

**MESLEKİ ALAN SEÇİMLERİNİN MAKİNE ÖĞRENMESİ  
ALGORİTMASI KULLANILARAK BELİRLENMESİ**

**Özkan ÜNSAL**

**YÜKSEK LİSANS TEZİ  
BİLGİSAYAR EĞİTİMİ**

**GAZİ ÜNİVERSİTESİ  
BİLİŞİM ENSTİTÜSÜ**

**HAZİRAN 2011  
ANKARA**

Özkan ÜNSAL tarafından hazırlanan MESLEKİ ALAN SEÇİMLERİNİN MAKİNE ÖĞRENMESİ ALGORİTMASI KULLANILARAK BELİRLENMESİ adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Halil İbrahim BÜLBÜL  
Tez Yöneticisi

Bu çalışma, jürimiz tarafından oy birliği / oy çokluğu ile Bilgisayar Eğitimi Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Başkan: : Prof. Dr. Ali Paşa AYDIN

Üye : Doç. Dr. Halil İbrahim BÜLBÜL

Üye : Yrd. Doç. Dr. Mehmet Akif OCAK

Tarih : 03 / 06 / 2011

Bu tez, Gazi Üniversitesi Bilişim Enstitüsü tez yazım kurallarına uygundur.

## **TEZ BİLDİRİMİ**

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada orijinal olmayan her türlü kaynağa eksiksiz atıf yapıldığını bildiririm.

Özkan ÜNSAL

# MESLEKİ ALAN SEÇİMLERİNİN MAKİNE ÖĞRENMESİ ALGORİTMASI KULLANILARAK BELİRLENMESİ

(Yüksek Lisans Tezi)

Özkan ÜNSAL

GAZİ ÜNİVERSİTESİ  
BİLİŞİM ENSTİTÜSÜ

Haziran 2011

## ÖZET

Günümüzde mesleki ve teknik eğitimin önemi gelişen teknolojiye paralel olarak her geçen gün artmaktadır. Mesleki ve teknik eğitimde, bireyler için meslek alanlarının en iyi şekilde belirlenebilmesi için bilgi ve iletişim teknolojilerinin sunduğu imkanlardan faydalanılması kaçınılmazdır. Bu kapsamda bir yapay zeka modeli olan ve geçmişteki tecrübelerden yararlanarak gelecekteki olayları tahmin eden “Makine Öğrenmesi” ile mevcut yöntemlere kıyasla daha verimli sistemler geliştirilebilmektedir. Bu çalışmada, mesleki teknik eğitim sürecinde başarılı ve başarısız olmuş bireylere ait toplanan verilere bir makine öğrenmesi tekniği olan “Naive Bayes” algoritması uygulanarak, sistemin geçmişte yapılan başarılı ve başarısız alan seçimlerini öğrenerek bu seçimler hakkında tecrübe kazanmasını sağlayan bir yazılım geliştirilmiştir. Geliştirilen yazılım ile mesleki alan seçim sürecinde bulunan bireyden alınan verilere göre birey için en uygun mesleki alanın sistem tarafından önerilmesi amaçlanmıştır.

**Bilim Kodu** : 702.1.014

**Anahtar Kelimeler** : makine öğrenmesi, mesleki alan seçimi, veri madenciliği

**Sayfa Adedi** : 72

**Tez Yöneticisi** : Doç. Dr. Halil İbrahim BÜLBÜL

**DETERMINING VOCATIONAL FIELD ELECTIONS BY USING MACHINE  
LEARNING ALGORITHMS**

**(M.Sc. Thesis)**

**Özkan ÜNSAL**

**GAZI UNIVERSITY  
INFORMATICS INSTITUTE**

**June 2011**

**ABSTRACT**

The importance of vocational and technical training is growing day by day in parallel to the developing technology. It is inevitable to utilise opportunities presented by information and communication technologies in order to determine vocational fields for individuals in vocational and technical training in the most efficient manner. In this respect, it is possible to develop more efficient systems compared to the current methods by utilising machine learning which is an artificial intelligence model that predicts events in the future depending on the past experiences. In the current study, a software is developed that ensures that the system gains experience and learns about the successful and unsuccessful choices made in the past by applying “Naive Bayes” algorithm, which is a machine learning algorithm, to the data collected concerning the individuals who turned out to be successful or unsuccessful in the vocational technical training process. In the software developed, it is aimed that the system recommends the most suitable vocational field for the individual by according to the data collected from the individual who is in the occupation selection process.

