de kalmıştır. Programlama Dilleri I ve Elektronik derslerinden kalan öğrencilerin 99'u Programlama Dilleri II dersinden de kalmıştır. Programlama Dilleri I dersinden kalan öğrencilerin 98'i Programlama Dilleri II dersinden de kalmıştır. Elektronik dersinden kalan öğrencilerin 97'si Programlama Dilleri II dersinden de kalmıştır. Veritabanı Yönetimi dersinden kalan öğrencilerin 98'i Programlama Dilleri II dersinden de kalmıştır.

Çizelge 5.14. Birliktelik kuralı ile ilk alıřta geemeyen ğrencilerin derslerinin tahmini

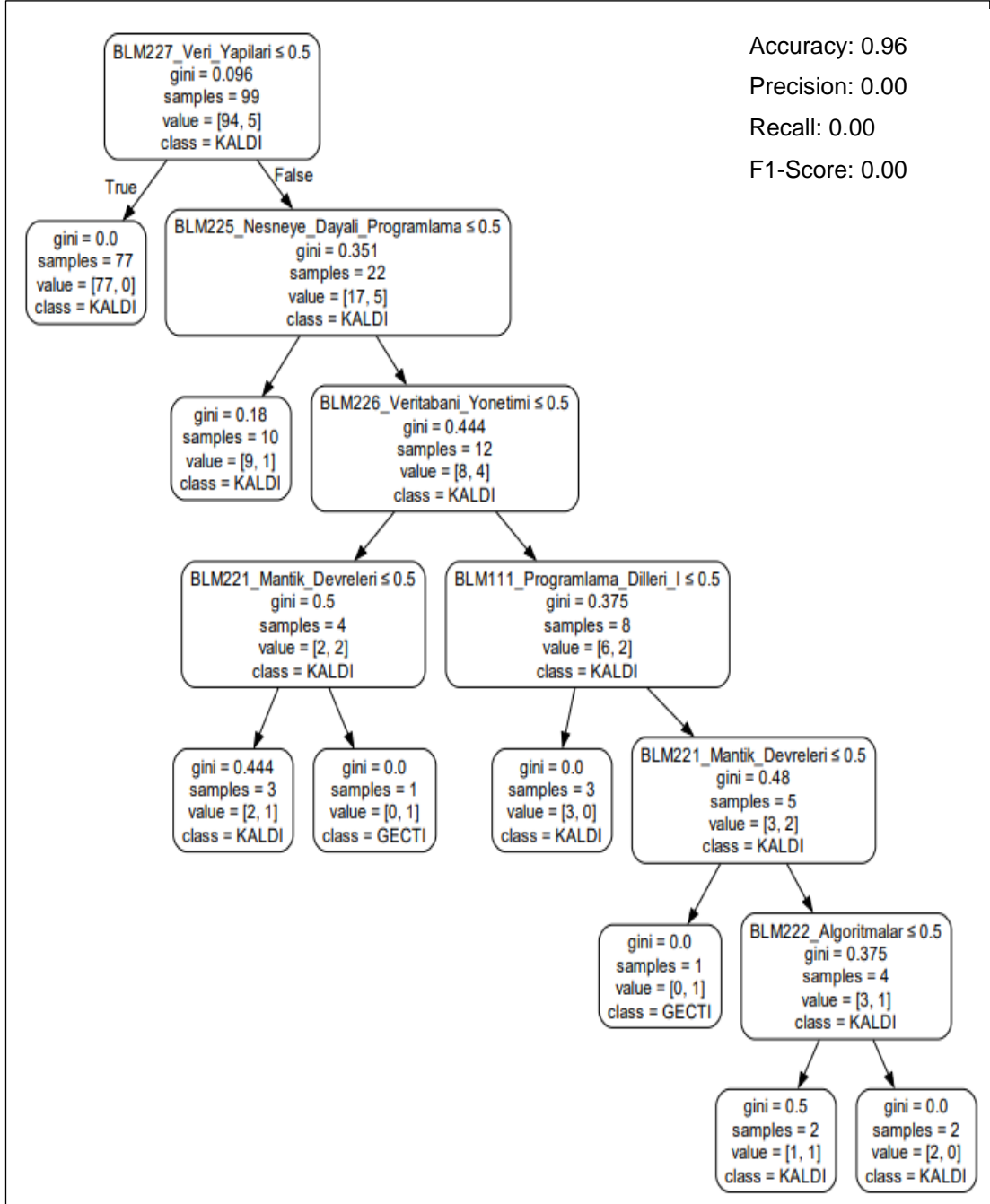
Tahmin Yöntemi	Tahmin Edener	Öğrenci Sayısı	Tahmin Edilen	Öğrenci Sayısı	Conf	Lift	Lev	Conv
Kalan	BLM227_Veri_Yapilari	96	BLM112_Programlama_Dilleri_II	96	1	1,04	0,03 [3]	3,87
Kalan	BLM111_Programlama_Dilleri_I	95	BLM112_Programlama_Dilleri_II	94	0,99	1,03	0,02 [2]	1,92
	BLM224_Elektronik							
Kalan	BLM111_Programlama_Dilleri_I	107	BLM112_Programlama_Dilleri_II	105	0,98	1,02	0,02 [2]	1,44
Kalan	BLM224_Elektronik	108	BLM112_Programlama_Dilleri_II	105	0,97	1,01	0,01 [1]	1,09
Kalan	BLM226_Veritabani_Yonetimi	88	BLM112_Programlama_Dilleri_II	86	0,98	1,02	0,01 [1]	1,18

