



T.C.

ULUDAĞ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ ANA BİLİM

DALI

EĞİTSEL VERİ MADENCİLİĞİ İLE ÖĞRENCİ PROFİLLERİNİN BELİRLENMESİ

YÜKSEK LİSANS TEZİ

Abdullah Ragıp Ersöz

BURSA

2017



T.C.

ULUDAĞ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR VE ÖĞRETİM TEKNOLOJİLERİ EĞİTİMİ ANA BİLİM

DALI

EĞİTSEL VERİ MADENCİLİĞİ İLE ÖĞRENCİ PROFİLLERİNİN BELİRLENMESİ

YÜKSEK LİSANS TEZİ

Abdullah Ragıp Ersöz

Danışman

Prof. Dr. Aysan Şentürk

BURSA

2017

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim.

Abdullah Ragıp Ersöz

03/02/2017



YÖNERGEYE UYGUNLUK ONAYI

“Eğitsel Veri Madenciliği İle Öğrenci Profillerinin Belirlenmesi” adlı Yüksek Lisans tezi, Uludağ Üniversitesi Eğitim Bilimleri Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ne uygun olarak hazırlanmıştır.

Tezi Hazırlayan

Abdullah Ragıp Ersöz

Danışman

Prof. Dr. Aysan Şentürk

Bilgisayar ve Öğretim Teknolojileri Eğitimi ABD Başkanı

Prof. Dr. Aysan ŞENTÜRK

T.C.

ULUDAĞ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Bilgisayar ve Öğretim Teknolojileri Eğitimi Ana bilim dalında 801490201 numaralı Abdullah Ragıp Ersöz'ün hazırladığı “Eğitsel Veri Madenciliği İle Öğrenci Profillerinin Belirlenmesi” konulu Yüksek Lisans çalışması ile ilgili tez savunma sınavı, 28/04/2017 günü 15:00 – 16:30 saatleri arasında yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin (**başarılı / başarısız**) olduğuna (**oy birliği / oy çokluğu**) ile karar verilmiştir.

Üye

Tez Danışmanı ve
Sınav Komisyonu Başkanı
Prof. Dr. Aysan Şentürk
Uludağ Üniversitesi

Üye

Prof. Dr. Sedat Yüksel
Uludağ Üniversitesi

Üye

Yard. Doç. Dr. Yusuf Levent Şahin
Anadolu Üniversitesi

Önsöz

Veri madenciliđi alanına beni yönlendiren deđerli danıřman hocam Prof. Dr. Aysan Őentürk'e, desteklerinden dolayı Arř. Gör. Burcu Nur Bařtürk'e, zor zamanlarda en büyük destekçim olan sevgili eřim Zeynep Ersöz ve tatlı kızım Zehra'ya teřekkürlerimi sunarım.



Özet

Yazar	: Abdullah Ragıp Ersöz
Üniversite	: Uludağ Üniversitesi
Ana Bilim Dalı	: Bilgisayar ve Öğretim Teknolojileri Eğitimi Ana Bilim Dalı
Bilim Dalı	: Bilgisayar ve Öğretim Teknolojileri Eğitimi Bilim Dalı
Tezin Niteliği	: Yüksek Lisans Tezi
Sayfa Sayısı	: XV + 60
Mezuniyet Tarihi	: 09 / 05 /2016
Tez	: Eğitsel Veri Madenciliği İle Öğrenci Profillerinin Belirlenmesi
Danışmanı	: Prof. Dr. Aysan Şentürk

EĞİTSEL VERİ MADENCİLİĞİ İLE ÖĞRENCİ PROFİLLERİNİN BELİRLENMESİ

Eğitimde internetin kullanımı, öğretim yazılımları ve e-uygulamalarda yaşanan artış sayesinde veri tabanlarında eğitimle alakalı birçok veri depolanmıştır. Bu verilerin arasında gizli, bilinmeyen, eğitimin nasıl yapılması gerektiğine ışık tutacak ve öğrencileri daha iyi anlamamızı sağlayacak ilişkiler olduğu söylenebilir. Bu çalışmada Eğitsel Veri Madenciliği yöntemleri kullanılarak öğrenci profillerinin oluşturulması amaçlanmıştır.

Araştırmanın örneklemini Uludağ Üniversitesi Eğitim Fakültesi öğrencilerinin tamamı oluşturmaktadır. Veriler ilk olarak Veri Ön İşleme sürecinden geçerek analize hazır hale getirilmiştir. Çalışmada Veri Madenciliği yöntemlerinden Sınıflandırma Analizi yöntemleri kullanılmıştır. Bulgular ışığında öğrenci başarılarının önem sırasına göre bölüm, akademik dönem, geliş şekli ve cinsiyet değişkenlerine göre farklılık gösterdiği tespit edilmiştir. Yapılan sınıflandırma analizi ile hazırlanan karar ağacında doğum yılı ve akademik dönem

değişkenleri düğüm noktaları olarak belirlenmiştir. Ayrıca veriler arasında bilinmeyen gizli kurallar oluşturulmuştur. Eğitsel Veri Madenciliği sayesinde ilk bakışta anlamsız görünen veri yığını içinden gizli ilişkiler, öğrencileri daha iyi tanımamızı sağlayacak bilgiler, öğrencileri yönlendirebileceğimiz kurallar çıkartılmıştır.

Anahtar Sözcükler: Eğitsel Veri Madenciliği, Sınıflandırma Analizi, Öğrenci Modelleme.



Abstract

Author	: Abdullah Ragıp Ersöz
University	: Uludag University
Field	: Computer Education and Instructional Technology
Branch	: Computer Education and Instructional Technology
Degree Awarded	: Master
Page Number	: XV + 60
Degree Date	: 09 / 05 /2016
Thesis	: Determining Student Profiles with Educational Data Mining
Supervisor	: Prof. Dr. Aysan Şentürk

DETERMINING OF STUDENT PROFILES WITH EDUCATIONAL DATA MINING

Many educational data accumulated in the databases with the increase the use of the internet in education, teaching softwares and e-applications. Among these data, it can be said that there are hidden, unknown, relationships that will shed light on how education should be done and how students learning. In this study, it is aimed to create student profiles by using Educational Data Mining methods.

The sample of the research consists all the students of Uludağ University Faculty of Education. The data was first made available to analyze through the Data Preprocessing process. Classification Analysis, Decision Trees, Rule Induction and Gain By Weight methods are used in the study. According to findings, student achievement differs according to department, academic period, arrival type and gender variables. Rule Indiction parameters and unknown secret rules were created between the data and the importance weights of the variables were determined by the Gain By Information Weight algorithm.

Through Educational Data Mining, hidden relationships from a seemingly insignificant amount of data at first sight, information that will help us better identify our students, and the rules we can direct students with are extracted through data mining analyzes.

Keywords: Educational Data Mining, Classification Analysis, Student Modeling.



İçindekiler

	Sayfa No
ÖNSÖZ	iv
ÖZET	v
ABSTRACT	vii
İÇİNDEKİLER	ix
ŞEKİLLER LİSTESİ	xii
TABLolar LİSTESİ	xiv
KISALTMALAR LİSTESİ	xv
1. BÖLÜM GİRİŞ	1
1.1 Problem Durumu	1
1.2 Amaç	2
1.3 Önem	2
1.4 Varsayımlar	3
1.5 Sınırlılıklar	3
1.6 Tanımlar	3
2. BÖLÜM ALANYAZIN	4
2.1 Büyük Veri	4
2.2 Veri Madenciliği	5
2.2.2 Veri Madenciliği yöntemleri	6
2.3 Eğitsel Veri Madenciliği	10
2.3.1 Eğitsel Veri Madenciliğinin hedefleri	12

2.3.2 Veri Çeşitleri	12
2.3.3 Kullanılan Yöntemler.....	13
2.3.4 Eğitsel Veri Madenciliği uygulama süreci.....	15
2.3.5 Eğitsel Veri Madenciliği değerlendirme süreci.	16
3. BÖLÜM YÖNTEM	17
3.1 Araştırmanın Modeli	17
3.2 Evren ve Örneklem	17
3.3 Verilerin Toplanması ve Çözümlemesi.....	17
3.3.1 Veri ön işleme.	17
3.3.2 Veri çözümleme teknikleri.....	19
3.3.3 Verilerin programa aktarılması.	20
3.3.4 Veri analizi.....	23
4. BÖLÜM BULGULAR	27
4.1 Öğrenci Modelleme (Student Modeling).....	33
4.1.1 Öğrenci başarısı cinsiyet ilişkisi.	33
4.1.2 Öğrenci başarısı doğum yeri ilişkisi.....	33
4.1.3 Öğrenci başarısı doğum yılı ilişkisi.	34
4.1.4 Öğrenci başarısı akademik dönem ilişkisi.	35
4.1.5 Öğrenci başarısı bölüm ilişkisi.....	36
4.1.6 Öğrenci başarısı geliş şekli ilişkisi.....	38
4.1.7 Öğrenci başarısı uyruk ilişkisi.....	40
4.1.8 Öğrenci başarısı öğretim tipi ilişkisi.	40

4.2 Sınıflandırma Analizi	41
4.3 Akademik Başarıya Etki Eden Değişkenlerin Önem Ağırlıkları	46
5. BÖLÜM TARTIŞMA VE ÖNERİLER.....	47
5.1 Değişkenler Bize Ne Anlatıyor?	47
5.1.1 Bölüm faktörü.	47
5.1.2 Yaş faktörü.	48
5.1.3. Geliş şekli faktörü.	49
5.1.4. Cinsiyet faktörü.....	50
5.2 Öğrenci Modelleme.....	51
5.3 Akademik Başarıya Etki Eden Değişkenlerin Önem Ağırlıkları	52
5.4 Öneriler	54
KAYNAKÇA.....	55
ÖZGEÇMİŞ	60

Şekiller Listesi

Şekil	Sayfa
1 Karar Ağacı Örneği.....	7
2 Kümeleme Analizi Örneği.....	9
3 Birliktelik Kuralı Örneği.....	10
4 Disiplinler arası bir alan olarak EVM.....	11
5 Uygulama Süreci.....	15
6 Ham Veriler.....	18
7 Verilerin aktarılması ilk işlem.....	20
8 Verilerin Aktarılması 1.....	21
9 Verilerin Aktarılması 2.....	22
10 Verilerin Aktarılması 3.....	23
11 Rapidminer Programı Ara Yüzü.....	24
12 Birinci Araştırma Sorusu Tasarım Ekranı.....	25
13 İkinci Araştırma Sorusu Tasarım Ekranı.....	25
14 Üçüncü Araştırma Sorusu Tasarım Ekranı.....	26
15 Tanımlayıcı İstatistikler.....	28
16 Bölüm Değişkeni.....	29
17 Not Ortalaması Değişkeni.....	30
18 Cinsiyet Değişkeni.....	31
19 Doğum Yılı Değişkeni.....	32
20 Öğrenci Başarısı Cinsiyet İlişkisi.....	33
21 Öğrenci Başarısı Doğum Yeri İlişkisi.....	34
22 Öğrenci Başarısı Doğum Yılı İlişkisi.....	35
23 Öğrenci Başarısı Akademik Dönem İlişkisi.....	36

24	Öğrenci Başarısı Bölüm İlişkisi	37
25	Öğrenci Başarısı Bölüm İlişkisi 2	38
26	Öğrenci Başarısı Geliş Şekli İlişkisi	39
27	Öğrenci Başarısı Geliş Şekli İlişkisi 2	39
28	Öğrenci Başarısı Uyruk İlişkisi	40
29	Öğrenci Başarısı Öğretim Tipi İlişkisi	41
30	Naive Bayes Yöntemi Basit Dağılım	42
31	Karar Ağacı	43
32	Karar Ağacı Dügüm Noktaları	43
33	Kural Çıkarımı	44
34	Değişkenlerin Önem Ağırlıkları	46
35	Değişkenlerin Önem Ağırlıkları 2	46
36	Bölüme Göre Ortalamalar	48
37	Geliş Şekline Göre Ortalamalar	50

Tablolar Listesi

<i>Tablo</i>		<i>Sayfa</i>
1	Değişken Kodlamaları.....	19



Kısaltmalar Listesi

EVM: Eğitsel Veri Madenciliđi



1. Bölüm

Giriş

1.1 Problem Durumu

Verilerin depolanmaya başlamasından günümüze gelen süreçte, her alanda giderek artan bir bilgi yığılması oluşmuştur. Bu veriler çıplak gözle bakıldığında bir anlam ifade etmeyen, yapılandırılmamış verilerdir. Veri Madenciliği sayesinde yapılan analizlerle ilk bakışta anlamsız gözükken veri yığınları içinden, anlamlı, birbiriyle ilişkisi olan, gizli bilgiler çıkarılabilir. Karmaşık verilerin düzenlenmesi ve anlamlı sonuçlar çıkarabilme açısından Veri Madenciliği son derece gerekli ve önemli bir ihtiyaçtır (Maimon & Rokach, 2010).

Dijital nesil olarak adlandırılan yeni nesil, eski nesle göre değişik yeteneklere ve görüşlere sahiptir (Prensky, 2011). Bu neslin nasıl daha iyi öğrendiğinin anlaşılması, hangi öğrenci verilerinin öğrenmeye katkısı olduğunun tespiti, hangi özelliklerde öğrencilerin aynı kümede toplandığının belirlenebilmesi Veri Madenciliği teknikleriyle mümkündür (Pena, 2013). Aynı zamanda bu veriler ışığında öğrenciler, yetenek ve ihtiyaçları doğrultusunda yönlendirilebileceklerdir.

Veri Madenciliği sağlık, savunma, ticaret gibi alanlarda yıllardır başarıyla uygulanmaktadır fakat eğitim alanındaki yani eğitsel Veri Madenciliği hakkındaki çalışmalar, hazırlanan makale ve tezler, özellikle ülkemizde sınırlı sayıdadır. YÖK Tez Arama Motorunda Eğitsel Veri Madenciliği anahtar kelimesi ile arama yapıldığında 7 teze ulaşılmakta olup bunların sadece 2 tanesi eğitim alanında yazılmış tezlerdir. Bu alanın yaygınlaşması, nitelikli uzmanların yetişmesi ve ilgili tez ve makalelerin sayısının artırılmasına ihtiyaç duyulmaktadır.

Öğrenci verilerinin kullanılarak öğrencilere danışmanlık hizmetinin verilmesi ve öğrencilerin yönlendirilmeleri dünya üzerinde araştırılan bir konudur. Veri Madenciliği analizleri sayesinde, okulu bırakacak öğrenciler daha önceden tespit edilebilir, öğrencilerin

nasıl daha iyi öğrendikleri belirlenebilir (The Guardian, 2013). Bu gelişmeler ışığında aşağıdaki gibi problem cümlesi belirlenmiştir.

Problem cümlesi: Öğrenci bilgi sisteminde yer alan eğitsel verilerin, Veri Madenciliği teknikleri ile analiz edilerek gizli ama anlamlı bilgilerin çıkarılıp çıkarılmayacağını araştırılmasıdır.

Bu bağlamda aşağıdaki araştırma soruları belirlenmiştir.

1. Demografik özellikleri ve ders başarılarına göre öğrenciler nasıl modellenir?
2. Benzer kişisel özellikleri taşıyan öğrenci grupları nasıl sınıflanmıştır?
3. Akademik başarıya etki eden değişkenlerin önem ağırlıkları nedir?

1.2 Amaç

Bu çalışmanın amacı, Uludağ Üniversitesi Eğitim Fakültesinde okuyan öğrencilerin, öğrenci bilgi sistemindeki verilerini, Veri Madenciliği teknikleri ile analiz ederek öğrenci profillerinin ortaya çıkarılmasını sağlama, bu karmaşık veri yığınları arasından gizli, birbiri ile ilişkili, anlamlı bilgiler ortaya çıkarmaktır.

Hazırlanan araştırma sorularına karşılık şu alt amaçlar belirlenmiştir;

1. Ders başarıları ve demografik özelliklerine göre öğrencilerin modellenmesi,
2. Benzer özelliklere sahip öğrencilere ait bir sınıflama modelinin oluşturulması,
3. Akademik başarıya etki eden değişkenlerin önem ağırlıklarının tespit edilmesi.

1.3 Önem

Eğitsel Veri Madenciliği ile yapılan çalışmalar ülkemizde emekleme aşamasındadır. Bu çalışma Veri Madenciliği tekniklerinin eğitimde uygulanması, Eğitsel Veri Madenciliğinin Türkiye’de yeni ve hızla gelişen bir alan olması açısından bu alanda yapılacak çalışmalar önem arz etmektedir.

2017 yılında YÖK tarafından 2000 öğrenciye doktora alanında verilecek burslar için belirlenen 100 tematik alan içinde Veri Madenciliği de yer almaktadır. Ayrıca Veri

Madenciliği TÜBİTAK projelerinde öncelikli alanlarda olup, bu alanda verilerin toplanması ve analizi ile ilgili büyük ölçekli projeler hazırlanmaktadır.

Eğitsel Veri Madenciliği ile yapılan çalışmalarda eğitim ortamlarından gelen verilerin analiz edilmesi, öğrencilerin daha iyi anlaşılması, hangi yollarla ve nasıl öğrendiklerinin belirlenmesi açısından önemlidir. Öğrencilerin eğitim ortamlarındaki davranışları sonucu kaydedilen verilerin ve demografik özelliklerinin analiz edilmesi, teknolojiyi hayatının her alanında kullanan yeni neslin beklenti ve ihtiyaçlarının belirlenmesinde etkisi olabilir (Baker & Yacef, 2009).

Öğrenci başarısı ve cinsiyet arasındaki ilişki örneğin t-testi ile belirlenebilir. Veri Madenciliği analizlerinin önemi verilerin programa aktarılmasının ardından tek tuş ile bize tüm değişkenler ile başarı arasındaki ilişkileri göstermesidir. Ayrıca geleneksel yöntemlerle ulaşamayacağımız verilerin arasındaki ilişkileri de kurallar oluşturarak kullanıcıya sunar.

1.4 Varsayımlar

Çalışmada kullanılan verilerin doğru olduğu, üzerinde herhangi bir işlem yapılmadığı, gerçek veriler olduğu kabul edilmiştir.

1.5 Sınırlılıklar

Çalışmada kullanılan veriler Uludağ Üniversitesi Eğitim Fakültesi Öğrenci Bilgi Sisteminden temin edilmiştir. Bilgi Sisteminde yer alan eksik bilgilerin fazlalığı sebebiyle bazı değişkenler çalışmaya dahil edilememiştir. Dolayısıyla çalışma Eğitim Fakültesi öğrencileri; sisteme kayıtlı olan ve analiz edilebilecek durumda olan veriler ile sınırlıdır.

1.6 Tanımlar

Eğitsel Veri Madenciliği: Öğrencileri daha iyi tanıma, nasıl daha iyi öğrendiklerini belirlemek için Veri Madenciliği analizlerinin eğitimde kullanılması.

2. Bölüm

Alanyazın

2.1 Büyük Veri

Son yıllarda internet üzerinde transfer edilen, paylaşılan, depolanan verilerle, sanal dünya bir bilgi çöplüğü haline gelmiştir. Teknolojinin ilerlemesiyle verilerin depolanması hızlanmış, daha çok veri depolanmaya başlanmıştır. Anlamalı ya da anlamsız sayısız miktarda veri, veri tabanlarında, sanal âlemde birikmiştir. Pennystocks, (2015) verilerine göre Facebook'ta sadece 1 saniye içinde 52196 beğeni gerçekleşmekte, Twitter'da 5700 tweet atılmakta, Youtube'da 2314 video yüklenmektedir. Benzeri sosyal ağlarda yaşananlarla birlikte internet üzerinde ortalama 20 saniye içerisinde 451 TB veri transfer edilmektedir. Dahası uzay bilimlerinde, tıp alanında ileri teknolojiyi kullanan aletler sayesinde şu ana kadar sahip olduğumuz bilgiler artık bir gün içinde üretilebilmektedir (Lepine & Shara, 2005). Bu da veri tabanlarında büyük bir bilgi yığılması anlamına gelmektedir. Yazılım şirketleri ve bilgisayar mühendisleri, bu bilgi yığılmasına Big Data yani Büyük Veri adını vermiştir.

Büyük Verinin tarihsel süreci ise veri tabanlarının ortaya çıkışına dayanmaktadır. Veri tabanları birbiriyle ilişkili verilerin depolandığı alanlardır. İlk olarak 1960 yıllarında ortaya çıkan veri tabanları, 70'lerde ilişkisel veri tabanların kurulması ile geliştirilmiş, yeni veri depolama araçlarıyla günümüze gelmiştir. Teknolojinin gelişmesiyle veri türleri ve kapasiteleri de artmış verilerin depolanması konusunda yeni teknikler geliştirilmiştir.

Bu veri tabanlarında zetabaytlarla ifade edilmeye çalışılan çok miktarda verinin büyük bir kısmı yapısal, işlenebilir değildir. Veriler karmaşıktır ve ilk bakıldığında bir anlam ifade etmemektedir. 90'lı yılların başına kadar bu veriler değersiz gözükmüş ama neden sonra bu verilerin çok değerli, kullanılabilir, önemli bilgiler taşıdığı fark edilmiştir. Bundan sonra yapılan teknolojiler, verilerin yorumlanması göz önünde bulundurularak geliştirilmiştir.

Örneğin, Hükümetler insanların yaptıkları işlemler sonucu ortaya çıkan bilgileri, demografik özellikleri depolamaktadırlar. Büyük miktarda paranın döndüğü bankalar müşterileri hakkında detaylı bilgiye sahip olmak için para hareketlerini saklamak zorundadır. Sosyal medya üzerinde kullanıcıların yaptıkları paylaşımlar, beğeniler ve diğer hareketler veri tabanlarında saklanmaktadır. Hastaneler müşterileri hakkında istatistikleri depolarlar. 2013 yılında Şili Atakama Çölü üzerinde kurulan ALMA teleskopunun, dünya genelinde şu ana kadar üretilen bilgi kadar bilgi ürettiği söylenmektedir (CNNTÜRK, 2014).

İlk bakıldığında bir anlam ifade etmeyen verilerin analizi sayesinde, tedavisi bulunamayan hastalıklara çözüm bulunması, suçluların önceden tespit edilmesi, uzay araştırmalarında bilinmeyen keşiflerin yapılması sağlanabilir. Büyük veriler ancak doğru analiz metotlarıyla yorumlandığında anlamlı ve doğru bilgilere ulaşılabilir. Bu amaçla Veri Madenciliği kavramı ortaya çıkmıştır.

2.2 Veri Madenciliği

Büyük veriler hakkında yorum yapabilmek için gerçek bilgilerin ortaya çıkarılması gerekir. Çeşitli yazılımlar yardımıyla yapılan bu işleme Veri Madenciliği denir. Veri Madenciliği basit olarak büyük miktarda veri bulunan veri tabanlarından, gelecekle ilgili tahminler yapılmasını sağlayacak anlamlı ilişkiler çıkarma, bilgiyi bulma, madenleme işidir (Şentürk, 2006).

Veri Madenciliği hakkında ilk çalışmalar 1960'lı yıllarda verilerin bilgisayar ortamında depolanmasıyla başlamıştır (Ledley, 1960). İlk olarak basit veri modelleri geliştirilmiş, daha sonra ihtiyaçlar doğrultusunda varlık-ilişki, ilişkisel ve nesne yönelimli veri modelleri oluşturulmuştur. 1990'lı yıllarda internetle birlikte daha da büyüyen veri tabanlarındaki verilerin analiz edilmesi için bilgisayar yazılımcıları tarafından Veri Madenciliği kavramı ortaya atılmıştır (Efron & Tibshirani, 1990). Veri Madenciliğinde geleneksel yöntemler yerine, profesyonel yazılımlar ve özel algoritmalar kullanılmıştır.