**Science Code : 702.1.014**

**Key Words : machine learning, vocational field elections, data mining**

**Page Number : 72**

**Adviser : Assoc. Prof. Dr. Halil İbrahim BÜLBÜL**

## TEŐEKKÜR

Çalıőmalarım boyunca deęerli yardım ve katkılarıyla beni yönlendiren tez danışmanım Doç. Dr. Halil İbrahim BÜLBÜL'e, makine öğrenmesi ve bayes teoremi hakkında beni bilgilendiren Dr. Hamdi Tolga KAHRAMAN'a, verilerin toplanma sürecinde yardımlarını esirgemeyen Beypazarı Kız Teknik ve Meslek Lisesi Müdürlüğü'ne ve arkadaşlarım Oęuzhan ÜNAL, Selçuk DEMİRCİ, Ali Osman YANAR'a ve beni her zaman destekleyen ve yanımda olan, başta babam Süleyman ÜNSAL'a ve aileme teşekkür ederim.

## İÇİNDEKİLER

	<b>Sayfa</b>
ÖZET.....	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER.....	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
2. KONU İLE İLGİLİ YAPILMIŞ ÇALIŞMALAR.....	5
3. MESLEKİ REHBERLİK VE MESLEK SEÇİMİ.....	8
3.1. Meslek.....	8
3.2. Meslek Seçimi ve Mesleki Eğitim.....	9
3.3. Mesleki Rehberlik.....	11
3.3.1. Mesleki rehberliğin dünya ve ülkemizdeki tarihi gelişim süreci.....	12
3.3.2. Mesleki rehberliğin amaçları.....	13
3.3.3. Mesleki rehberlikte özellik-faktör yaklaşımı.....	15
3.3.4. Mevcut mesleki yönlendirme teknikleri.....	16
4. VERİ MADENCİLİĞİ VE MAKİNE ÖĞRENMESİ.....	18
4.1. Veri Madenciliği.....	18
4.1.1. Veri madenciliği ve veritabanı sistemlerinin tarihi gelişimi.....	20
4.1.2. Veri madenciliği uygulama alanları.....	22

	<b>Sayfa</b>
4.1.3. Veri madenciliği uygulama süreci .....	24
4.2. Makine Öğrenmesi .....	30
4.2.1. Makine öğrenmesi uygulama alanları .....	31
4.2.2. Makine öğrenmesi modeli.....	32
4.2.3. Makine öğrenmesi teknikleri .....	33
4.3. Bayes Teoremi .....	37
4.3.1. Bayes sınıflandırma tekniği ve Naive Bayes algoritması .....	38
5. MATERYAL VE METOD .....	44
5.1. Mesleki Yönlendirmede Kullanılan Kriterlerin ve Algoritmanın Seçimi.....	44
5.2. Verilerin Elde Edilme Süreci .....	44
5.3. Geliştirilen Mesleki Yönlendirme Yazılımı.....	45
5.3.1. Yazılımın geliştirilmesinde kullanılan araçlar .....	45
5.3.2. Veritabanı tasarımı.....	45
5.3.3. Yazılım ara yüz tanıtımı.....	45
6. MESLEKİ ALAN SEÇİMLERİNİN MAKİNE ÖĞRENMESİ ALGORİTMASI KULLANILARAK BELİRLENMESİ.....	51
6.1. Mesleki Yönlendirme İçin Naive Bayes Algoritmasının Uygulanması.....	51
6.2. Mesleki Yönlendirme Yazılımının Kullanılması .....	52
7. SONUÇ VE ÖNERİLER .....	57
KAYNAKLAR .....	60
EKLER .....	64
EK-1. İnternet üzerinde yapılan anket formu.....	65
EK-2. Eğitim ve test veri setlerinin oluşturulmasında kullanılan t-sql komutları .....	69
EK-3. Algoritmanın uygulanmasında kullanılan C#.NET dili kodları .....	70
ÖZGEÇMİŞ .....	72



**ÇİZELGELERİN LİSTESİ**

<b>Çizelge</b>	<b>Sayfa</b>
Çizelge 4.1. Eğitim veri seti.....	41
Çizelge 4.2. Test verisine ait kriterlerin frekansları.....	41
Çizelge 4.3. Sıfır sorunu için test verisine ait kriterlerin frekansları .....	42
Çizelge 7.1. Test verileri için yapılan başarı tahmini sonuçları.....	59

## ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 4.1. Veri madenciliğinin tarihsel perspektifi .....	20
Şekil 4.2. Veri madenciliği uygulama süreci .....	24
Şekil 4.3. Makine öğrenmesi şeması.....	33
Şekil 4.4. Karar ağacı modeli örneği.....	35
Şekil 5.1. Ana form .....	46
Şekil 5.2. Mesleki alan düzenleme formu.....	46
Şekil 5.3. Veli türleri düzenleme formu.....	47
Şekil 5.4. Ders kayıtları düzenleme formu .....	47
Şekil 5.5. Ekonomik durum kayıtları düzenleme formu .....	48
Şekil 5.6. Öğrenim durum kayıtları düzenleme formu .....	48
Şekil 5.7. Meslek kayıtları düzenleme formu .....	49
Şekil 5.8. Öğrenci kayıtları ve alan başarı tahmin formu.....	50
Şekil 6.1. Öğrenci genel bilgileri sekmesi .....	52
Şekil 6.2. Öğrenci veli bilgileri sekmesi .....	53
Şekil 6.3. Öğrenci alan tercihleri sekmesi.....	54
Şekil 6.4. Öğrenci kültür ders bilgileri sekmesi.....	54
Şekil 6.5. Öğrenci mesleki ders bilgileri sekmesi .....	55
Şekil 6.6. Alan başarı tahmin sonuç mesajı .....	56
Şekil 6.7. Başarı tahmini kriter seçim formu .....	56

## SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

<b>Simgeler</b>	<b>Açıklama</b>
$\sigma$	Standart sapma
<b>Kısaltmalar</b>	<b>Açıklama</b>
<b>ABD</b>	Amerika Birleşik Devletleri
<b>BİLDEMER</b>	Bilgisayar Destekli Meslek Rehberliği
<b>E-CRM</b>	Elektronik Müşteri İlişkileri Yönetimi
<b>E-OKUL</b>	E-Okul Yönetim Bilgi Sistemi
<b>GB</b>	Gigabyte
<b>MAP</b>	En Büyük Sonrasal Sınıflandırma Yöntemi (Maximum A Posteriori Classification)
<b>MBS</b>	Ulusal Mesleki Bilgi Sistemi
<b>MEB</b>	Milli Eğitim Bakanlığı
<b>MSSQL</b>	Microsoft Sql Veritabanı Sunucusu
<b>MYSQL</b>	Mysql Veritabanı Sunucusu
<b>OLAP</b>	Çevrimiçi Analitik İşleme (Online Analytic Processing)
<b>OLTP</b>	Çevrimiçi Hareket İşleme (Online Analytic Processing)
<b>ÖSYM</b>	Öğrenci Seçme Ve Yerleştirme Merkezi
<b>PDR</b>	Psikolojik Danışma Ve Rehberlik
<b>PHP</b>	Hypertext Preprocessor
<b>SVM</b>	Destek Vektör Makinesi (Support Vector Machine)

**Kisaltmalar****Açıklama****SQL**Yapısal Sorgulama Dili  
(Structured Query Language)**TDK**

Türk Dil Kurumu

**XML**Genişletilebilir İşaretleme Dili  
(Extensible Markup Language)

## 1. GİRİŞ

Meslek, insanın eğitim hayatı boyunca aldığı birikimi maddi kazanç elde etmek için yaptığı, kendisini sürekli geliştirdiği ve sosyal yaşamda kendisini tanımlayan bir unvan olarak kullandığı, belirli kurallar ve çeşitli disiplinlerden yararlanılarak ortaya çıkarılmış bir kavramdır. Meslek, insana hayatını sürdürmesi için gerekli olan maddi kazancı sağlaması yanı sıra kişinin yaşam tarzını belirler ve bireylere yetişkinlik döneminde psikolojik olarak elde etmesi gereken doyumu kazandırır. Meslek bir insanın mutluluğundan yaşam standartlarına kadar her açıdan insanın hayatını etkilemektedir [1].

Günümüzde artan nüfus ve güçleşen ekonomik koşullar, bireylerin iş bulmalarını giderek zorlaştırmaktadır. Teknolojik gelişmeler ve sanayileşme bireylerin mesleki yaşantılarını ve geleceğini planlamayı oldukça güçleştirmiştir [2]. Meslek seçimi insanların hayatlarında verdikleri en önemli kararlardan bir tanesidir. Bu karar kişinin gelecekteki yaşamını bir ölçüde belirlemektedir. Bireylerin doğru mesleği seçmeleri ve bu yönde eğitim almaları hem eğitim hayatlarının verimli geçmesine hem de sektörün ihtiyaç duyduğu uygun bireyin yetişmesini sağlamaktadır. Doğru mesleki alanı seçen bireyler hem iş bulmakta sorun yaşamazlar hem de yaptıkları işten aldıkları doyum ile başarılı ve mutlu bir hayat geçirebilirler. Kendine uygun meslekte olan bireyler toplumda işlerini isteyerek yapar, mesleğinde ilerler, mutlu ve verimli olarak yaşamlarını sürdürürler [3]. Yanlış meslek seçimleri sonucunda bireyler mesleki hayatlarında başarısız ve mutsuz olmaktadır [4]. Kariyer Dergisi tarafından yapılan bir iş memnuniyeti araştırmasının sonuçlarına göre, katılımcıların 5'te 1'i işlerinden memnun olmadıklarını ifade ediyor. Araştırma; 16-35 yaş arasındaki grubun 4'te 1'i ise, iş doyumu eksikliği yaşadıklarını belirtiyor [5].