5.2.2.3. Karar Ağacı ile Elde Edilen Sonuçlar

Dersleri ilk alıřta geemeyen ğrencilerin dersleri arasındaki iliřkiye yönelik Karar ağacı ile yapılan analiz sonuçları Çizelge 5.15’te verilmiřtir.

Çizelge 5.15. Karar ağacı ile ilk alıřta dersi geemeyen ğrencilerin ders tahmini (bütün derslere göre)

Ders	Accuracy	Depth
BLM112_Programlama_Dilleri_II	0.96	6
BLM111_Programlama_Dilleri_I	0.88	4
BLM224_Elektronik	0.84	5
BLM225_Nesneye_Dayali_Programlama	0.80	4
BLM227_Veri_Yapilari	0.76	5
BLM223_Devre_Analizi	0.72	4
BLM221_Mantik_Devreleri	0.68	5
BLM222_Algoritmalar	0.56	4
BLM209_Ayrik_Matematik	0.52	5
BLM226_Veritabani_Yonetimi	0.52	4

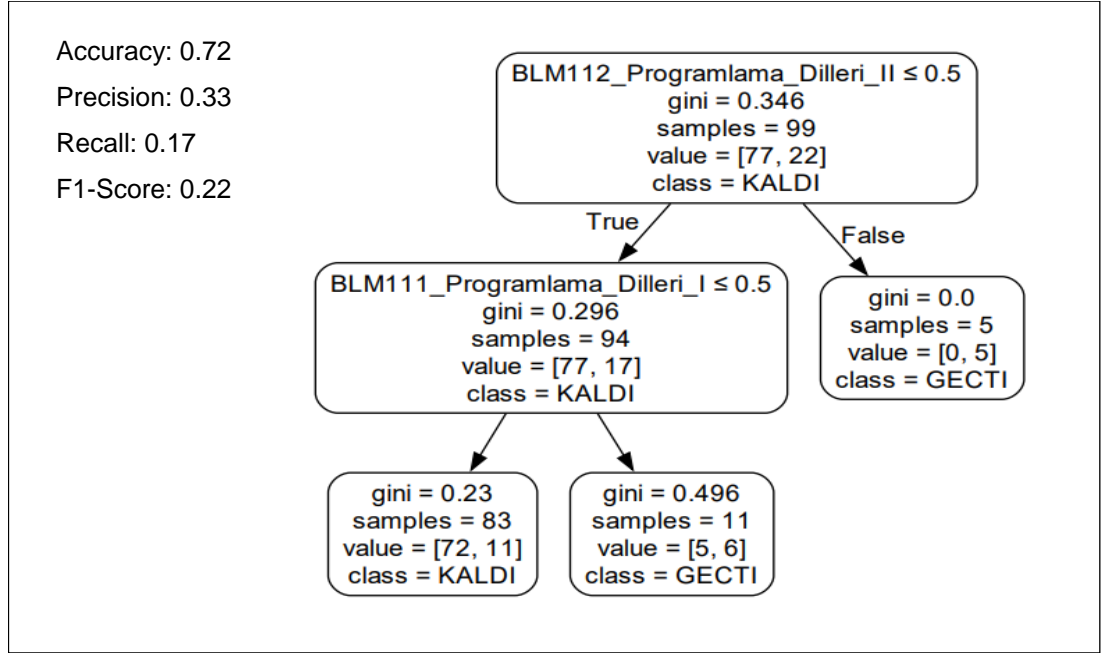


Şekil 5.4. (BLM112) Programlama Dilleri II dersi karar ağacı

Sınıf farkı bakılmaksızın yapılan karar ağacı ilişkisinde en yüksek başarı %96 ile Programlama Dilleri II’de elde edildi.