Veri Madenciliğinde veriler bilgisayar programları kullanılarak analiz edilir. Başlıca kullanılan programlar, R, Rapidminer, WEKA, Orange yazılımlarıdır. Programların bazıları tüm kamu ve bireylerin kullanılabileceği açık kaynak kodlu, içerisinde pek çok algoritma bulunduran ücretsiz yazılımlardır.

2.2.2 Veri Madenciliği yöntemleri. En çok kullanılan Veri Madenciliği yöntemleri Tahmin Edici Yöntemler; Sınıflandırma (Classification), İstatistiksel Tahmin modelleri (Statistical Modeling) ve Tanımlayıcı Yöntemler: Birliktelik Kuralı (Association Rule Mining) ve Kümeleme (Clustering) analizleridir.

Tahmin edici yöntemler sonuçları bilinen bir veri kümesinden henüz sonuçları bilinmeyen benzer ama daha büyük veri kümelerinin sonuçlarını tahmin etmemizi sağlayan modeller geliştirir (Han, Pei & Kamber, 2011). Bir GSM operatörünün hangi kullanıcısının başka bir operatöre geçeceğini önceden tahmin etmesi örnek olarak verilebilir. Eğitim açısından bakarsak okulu bırakacak bir öğrencinin önceden tespiti ya da genel ağırlıklı not ortalamasının önceden tahmin edilmesi tahmin edici analizler ile yapılabilir. Tanımlayıcı yöntemler ise mevcut verilerin, aralarındaki örüntülerin, benzerliklerin tanımlanmasını sağlayan modellerdir (Zhong & Zhou, 1999).

En çok kullanılan Veri Madenciliği yöntemleri, kısaca tanımları ve analizlerde kullanılan algoritmalar, kolay anlaşılabilmesi için örneklerle birlikte aşağıda anlatılmıştır.

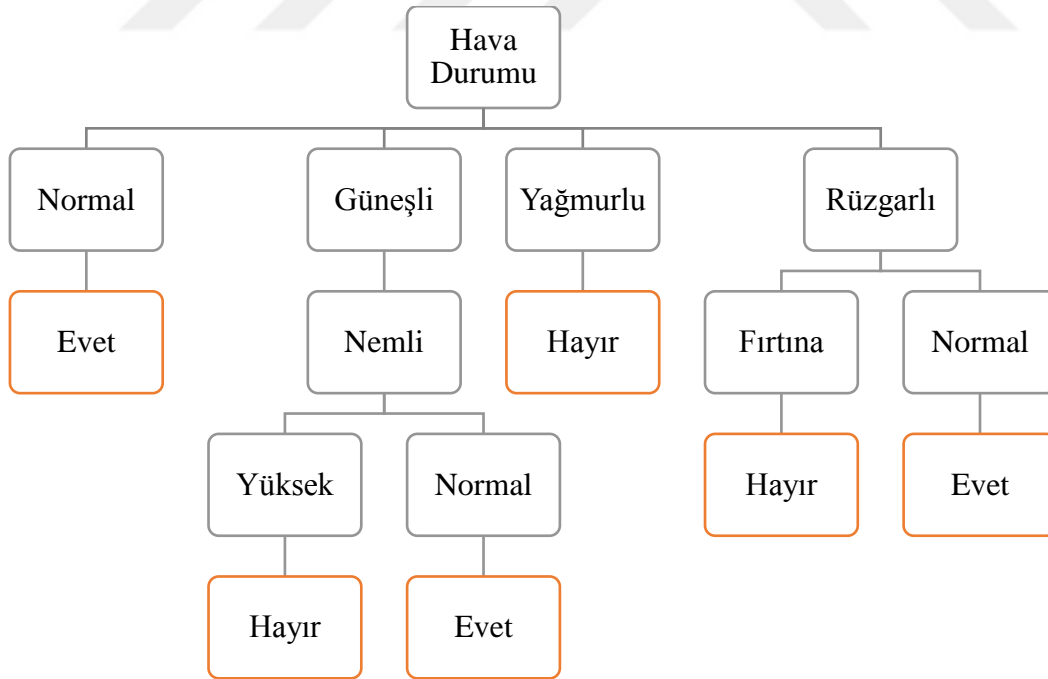
2.2.2.1 Sınıflandırma analizi. Sınıflandırma Analizi veri setinin tamamını açıklayacak verileri belli sınıflara ayıran ve hangi verinin hangi sınıfa ait olduğunu belirleyen bir model oluşturma sürecidir (Han ve diğerleri, 2011). Karar Ağaçları, Yapay sinir Ağları ve Genetik Algoritmalar en çok kullanılan sınıflandırma tekniklerindedir. Bu çalışmada kullanılan sınıflandırma analizleri Karar Ağaçları, Bayes Sınıflandırma Yöntemi, Kural çıkarımı ve Bilgi kazanç oranına göre ağırlık hesaplama algoritmasıdır.

2.2.2.1.1 *Karar ağaçları*. Karar Ağaçları verilerin sınıflandırılmasında kullanılır. İlk olarak Quinlan (1986) tarafından ortaya atılmış, zamanla yeni algoritmalarla güçlenmiştir. Verilerin küme içindeki önem derecelerini tespit eder ve ağaca benzer bir şekil oluşturur. Karar ağaçlarında düğüm noktaları vardır ve bu noktalarda veri kümesi alt sınıflara ayrılır ve dallanma bu şekilde veri kümesinin hiç elemanı kalmayana dek devam eder. C4.5, Id3, Chart karar ağacı algoritmalarına örnek olarak verilebilir.

Şekil 1’de bir kişinin hava durumuna göre dışarıya çıkmaya karar vermesiyle ilgili örnek bir karar ağacı oluşturulmuştur. Düğüm noktaları ve ağaç şeklinin yapısı şekilde gösterilmiştir. Buna göre hava durumu normalse birey dışarıya çıkabilir; eğer güneşliyse havanın nemli olup olmadığına bakılır. Yüksek nem varsa çıkamaz ama normal nem varsa dışarıya çıkabilir şeklinde ağaç yapısı devam etmektedir.

Şekil 1

Karar Ağacı Örneği



2.2.2.1.2 *Bayes sınıflandırma metodu*. Naive Bayes Sınıflandırıcısı Bayes teoreminden esinlenmiş, tahmin edici bir sınıflandırmadır. Bir sınıfta bulunan belirli bir değişkenin varlığının veya yokluğunun, başka bir değişkenin varlığı veya yokluğu ile ilgisi olmadığını

varsayar. Değişkenler arasındaki matematiksel ilişkileri kodlayan bir modeldir (Heckerman, 1996).

Bayes Teoreminden etkilenecek hazırlanmış Naive Bayes Sınıflandırıcısı, tüm değişkenler arasındaki ilişkileri kodlar (Bernardo & Smith, 1994). Bhardjaw ve Pal (2012) yılında öğrenci başarıları arasındaki farkı tanımlamak için yaptıkları çalışmada Naive Bayes yöntemini kullanarak bir tahmin modeli oluşturmuş ve öğrenci başarısındaki etkili faktörlerin annelerin niteliği, öğrencilerin alışkanlıkları, öğrencilerin aile statüleri ve ailelerin yıllık geliri değişkenleri olduğu sonucuna varmışlardır.

2.2.2.1.3 Kural Çıkarımı. Veri Madenciliğinde en çok kullanılan yöntemlerden birisidir. Veriler arasında gizlenen ilişkilerin kurallarla ifade edilmesidir (Maimon & Rokach, 2010). Bu yöntemde LEM1, LEM2, Lers gibi algoritmalar kullanılır. Kurallar genel olarak tablo benzeri yapıda bir form şeklinde sunulur. Genel olarak kullanılan form türü:

Eğer (Özellik 1, Değer 1) ve (Özellik 2, Değer 2) ve (Özellik 3, Değer 3)... ve (Özellik n, Değer n) sonra (Karar, Değer) şeklindedir.

Kural çıkarımı, yaygın sınıflardan başlayarak, olumlu bir örnek kalmayana veya hata oranı % 50'den büyük olana dek algoritmayı devam ettirir, tekrar tekrar geliştirerek kuralları oluşturur (Cohen, 1995).

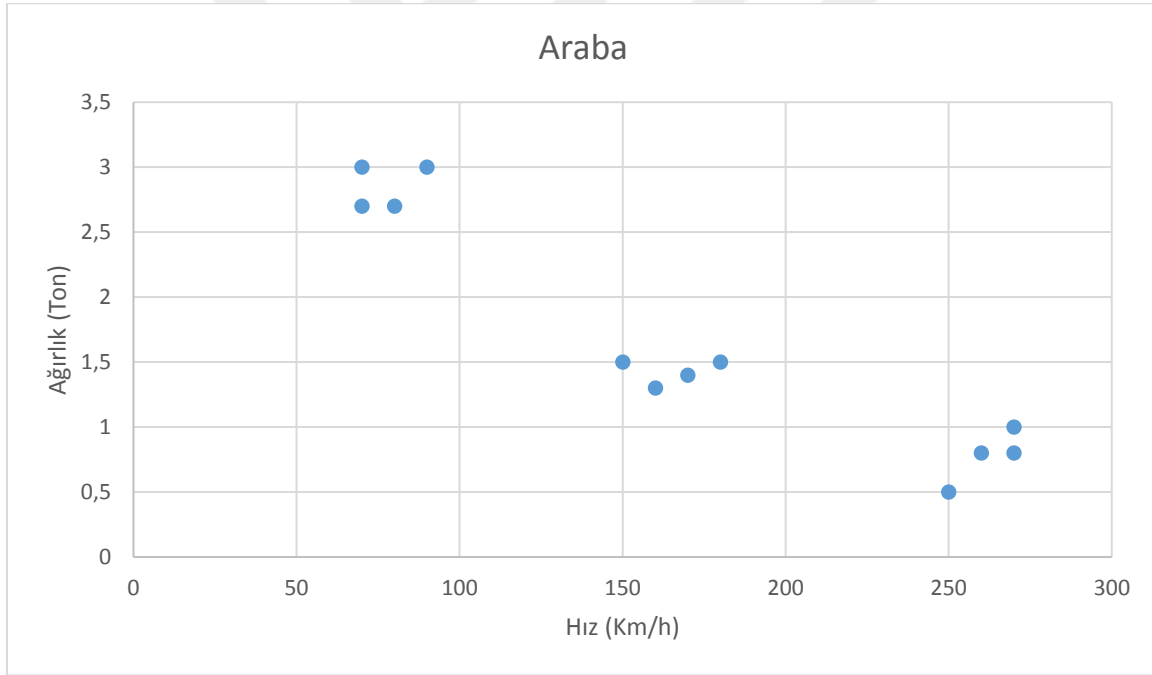
2.2.2.1.4 Bilgi kazanç oranına göre ağırlık (The Weight By Information Gain Ratio). Bu algoritma sınıf özneliğine göre değişkenlerin ağırlığını hesaplar. Bir özelliğin ağırlığı ne kadar yüksekse o kadar önemlidir. Sadece nominal değişkenlere uygulanabilir. Akça (2014) Fen Fakültesi öğrencilerine ait bir öğrenci profili ortaya çıkarmak amacıyla yazdığı yüksek lisans tezinde, 4 farklı sınıflandırma algoritmasını kullanarak öğrencilerin not ortalamaları üzerinde en önemli değişkenin cinsiyet; bölüm tercihinde en önemli değişkenin mezun olduğu lise türü ve ÖSYM puanı olduğu sonuçlarına ulaşmıştır.

2.2.2.2 Kümeleme analizi. Veri kümesindeki verilerin benzerliklerine göre gruplara ayrılma sürecidir (Klößgen & Zytow, 2002). K-means, X-means, en bilinen kümeleme analizleridir. Veri kümesi istenilen sayı kadar kümeye bölünebilir. Hangisinde daha başarılı bir kümeleme yapıldıysa o sayıyı kullanmak gerekir.

Şekil 2’de ağırlık ve hız değerlerine göre arabalar belli gruplar halinde kümelendiği görülmüştür. Hızı düşük ama ağırlığı çok olanlar ağır taşıtları; ağırlığı ve hızı normal seviyede olanlar orta düzey arabaları; hızı yüksek ve ağırlığı az olan araçlar ise spor arabaları temsil eder. Temel olarak kümelemenin mantığı budur.

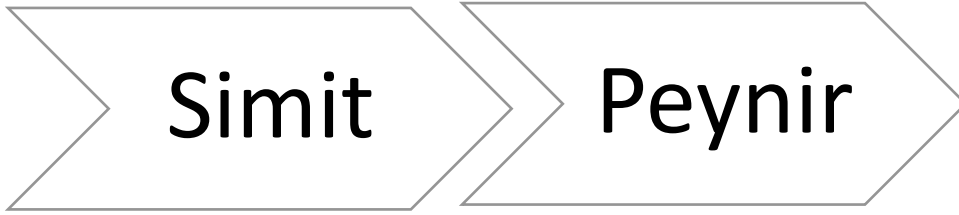
Şekil 2

Kümeleme Analizi Örneği



2.2.2.3 Birliktelik kuralı. Birliktelik Analizi aynı zamanda gerçekleşen olayların birbirleri arasındaki ilişkilerin tanımlanmasında kullanılır (Liu, Hsu & Ma, 1998). Veri kümesinde yer alan önemli önemsiz tüm kuralları oluşturur. Bir kafede simit alan öğrencilerin %80’in yanına peynir alması, bankadan kredi çeken müşterilerin bir sonraki yıl % 70 tekrar kredi çekmesi gibi analizler Birliktelik kurallarına örnek olarak gösterilebilir.

Şekil 3

Birliktelik Kuralı Örneği

Okul kantinlerinde, simitçilerde simit ile peynirin yan yana satılıyor olması bir birliktelik kuralı analizi sonucudur. (Şekil 3)

Günümüzde Veri Madenciliği çeşitli alanlarda yukarıda anlatılan ve benzeri yöntemler ile kullanılmaktadır. Veri Madenciliği, satış tahminleri ile müşteri değerlendirmesi, Pazar Sepet Analizleri gibi konularla pazarlama alanında; kredi taleplerinin değerlendirilmesi, müşteri portföylerinin belirlenmesi konularında bankacılık sektöründe; çeşitli hastalıkların tedavi sürecinde, hastalıklı hücrelerin yayılma şekillerinin izlenmesinde ve laboratuvar testlerinde hata tespiti gibi konularda sağlık alanında; toplumun güvenliğini tehdit eden hırsız, katil, dolandırıcı vb. kişilerin tespitinde güvenlik ve istihbarat alanlarında başarılı bir şekilde uygulanmaktadır (Pena, Dominguez ve Medel, 2009). Veri Madenciliği geniş bir kullanım alanına sahiptir ve uygulamalar giderek daha da yaygınlaşmaktadır.

2.3 Eğitsel Veri Madenciliği

Veri Madenciliği tekniklerinin eğitimde kullanılma ihtiyacı eğitsel Veri Madenciliği kavramını doğurmuştur. Öğrencileri daha iyi anlama, nasıl daha iyi öğrendiklerinin belirlenmesi amacıyla Veri Madenciliği teknikleri kullanılarak yapılan işleme Eğitsel Veri Madenciliği denir (Romero & Ventura, 2007).

Eğitsel Veri Madenciliği, içinde buldukları çok büyük veri hacminden dolayı analiz edilmesi zor veya imkansız olan geniş eğitsel verilerin arasındaki ilişkileri tespit etmek için bilgisayarlı yöntemler geliştirmek, araştırmak ve uygulamakla ilgilenmektedir (Romero & Ventura, 2013). Eğitsel Veri Madenciliği üniversiteler, okullar, e- öğrenme ortamları ve diğer

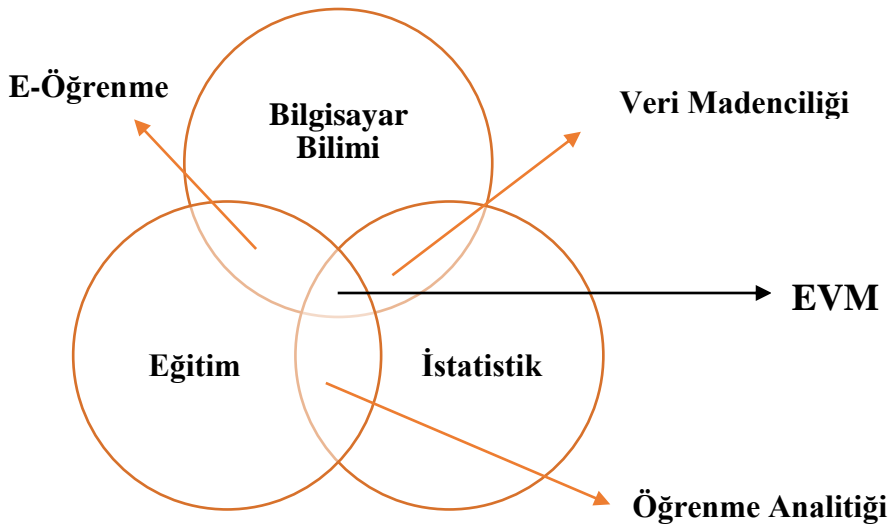
akademik çalışmalardan gelen eğitsel verilerin Veri Madenciliği yöntemleriyle analiz edilmesidir.

Aslında bilgisayarın ve verilerin depolanmasının ortaya çıkışıyla başlayan bu analiz yöntemi, 2000- 2007 yıllarında araştırma sahasında popüler bir konu olmuş, 2007 yılında bu alanda yapılan çalışmalar Romeo ve Ventura tarafından bir makalede toplanmıştır. Aynı zamanda alanla ilgili uluslararası bir konferans yapılmaya başlanmıştır. Ayrıca Journal of Educational Data Mining adında bir dergi yalnızca bu alanda yazılan makaleleri yayınlamaya başlamıştır.

Eğitsel Veri Madenciliği Öğrenme Analitiği alanıyla karıştırılmaktadır ama iki kavram aynı şey değildir. Eğitsel Veri Madenciliği, veri kümesindeki yenilikleri keşfeder ve yeni algoritmalar üretirken; Öğrenme Analitiği teknik ve sosyo-pedagojik kısımlarla ilgilidir (Baker & Inventado, 2014). Türkiye’de yanlış çeviriden kaynaklanan, Veri Madenciliği tekniklerinin öğretilmesi, Veri Madenciliği eğitimi gibi kavramlar ile de karıştırılmaktadır. Eğitsel Veri Madenciliği Öğrenme Analitiği, Veri Madenciliği ve E-Öğrenme Alanlarının hepsini kapsayan disiplinler arası bir alandır (Şekil 4).

Şekil 4

Disiplinler arası bir alan olarak EVM



2.3.1 Eğitsel Veri Madenciliğinin hedefleri. Eğitsel Veri Madenciliği hedefleri öğrenciler, eğitimciler ve araştırmacılar açısından 3 ana başlıkta toplanabilir (Pena, 2013);

- Öğrencilerin öğrenim sürecinde akıllı ve uyumlu geribildirim almalarını sağlamak, öğrenci davranışlarını desteklemek;
- Eğitimciler ve yöneticiler açısından öğrencilerin daha iyi anlaşılmasını sağlamak, nasıl daha iyi öğrendiklerini belirlemek;
- Araştırmacılar açısından eğitsel Veri Madenciliği tekniklerini karşılaştırmak ve geliştirmektir.

2.3.2 Veri Çeşitleri. Eğitsel Veri Madenciliğinde temel olarak iki çeşit veri kullanılır, nominal ve sayısal (nümerik) veriler (Witten & Frank, 2005). Ayrıca metin, tarih, kayıp veri gibi değişkenler de Eğitsel Veri Madenciliği için önemlidir. Verilerin türü yapılacak analiz için temel teşkil eder. Verilerin analiz kısmında hazırlanan tasarım ekranları ve kullanılan parametreleri çeşitlerine göre farklılık gösterir.

2.3.2.1 Nominal veri. Bölüm, cinsiyet, uyruk gibi kategorik değişkenlere nominal veri denir. Cinsiyet gibi iki değişkene sahip veriler binominal, birden fazla değişkene sahip veriler ise polinomial olarak adlandırılır. Veri Madenciliğinde sınıflandırma ve kümeleme analizlerinde en çok kullanılan veri türleridir.

2.3.2.2 Nümerik veri. Genel ağırlıklı not ortalaması, öğrenci notları gibi sayısal verilerdir. Integer, real gibi çeşitleri vardır. Veri Madenciliğinde kullanılan analizler nümerik verilerin nominal olarak kullanılmasına izin verir.

2.3.2.3 Kayıp veriler. Veri Madenciliğinde kayıp veriler büyük önem taşır. Bir çalışma için veriler orijinal olarak her toplandığında mutlaka, eksik, boş sütun ve satırlar olur. Veri isimlerinin yanlış yazımından doğan hatalardan dolayı da eksik veriler oluşabilir. Verilerin

analiz kısmında kullanılan program mutlaka kayıp veriler olup olmadığına bakar ve eksik olarak yazılan verilerin düzenlenmesini ister (Witten & Frank, 2005).

Her zaman hazır veriler üzerinden hareket edilmez. Araştırmaların bazılarında hazırlanan ara yüz sayesinde bilgiler depolanır ve ardından analiz edilerek hem bilgi toplama hem de Veri Madenciliği süreçleri gerçekleştirilmiş olur. Gürüler, İstanbullu ve Karahasan (2010) MUSKUP adlı bir yazılım geliştirerek, ilk önce öğrencilerin bilgilerini toplamış, sonra karar ağacı ve kümeleme analizlerini yaparak öğrencileri modellemişlerdir.

2.3.3 Kullanılan Yöntemler. Bu bölümde özellikle Eğitsel Veri Madenciliğinde kullanılan yöntemler ve bu yöntemler ile ilgili yapılmış çalışmalar anlatılmıştır.

2.3.3.1 Tahmin yöntemleri. Tahmin yöntemlerinde amaç, veri kümesindeki verilerden tahmin edilen değişkeni çıkaran bir model inşa etmektir. Tahmin yöntemleri eğitsel Veri Madenciliğinde iki önemli kapıyı açar (Baker, 2010). Bunlardan birincisi ve yaygın olarak kullanılan öğrencilerin çıktılarını önceden tahmin etmeyi sağlayan yöntemlerdir. İkincisi ise öngörülemeyecek olaylarda veri kümelerinin çıktı değerlerini tahmin etmektir. Tahmin yöntemleri genelde sınıflandırma metotları, regresyon ve yoğunluk tahmini şeklindedir. Akçapınar (2014), öğrencilerin dönem sonundaki akademik başarısının daha önceki haftalarda tahmin edilip edilemeyeceğini, bu konuda hangi algoritmanın en iyi sonuçları verdiğini ve benzer davranış gösteren öğrenci gruplarını belirlemeyi amaçladığı doktora tezinde, 76 üniversite 2. Sınıf öğrencisinin 14 hafta boyunca çevrimiçi öğrenme ortamındaki etkileşim verilerini incelemiştir. Çalışma sonunda etkileşim verileri kullanılarak dönem sonundaki akademik performanslarının, 3. hafta gibi kısa bir sürede % 74 oranında başarıyla tahmin edilebileceği tespit edilmiştir. Öğrenciler, derslerde yapmış oldukları aktivitelere göre 3 ayrı grupta kümelenebilir, bunun sonucunda aktif olan öğrencilerin daha başarılı olduğu sonuçlarına varılmıştır.

2.3.3.2 Kümeleme analizi. Doğal olarak birlikte gruplanmış örnekleri ya da bir veri kümesini kategorilere ayırmak için kullanılır. Öğrencileri öğrenme biçimlerine göre belli gruplarda toplamak kümeleme analizine örnek olarak verilebilir. Gürsoy (2012) Uygulamalı Veri Madenciliği Yöntemleri kitabında bir üniversitede bir yüksek lisans programına kayıt yaptıran öğrencilerin verilerini kümeleme analizi yöntemiyle incelemiş, analiz sonucunda birbirinden farklı 3 küme belirlemiştir.

Aksoy (2014)'te yayınladığı yüksek lisans tezinde, kümeleme analizi ile öğrencilerin üstün yetenekli olma durumlarını incelemiş, kümeleme oluşumunda kişilik tipi ve çoklu alanı değişkenlerinin anlamlı bir etkisi olduğu sonucuna ulaşmıştır.

2.3.3.3 Birliktelik kuralları. Bir veri kümesindeki veriler arasındaki ilişkileri belirlemek için belirlenen kurallardır. En çok bilinen birliktelik kuralı, Almanya da bir market üzerinde yapılan analizde, bebek bezi alan müşterilerin aynı zamanda en çok tercih ettikleri 2. ürünün bira olduğunun belirlenmesi ve bunun sonucu marketin iki ürünü yan yana koyarak satış politikasını Veri Madenciliği analizi sonucu değiştirmesidir.

Birliktelik Kurallarının eğitimde kullanımı ile de çeşitli çalışmalar yapılmaktadır. Romeo, Ventura, Pechenizkiy, Baker ve Eds (2010) öğretmenlerin öğrencilerin durumları hakkında karar verirken Birliktelik Kurallarından faydalanabileceklerini ve çeşitli aktivitelerle öğrencilerin eğitimlerini geliştirebileceğini söylemiştir. Hark (2013), yüksek lisans tezinde anket yoluyla hazırlanan veriler üzerinde birliktelik kuralları uygulayarak, akıllı tahta kullanımına yönelik öğrenci tutumları incelenmiştir.

2.3.3.4 Sosyal medya analizi. Ağ içindeki sosyal ilişkileri, dostluk ve benzeri bağlantıları içeren analizlerdir. Bugün Facebook, Twitter gibi büyük sosyal medyalarda yapılan analizler ülke güvenliğine katkı sağlamaktadır. İşbirliği içinde yapılan görevlerde ve iletişim araçları sayesinde kurulan etkileşimlerin incelenmesinde Veri Madenciliği analizleri büyük önem taşır. Beyhan (2014), Sosyal Medya üzerinden Metin Madenciliği ve Duygu

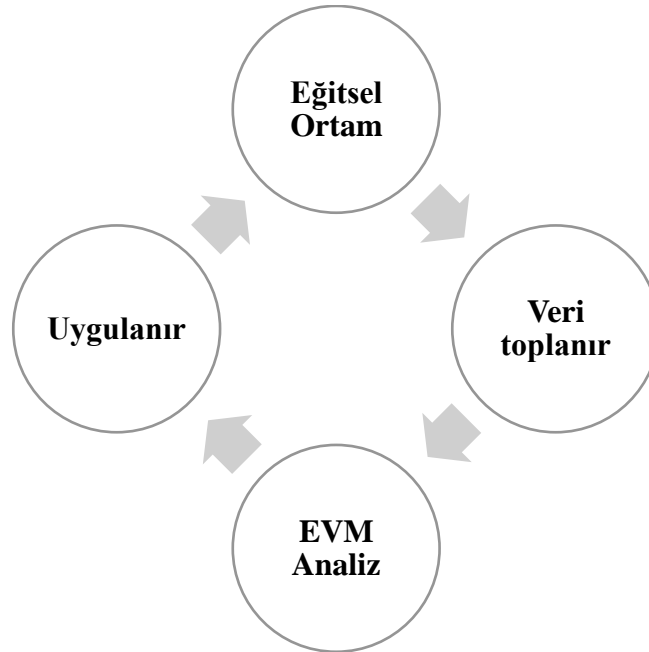
Analizi ile Pazar Değerlendirme adlı yüksek lisans tezinde, GSM sektöründe yer alan üç firmanın Twitter’da adının geçtiği bütün tweetleri, kümeleme yöntemiyle analiz etmiş, çıkan sonuçları firmaların satış verileri ve yürüttükleri pazarlama stratejileri göz önünde bulundurularak değerlendirmiştir.

2.3.3.5 Metin madenciliği. Veri Madenciliğinin metin üzerinde odaklanan bir uzantısıdır. Metinlerin, tartışma panolarının, forumların, sohbetlerin incelenmesinde içeriğin analizinde kullanılır. He (2013)’te yayınlanan makalesinde, çevrimiçi bir eğitim sistemi tarafından otomatik kaydedilen çevrimiçi soruları ve mesajları metin madenciliği kullanarak analiz etmiş, öğrenci ve akran etkileşimi, arasındaki tutarsızlıkları ve benzerlikleri bulmuştur.

2.3.4 Eğitsel Veri Madenciliği uygulama süreci. Eğitsel Veri Madenciliğinde analiz süreci Şekil 5’te olduğu gibi işler. Veri ön işleme sonrası uygun analiz yöntemleri seçilir. Ortaya çıkan sonuçlar yorumlanır ve eğitsel ortamlara geri aktarılır.

Şekil 5

Uygulama Süreci



2.3.5 Eğitsel Veri Madenciliği değerlendirme süreci. Eğitsel Veri Madenciliği uygulamaları sonucu ortaya çıkan sonuçlar yorumlanarak eğitim ortamlarına geri aktarılır. Eğitsel Veri Madenciliğinde değerlendirme aşaması, benzer yaklaşımlarla uğraşan insanlar için bir motivasyon kaynağıdır. Ayrıca en iyi sonuç veren yöntemlerin belirlenmesi için yapılan analizlerin yorumlanıp mutlaka eğitim ortamlarına aktarılmalıdır (Schabacker & Wohlbold, 2002).

Çalışmada öğrenciler durum değişkeni etrafında modellenmiştir. Durum değişkeni başarılı ve başarısız olmak üzere binominal bir değişkendir. Bir öğrencinin okul başarısı, bulunduğu okul, sınıf, ders, demografik ve kişilik özellikleri ve tutumlarına göre belirlenmiş sonuçlara ulaşmada göstermiş olduğu ilerlemedir. Başarısızlık kavramı ise öğrencilerin gelişim düzeylerinin ve yeteneklerinin altında başarı göstermesi ve bu başarısızlığı telafi edememesi olarak tanımlanmıştır (Milli Eğitim Bakanlığı, 2013).

3. Bölüm

Yöntem

Bu bölümde araştırmanın modeli, evren ve örneklem, verilerin toplanması ve çözümlenmesi, Veri Madenciliğinde olması gereken adımlar, veri analizinin nasıl yapıldığı ve tasarım ekranları tanımlanmıştır.

3.1 Araştırmanın Modeli

Bu çalışmada Veri Madenciliği teknikleri kullanılmıştır. Araştırma var olan bir durumu betimlemeye çalışıldığından tarama modelindedir.

3.2 Evren ve Örneklem

Araştırma örneklemini Uludağ Üniversitesi Eğitim Fakültesinde 2015-2016 Öğretim yılında devamlı öğrenci statüsünde bulunan 5363 öğrenciden, Veri önışleme tekniklerinin yapılmasıyla elde edilmiş 4992 öğrenciden oluşmaktadır. Örneklem evrenin tamamını kapsamaktadır.

3.3 Verilerin Toplanması ve Çözümlenmesi

Çalışmada Uludağ Üniversitesi Eğitim Fakültesi 2016 yılı öğrencilerine ait veriler kullanılmıştır. Öğrenci profillerini belirlemek amacıyla seçilen değişkenler; isim, cinsiyet, bölüm, akademik dönem, doğum yılı, doğum yeri, öğretim tipi, geliş şekli, uyruk, durum ve genel ağırlıklı not ortalamasıdır. Çalışmada kullanılacak veriler, gerekli izinler alındıktan sonra 20955 satır ve 37 sütundan oluşan bir Excel tablosu olarak Öğrenci İşlerinden temin edilmiştir.

Veri Madenciliği analizi yapmadan önce ham veriler düzenlenmeli, eksik veriler silinmeli, değişkenler yapılarına göre değerlendirilmelidir (Maimon & Rokach, 2005). Bu işlemlere Veri önışleme adı verilir. Veri önışleme kısmında yapılanlar aşağıda özetlenmiştir.

3.3.1 Veri önışleme. Veri önışleme verinin analiz edilmesini engelleyecek problemlerin temizlenmesi, veri yapısının anlaşılmasını daha kolay kılacak işlemlerdir.

Verilerin ilk halinin bir kısmı Şekil 6’da gözükmektedir. İlk bakıldığında eksik ve anlamsız görünen bu veriler, analiz edilmeden önce bazı aşamalardan geçmektedir.

Şekil 6

Ham Veriler

C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Akademik Donem	ArtıkYil	AyrılışTa rihi	BireyTipi	Bölüm/Pr ogram	BursDuru mu	Cinsiyet	Danışma n	Diploma No	DoğumTa rihi	DoğumYe ri	DoğumYıl	Gano	Gelişme ni	GirisPua rumu	HazırlıkD li	HazırlıkYı l	İletişimBi lgi	İl(Nc)	İç(Nc)	Kayıt hi
1			Öğrenci	Öğretme		Kadın			*****	AZI	1993		ÖSS	0	Seçiniz			SIVAS	I	****
0			Öğrenci	Öğretme		Erkek			*****	KADIKÖY	1992		ÖSS	0	Seçiniz			L	E	****
0			Öğrenci	Öğretme		Kadın			*****	MIDYAT	1984		ÖSS	0	Seçiniz			MARDIN	T	****
1			Öğrenci	Öğretme		Erkek			*****	Ğ	1994		ÖSS	0	Seçiniz			NU	TOSYA	****
1			Öğrenci	Öncesi		Erkek			*****	BELEN	1996		ÖSS	0	Seçiniz			HATAY	BELEN	****
0			Öğrenci	Öğretme		Erkek			*****	ANTAKYA	1994		ÖSS	0	Seçiniz			SIVAS	GÜRÜN	****
1			Öğrenci	Engelliler		Kadın			*****	GÜLNAR	1991		ÖSS	0	Seçiniz			ORDU	UNYE	****
0			Öğrenci	Öğretme		Kadın			*****	YILDIRIM	1995		ÖSS	0	Seçiniz			BURSA	YILDIRIM	****
0			Öğrenci	Öğretme		Erkek			*****	BURSA	1985		ÖSS	0	Seçiniz			Sakarya	Geyve	****
7		*****	Öğrenci	Öğretme	-----	Kadın		20774	*****	ANKARA	1986	2,18	Gelenler	6111	Seçiniz			Ankara	Sincan	****
8		*****	Öğrenci	Öncesi	-----	Kadın			*****	KOCAELI	1987	2,97	Yandal		Seçiniz			Balıkesir	Balıkesir	****
7	False	*****	Öğrenci	Öğretme	-----	Erkek		17272	*****	BERLIN	1984	2,15	ÖSS	363,227	(Hazırlık			HATAY	I	****
4		*****	Öğrenci	Programı		Erkek			*****	BURSA	1989	0,43	-----	0	(Hazırlık					****
4		*****	Öğrenci	Teknoloji		Erkek			*****	L	1985	2	-----	0	(Hazırlık					****
4		*****	Öğrenci	ve	-----	Erkek			*****	ANKARA	1987	1,14	ÖSS	343,636	Seçiniz			BURDUR	MERKEZ	****
4		*****	Öğrenci	k	-----	Kadın			*****				-----	0	Seçiniz					****
4		*****	Öğrenci	Öğretme	-----	Kadın		13650	*****	ANE	1985	2,71	ÖSS	367,93	(Hazırlık			ANE	MERKEZ	****
4		*****	Öğrenci	Sosyoloji		Kadın			*****				-----	0	Seçiniz					****
4		*****	Öğrenci	Öğretme	-----	Kadın			*****	BOR	1986		ÖSS	357,198	Seçiniz			KAYSERİ	INCESU	****
4		*****	Öğrenci	k	-----	Kadın			*****	BİGA	1987		-----	0	(Hazırlık	2005				****
4		*****	Öğrenci	Bilgisi	-----	Erkek			*****	AN	1986		ÖSS	330,563	Seçiniz			KAYSERİ	AN	****

Bu bağlamda Öğrenci bilgi sisteminde çok miktarda eksik olan birey tipi, burs durumu vb. değişkenlerin analize dâhil edilmesi uygun görülmemiş ve çıkartılmıştır. Ayrıca 6698 sayılı Kişisel Verilerin Kullanılması Kanunu (2016) dikkate alınarak TC, iletişim, isim, soy isim ve adres bilgileri silinmiştir.

İkinci öğretimi olan bölümler birinci öğretimle aynı başlıkta birleştirilmiştir. Öğrencilerin öğretim tipi ayrı bir değişkende saklanmıştır.

Eğitim Fakültesinden mezun olmuş ya da ayrılmış kişilerin verileri silinmiş sadece şu anda kayıtlı olan öğrenciler analize tabi tutulmuştur. Burs Durumu, Danışman, Diploma No ve Ayrıntılı doğum tarihi değişkenleri silinmiştir. 1 kişi devamlı gözükmemesine rağmen ayrıldığı için silinmiştir. Veri ön işleme teknikleri sonucu, 5363 öğrenci sayısı 4992’ye düşmüştür.

Değişkenlere analizin daha sağlıklı yapılabilmesi için bazı kodlamalar yapılmış, bu işlemler Tablo 1’de belirtilmiştir.

Tablo 1

Değişken Kodlamaları

Değişken İsmi	Kodlama
Cinsiyet	Kadın, Erkek
Bölüm	Almanca Öğretmenliği, Beden Eğitimi ve Spor Öğretmenliği, Bilgisayar ve Öğretim Teknolojileri Eğitimi Öğretmenliği, Fen Bilgisi Öğretmenliği Fransızca Öğretmenliği, İlköğretim Matematik Öğretmenliği, İngilizce Öğretmenliği Müzik eğitimi Anabilim Dalı, Okul Öncesi Öğretmenliği, Rehberlik ve Psikolojik Danışmanlık Resim-İş Öğretmenliği, Sınıf Öğretmenliği Sosyal Bilgiler Öğretmenliği, Türkçe Öğretmenliği Zihin Engelliler Öğretmenliği
Uyruk	Türkiye Cumhuriyeti=1, Yabancı Uyruklu=2
Doğum Yılı	1990 öncesi doğanlar = 1 1990 sonrası doğanlar = 2
Durum	GANO<2 = Kötü, GANO>2 = İyi
Doğum Yeri	Bursa, Bursa Dışı

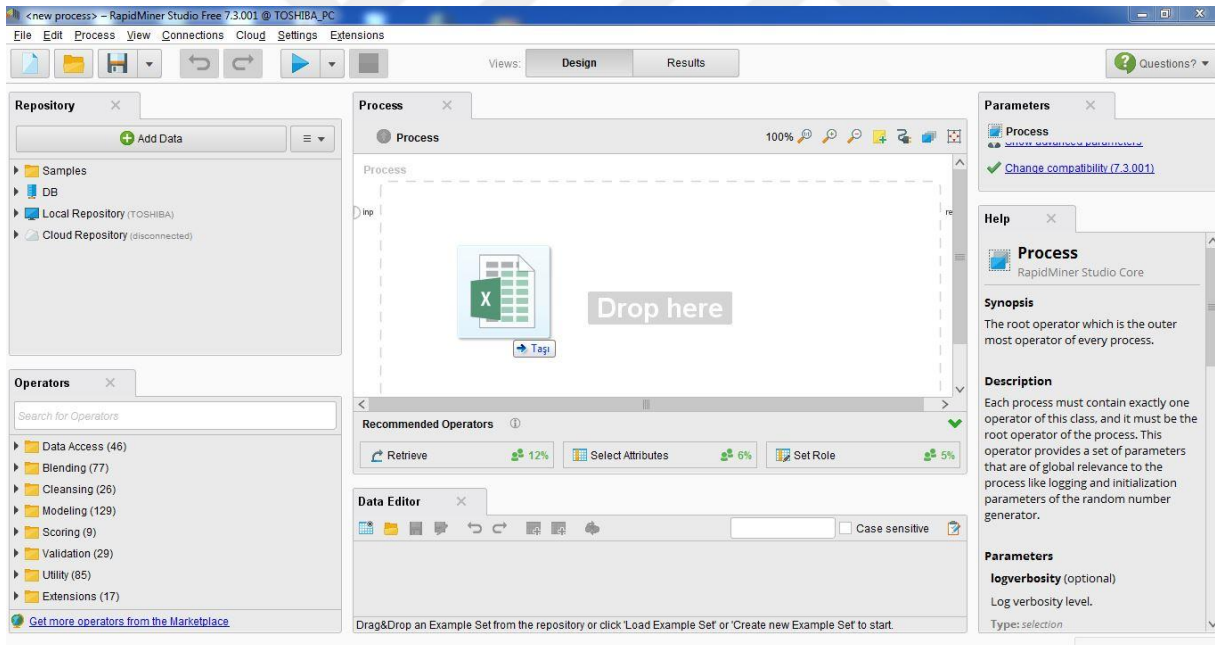
3.3.2 Veri çözümlene teknikleri. Veri Madenciliği bilgisayar programları yardımıyla yapılır. Dünyada en çok kullanılan Veri Madenciliği yazılımları R, Rapidminer, Orange ve WEKA'dır (Akça, 2014). Verilerin analizinde internet üzerinden kolayca indirilen, belli bir sınır içinde kullanımı ücretsiz olan, kullanımı kolay olan, Excel uzantılı dosyaların analizine

izin veren ve en çok kullanılan programlar arasında olan Rapidminer programının son sürümü Rapidminer Studio 7.3 tercih edilmiştir.

3.3.3 Verilerin programa aktarılması. Rapidminer programının yeterince tanınmaması ve analiz yapmak isteyenler için örnek olması açısından verilerin programa aktarılması kısmı resimlerle ve basamaklar şeklinde anlatılmıştır. Rapidminer, Excel verilerini doğrudan kabul eden bir programdır. Örneğin WEKA yazılımında verilerin analiz edilebilmesi için ARFF formatında kaydedilmesi gerekir. Rapidminer yazılımında ara yüze Excel dosyasını sürüklemekle verilerin aktarılma işlemleri başlar. Aşağıda verilerin aktarılma aşamaları verilmiştir.

Şekil 7

Verilerin aktarılması ilk işlem



İlk olarak Veri ön işleme teknikleri sonucu hazırlanan Excel dosyası program ekranının ara yüzüne sürüklenmiştir (Şekil 7). Programın ara yüzü hakkında bilgiler veri analizi kısmında anlatılmıştır.

Şekil 8

Verilerin Aktarılması 1

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sayfa1 Cell range: A:K Select All Define header row: 1

	A	B	C	D	E	F	G	H	I	J
1	Ad	Akademi...	Bölüm/P...	Cinsiyet	DoğumY...	DoğumYılı	Gano	GelişŞekli	Öğretim...	Uyruk
2	BİRGÜL	0.000	Fransızc...	Kadın	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
3	SEMA	0.000	Resim-İ...	Kadın	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
4	ELLEN ...	0.000	Almanca...	Kadın	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
5	MUSA	0.000	Fransızc...	Erkek	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
6	RIDVAN	0.000	Beden E...	Erkek	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
7	ENİS	8.000	Sınıf Öğr...	Erkek	BURSA ...	1.000	2.190	Af İle Gel...	İkinci öğr...	1.000
8	NAZİM	0.000	Fransızc...	Erkek	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
9	AYTEN	0.000	Rehberli...	Kadın	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
10	ESRA	0.000	Fransızc...	Kadın	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
11	KASIM	0.000	İlköğreti...	Erkek	BURSA ...	1.000	0.000	Af İle Gel...	İkinci öğr...	1.000
12	TURGAY	0.000	Fransızc...	Erkek	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
13	FAHRET...	0.000	Almanca...	Erkek	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000
14	HALE	0.000	Müzik Eğ...	Kadın	BURSA ...	1.000	0.000	Af İle Gel...	Örgün ö...	1.000

← Previous → Next ✖ Cancel

Ardından programın verdiği talimatlar izlenir. İlk olarak veriler program tarafından gelen ekranda (Şekil 8) ön izleme olarak gösterilir. Burada isteğe bağlı olarak ilk satır başlık satırı olarak seçilebilir. Veri kümesinin tamamı yerine belli bir kısmı analiz edilmek istenirse, cell range kısmından ayarlanabilir. Ayrıca Excel dosyasında birden fazla sayfa varsa sheet kısmından bu ayarda yapılabilir. Daha sonra ileri tuşuna basılır.

Şekil 9

Verilerin Aktarılması 2

Import Data - Format your columns.

Format your columns.

Date format: MMM d, yyyy h:mm:ss a z Replace errors with missing values

	Ad <i>polynomial</i>	AkademikD... <i>integer</i>	Bölüm/Pro... <i>polynomial</i>	Cinsiyet <i>polynomial</i>	DoğumYeri <i>polynomial</i>	DoğumYılı <i>integer</i>	Change Role
1	BİRGÜL	0	Fransızca Öğret...	Kadın	BURSA DIŞI	1	(^)
2	SEMA	0	Resim-İş Öğretm...	Kadın	BURSA DIŞI	1	(^)
3	ELLEN NEVİN	0	Almanca Öğretm...	Kadın	BURSA DIŞI	1	(^)
4	MUSA	0	Fransızca Öğret...	Erkek	BURSA DIŞI	1	(^)
5	RIDVAN	0	Beden Eğitimi ve ...	Erkek	BURSA DIŞI	1	(^)
6	ENİS	8	Sınıf Öğretmenliği	Erkek	BURSA DIŞI	1	(^)
7	NAZİM	0	Fransızca Öğret...	Erkek	BURSA DIŞI	1	(^)
8	AYTEN	0	Rehberlik ve Psik...	Kadın	BURSA DIŞI	1	(^)
9	ESRA	0	Fransızca Öğret...	Kadın	BURSA DIŞI	1	(^)
10	KASIM	0	İlköğretim Matem...	Erkek	BURSA DIŞI	1	(^)
11	TURGAY	0	Fransızca Öğret...	Erkek	BURSA DIŞI	1	(^)
12	FAHRETTİN	0	Almanca Öğretm...	Erkek	BURSA	1	(^)

no problems.