Çizelge 5.16. Karar ağacı ile ilk alıştırma dersini geçemeyen öğrencilerin ders tahmini (1.sınıf derslerine göre 2.sınıf dersleri)

Ders	Accuracy	Depth
BLM224_Elektronik	0.88	2
BLM209_Ayrik_Matematik	0.72	2
BLM225_Nesneye_Dayali_Programlama	0.72	2
BLM227_Veri_Yapilari	0.72	2
BLM221_Mantik_Devreleri	0.68	2
BLM226_Veritabani_Yonetimi	0.68	2
BLM223_Devre_Analizi	0.64	2
BLM222_Algoritmalar	0.60	2



Şekil 5.5. (BLM227) Veri Yapıları karar ağacı

2. Sınıf derslerinin sınıflandırılmasında 1. Sınıf derslerinden faydalandığında en yüksek başarı %88 ile elde edildi. Birbiriyle daha ilişkili gördüğümüz Veri Yapılarının karar ağacında tahminde Programlama Dilleri I ve II yer almaktadır. Bu dersin başarı yüzdesi ise %72 olarak gerçekleşmiştir.

BÖLÜM 6

DEĞERLENDİRMELER VE ÖNERİLER

Bu bölümde yapılan araştırmada çıkan sonuçlar ve öneriler sunulmuştur.

6.1. DEĞERLENDİRMELER

Bu çalışmada üniversitede aynı bölüm farklı programlardaki öğrencilerin derslerdeki başarıları, farklı açılardan değerlendirilerek örnek olarak bir çalışma yapılmış ve incelenmiştir. Literatür taraması sonucunda program bazında tüm derslerin birlikte alındığı kapsamlı bir çalışma olmadığı görülmüştür. Bu çalışma da Bilgisayar mühendisliği bölümünün 5 farklı programında öğrenim gören öğrencilerin yaş, ders, dönem ve programları gibi özellikler arasındaki başarı ilişkisi incelenmiştir. Sonuçlar için 2009-2014 kayıt tarihleri arasında dersi alan tüm öğrenciler baz alınmıştır.

Yapılan çalışmada aşağıdaki sonuçlar bulunmuş ve uygun öneriler sunulmuştur:

- 1) Kayıt tarihi 2009-2014 yılları için öğrenci sayılarına bakıldığında ders alan öğrenci sayısının en fazla Uzaktan Eğitim programında olduğu görülmektedir.
- 2) Programlar arasında genel ders başarı oranlarına bakıldığında en az başarılı öğrencilerin Uzaktan Eğitim programı en başarılı programların ise örgün eğitim programları olduğu görülmektedir. Öğrencilerin yüz yüze eğitim almasının başarıyı etkilediği söylenebilir.
- 3) Programlardaki derslere bakıldığında “Oyun Programlama”, “Paralel Programlama Esasları”, “Programlama Dilleri II” derslerinin öğrencilerin en az başarılı olduğu dersler arasında olduğu görülmektedir. Bunun sebepleri

olarak dersin seçmeli olup tekrar açılmaması, dersin zorluğu, sınavların zorluk derecesi, ilgili ders hocalarının etkisi gibi nedenler gösterilebilir.