Previous Next Cancel

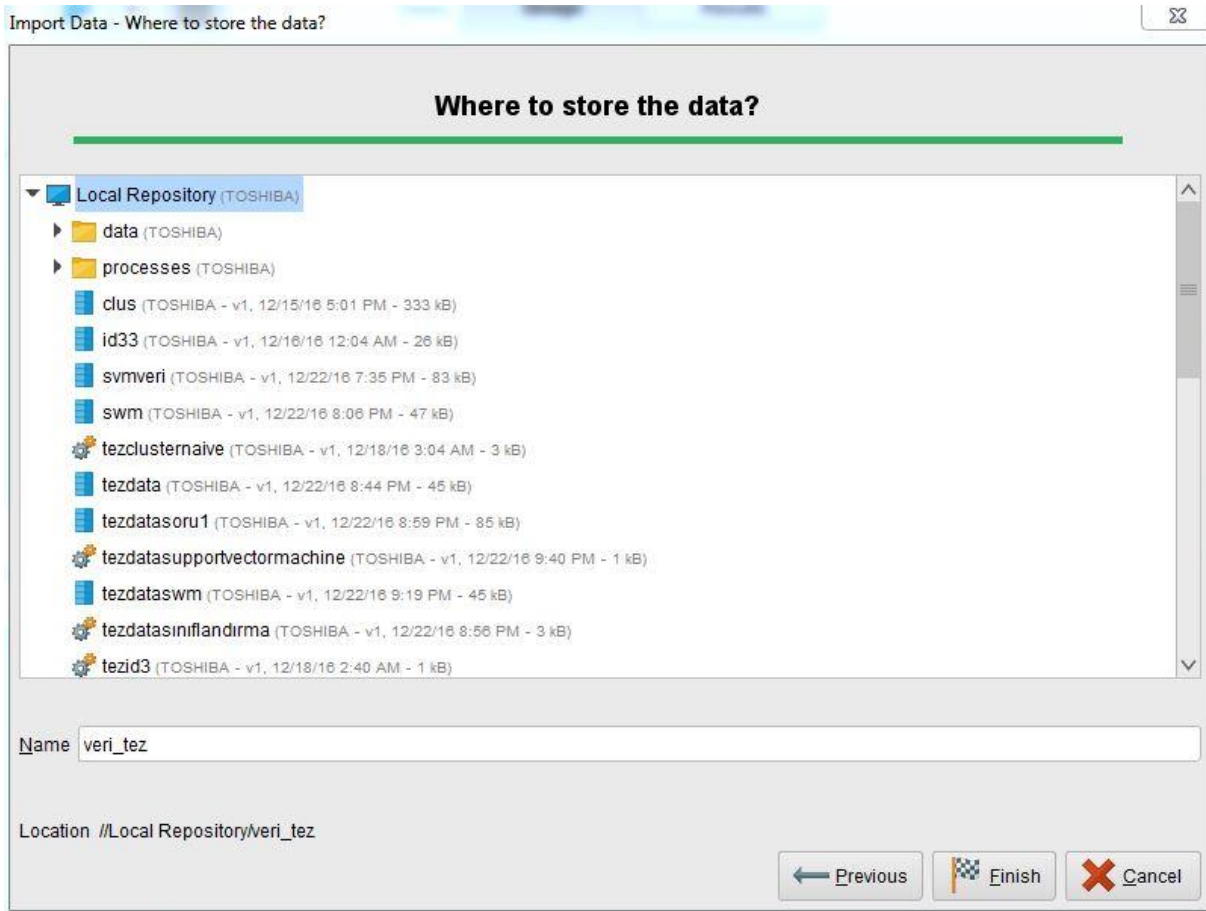
Karşımıza gelen ekranda (Şekil 9), yapılan analize göre genel ağırlıklı not ortalaması ya da durum değişkeni change role kısmından label olarak değiştirilmiştir. Kümeleme ve sınıflandırma gibi analizlerde verileri hangi değişkene göre sınıflandırmak isteniyorsa, o değişken label olarak belirlenmektedir.

Bununla birlikte isim değişkeni kaldırılmış, cinsiyet, doğum yeri, doğum yılı, uyruk ve durum değişkenleri binominal; bölüm, akademik dönem, geliş şekli değişkenleri de polinomial değişken olarak düzenlenmiştir. Değişkenlerin tamamının nominal değişken olarak düzenlenmesinin sebebi, Yapılan sınıflandırma analizinin nümerik değişken kabul etmemesidir.

İkinci araştırma sorusunda kullanılan analizde genel ağırlıklı not ortalaması değişkeni nümerik olarak kalmış ve label değişken olarak belirlenmiştir.

Şekil 10

Verilerin Aktarılması 3



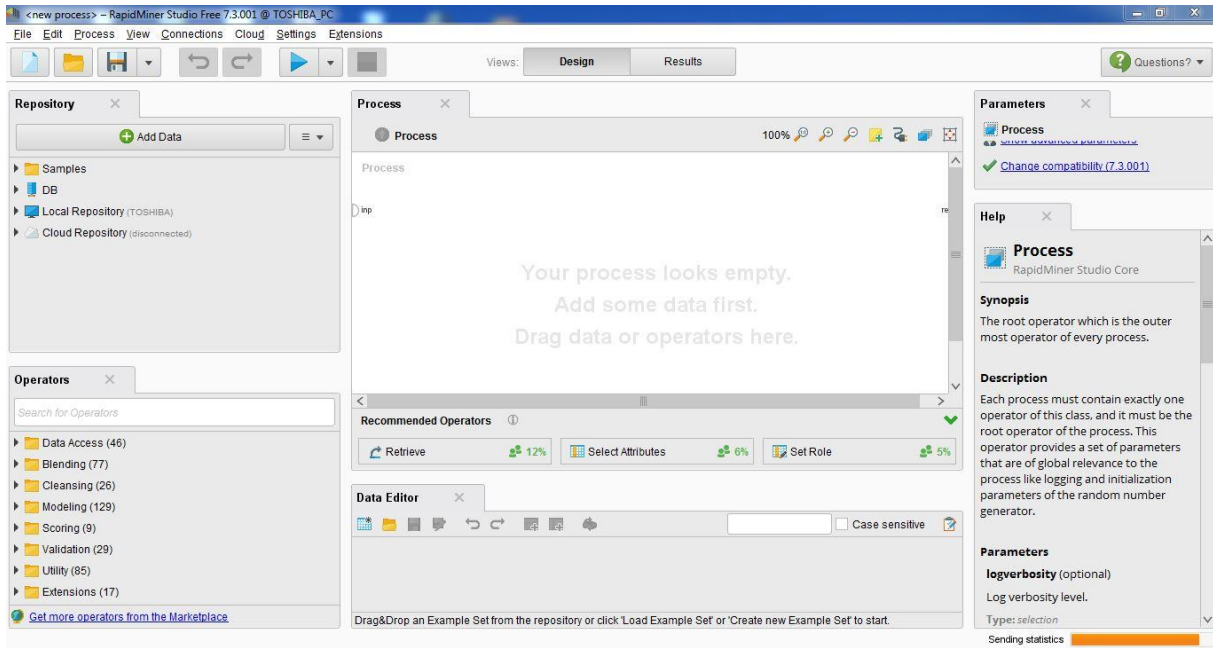
Finish denilerek verilerin programa aktarılma süreci bitirilir.

3.3.4 Veri analizi. Çalışmada Veri Madenciliği yöntemlerinden Naive Bayes Sınıflandırıcısı, Karar Ağaçları ve kural çıkarımı, değişkenlerin önem ağırlıklarının tespitinde bilgi kazanç oranına göre ağırlık hesaplama algoritması kullanılmıştır. Hazırlanan araştırma soruları için yapılan analizlerin hazırlık süreci, uygulamanın ara yüzü ayrıntılı olarak aşağıda paylaşılmıştır.

Rapidminer programının ara yüzü Şekil 11’de gösterilmiştir. Sol tarafta repository kısmında veriler, örnek ve hazır taslaklar; hemen altında operatörler kısmında analizler yer almaktadır. Sağ tarafta ise yapılan analizle ilgili parametreler ve analizin tanımı bulunmaktadır.

Şekil 11

Rapidminer Programı Ara Yüzü

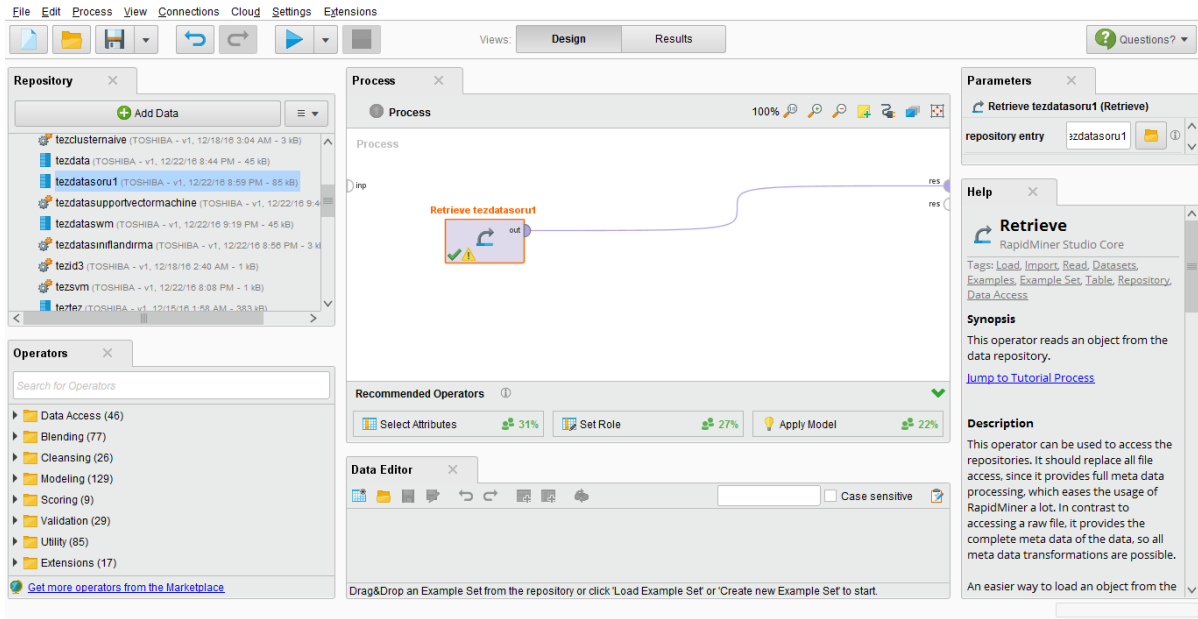


Yapılacak analizler Ortada bulunan process bölümünde hazırlanmaktadır. İlk olarak veri Process bölümüne sürüklenir ardından yapılacak analiz ve gerekli parametreler eklenir bu parametreler birbirine bağlanır. Play tuşuna basılarak analiz gerçekleşir.

3.3.4.1 Birinci araştırma sorusu analizi. Birinci araştırma sorusunun analizi, Excel dosyası ara yüze sürükledikten sonra result sonucuna bağlanarak hazırlanmıştır (Şekil 12).

Şekil 12

Birinci Araştırma Sorusu Tasarım Ekranı

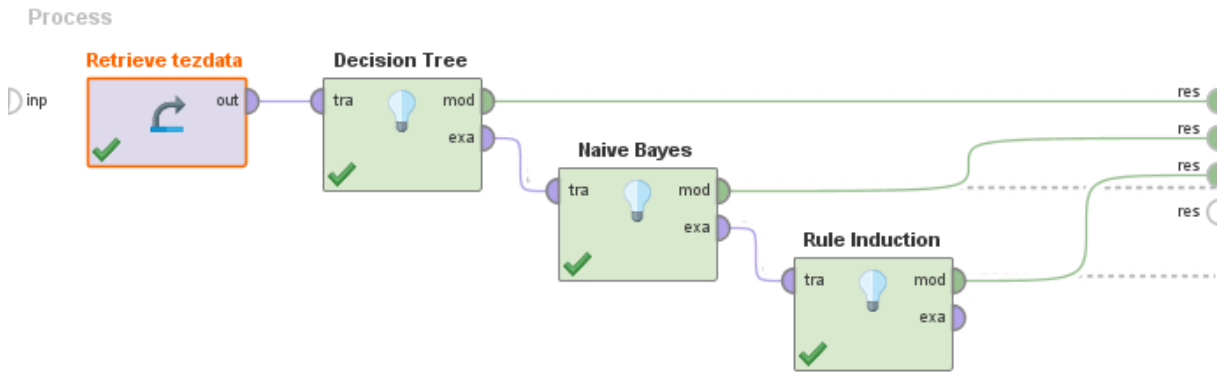


3.3.4.2 İkinci araştırma sorusu analizi. Öğrenci verilerinin sınıflandırılması için

Sınıflandırma yöntemlerinden Naive Bayes Sınıflandırıcısı yöntemi tercih edilmiştir.

Şekil 13

İkinci Araştırma Sorusu Tasarım Ekranı



Yapılan Naive Bayes sınıflandırıcısının tasarım ekranı Şekil 13'te paylaşılmıştır.

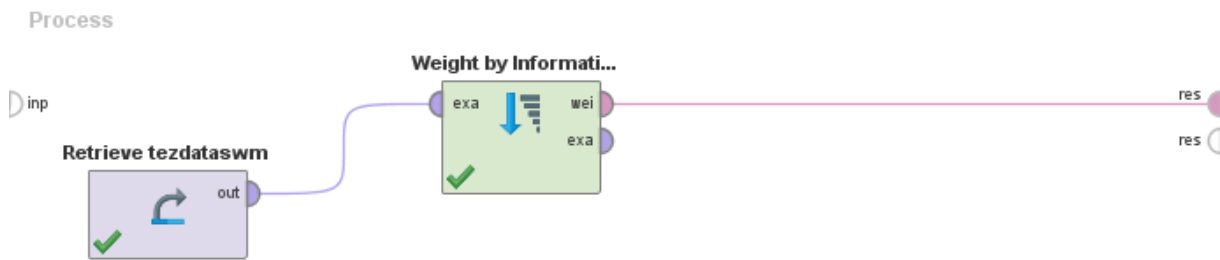
Hazırlanan veri kümesi öncelikle bir karar ağacına bağlanmıştır. Bu analiz sonrası bize bir

karar ağacı çizmesini sağlar. Ardından Naive Bayes parametresi ile Kural Çıkarımı operatörü bağlanarak analiz gerçekleştirilmiştir.

3.3.4.3 Üçüncü araştırma sorusu analizi. Değişkenlerin genel ağırlıklı not ortalaması üzerindeki ağırlığının tespit edilmesi amacıyla bilgi kazanç oranına göre ağırlık algoritması kullanılmıştır.

Şekil 14

Üçüncü Araştırma Sorusu Tasarım Ekranı



Veri Şekil 14'te görüldüğü gibi parametreye ve ardından sonuç kısmına bağlanarak analiz gerçekleştirilmiştir.

4. Bölüm

Bulgular

Veri Madenciliği yaptığı istatistiksel analizlerle birlikte grafiksel, görsel bir sonuç sağlar. Verilerin programa aktarılmasının ardından program verileri ilgili birtakım tanımlayıcı istatistikler sunar. Araştırma sorularının cevaplarını teşkil eden bulgulara geçmezden evvel öğrenci verileri ilgili genel bilgiler verilmiştir.

Bulgular kısmının birinci bölümünde birinci araştırma sorusuna cevap teşkil eden demografik özellikleri ve ders başarısı arasındaki ilişkileri gösteren analizlerin bulguları paylaşılmıştır.

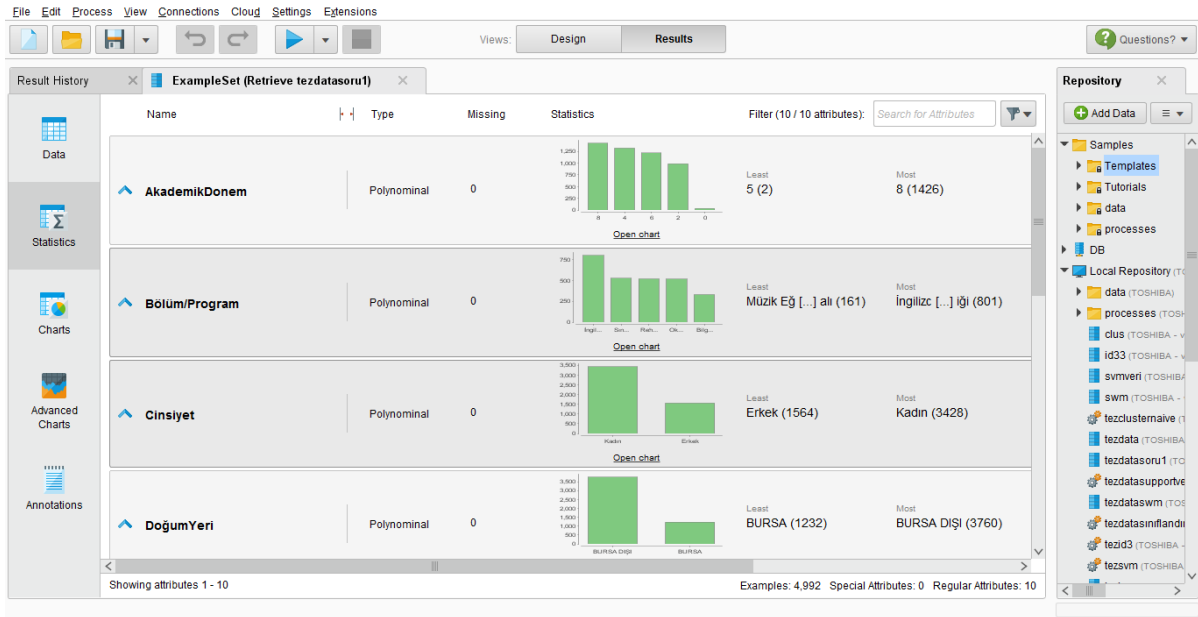
Bulgular kısmının ikinci bölümünde sınıflandırma analizinin sonuçları yer almaktadır. Bu bölüm ikinci araştırma sorusuna cevap teşkil etmiş ve bulunan ilginç bulgular ve ortaya çıkan kurallar paylaşılmıştır.

Bulgular kısmının üçüncü kısmında akademik başarıya etki eden değişkenlerin önem ağılıkları yer almaktadır.

Rapidminer programı veri kümesine ait minimum ve maksimum miktarların yer aldığı ve genel olarak verilerle ilgili fikir edinebileceğimiz tanımlayıcı istatistikler sunar (Şekil 15)

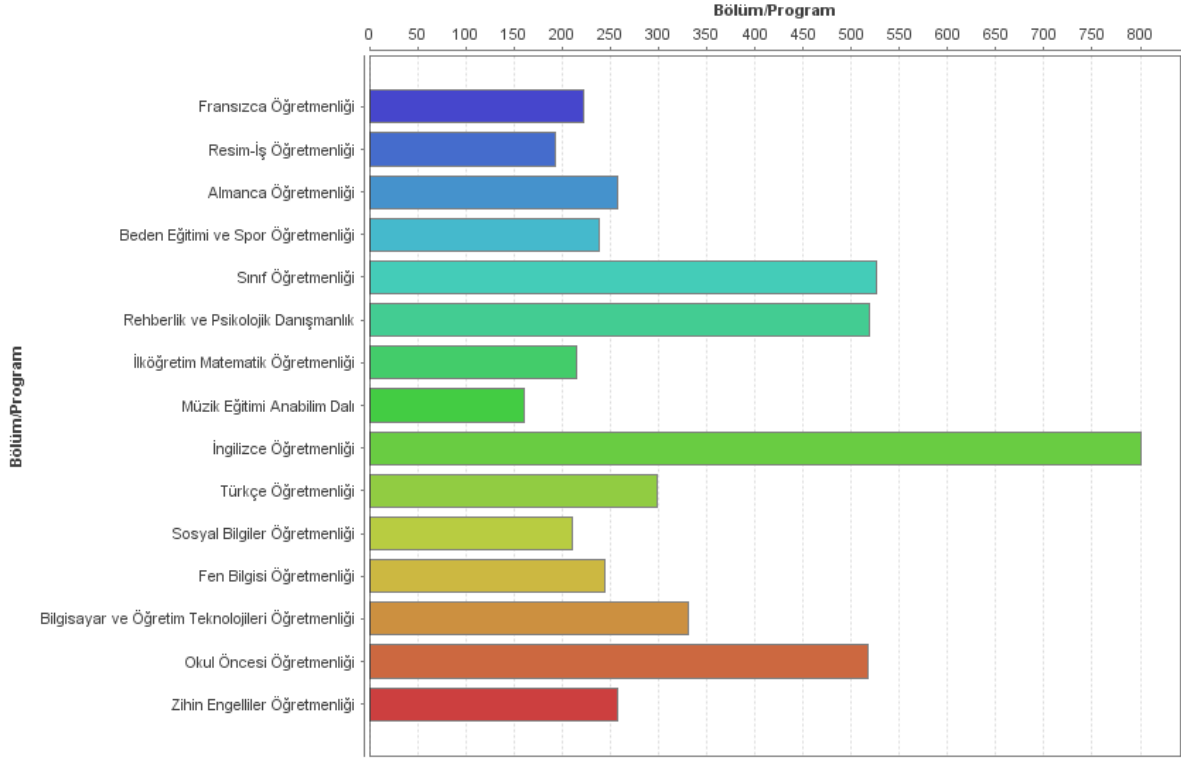
Şekil 15

Tanımlayıcı İstatistikler



Veriler hakkındaki bilgilerin yanı sıra, ekranın solunda görülen Charts kısmından verilerin birbirleriyle ilgili karşılaştırmalar yapılabilir.

Şekil 16

Bölüm Değişkeni

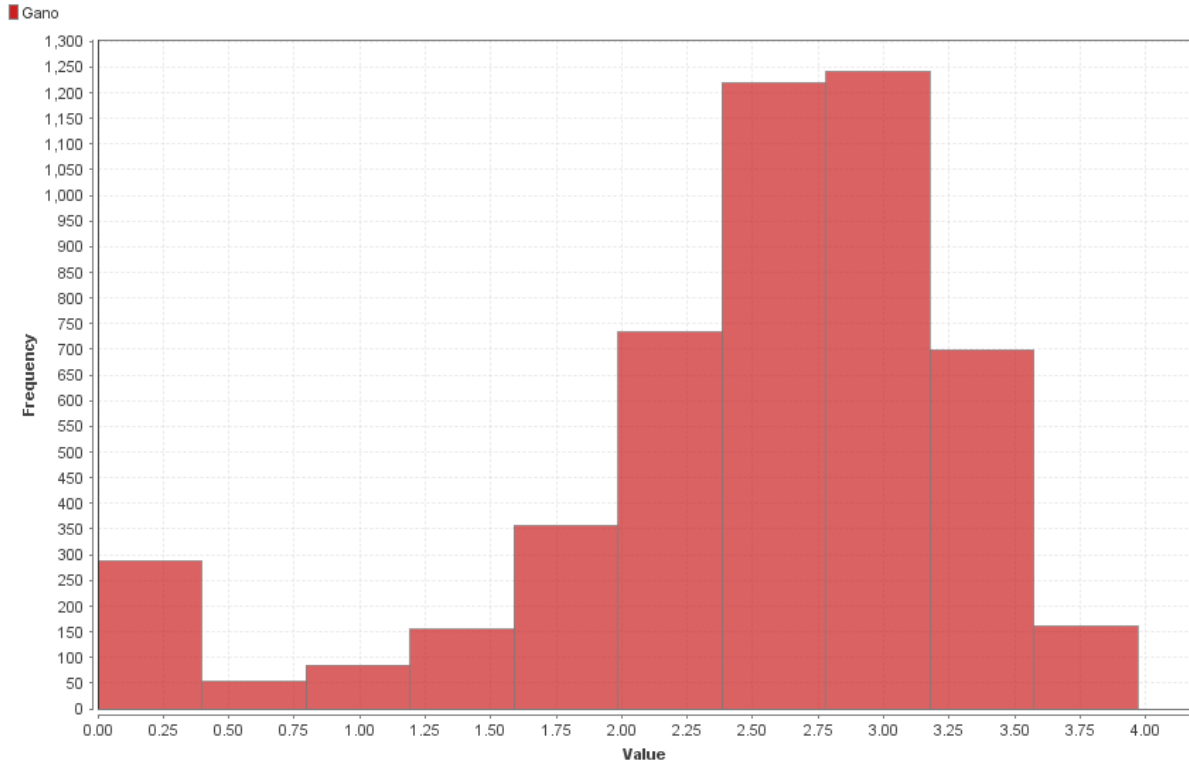
Fakültede bulunan 15 bölüm içerisinde en fazla öğrenci İngilizce Öğretmenliği (801)

bölümünde, en az öğrenci ise Müzik Eğitimi Anabilim dalında (161) yer almaktadır (Şekil

16). Fakültede bir bölümde ortalama 332 öğrenci öğrenim görmektedir.

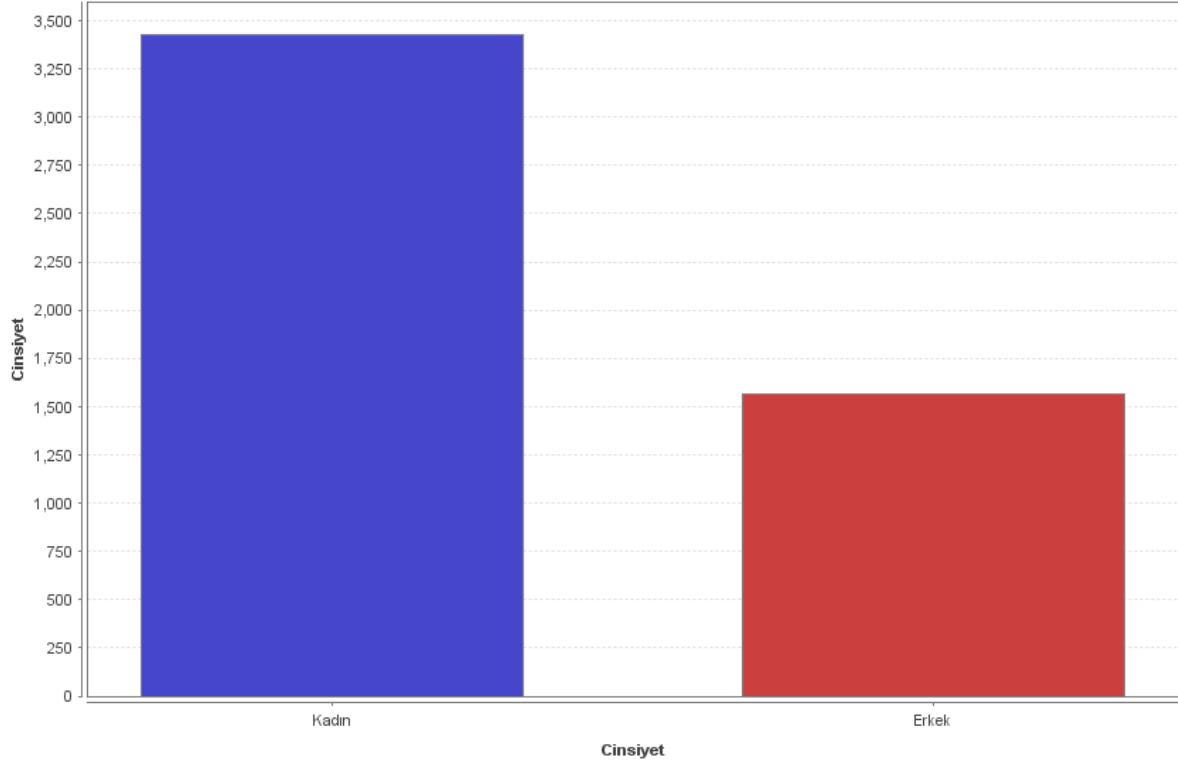
Şekil 17

Not Ortalaması Değişkeni



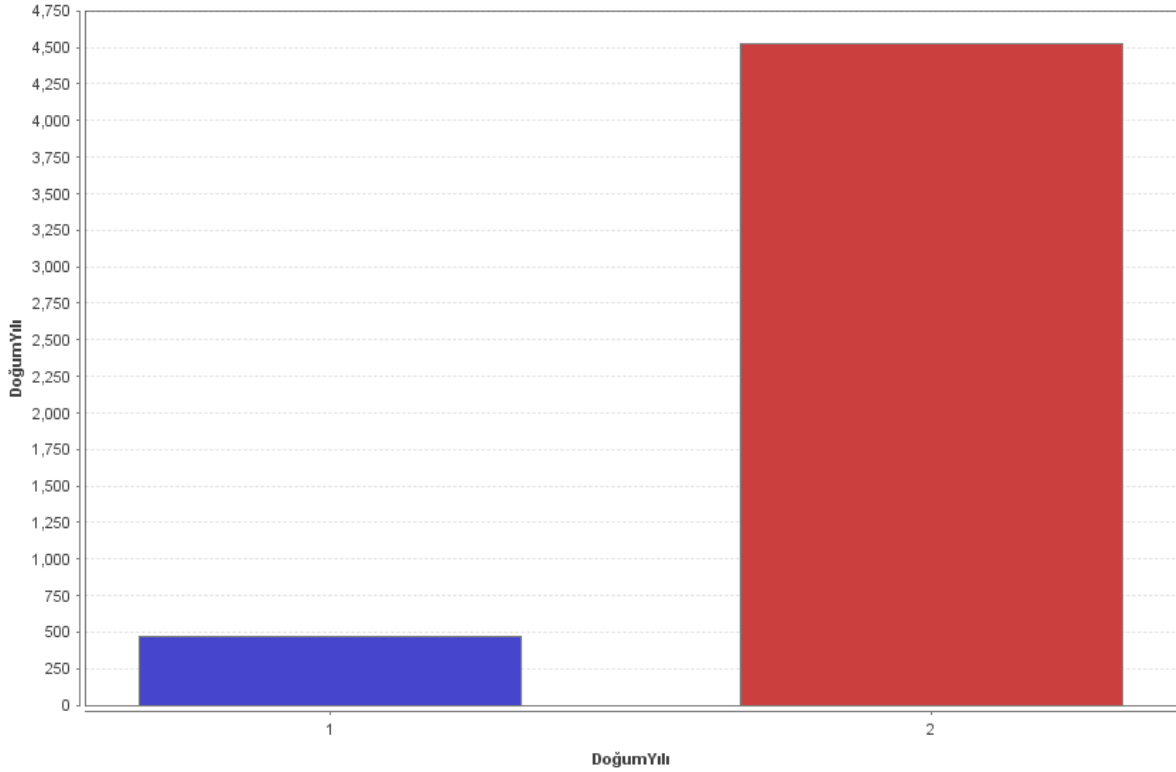
Fakülte de en yüksek ortalama 3,97 ile bir kadın öğrenciye aittir. Tüm öğrencilerin ortalama genel ağırlıklı not ortalaması 2,479 dur. Genel ağırlıklı not ortalamasının fakültedeki dağılımını gösteren Şekil 17’de 254 öğrencinin ortalaması sıfır gözükmemektedir.

Şekil 18

Cinsiyet Değişkeni

Okulda 3478 kadın, 1564 erkek öğrenci bulunmaktadır. Çalışmada bulunan kadın öğrencilerin sayısı Şekil 18’de görüldüğü gibi, erkek öğrencilerin iki katı civarında olduğu söylenebilir.

Şekil 19

Doğum Yılı Değişkeni

Öğrencilerin 4527 adet öğrenci 27 yaşın altındadır. 465 öğrenci ise 1990 öncesi senelerde doğmuştur yani 27 yaşın üzerindedir (Şekil 19).

Diğer değişkenlere bakmak gerekirse, öğrencilerin % 76'sı ÖSS sınavıyla, kalan % 24'ü diğer sınavlarla üniversiteye kayıt yaptırmıştır.

Öğrencilerin % 2'si yabancı uyruklu olup, sırasıyla Bulgaristan, Türkmenistan, Yunanistan ve Almanya uyruklu öğrenciler en fazla öğrenciye sahip ülkelerdir.

İkinci öğretimde 606 öğrenci okumaktadır.

Fakültede en fazla bulunan isim 126 ile "Büşra" dır. Sırayı Merve, Elif, Kübra, Fatma ve Esra isimleri takip etmektedir.

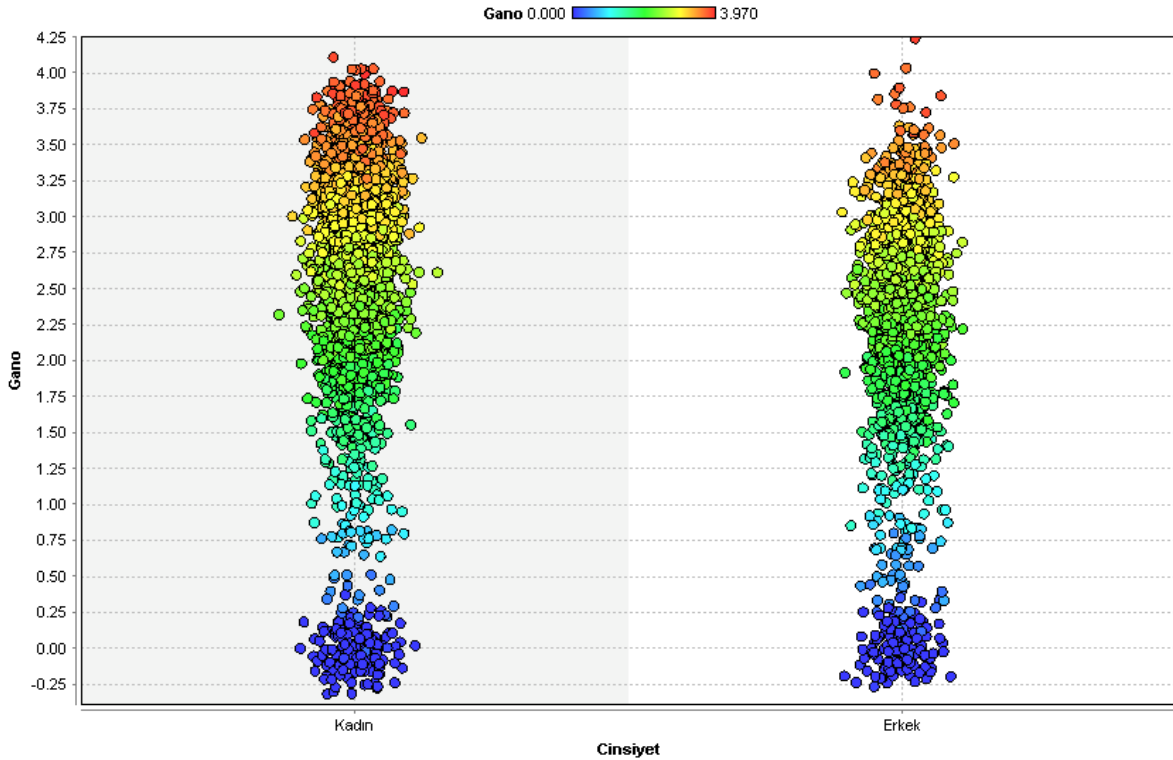
4.1 Öğrenci Modelleme (Student Modeling)

Bu bölümde birinci araştırma sorusuna cevap teşkil eden, tasarım ekranı Veri Analiz bölümünde paylaşılan ve bize öğrencilerin demografik özellikleriyle genel ağırlıklı not ortalamaları arasında nasıl ilişkiler bulunduğunu gösteren analize ait bulgular paylaşılmıştır.

4.1.1 Öğrenci başarısı cinsiyet ilişkisi. Tüm öğrencilerin yer aldığı Şekil 20’de cinsiyet değişkenine göre öğrenci notları maviden kırmızıya renklendirilmiş olarak yer almaktadır. Kız öğrencilerin ortalamaları 3 etrafında toplanırken, erkek öğrencilerin notları 2,5 etrafında toplanmıştır. Sıfır etrafında toplanan öğrencilerde görülmektedir.

Şekil 20

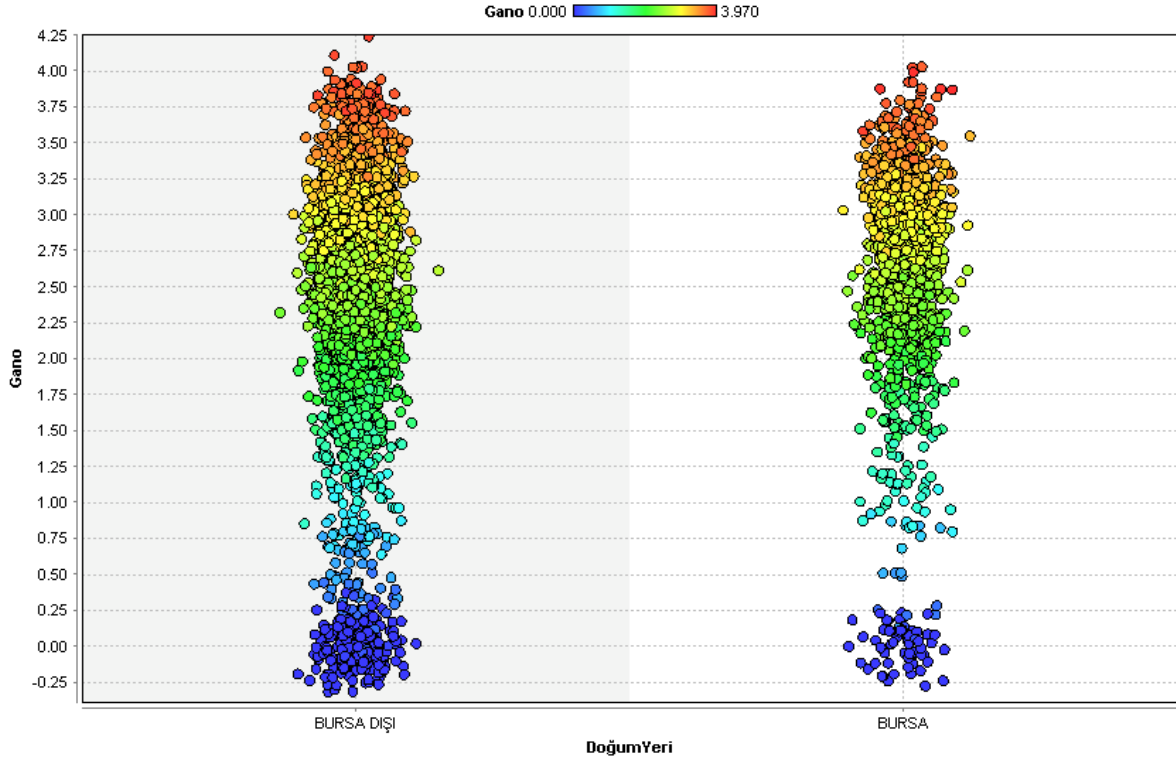
Öğrenci Başarısı Cinsiyet İlişkisi



4.1.2 Öğrenci başarısı doğum yeri ilişkisi. Bursa ve Bursa dışı doğumlu öğrencilerin not ortalamaları ile olan ilişki Şekil 21’de görülmektedir. Örneklem geneli öğrencilerin ortalaması 3 etrafında toplanırken, Bursa içi öğrencilerin bulunduğu grafikte bir not ortalaması etrafında yığılma olmamıştır.

Şekil 21

Öğrenci Başarısı Doğum Yeri İlişkisi

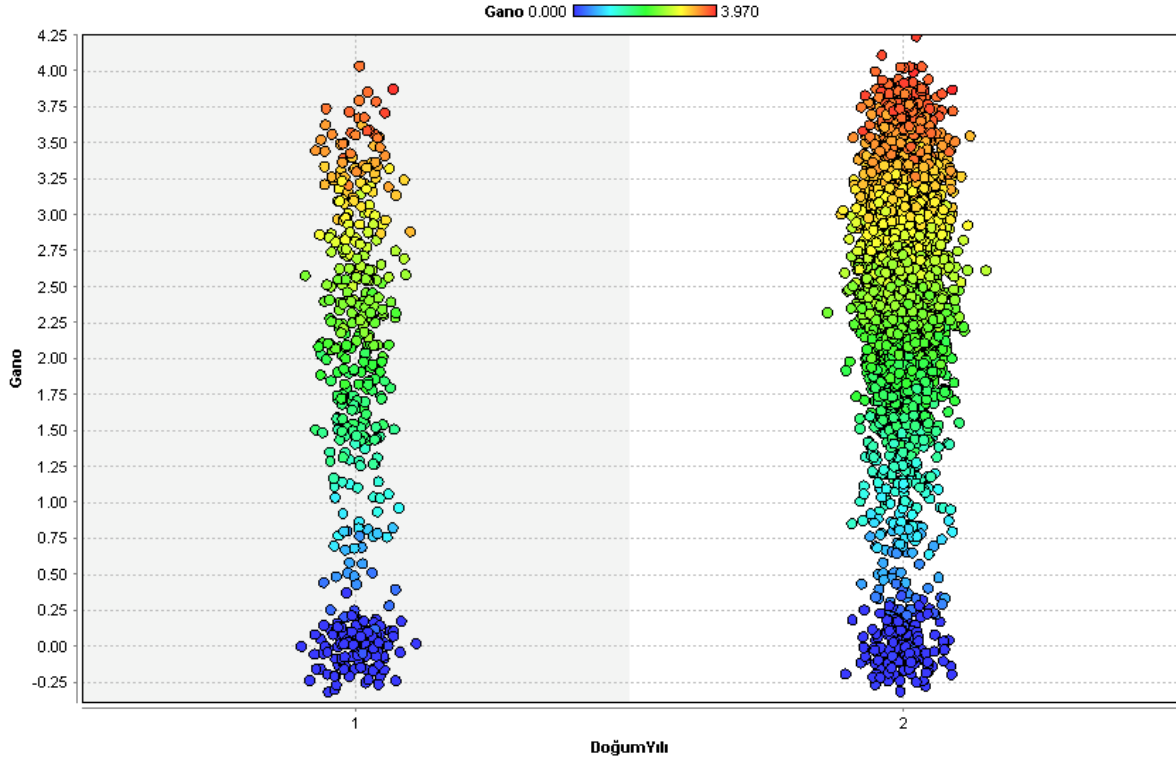


4.1.3 Öğrenci başarısı doğum yılı ilişkisi. 27 yaşın altındaki öğrencilerin yer aldığı

ikinci grupta, not ortalamaları 2,5'in üstünde yoğunlaşırken, 27 yaşın üstünde olan öğrencilerin yer aldığı birinci grupta bir not etrafında yığılma olmamıştır. 2 numaralı öğrencilerin tamamının ortalaması 2,5 iken, 1 numaralı öğrencilerin not ortalaması 1,7'dir. 27 yaş altındaki öğrencilerin büyük çoğunluğu başarılı öğrenci olarak gruplanmasına karşılık, 27 yaşın üstündeki öğrencilerin her başarı seviyesinde benzer sayıda öğrencisi olduğu görülmektedir.

Şekil 22

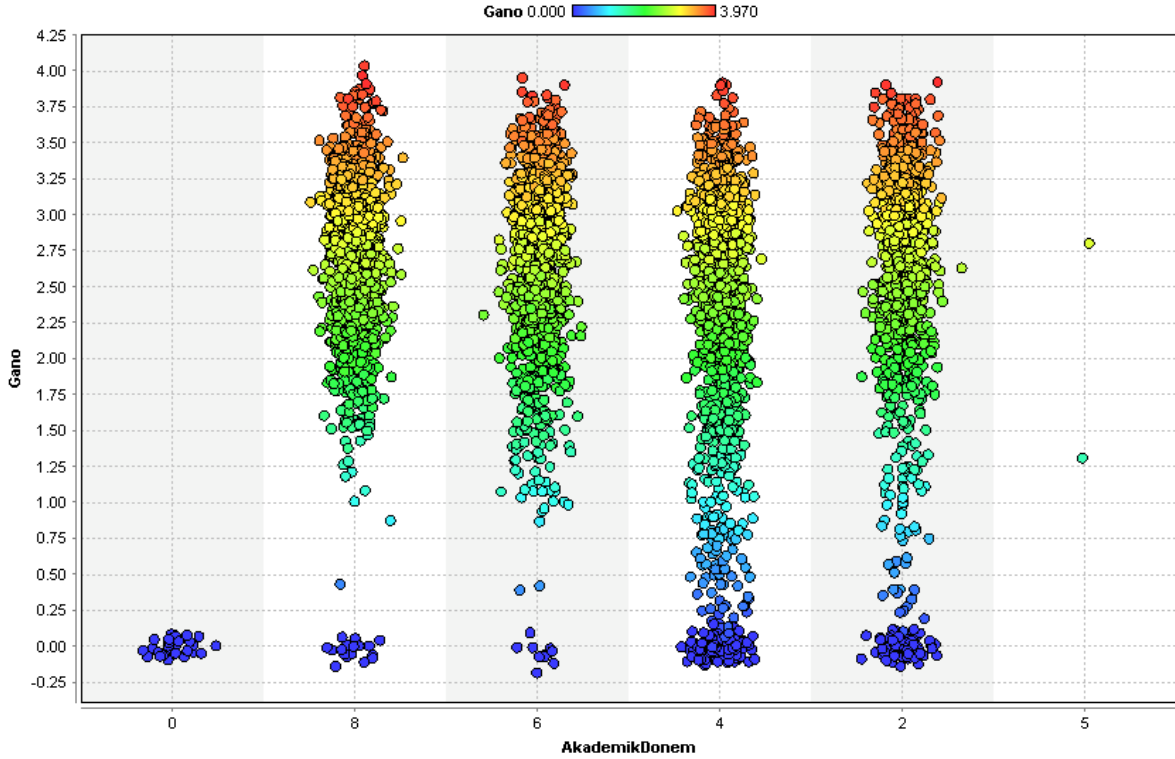
Öğrenci Başarısı Doğum Yılı İlişkisi



4.1.4 Öğrenci başarısı akademik dönem ilişkisi.

Akademik Dönem Öğrenci başarısı arasındaki ilişkiyi gösteren Şekil 23'te, 2 değerinin altında olan öğrencilerin çoğu 1. ve 2. sınıflarda yoğunlaşmış, 3. ve 4. sınıftaki öğrenciler ise genel olarak 2 not ortalamasının üzerinde olup, durum değişkeni iyi olarak sınıflandırılmışlardır. Akademik dönem olarak bakıldığında fakültede en fazla başarısız öğrencinin bulunduğu dönem 2. sınıftır. 1. ve 2. sınıflarda öğrencilerin ortalaması 2- 2,25 etrafında, 3. ve 4. sınıfta okuyan öğrencilerin ortalamaları 2,75 etrafında toplanmaktadır. Yeni kayıt yaptıran öğrencilerin akademik dönemi ve ortalamaları sıfır olarak veri tabanına kaydedildiği için sıfır gözükmemektedir.

Şekil 23

Öğrenci Başarısı Akademik Dönem İlişkisi**4.1.5 Öğrenci başarısı bölüm ilişkisi.** Öğrenci başarısı ve bölüm arasındaki ilişkiyi

gösteren grafikler Şekil 24 ve Şekil 25'te görülmektedir. Bölüm bazında bakıldığında tüm bölümlere aynı anda görsel bir şekilde bakabildiğimiz grafikte, ilginç bulgulara ulaşılmıştır.