- 4) Öğrencilerin ders başarılarının dönemlerine göre bakıldığında “Bahar” döneminin genel olarak en başarılı dönem olduğu görülmektedir. Sadece 100 İngilizce programlarının “Güz” döneminde farklılık görülmektedir. Müfredattaki derslerin zorluk kolaylık derecesi, Hava şartları ve ilgili ders hocalarının etkisi gibi sebeplerden dolayı dönem bazlı başarı oranının etkilenildiği düşünülmektedir.
- 5) Tüm programdaki öğrencilerin ders gruplarına bakıldığında “Sosyal” ders grubu içerisindeki derslerin iki dönemde de başarı oranlarının yüksek olduğu görülmüştür. Veri setindeki bilgilere bakıldığında bu derslerin kredisi az olan ve yoruma dayalı dersler olduğu düşünülmektedir. Ayrıyeten tablodaki bilgilere bakıldığında Uzaktan Eğitim programı hariç diğer programların “Ortak” ders grubundaki ders başarı oranının yüksek olduğu görülmektedir. Dersler genel anlamda incelendiğinde Uzaktan eğitim programında teknik derslerin ortak derslere göre daha başarılı olduğu görülmektedir.
- 6) Ön koşullu derslerin tüm programlara ve derslere göre ilk alıfta geçen öğrencilerin başarı oranlarına bakıldığında Devre Analizi dersini geçenlerin Elektronik dersinden aynı performansı göstermeyip başarı oranının 50 nin altında kaldığı görülmüştür.
- 7) Programlara göre bakıldığında sonuç olarak Bilgisayar Mühendisliği (U.E)’nin ön koşullu derslerde başarı oranının 50 nin altığında olduğu görülmüştür. İlk alıfta ilk dersten geçen öğrencilerin sayısının ikinci derste aynı başarıya sahip olmadığı görülmüştür. Bu durumda bu bölüm için birinci dersi geçen öğrencilerin çoğunun devamı olan ikinci dersi geçemediği görülmüştür.
- 8) Tüm programlarda en fazla öğrenci sayısının 17-22 yaş aralığında olduğu görülmektedir. Kayıt tarihine göre tüm yaşlar baz alındığında 18 ve 19 yaşında olan öğrencilerin sayısı diğer öğrencilerin sayısından daha çok olduğu tespit edilmiştir. 33-47 yaş aralığında ise en fazla Uzaktan Eğitim

programının tercih edildiği görülmüştür. Bu bilgilere göre 33-47 yaş aralığındaki öğrencilerin Uzaktan Eğitim programını tercih etme sebepleri arasında öğrencilerin işte çalışıp çalışmama durumu, ilk üniversiteleri olmadığı savunulabilir.

- 9) Apriori algoritması 1. Veri kümesi olan 19 ortak ders dikkate alınarak uygulandığında dersleri diğer derslere bağlı olarak çok yüksek yüzdeyle tahmin edebildiği görülmüştür. Örneğin Programlama Dilleri II %95, Elektronik dersi %91 ile tahmin edilebilmiştir. Aynı algoritmayı 2. Veri kümesini oluşturan dersin ikinci alıştırma geçme ve kalma uyguladığımızda ise daha az olan bu veri setinde Programlama Dilleri II dersi %100 doğrulukla tahmin edilebilmiştir.
- 10) Karar ağacı algoritması 2 farklı veri setine uygulanarak diğer derslere bağlı olarak tahmin edilebilmesi incelenmiştir. Burada çıkan ilişkilerin dersler arasındaki etkiyi analiz edebilmek boyutundan da değerlendirilmiştir. Birliktelik analizine göre daha düşük bir doğruluk yüzdesi elde edilmiştir. Bu başarı sınıf farkı gözetilmeden veri setini kullanılması durumunda daha yüksek değerler oluşturmaktadır.
- 11) Elde edilen bütün sonuçlar değerlendirildiğinde veri setinin getirmiş olduğu kısıtlar bulunmasından dolayı daha somut çıkarımlar elde edilmesinde zorlukla karşılaşılmıştır. Özellikle veri setinin kapsadığı dönemde kayıtlı olduğu halde ders almayan öğrenciler bulunabilmektedir.

6.2. ÖNERİLER

Bu çalışmada Bilgisayar Bölümündeki öğrencilere ait veriler üzerinde çalışmalar yürütülmüş olup anlamlı sonuçlar elde edilmeye çalışılmıştır. Benzer çalışmalar farklı bölümlerde de gerçekleştirilebilir. Her bölümün kendi ders çeşidi dikkate alınarak farklı değerlendirmeler yapılabilir. Çalışmada kullandığımız veri seti belli dönemi kapsayan verileri içermektedir. Veri setinin belirli aralıkta kayıt yapıp mezun olana ya da okuldan ilişkisi kesilene kadarki aldığı tüm dersleri kapsar şekilde olması

farklı deęerlendirmelerin yapılmasına fırsat sağlayabilir. KVKK kapsamında bazı verilerin elde edilememiş olması deęerlendirmelerin kısıtlı kalmasına sebep olmuştur. Bu çerçevede Dersleri veren hocaların başarıya etkisi deęerlendirilememiştir. Aynı şekilde coęrafik bölgelerin ve mezun olunan okul tipinin başarıya etkisi de gerekli veriler olmadığı için gerçekleştirilememiştir. Bu verilerin anonim olarak sağlanması durumunda bu tür deęerlendirmeler de yapılabilecektir. Kullandığımız veri setinin kapsadığı dönemde ders müfredatında deęişiklikler olmuştur. Daha stabil döneme denk gelen bir veri setinin deęerlendirilmesinde fayda görmekteyiz.