Not ortalaması temel alındığında öğrenci başarılarına göre bölümler arasında gözle görülen farklılıklar vardır.

Grafiğe göre Fransızca bölümü diğer bölümlerden farklı şekilde ayrılmakta ve öğrenci başarılarının kötü seviyede toplandığı görülmektedir. Genel olarak 2 not ortalamasının altında yoğunlaşan grafikte, 3 not ortalamasının üzerinde çok az öğrenci bulunmaktadır.

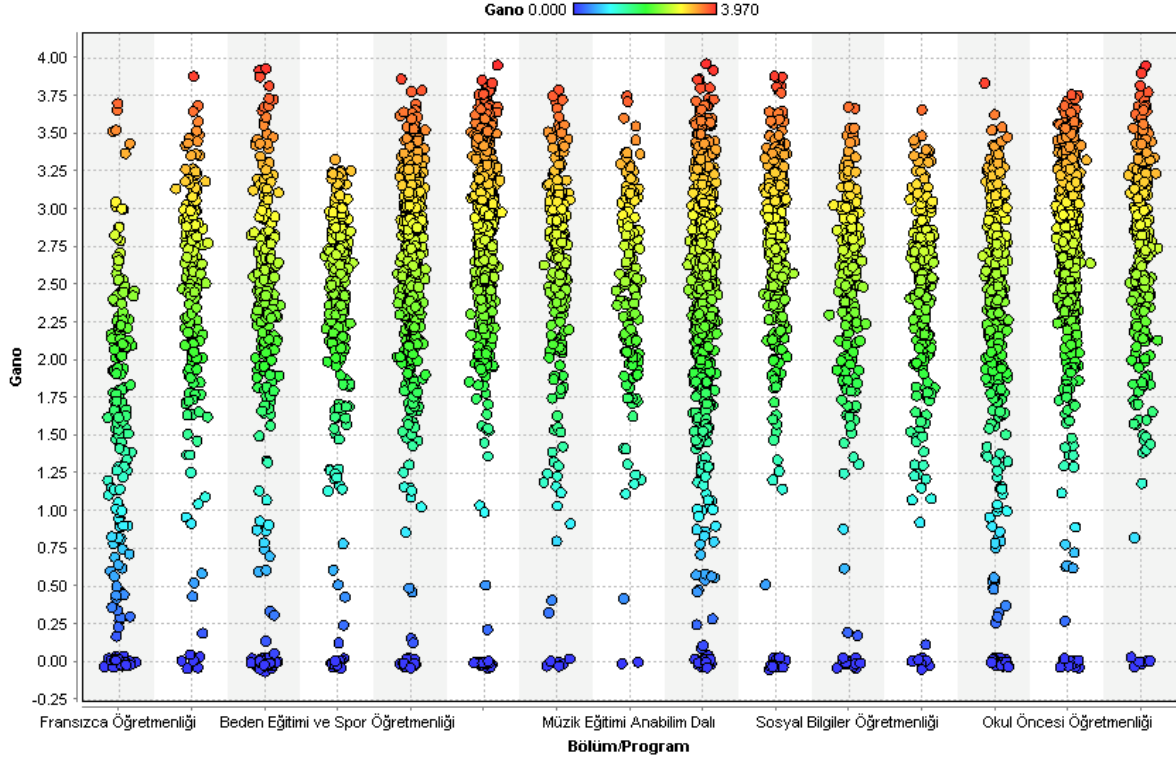
Rehberlik ve Psikolojik Danışmanlık bölümü öğrencileri, başarılı, yüksek ve artan bir grafik seyretmektedir.

Beden eğitimi ve Spor Öğretmenliği bölümünde 3,25'in üzerinde başarıya sahip hiç öğrenci yoktur.

Grafikte bazı öğrenciler farklı şekilde ayrılarak, kendilerini göstermektedirler.

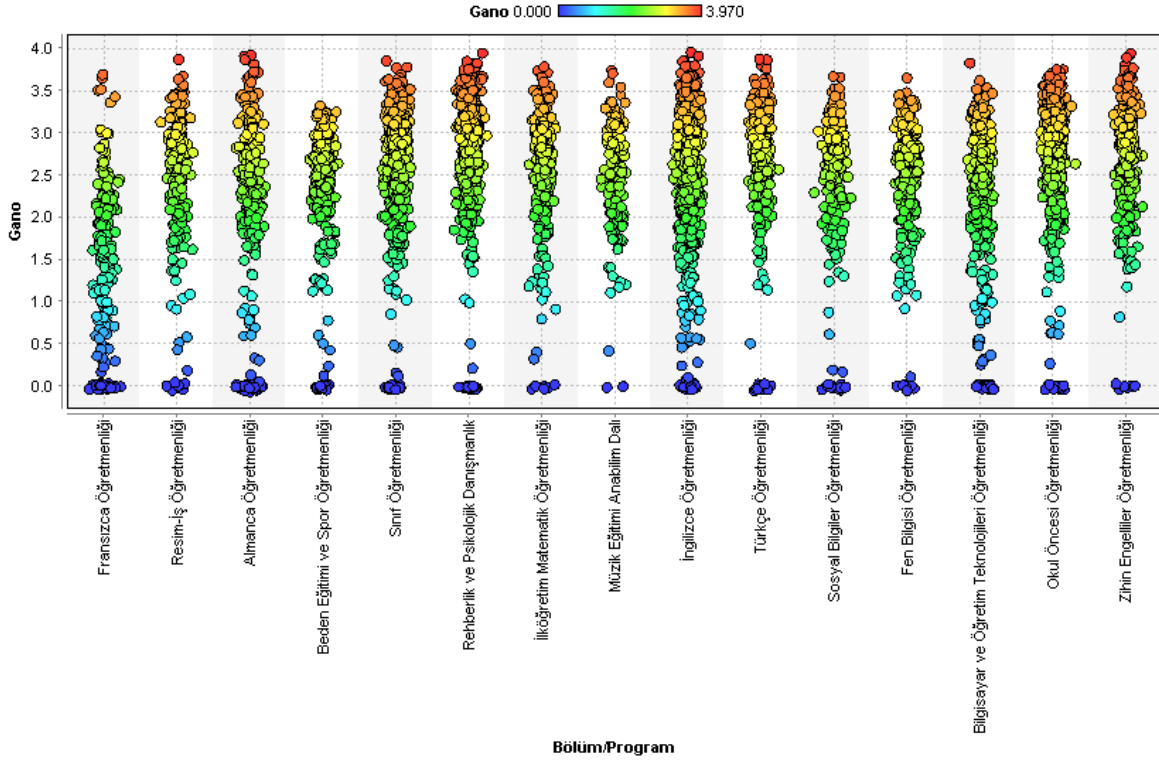
Şekil 24

Öğrenci Başarısı Bölüm İlişkisi



Şekil 25

Öğrenci Başarısı Bölüm İlişkisi 2

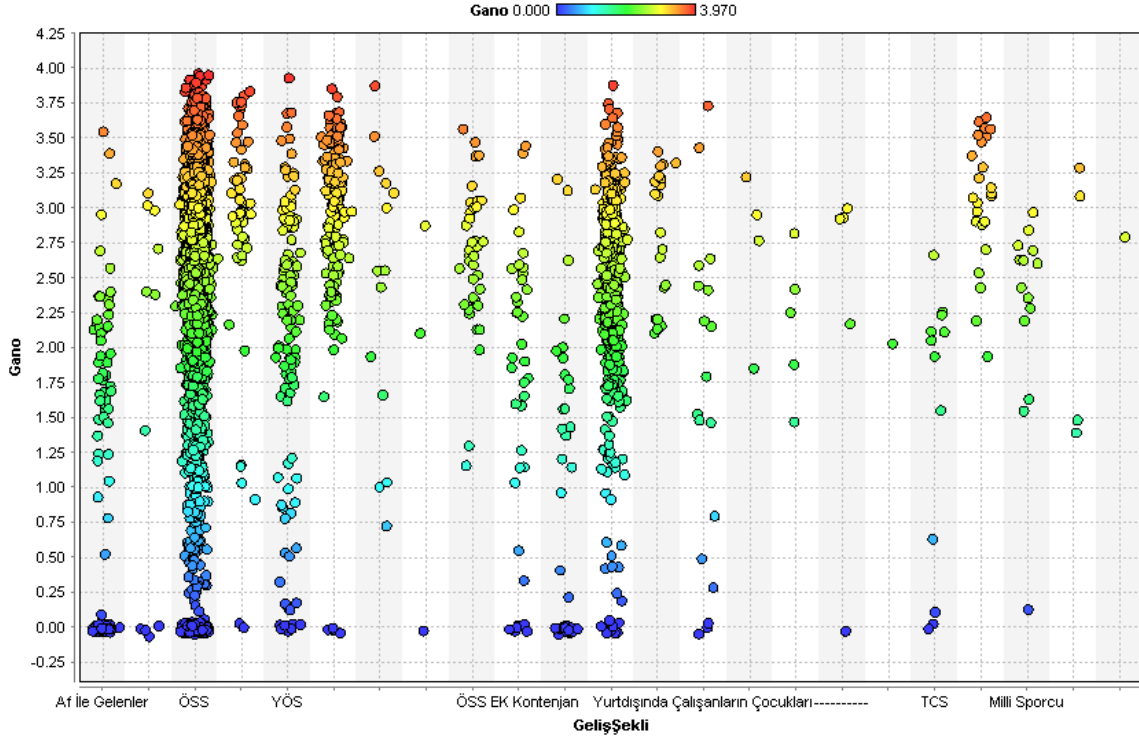


4.1.6 Öğrenci başarısı geliş şekli ilişkisi. Geliş şekli değişkeni öğrencilerin

üniversiteye nasıl kayıt yaptırdığını göstermektedir. Öğrenci başarısı geliş şekli arasındaki ilişkiyi gösteren Şekil 26 ve Şekil 27'de Çift Ana dal yapan, yatay ve dikey geçişle gelen öğrencilerin ortalamaları yüksek, af ile gelen öğrencilerin başarıları alt seviyelerde toplandığı görülmektedir. Öğrencilerin başarıları farklı geliş şekillerine göre değişkenlik göstermektedir.

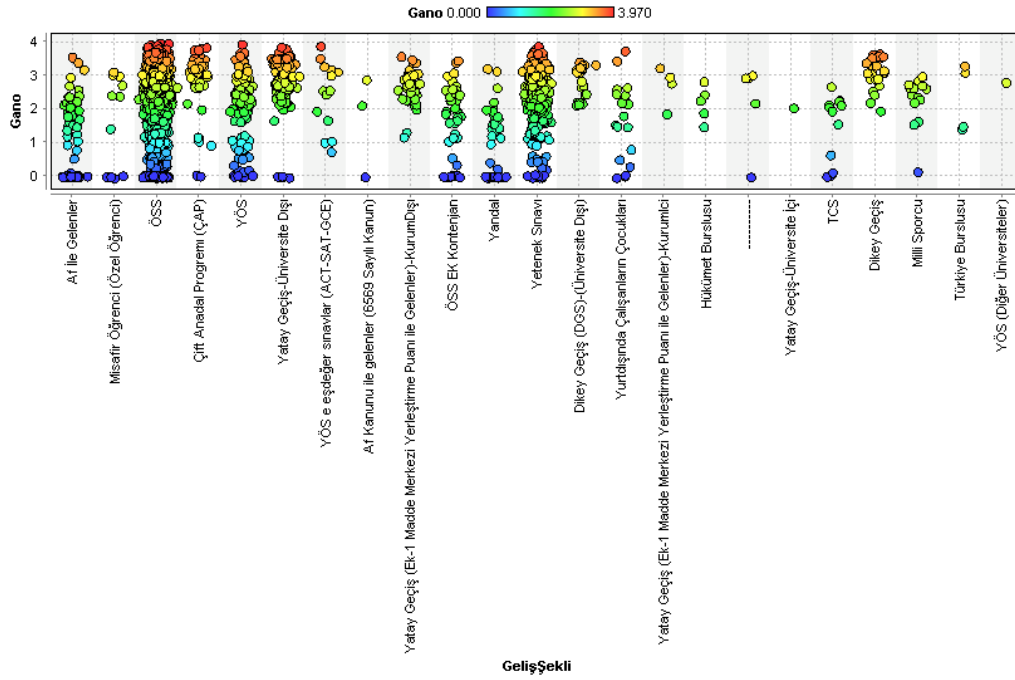
Şekil 26

Öğrenci Başarısı Geliş Şekli İlişkisi



Şekil 27

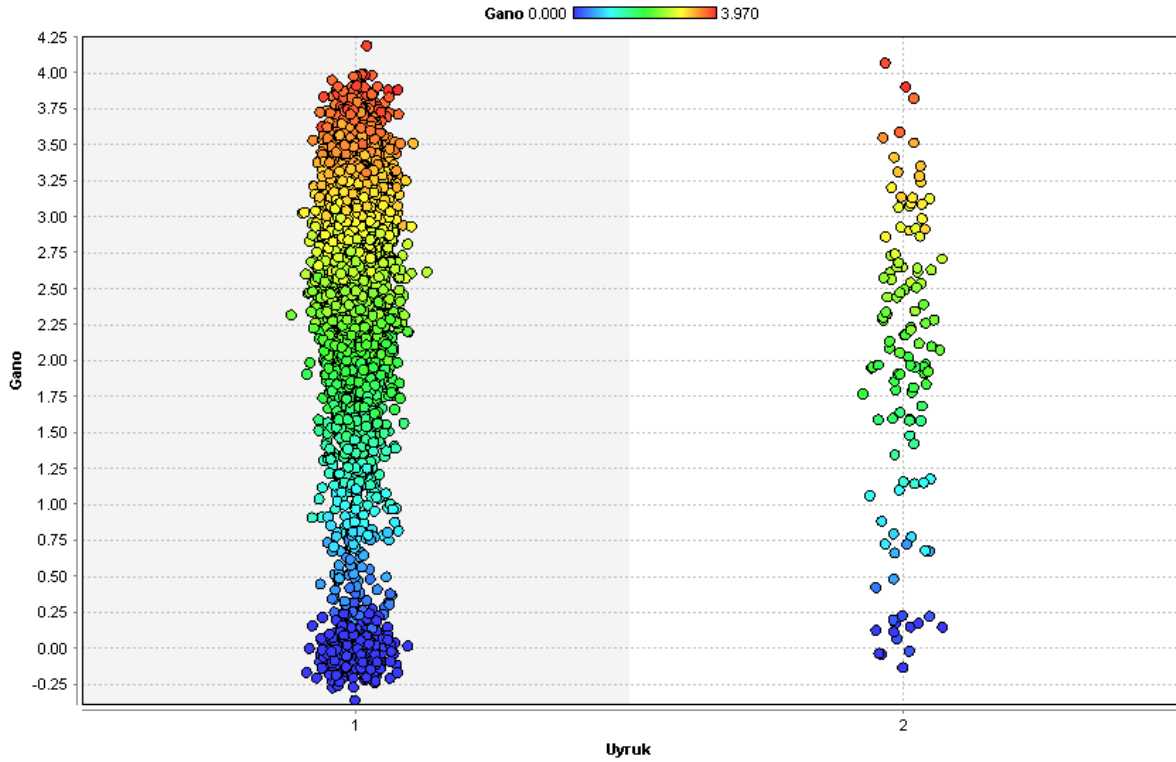
Öğrenci Başarısı Geliş Şekli İlişkisi 2



4.1.7 Öğrenci başarısı uyruk ilişkisi. Öğrenci başarısı uyruk arasındaki ilişkiyi gösteren Şekil 28’de T.C. vatandaşı olan öğrencilerin ortalamaları 2,5 etrafında toplanırken, Yabancı uyruklu öğrencilerin ortalamaları 2 etrafında toplanmış ve normal seyreden bir grafik oluşmuştur.

Şekil 28

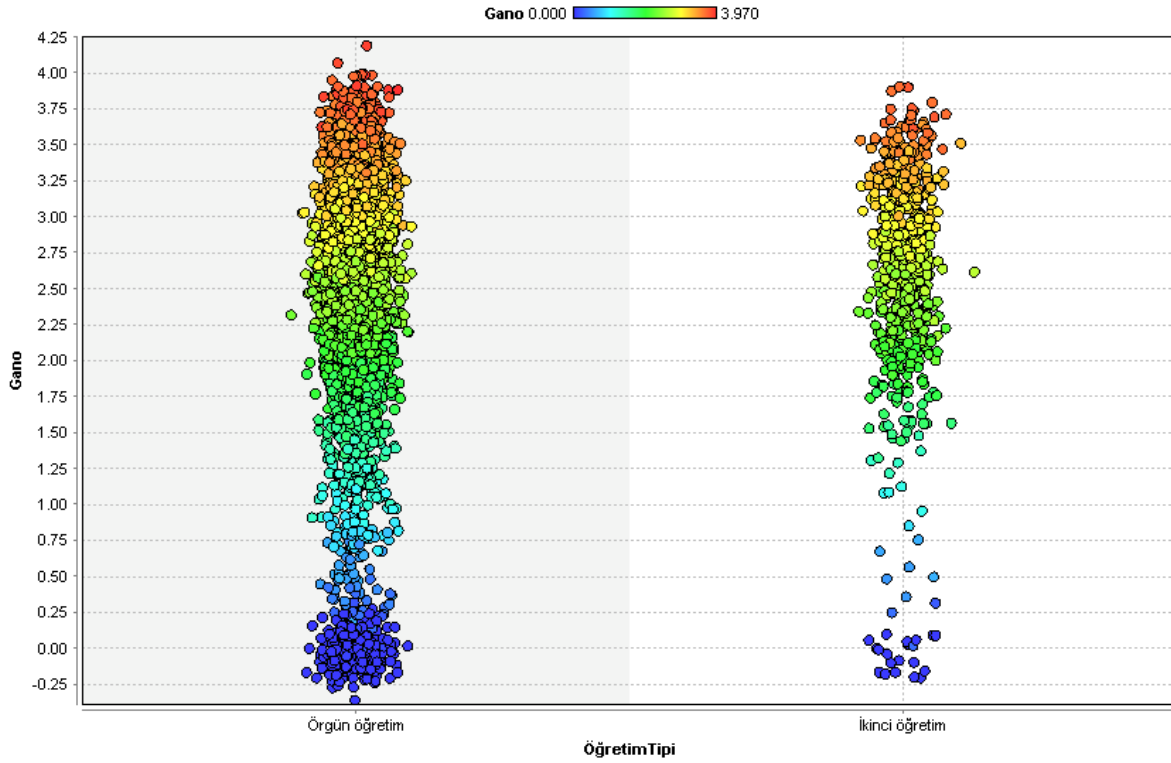
Öğrenci Başarısı Uyruk İlişkisi



4.1.8 Öğrenci başarısı öğretim tipi ilişkisi. Öğrenci başarısı açısından 1. Ve 2. öğretim öğrencileri arasında gözle görülen bir farklılık gözükmemektedir (Şekil 29). 2. Öğretim öğrencilerinde başarı artan bir grafik olarak ilerlememektedir ve belirli bir not etrafında yoğunlaşmamaktadır. Ayrıca 2. öğretimde okuyanların ortalamaları 1. öğretime oranla daha düşük bir ortalama etrafında toplanmaktadır.

Şekil 29

Öğrenci Başarısı Öğretim Tipi İlişkisi

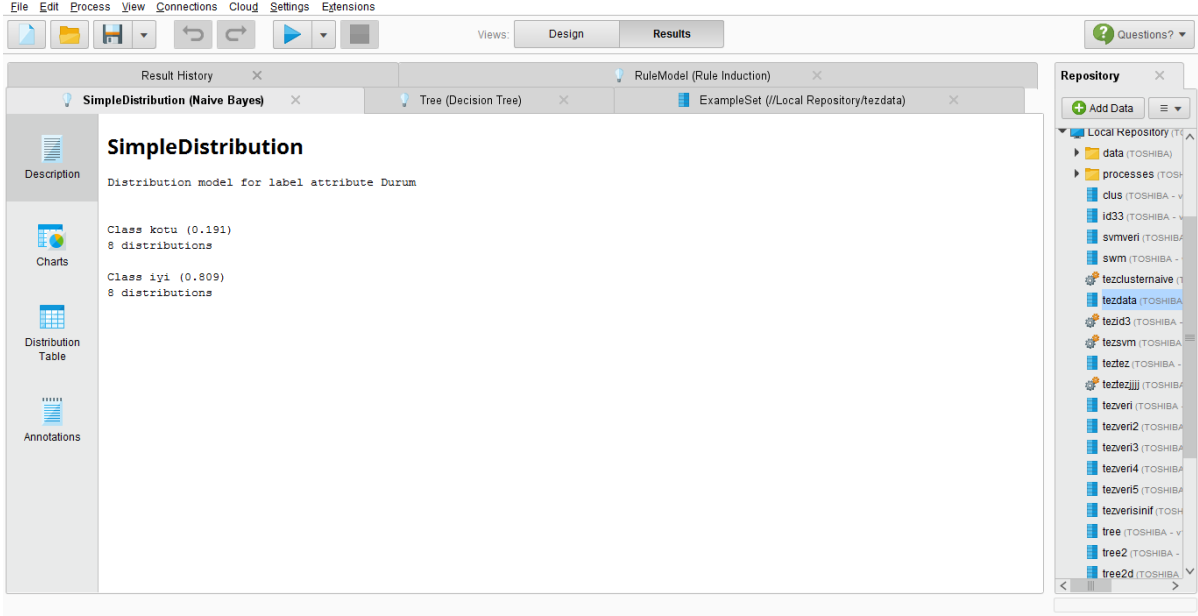


4.2 Sınıflandırma Analizi

Bu bölümde Naive Bayes Sınıflandırma Analizi sonucu ulaşılan bulgulara yer verilmiştir. Bayes algoritmasıyla öğrenciler durum değişkeni altında iyi ve kötü olmak üzere 2 sınıfa ayrılmış ve bu değişkene göre modellenmiştir. Bu analizde basit bir karar ağacı çizilmiş ve önceden bilinmeyen, veriler arasında bulunan gizli ilişkiler Kural Çıkarımı yöntemiyle ortaya çıkarılmıştır.

Öğrencilerin yaklaşık %81'i iyi, %19'u kötü yani 2 not ortalamasının altında sınıflandırılmıştır (Şekil 30).

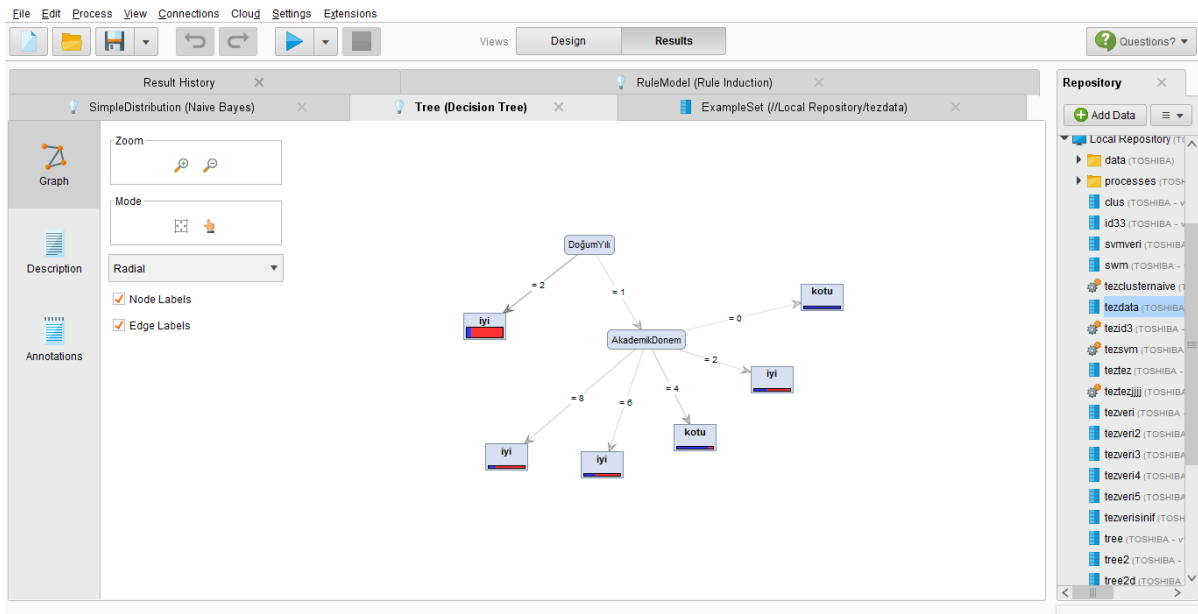
Şekil 30

Naive Bayes Yöntemi Basit Dağılım

Öğrencilerin durumlarının iyi ya da kötü olduğunun belirlenmesi için yapılan analiz sonucu bir karar ağacı oluşturulmuştur. Şekil 31’de görüldüğü gibi karar ağacında iki adet düğüm noktası belirlenmiş olup, bunlardan ilki doğum yılı, ikincisi ise akademik dönem değişkenidir.

Şekil 31

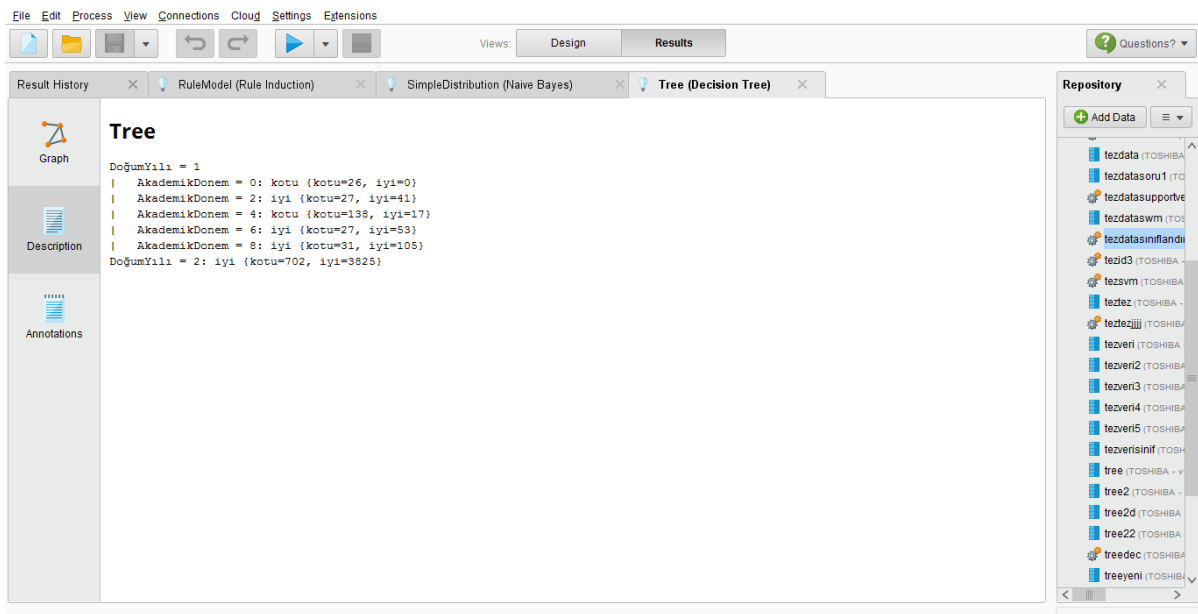
Karar Ağacı



Karar Ağacında belirlenen düğüm noktalarına göre iyi ve kötü öğrenci sayılarını gösteren grafik Şekil 32’de gösterilmiştir.

Şekil 32

Karar Ağacı Düğüm Noktaları



Sınıflandırma Analizi tasarım aşamasında eklenen Kural Çıkarımı parametresi sayesinde birtakım kurallar ortaya çıkmıştır (Şekil 33). Veri kümesi % 50 hata payı olana dek analiz edilmiş ve veriler arasında ilginç bilinmeyen bulgular bulunmuştur.

Şekil 33

Kural Çıkarımı

The screenshot shows a software interface with a menu bar (File, Edit, Process, View, Connections, Cloud, Settings, Extensions) and a toolbar. The main window is titled 'RuleModel (Rule Induction)' and contains the following text:

```

RuleModel
Description
if Cinsiyet = Kadın and DoğumYılı = 2 then iyi (342 / 2815)
if AkademikDonem = 8 and DoğumYılı = 2 then iyi (46 / 332)
if DoğumYılı = 2 and AkademikDonem = 6 then iyi (69 / 259)
if Gelişekli = ÖSS and DoğumYılı = 2 then iyi (188 / 374)
if AkademikDonem = 4 and DoğumYılı = 1 then kötü (138 / 17)
if AkademikDonem = 8 and Gelişekli = ÖSS then iyi (8 / 44)
if Cinsiyet = Kadın and AkademikDonem = 8 then iyi (11 / 38)
if AkademikDonem = 6 and Cinsiyet = Kadın then iyi (12 / 35)
if AkademikDonem = 0 then kötü (29 / 0)
if DoğumYılı = 1 and Cinsiyet = Kadın then iyi (15 / 32)
if Uyrak = 1 and AkademikDonem = 8 then iyi (9 / 22)
if Uyrak = 2 and DoğumYeri = BURSA DIŞI then kötü (21 / 8)
if Gelişekli = Yatay Geçiş (Ek-1 Madde Merkezi Yerleştirme Puanı ile Gelenler)-KurumDışı then iyi (0 / 5)
if Bölüm/Program = Beden Eğitimi ve Spor Öğretmenliği and AkademikDonem = 4 then kötü (12 / 5)
if DoğumYeri = BURSA and Gelişekli = Yetenek Sınavı then iyi (4 / 9)
if Bölüm/Program = Fransızca Öğretmenliği and DoğumYeri = BURSA DIŞI then kötü (4 / 0)
else iyi (38 / 41)

correct: 4210 out of 4982 training examples.

```

The interface also shows a 'Repository' panel on the right with a list of data sources, including 'Local Repository (TOSHIBA)', 'data (TOSHIBA)', 'processes (TOSHIBA)', 'cius (TOSHIBA - v)', 'id33 (TOSHIBA - v)', 'swm (TOSHIBA - v)', 'tezclusternaive (TOSHIBA - v)', 'tezdata (TOSHIBA - v)', 'tezd3 (TOSHIBA - v)', 'teztz (TOSHIBA - v)', 'teztz3 (TOSHIBA - v)', 'teztz4 (TOSHIBA - v)', 'teztz5 (TOSHIBA - v)', 'teztz6 (TOSHIBA - v)', 'teztz7 (TOSHIBA - v)', 'teztz8 (TOSHIBA - v)', 'teztz9 (TOSHIBA - v)', 'teztz10 (TOSHIBA - v)', 'teztz11 (TOSHIBA - v)', 'teztz12 (TOSHIBA - v)', 'teztz13 (TOSHIBA - v)', 'teztz14 (TOSHIBA - v)', 'teztz15 (TOSHIBA - v)', 'teztz16 (TOSHIBA - v)', 'teztz17 (TOSHIBA - v)', 'teztz18 (TOSHIBA - v)', 'teztz19 (TOSHIBA - v)', 'teztz20 (TOSHIBA - v)', 'teztz21 (TOSHIBA - v)', 'teztz22 (TOSHIBA - v)', 'teztz23 (TOSHIBA - v)', 'teztz24 (TOSHIBA - v)', 'teztz25 (TOSHIBA - v)', 'teztz26 (TOSHIBA - v)', 'teztz27 (TOSHIBA - v)', 'teztz28 (TOSHIBA - v)', 'teztz29 (TOSHIBA - v)', 'teztz30 (TOSHIBA - v)', 'teztz31 (TOSHIBA - v)', 'teztz32 (TOSHIBA - v)', 'teztz33 (TOSHIBA - v)', 'teztz34 (TOSHIBA - v)', 'teztz35 (TOSHIBA - v)', 'teztz36 (TOSHIBA - v)', 'teztz37 (TOSHIBA - v)', 'teztz38 (TOSHIBA - v)', 'teztz39 (TOSHIBA - v)', 'teztz40 (TOSHIBA - v)', 'teztz41 (TOSHIBA - v)', 'teztz42 (TOSHIBA - v)', 'teztz43 (TOSHIBA - v)', 'teztz44 (TOSHIBA - v)', 'teztz45 (TOSHIBA - v)', 'teztz46 (TOSHIBA - v)', 'teztz47 (TOSHIBA - v)', 'teztz48 (TOSHIBA - v)', 'teztz49 (TOSHIBA - v)', 'teztz50 (TOSHIBA - v)', 'teztz51 (TOSHIBA - v)', 'teztz52 (TOSHIBA - v)', 'teztz53 (TOSHIBA - v)', 'teztz54 (TOSHIBA - v)', 'teztz55 (TOSHIBA - v)', 'teztz56 (TOSHIBA - v)', 'teztz57 (TOSHIBA - v)', 'teztz58 (TOSHIBA - v)', 'teztz59 (TOSHIBA - v)', 'teztz60 (TOSHIBA - v)', 'teztz61 (TOSHIBA - v)', 'teztz62 (TOSHIBA - v)', 'teztz63 (TOSHIBA - v)', 'teztz64 (TOSHIBA - v)', 'teztz65 (TOSHIBA - v)', 'teztz66 (TOSHIBA - v)', 'teztz67 (TOSHIBA - v)', 'teztz68 (TOSHIBA - v)', 'teztz69 (TOSHIBA - v)', 'teztz70 (TOSHIBA - v)', 'teztz71 (TOSHIBA - v)', 'teztz72 (TOSHIBA - v)', 'teztz73 (TOSHIBA - v)', 'teztz74 (TOSHIBA - v)', 'teztz75 (TOSHIBA - v)', 'teztz76 (TOSHIBA - v)', 'teztz77 (TOSHIBA - v)', 'teztz78 (TOSHIBA - v)', 'teztz79 (TOSHIBA - v)', 'teztz80 (TOSHIBA - v)', 'teztz81 (TOSHIBA - v)', 'teztz82 (TOSHIBA - v)', 'teztz83 (TOSHIBA - v)', 'teztz84 (TOSHIBA - v)', 'teztz85 (TOSHIBA - v)', 'teztz86 (TOSHIBA - v)', 'teztz87 (TOSHIBA - v)', 'teztz88 (TOSHIBA - v)', 'teztz89 (TOSHIBA - v)', 'teztz90 (TOSHIBA - v)', 'teztz91 (TOSHIBA - v)', 'teztz92 (TOSHIBA - v)', 'teztz93 (TOSHIBA - v)', 'teztz94 (TOSHIBA - v)', 'teztz95 (TOSHIBA - v)', 'teztz96 (TOSHIBA - v)', 'teztz97 (TOSHIBA - v)', 'teztz98 (TOSHIBA - v)', 'teztz99 (TOSHIBA - v)', 'teztz100 (TOSHIBA - v)'. The 'teztz100' item is selected.

Bu kurallar;

- Eğer bir öğrenci kadınsa ve 1990 sonrası doğmuşsa durum değişkeni iyi olarak sınıflandırılmıştır.
- Eğer bir öğrenci 4. Sınıfta okuyorsa ve 1990 sonrası doğmuşsa durum değişkeni iyi olarak sınıflandırılmıştır.
- Eğer bir öğrenci 1990 sonrası doğmuşsa ve 3. sınıfta okuyorsa durum değişkeni iyi olarak sınıflandırılmıştır.
- Eğer bir öğrenci ÖSS ile yerleşmişse ve 1990 sonrası doğmuşsa durum değişkeni iyi olarak sınıflandırılmıştır.
- Eğer bir öğrenci ikinci sınıftaysa ve 1990 öncesi doğmuşsa durum değişkeni kötü olarak sınıflandırılmıştır.

- Eđer bir öđrenci son sınıfta okuyorsa ve ÖSS ile kayıt yaptırmıřsa durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrenci kadınsa ve son sınıfta okuyorsa durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrenci 3. sınıfta okuyorsa ve kadında durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrencinin akademik dönem deđiřkeni sıfırda durum deđiřkeni kötü olarak sınıflandırılmıřtır.
- Eđer bir öđrenci 1990 öncesi doğmuřsa ve kadınsa durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrenci Türkiye Cumhuriyeti vatandařı ve son sınıfta durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrenci yabancı uyrukluysa ve doğum yeri Bursa dıřı ise durum deđiřkeni kötü olarak sınıflandırılmıřtır.
- Eđer bir öđrenci yatay geçiřle kayıt yaptırmıřsa durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrenci Beden Eđitimi ve Spor Öđretmenliđi Bölümünde okuyorsa ve 2. Sınıfta okuyorsa durum deđiřkeni kötü olarak sınıflandırılmıřtır.
- Eđer bir öđrenci Bursa doğumlu ve yetenek sınavı ile kayıt yaptırdıysa durum deđiřkeni iyi olarak sınıflandırılmıřtır.
- Eđer bir öđrenci Fransızca Öđretmenliđi bölümünde okuyorsa ve Bursa dıřında doğmuřsa kötü deđilse iyi olarak sınıflandırılmıřtır.

4.3 Akademik Başarıya Etki Eden Değişkenlerin Önem Ağırlıkları

Bu bölümde bilgi kazanç oranına göre ağırlık algoritması ile tespit edilen, öğrencilerin iyi ya da kötü olarak sınıflandırılmasında verilerin önem ağırlıkları paylaşılmıştır. Şekil 35'te görüldüğü gibi genel ağırlıklı not ortalamasına etki eden en önemli değişken bölüm; ardından akademik dönem ve sırasıyla geliş şekli, doğum yılı, cinsiyet, uyruk, doğum yeri ve öğretim tipi değişkenleridir.

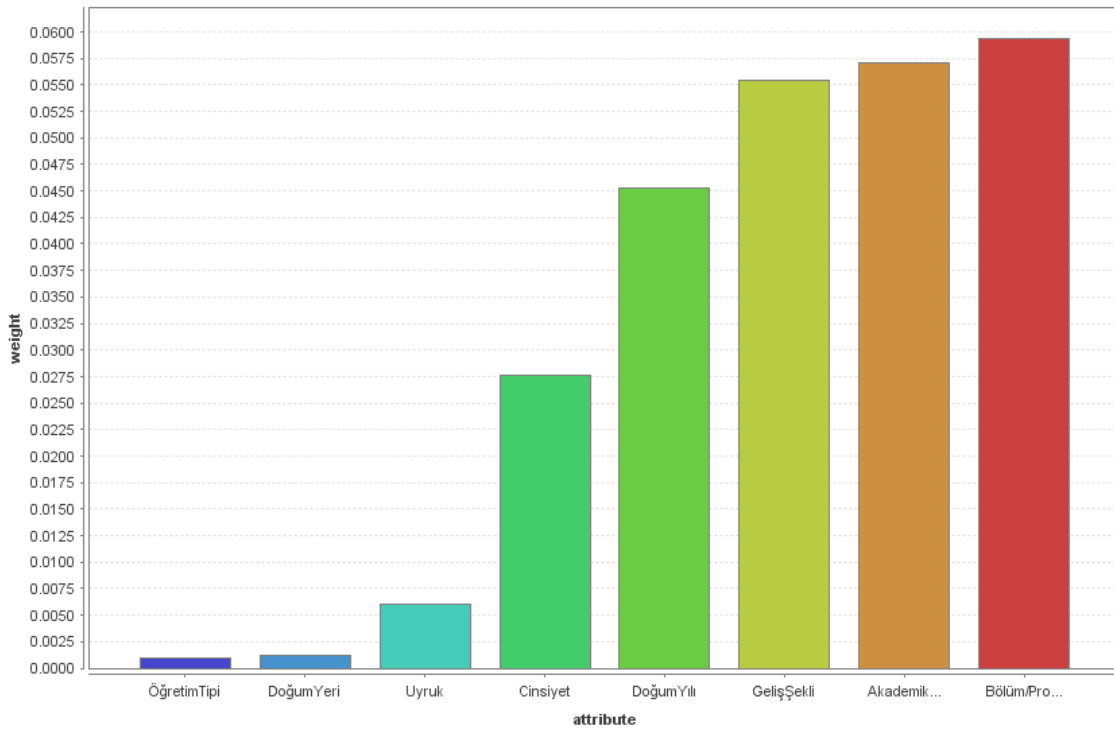
Şekil 34

Değişkenlerin Önem Ağırlıkları

attribute	weight
ÖğretimTipi	0.001
DoğumYeri	0.001
Uyruk	0.006
Cinsiyet	0.028
DoğumYılı	0.045
GelişŞekli	0.055
AkademikDonem	0.057
BölümüProgram	0.059

Şekil 35

Değişkenlerin Önem Ağırlıkları 2



5. Bölüm

Tartışma ve Öneriler

Çalışmanın başında belirlenen, öğrencilerin profillerinin ortaya çıkarılması, sınıflandırma modelinin oluşturulması ve değişkenlerin önem ağırlıklarının tespit edilmesi amaçlarına ulaşılmaya çalışılmıştır. İlk bakışta anlamsız gözükten ve yapılandırılmamış şekilde bulunan veriler düzenlenmiş, verilerin içinden anlamlı, gizli ilişkiler çıkarılmıştır. Eğitsel Veri Madenciliği analizlerinin nasıl yapıldığı tasarım ekranları da paylaşarak anlatılmaya çalışılmıştır.

Bu bölümde Veri Madenciliği analizlerine ilişkin bulgulardan yola çıkarak ortaya çıkan sonuçlar, alan yazınla yapılan karşılaştırmalar ve öneriler yer almaktadır. Uygulanan Veri Madenciliği yöntemlerine karşılık yorumlanan sonuçlar 3 ana başlıkta sunulmuştur.

5.1 Değişkenler Bize Ne Anlatıyor?

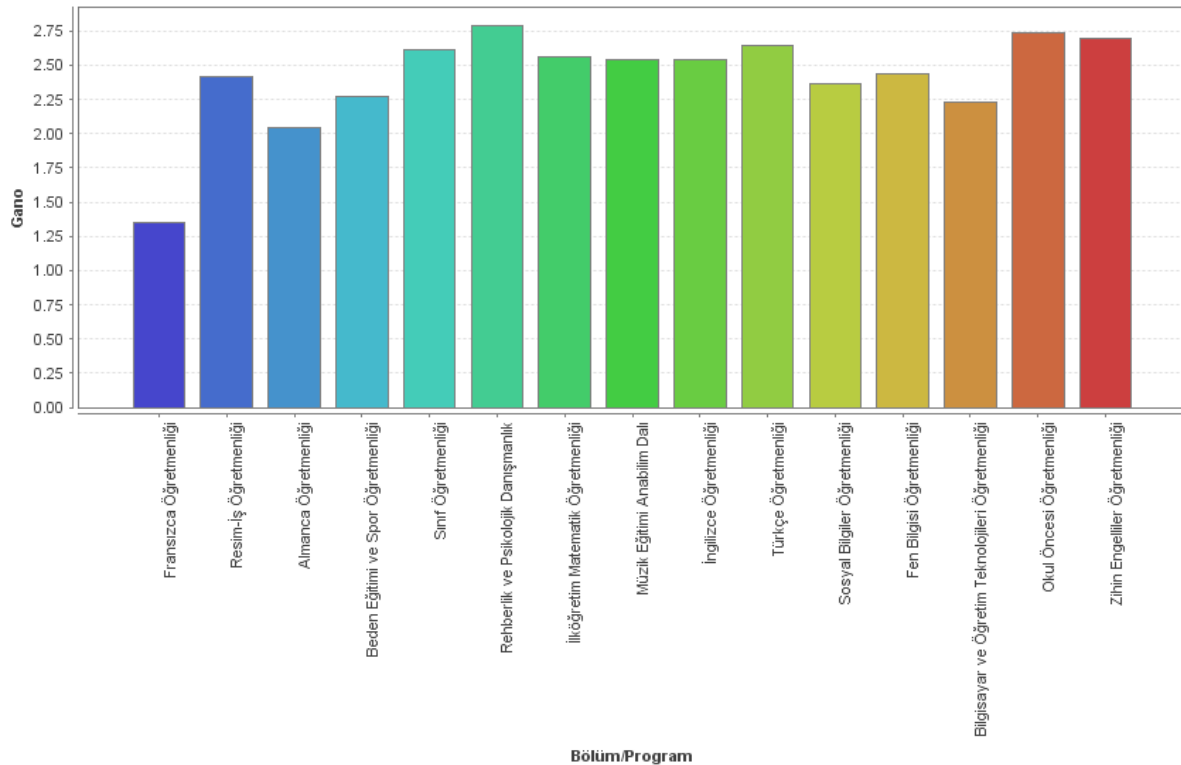
5.1.1 Bölüm faktörü. Bölümlerdeki öğrencilerin başarıları tek bir grafikte, Veri Madenciliği yöntemi sayesinde Şekil 36'da görülmektedir. Yapılan analiz ışığında, bölüm değişkenine göre öğrenci başarıları farklılık göstermektedir. Grafiğe bakıldığında Fransızca bölümü gözle görülebilir şekilde başarı olarak diğer bölümlere oranla düşük bir seviyede seyretmektedir. Bu analiz başarısız not ortalamasının altında yatan nedenleri vermez. Bu bölümdeki hocaların düşük notlar vermesi, öğrencilerin derslere çalışmaması ya da bölümün çok zor olması olabilir. Ama fakültenin genel anlamda başarısını, öğrencilerin motivasyonunu etkileyen ve düzeltilmesi gereken bir problemdir.

Bölüm ortalamaları temel alınarak hazırlanmış farklı bir grafikte de görülmektedir. Fakültenin en başarılı bölümü 2,75 ortalamaıyla Rehberlik ve Psikolojik Danışmanlık Bölümü, ardından Okul Öncesi Öğretmenliği ve Zihin Engelliler Öğretmenliğidir.

Araştırmacılar tarafından Fransızca bölümünde öğrenci başarılarını etkileyen faktörlerin araştırılması, yöneticiler tarafından ilgili eğitim politikasının gözden geçirilmesi ve öğrencilerin yönlendirilmesi gerekmektedir.

Şekil 36

Bölüme Göre Ortalamalar



5.1.2 Yaş faktörü. Yaş değişkeni ile ilgili sunulan genel grafik bize hangi neslin şu an okulda yoğun olarak bulunduğunu göstermektedir. Grafığe göre öğrenciler ağırlıklı olarak 2000’li yıllarda doğmuştur. Okulda G nesli olarak adlandırılan öğrenme biçimleri farklılık gösteren yeni nesle eğitim verilmektedir (Palfrey & Gasser, 2013).