Bu tespitleri dikkate alınarak gerçekleştirilecek çalışmalarla elde edilecek daha detaylı sonuçlar gerek öğrencilerin gerekse ders veren hocaların durumlarının tespit edilmesine fırsat tanıyacaktır. Bu tespitler çerçevesinde de ders başarısını arttırmaya yönelik önlem ve çözümlerin elde edilmesi sağlanabilir.

KAYNAKLAR

Abbasođlu, B., “Ortaokul öđrencilerinin akademik başarılarının eğitsel veri madenciliđi yöntemleri ile tahmini”, *Veri Bilim Dergisi*, 3(1): 1-10 (2020).

Afacan-Adanır, G., “E-sınav sistemi ile ilgili öđrenci görüşleri ve yaşanan sorunların araştırılması”, *Organizasyon ve Yönetim Bilimleri Dergisi*, 13(2): 189-207 (2021).

Akçay, H., Özyurt, B. B., Bezir-Akçay, B., “Çoklu yazma etkinliklerinin fen ve teknoloji dersi öğretiminde kullanılmasının öđrenci başarısı ve kavram öğrenmeye etkisi. *Bayburt Üniversitesi Eğitim Fakültesi Dergisi*, IX(II): 15-31 (2014).

Akgün, E., Maral-Karanfil, Ö., “Uzaktan eğitimde veri madenciliđi yöntemi kullanılarak yapılmış arařtırmalarda öğrenme çıktıları üzerine sistematik bir inceleme”, *Bilgi ve İletişim Teknolojileri Dergisi*, 4(2): 197-226 (2022).

Akgün, K., Bulut-Özek, M., “Eđitsel veri madenciliđi yöntemi ile ilgili yapılmış çalışmaların incelenmesi: içerik analizi”, *Uluslararası Eğitim Bilim ve Teknoloji Dergisi*, 6(3): 197-213 (2020).

Akın, E., Kudak, H., “2023 eğitim vizyon belgesiyle oluşturulan drama ve eleştirel düşünce atölyelerinin Türkçe dersi akademik başarısı üzerine etkisi”, *Avrasya Sosyal ve Ekonomi Arařtırmaları Dergisi*, 7(6): 58-74 (2020).

Aksu, G., Dođan, N., “Veri madenciliđinde kullanılan öğrenme yöntemlerinin farklı koşullar altında karşılaştırılması”, *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 51(3): 71-100 (2018).

Alan, M. A., “Veri madenciliđi ve lisansüstü öđrenci verileri üzerine bir uygulama”, *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (33): 165-174 (2012).

Algarni, A. “Data mining in education”, *International Journal of Advanced Computer Science and Applications*, 7(6): 456-461 (2016).

Altun, M., Kayıkçı, K., Irmak, S., “Sınıf öğretmenliđi öđrencilerinin mezuniyet notlarının regresyon analizi ve yapay sinir ađları yöntemleriyle tahmini”, *E-Uluslararası Eğitim Arařtırmaları Dergisi*, 10(3): 29-43 (2019).

Alyahyan, E. ve Düşteğör, D., “Predicting academic success in higher education: Literature review and best practices”, *International Journal of Educational Technology in Higher Education*, 17(3): 1-21 (2020).

Arcinas, M. M., Sajja, G. S., Asif, S., Gour, S., Okoronkwo, E., Naved, M., “Role of data mining in education for improving students performance for social change”, *Turkish Journal of Physiotherapy and Rehabilitation*, 32(3): 6519-6526.

Aslanargun, E., “Okul-aile işbirliği ve öğrenci başarısı üzerine bir tarama çalışma”, *Sosyal Bilimler Dergisi*, (18): 119-135 (2007).

Aydemir, E., Kaysi, F., Gülseçen, S., “Üniversite öğrencilerinin Türk dili dersi sınav sonuçlarının sınava hazırlık düzeylerine göre tahminlenmesi”, *The Journal of Operations Research, Statistics, Econometrics and Management Information Systems*, 7(2): 351-356 (2019).