Öğretim yapılan neslin kimler olduğunu tanımak, beklenti ve ihtiyaçlarını, nasıl daha iyi öğrendiklerini bilmek gerekir (Mitra, 2013). Yaş değişkeni baz alınarak öğretim politikaları güncellenmeli ve öğretim modellerinde değişikliğe gidilmelidir. 20 yıl önce kullandığımız metotlar, hayatın her alanında teknolojiyi kullanan öğrencilerin beklenti ve ihtiyaçlarını karşılamayacaktır. Yaş değişkenine dikkat edilmesi, beraberinde yaş ile ilgili;

devamsızlık, okulu bırakma, kuşak çatışması, yaş ile ilgili hastalıkların önceden tespit edilmesi gibi problemlere önceden müdahale edilmesini sağlamaktadır.

İlgili grafikte 27 yaş altı öğrencilerin yer aldığı 2 numaralı grubun ortalaması 27 yaş üstü öğrencilerin ortalamasında yaklaşık 1 puan fazladır. Yaş arttıkça genel ağırlıklı not ortalaması düşmektedir. Yaş olarak büyük öğrencilerin kümeden dikkat çekecek şekilde ayrıldığı görülmektedir.

Farklı yaş gruplarına farklı neslin özelliklerini taşıdıkları için farklı yöntemler uygulanması bu yöndeki problemleri giderebilir. PISA (2012) raporunda yaşa bağlı olarak öğrencilerin derslerle alakalı problemlerinin değişkenlik gösterdiği sunulmuştur.

5.1.3. Geliş şekli faktörü. Öğrencilerin not ortalamaları geliş şekli açısından farklılar göstermektedir. Bu da bize Yatay Geçiş, DGS, af ile öğrencilerin geri kazandırılması gibi eğitim politikalarının başarılı olup olmadığı ve nasıl grafik sergilediği sonuçlarını göstermektedir.

Geliş şekline göre öğrencilerin ortalama başarıları Şekil 37’de yer almaktadır. Bu grafik ışığında fakültenin en başarılı öğrencileri DGS ile kayıt yaptıran öğrencilerdir. Daha sonra sırayla üniversite dışından yatay geçişle gelenler, çift anadal yapanlar ve üniversite dışından DGS ile kayıt yaptıranlar en başarılı öğrenci sıralamasında yer almaktadır.

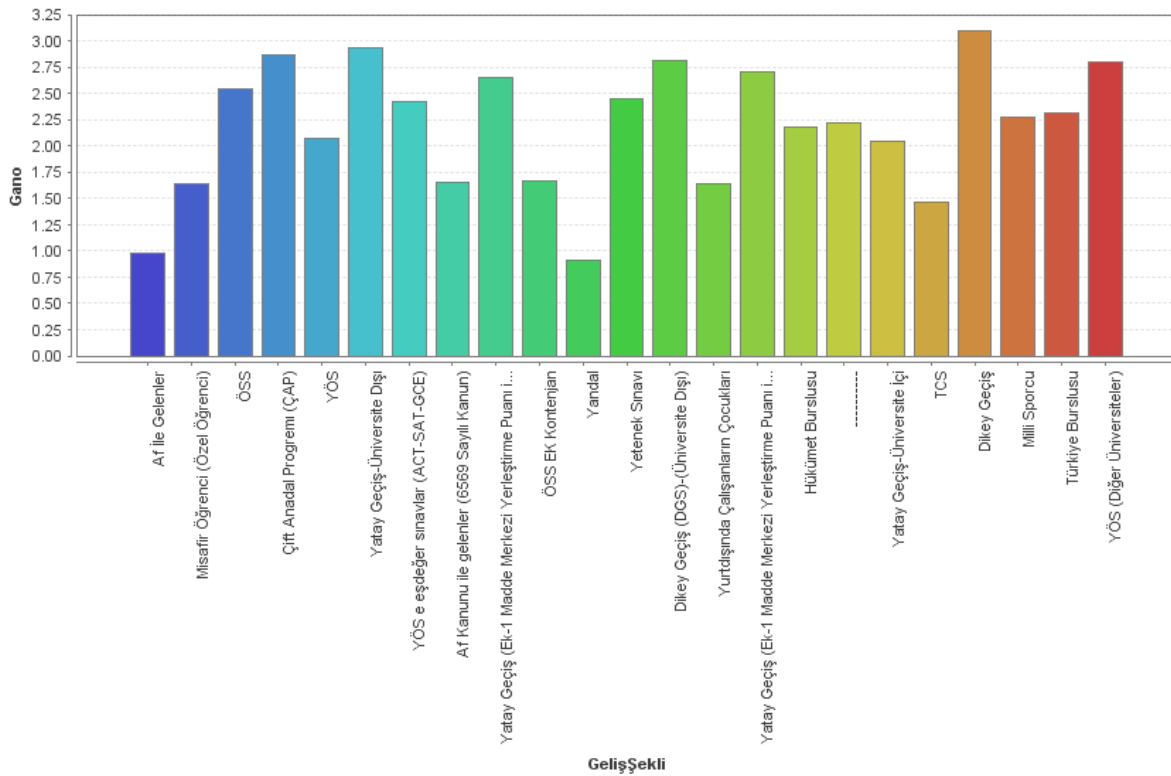
Fakültenin en başarısız öğrencileri ise af ile gelen öğrenciler ve yandal yapan öğrencilerdir. Bu öğrencilerin ortalamaları 1’in altındadır. Buradan af ile gelen öğrencilerde verim alınmadığı, öğrencilerin iyi takip edilemediği ve başarılı olamadığı sonuçlarına varılabilir. Af ile öğrencilerin okula geri döndürülmesi politikası başarısızlıkla sonuçlanmıştır diyebiliriz.

6569 sayılı kanun ile af ile gelen öğrenciler normal af ile gelen öğrencilerden daha başarılı fakat yine 2 not ortalamasının altında yer almaktadırlar.

Üniversite içi ve üniversiteler arası uygulanan Çift Anadal, Yatay Geçiş, DGS ve YÖS uygulamaları çok başarılı uygulamalardır ve fakültenin en başarılı öğrencileri bu yollar ile üniversiteye devam etmektedirler. Benzer uygulamaların artması fakültenin genel başarısına olumlu katkıda bulunacaktır.

Şekil 37

Geliş Şekline Göre Ortalamalar



5.1.4. Cinsiyet faktörü. Öğrencilerin not ortalamaları cinsiyet faktörüne göre farklılık göstermektedir denilebilir. Fakültede kız öğrencilerin sayısının erkek öğrencilerin sayısının iki katı olmasının yanında, kız öğrencilerin daha başarılı olduğu sonucuna varılmıştır. Kız öğrencilerin ortalamaları erkek öğrencilerin yarım puan üstündedir.

Literatüre bakıldığında cinsiyet faktörünün okul başarısını etkilediği görülmekte olup, yapılan araştırmaların genelinde kız öğrencilerin erkek öğrencilerden daha başarılı oldukları sonucu ortaya çıkmıştır (Akça, 2014).

PISA 2012 raporunda, bazı derslerde mesela matematikte erkek öğrencilerin daha başarılı, genel anlamda ise kız öğrencilerin daha başarılı oldukları sonucu ortaya çıkmıştır. Kız öğrencilerin kariyerleri konusunda erkeklere göre daha çok beklentileri olduğu ve daha çok çalıştıkları da raporda yer almaktadır.

Stoet ve Geary (2015) yaptıkları çalışmada, öğrenci başarısında erkeklerin ve kızların farklı dağılımlar sergilemesinin sebeplerinin, politik, ekonomik ve sosyal eşitlik ile ilgili olmadığı sonucuna varılmıştır.

Eğitsel Veri Madenciliğinin sonunda yapılan değerlendirmelerin tekrar eğitim ortamına aktarılması aşamasında cinsiyet faktörü açısından farklı derslerde cinsiyete bağlı ders başarısındaki ortalamaları düzeltmek için farklı değerlendirmeler yapılabilir. Ayrıca öğrencilerin mesleki yönlendirmelerinde cinsiyet faktörüne ve kadın-erkek eşitliğine dikkat edilmesi önemli bir faktördür. Avrupa Komisyonu (2009) raporuna göre kız ve erkeklerin gelecekteki mesleki hayatları cinsiyet rolleri ile şekillenmektedir.

5.2 Öğrenci Modelleme

Bu bölümde, Veri Madenciliği analizleri sayesinde çıkarılan, veriler arasındaki gizli ilişkiler ve oluşturulan kurallar yorumlanmıştır.

Yapılan analiz sonucunda basit bir karar ağacı oluşturulmuş ve akademik başarıya etki eden değişkelerin incelenmesinde 27 yaşın altındaki öğrencilerin genel olarak daha başarılı olduğu; 27 yaşın üstündeki öğrencilerin başarı durumlarının ise akademik dönemlerine bakılarak şekillendiği tespit edilmiştir. Bu karar ağacı veriler ışığında çizilmiş ve basit bir algoritmayla ortaya konmuştur. Burada akademik dönem değişkeni sıfır olan ve ikinci sınıfta yer alan öğrencilerin genel olarak başarısız olarak sınıflandırıldığı görülmektedir. Burada yer alan bulgu yanıltıcı olabilir. Değeri sıfır olan bilgiler üniversitenin veri toplama aşamasında yeterli ve doğru bilgi toplayamamış olmasından kaynaklanabilir. Dolayısıyla kayıp veriler her zaman başarısızlığı değil, kimi zamanda veri toplamadaki eksiklikleri işaret ediyor olabilir.

Öğrencilerin başarılı ve başarısız şekilde sınıflandığı bu analizde, hazırlanan kurallar sayesinde öğrenciler yönlendirilebilir ve yöneticilerin politikaları bu kurallar ışığında düzenlenebilir. Hangi öğrencilerin akademik, hangi öğrencilerin okulu bırakmaya meyilli olduğu anlaşılabilir. Kural Çıkarımı yönteminin uygulanması sonucu ortaya çıkan tabloda cinsiyet faktörü kadın olup, 27 yaşın altındaki öğrencilerin durum değişkeni iyi olarak sınıflandırılmış, başarılı olarak sınıflanan ve en yüksek orana sahip öğrenci topluluğunu oluşturmuşlardır. Mesela bir akademisyen fakültede yapacağı bir çalışmada fakültedeki en geniş ve başarılı öğrenci kitlesini örneklem olarak seçmek isterse bu sınıfı kullanabilir. Fakültede 2815 öğrenci bu sınıfa girmektedir.

Fakültenin son sınıfında olan ve mezuniyete hazırlanan aynı zamanda 1990 sonrası doğmuş olan bireylerin başarılı olarak sınıflandırılması da bu analizde ortaya çıkan en güçlü ikinci kuraldır. Bu beklenen bir sonuçtur.

Eğer bir öğrenci 27 yaşın üzerinde ve ikinci sınıfta okuyor ise başarısız olma ihtimali yüksektir. Fakültenin 2. sınıfında okuyan öğrencilerin diğer analizlerde de başarısız oldukları ortaya çıkmıştır. 2. sınıflardaki başarısızlığın araştırılması ve problemin çözülmesi önemli bir ihtiyaçtır. Oğuz (2012)'de yapmış olduğu çalışmada, öğrencilerin son sınıfta öz yeterlilik inançlarının güçlendiğini, 1. ve 2. sınıfta bu algının daha az olduğunu tespit etmiştir. Fakültelerin 2. sınıfında ders başarısının öneminin vurgulanarak öğrenci başarısının çeşitli etkinliklerle artırılması sağlanmalıdır.

5.3 Akademik Başarıya Etki Eden Değişkenlerin Önem Ağırlıkları

Bu bölümde 3. Araştırma sorusuna cevap teşkil eden, değişkenlerin akademik başarı üzerindeki önem ağırlıklarını tespitine dair değerlendirmeler paylaşılmıştır.

Değişkenlerin önem ağırlıkları çeşitli algoritmalarla tespit edilebilir. Burada seçilecek algoritma değişkenlerin ne türde olduklarına bağlıdır.

Akademik başarıya etki eden değişkenler içinde en yüksek çıkan değer bölüm değişkenidir. Yapılan benzer bir çalışmaya baktığımızda öğrencilerin akademik başarıları üzerindeki en önemli etkenin cinsiyet, ardından bölüm olarak tespit edildiği görülmüş (Akça, 2014).

Üniversiteye girmeden önce not ortalamasını yüksek tutmayı hedefleyen bir öğrenci, seçeceği bölümü iyi seçmelidir. Öğrenci bilgi sisteminde kayıtlı olan verilerden en önemli değişken bölümdür.

Değerler Uludağ Üniversitesi Eğitim Fakültesi öğrencileri için bu şekilde çıkmıştır. Farklı bir eğitsel veri setinde değişik sonuçlar çıkabilir ve yöneticiler tarafından ilgili değerlendirmeler yapılabilir. Uludağ Üniversitelerinin akademik başarılarında en önemli 2. değişken akademik dönemdir. Öğrencilerin başarıları hangi sınıfta olduklarına bağlıdır.

Eğitim Fakültesi öğrencileri en çok ikinci sınıfta genel anlamda başarısız oldukları ilgili grafiklerde görülmektedir. Dolayısıyla üniversitenin ilk yıllarında öğrenci başarısı ile farkındalığın artırılması, çeşitli etkinliklerle bu soruna çözüm bulunulması gerekmektedir. Öğrencilerin son sınıfa doğru olası başarısızlıklarını önceden tahmin etmek, öğrenci davranışlarının tespit edilmesi, dersi bırakma eğilimi olan öğrencilerin belirlenmesi açısından büyük önem taşır (Berland ve diğerleri, 2014).

Veri Madenciliği uygulamalarının değerlendirmesini yaparken, öğrenciler hakkında teşhis yaptığımız için doktorlar ile hastaları arasındaki ilişki eğitime uyarlanabilir. Mesela bacak ağrısı şikâyetinin dünya genelindeki en çok rastlanılan sebebi kas ağrısıdır. Dolayısıyla bu şikâyetle gelen bir hastaya sorulacak ilk soru bu olursa, hastalık tanısı o kadar erken konur ve tedavi süreci hemen başlatılarak hastanın daha kötü bir duruma gelmesi engellenir.

Bu örneği çalışmamız açısından uyarlırsak Uludağ Üniversitesi Eğitim Fakültesinde okuyan ve derslerinde başarısız olan bir öğrenciyle ilgili değerlendirme yaparken akla gelecek

ilk deęişken bölüm, ardından akademik dönemi ardından sırayla tespit edilen dięer deęişkenler gelmelidir.

5.4 Öneriler

Eđitsel Veri Madencilięi sayesinde ilk bakışta anlamsız görünen veri yığıını içinden gizli ilişkiler, öğrencileri daha iyi tanımamızı sağlayacak bilgiler, öğrencileri yönlendirebileceğimiz kurallar Veri Madencilięi analizleriyle çıkartılmıştır.

Öğrencilere ait verilerin artması onları tanımamızı, nasıl daha iyi öğrendiklerini bilmemizi de hızlandıracaktır. Bu sebeple bu çalışmadan çıkarılacak ilk sonuç öğrencilere ait daha fazla eđitsel verinin, eksik veri olmadan, büyük bir özenle depolanması gerektięidir. Başar ve Aslay (2011) yapmış oldukları çalışmada çeşitli deęişkenlerin öğrenci başarısını etkiledięi ve ama daha etkili sonuçlar alınabilmesi için daha yüksek performanslı bir yazılımın olması gerektięini savunmuşlardır.

Ne kadar fazla veri varsa o kadar iyidir. Kanser hastalığına çözüm bulunabilmesi için ne kadar çok veri toplanırsa, hastalığın ilerleme haritası o kadar iyi çizilebilir ve çözüm daha kolay bulunabilir. Sağlık bakanlığı 2015 yılında obezite sorununa çözüm bulabilmek adına ulusalobeziteveritabani.com web sitesini açmıştır. Buraya girilen hasta bilgilerine yapılacak analizler sayesinde, sorunun sebeplerinin tespit edilerek çözülebileceęi görüşündedir.

Aynı şekilde eğitimle alakalı toplanacak verilerin ve kalitesinin artması ve bu verilerin Eđitsel Veri Madencilięi yöntemleri ile analiz edilerek eğitim ortamlarına aktarılması, eđitsel problemlerin giderilmesini, öğrencilerin başarılarının artmasını ve onları daha iyi tanımamızı sağlayacaktır. Aynı zamanda ülkemizin genel anlamda eğitimin kalitesini artırarak bilimsel açıdan daha zengin, daha kaliteli ve dünyada aranan, örnek alınan bir medeniyet olmasının önünü açacaktır.

Kaynakça

- Akça F. (2014). *Veri Madenciliği ile Fen Fakültesi Öğrenci Profillerinin İncelenmesi: Gazi Üniversitesi Örneği* (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara
- Akçapınar G. (2014). *Çevrimiçi Öğrenme Ortamındaki Etkileşim Verilerine Göre Öğrencilerin Akademik Performanslarının Veri Madenciliği Yaklaşımı ile Modellenmesi* (Yayımlanmamış doktora tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara
- Aksoy E. (2014). *Matematik Alanında Üstün Yetenekli ve Zekalı Öğrencilerin Bazı Değişkenler Açısından Veri Madenciliği ile Belirlenmesi* (Yayımlanmamış yüksek lisans tezi). Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü, İzmir
- Avrupa Komisyonu, (2009). *Eğitim Çıktılarında Cinsiyet Farklılıkları: Avrupa'da alınan tedbirler ve Mevcut Durum*
http://sgb.meb.gov.tr/eurydice/kitaplar/Egitim_ciktilarinda_cinsiyet_farkliliklari/Egitim_ciktilarinda_cinsiyet_farkliliklari.pdf'den alınmıştır.
- Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *JEDM-Journal of Educational Data Mining, 1*(1), 3-17.
- Baker, R. S. J. D. (2010). Data mining for education. *International encyclopedia of education, 7*, 112-118.
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61-75).Springer New York.

- Başar, M. S., & Aslay, F. (2011). Yazılım Ergonomisi: Atatürk Üniversitesi Öğrenci Bilgi Sisteminin Ergonomisinin İncelenmesi. *Journal of Graduate School of Social Sciences, 15*(1).
- Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning, 19*(1-2), 205-220.
- Bernardo, J. M., & Smith, A. F. M. (1994). Bayesian Theory Wiley. *New York*.
- Beyhan H. D. (2014). Sosyal Medya Üzerinden Metin Madenciliği ve Duygu Analizi ile Pazar Değerlendirme (Yayımlanmamış yüksek lisans tezi). İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul
- Cohen, W. W. (1995, July). Fast effective rule induction. In *Proceedings of the twelfth international conference on machine learning* (pp. 115-123).
- Çapa E. (2014, 21 Kasım). Veri Madenciliği [Video]. <http://www.cnnturk.com/video/hazir-cevap/dunyanin-1001-hali/bilginin-ufuklari/veri-madenciligi> adresinden alınmıştır
- Efron, B., & Tibshirani, R. (1990). *Statistical data analysis in the computer age*. University of Toronto, Department of Statistics.
- Guruler, H., Istanbulu, A., & Karahasan, M. (2010). A new student performance analysing system using knowledge discovery in higher educational databases. *Computers & Education, 55*(1), 247-254.
- Gürsoy, U. T. Ş. (2012). *Uygulamalı Veri Madenciliği Sektörel Analizler*. Pegem Akademi.
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior, 29*(1), 90-102.
- Heckerman, D. (1996). Bayesian Networks for Knowledge Discovery. *Advances in knowledge discovery and data mining, 11*, 273-305.

- Kişisel Verilerin Korunması Kanunu (2016), T.C. Resmi Gazete, 6698, 24 Mart 2016
- Klößgen, W., & Zytkow, J. M. (2002). *Handbook of data mining and knowledge discovery*. Oxford University Press, Inc..
- Ledley, R. S. (1960). Using electronic computers in medical diagnosis. *IRE transactions on medical electronics*, 4(ME-7), 274-280.
- Lépine, S., & Shara, M. M. (2005). A Catalog of Northern Stars with Annual Proper Motions Larger than 015 (LSPM-NORTH Catalog) Based on data mining of the Digitized Sky Surveys (DSSs), developed and operated by the Catalogs and Surveys Branch of the Space Telescope Science Institute (STScI), Baltimore. Developed with support from the National Science Foundation (NSF), as part of the NASA/NSF NStars program. *The Astronomical Journal*, 129(3), 1483.
- Liu B., Hsu W. & Ma Y. (1998). Integrating classification and association rule mining. In *Proceedings of the fourth international conference on knowledge discovery and data mining*. New York: AAAI Press
- Maimon, O., & Rokach, L. (Eds.). (2005). *Data mining and knowledge discovery handbook* (Vol. 2). New York: Springer.
- Milli Eğitim Bakanlığı, (2013) Okul Başarısını Etkileyen Faktörler http://mebk12.meb.gov.tr/meb_iys_dosyalar/16/15/973218/dosyalar/2013_05/080733_16_okulbaar%C4%B1s%C4%B1n%C4%B1etkileyenfaktrler.doc ‘ den alınmıştır. Erişim Tarihi: 25.12.2016
- Mitra, S. (2013). Beyond the hole in the wall: Discover the power of self-organized learning. Ted Conferences.
- Oğuz, A. (2012). Sınıf öğretmeni adaylarının akademik öz yeterlik inançları. *Anadolu Journal of Educational Sciences International*, 2(2).

- Palfrey, J., & Gasser, U. (2013). *Born digital: Understanding the first generation of digital natives*. Basic Books.
- Peña-Ayala, A. (Ed.). (2013). *Educational Data Mining: Applications and Trends* (Vol. 524). Springer.
- Pena, A., Domínguez, R., & Medel, J. D. J. (2009). Educational data mining: a sample of review and study case. *World Journal On Educational Technology*, 1(2), 118-139.
- PISA, (2012) Ulusal rapor the ABC of gender equality in education
<https://www.oecd.org/pisa/keyfindings/pisa-2012-results-gender-eng.pdf> den alınmıştır. Erişim Tarihi: 10.01.2013
- Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the horizon*, 9(5), 1-6.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.
- Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, 33(1), 135-146.
- Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (Eds.). (2010). *Handbook of educational data mining*. CRC Press.
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.
- Schabacker, M., & Wohlbold, L. (2002, June). Benefit Evaluation of EDM/PDM Systems. In *Proceedings of the 8th International Conference on Concurrent Enterprising* (pp. 17-19).
- Shave C. (2015, 6 Temmuz) Should universities collect personal data to monitor their students? – live chat. The Guardian. Retrieved from
<https://www.theguardian.com/higher-education-network/2015/jul/06/should-universities-collect-personal-data-to-monitor-their-students-live-chat>

- Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence*, 48, 137-151.
- Şentürk, A. (2006). *Veri Madenciliği kavram ve teknikler*. Ekin Kitabevi, Bursa
- Team Techansar (2015, 19 Ağustos). Internet data real time. Techansar. Retrieved from <http://techsansar.com/internetworking/internet-data-real-time>.
- Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Zhong, N., & Zhou, L. (Eds.). (2003). *Methodologies for Knowledge Discovery and Data Mining: Third Pacific-Asia Conference, Beijing, China*: Springer.

Özgeçmiş

Doğum Yeri ve Yılı : Hakkari – 1986

Lisans : Balıkesir Üniversitesi Bilgisayar ve Öğretim
Teknolojileri Eğitimi (2010 - 2013)

Yüksek Lisans : Uludağ Üniversitesi Bilgisayar ve Öğretim
Teknolojileri Eğitimi (2014 - 2017)

Bildiği Yabancı Diller ve Düzeyi : İngilizce – Çok iyi

Çalıştığı Kurumlar : Uludağ Üniversitesi (2014 - ...)

01.05.2017

Abdullah Ragıp Ersöz