Aydın, S., Takkaç, M., “İngilizceyi ikinci dil olarak öğrenenlerde sınav kaygısının cinsiyet ile ilişkisi”, *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(1): 259-266 (2007).

Ayık, Y. Z., Özdemir, A., Yavuz, U., “Lise türü ve lise mezuniyet başarısının, kazanılan fakülte ile ilişkisinin veri madenciliği tekniği ile analizi”, *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(2): 441-454 (2007).

Baltacı, A., “Veri madenciliği: İlahiyat fakültesi öğrencilerinde akademik başarının veri madenciliği ile ölçülmesi”, *Din ve Bilim-Muş Alparlan Üniversitesi İslami İlimler Fakültesi Dergisi*, 1(1): 1-23 (2018).

Barczak, T., Jain, H., White, K., “Mining high school data to predict and increase student success in college”, *Proceedings of the Southern Association for Information Systems Conference*, September 11th–12th, 1-2 (2020).

Delavari, N., Phon-Amnuaisuk, S., Beikzadeh, M. R., “Data mining application in higher learning institutions”, *Informatics in Education*, 7(1): 31-54 (2008).

Dragicevic, M., Bach, M. P., Simicevic, V., “Improving university operations with data mining: Predicting student performance. *World Academy of Science, Engineering and Technology: International Journal of Economics and Management Engineering*, 8(4): 1101-1106 (2014).

Erickson, V., “Identifying at-risk computer science students early in semester utilizing data-driven models”, *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, pp. 865-870, (2019).

Indahyanti, U. Azizah, N. L., Setiawan, H., “Educational data mining on student academic performance prediction: A survey”, *Procedia of Sciences and Humanities: Proceedings of the 1st SENARA*, Universitas Muhammadiyah Sidoarjo, 1442-1447 (2022).

Kovacic, Z. J., “Predicting student success by mining enrolment data. *Research in Higher Education Journal*, 10: 647-665 (2010).

Kuzilek, J., Zdrahal, Z., Fuglik, V., “Student success prediction using student exam behaviour”, *Future Generation Computer Systems*, 125: 661-671 (2021).

Mariame, O., Khoulijand, S., Kerkep, M. L., “Feature engineering, mining for predicting student success based on interaction with the virtual learning environment

using artificial neural network”, *Annals of the Romanian Society for Cell Biology*, 25(6): 12734-12746.

Martínez, D. L. R., Gómez, K. E. P., “Contributions from data mining to study academic performance of students of a tertiary institute”, *American Journal of Educational Research*, 2(9): 713-726 (2014).

Merwe, C. V. D., Merwe, S. V. D., “Student success: Data mining measures what matters”, *Journal for New Generation Sciences*, 7(2): 284-302 (2009).

Natek, S., Zwillling, M., “Student data mining solution–knowledge management system related to higher education institutions”, *Expert Systems with Applications*, 41(14): 6400-6407 (2014).

Ndou, N., Ajoodha, R., Jadhav, A., “Educational data-mining to determine student success at higher education institutions”, *2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, p.1-8, 25-27 November 2020.

Osmanbegovic, E., Suljic, M., Data mining approach for predicting student performance”, *Economic Review – Journal of Economics and Business*, X(1): 3-12 (2012).

Puarungroj, W., Boonsirisumpun, N., Pongpatrakant, P., Phromkhot, S., “Application of data mining techniques for predicting student success in english exit exam”, *IMCOM '18: Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication*, (25): 1-6 (2018).

Shahiri, A. M., Husain, W., Rashid, N. A., “A review on predicting student’s performance using data mining techniques”, *Procedia Computer Science*, 72: 414-422 (2015).

Tinto, V., “Enhancing student success: Taking the classroom success seriously”, *The International Journal of the First Year in Higher Education*, 3(1): 1-8 (2012).

ÖZGEÇMİŞ

Fatma Tuğçe TÜRK ilk ve orta öğrenim eğitimini Antalya'da tamamladı. Aydın Karpuzlu Lisesi'nden mezun oldu. 2010 yılında Karabük Üniversitesi Bilgisayar Mühendisliği Bölümü'nde öğrenime başlayıp 2017 yılında mezun oldu. 2019 yılında Karabük Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda başlamış olduğu Bilgisayar Mühendisliği yüksek lisans programını 2023 yılında tamamladı.