Bireylerin yaşamında eğitim ve iş hayatı önemli bir yer tutmaktadır. Bütün bireyler belirli bir seviyeye kadar aynı temel eğitimden geçerler fakat sonrasında hayatlarını sürdürebilmeleri için yapmaları gereken iş doğrultusunda farklı mesleki alanlarda eğitim almaktadırlar. Bireylerin mesleki hayatlarının başarılı olmasını, aldıkları

eğitimin kalitesi kadar kendilerine uygun mesleki alanı seçmeleri de doğrudan etkilemektedir.

Mesleki eğitim sürecinin erken yaşlarda başlaması bireyin iş hayatına mesleki açıdan daha tecrübeli ve daha iyi yetişerek katılmasını sağlamaktadır. Bu açıdan mesleki eğitim veren ortaöğretim kurumları büyük önem kazanmaktadır. Mesleki eğitim veren kurumlar ara eleman yetiştirme konusunda da çok önemli bir görev üstlenmiştir. Lise düzeyinde meslek öğrenmeye başlayan öğrencilerin, üniversite düzeyinde meslek öğrenmeye başlayan öğrencilere göre sahip olduğu avantajlar ve mesleki gelişim hızları arasındaki farklar önemli ölçüde görülmektedir. Bireylerin mesleki hayata erken yaşta girmeleri ne kadar önemli ise doğru mesleki alanda eğitim görmeleri de bir o kadar önemlidir. Mesleki gelişim uzun süreli bir dönemi içermektedir. Çocukluktan başlayarak, eğitim-öğretim yaşantıları süresince devam eden ve bir iş elde edip, emekliliğe kadar süren bir dönemi kapsamaktadır [2].

Günümüzde bütün gelişmiş ülkelerde olduğu gibi ülkemizde de meslek seçimi temel eğitim sonrasında gencin bilinçli ve özgürce karar vermesi ilkesine dayandırılmıştır [1]. Her insan sevdiği ve yapmak istediği mesleki alanda eğitim almak ister. Bu konuda kararı en iyi bireyin kendisi verebilir. Fakat erken yaşta mesleki eğitime başlayacak olan bireyler doğru alanları seçebilecek yeterli olgunluk ve deneyime sahip değildirler. Bu nedenle bireylerin bu seçimi yaparken mutlaka yardıma ihtiyaçları vardır.

Bireyleri ilgi ve becerilerine uygun doğru mesleki alana yöneltme için mevcut birçok çalışma (meslek envanterleri, anketler, tutum ölçekleri vb.) bulunmaktadır. Mevcut çalışmaların ortak özelliği, bireylerin özelliklerini ayırt etmeksizin tüm kişilere standart olarak aynı şekilde uygulanmasıdır. Fakat bu testlerin uygulandığı andaki bireyin tutum ve düşünceleri mesleki alan seçimi yapıldıktan sonra farklılıklar gösterebilmektedir. Mevcut yöntemlerle yapılan mesleki yönlendirmenin uygun olarak yapılamaması kişilerin başarılı olamayacağı bir mesleği de seçmelerine sebep olabilmektedir. Yanlış meslek seçiminin kişinin iş hayatı boyunca mutsuz ve mesleğinde başarısız olmasına da yol açmaktadır. Bu çalışmalarda mesleki alan

kavramından çok sayısal, sosyal, bedensel, sanatsal gibi genel sınıflama kavramları kullanılmaktadır. Bu kavramların kapsadığı birçok mesleki alan mevcuttur ve birey alan seçimi sırasında bu tercihi yapmak durumundadır. Mesleki alan seçimi yapacak bireyin kendi tercih ve ilgileri kadar o alanda daha önceden seçim yapmış başarılı ve başarısız olan bireylerin özellikleri de önemlidir. Dolayısıyla bireyin seçmek istediği alanda yapılmış önceki tercihlerden edinilen tecrübelerde önemli ölçüde yönlendirme çalışmalarına yardımcı olmaktadır.

Veri Madenciliği; veri ambarlarındaki tutulan, çok çeşitli ve çok miktarda veriye dayanarak daha önce keşfedilmemiş bilgileri ortaya çıkarmak, bunları karar verme ve eylem planını gerçekleştirmek için kullanma sürecidir [6]. Veri madenciliği bilgi topluluklarından varılmak istenen hedef doğrultusunda anlamlı bilgilerin çıkarılmasını ve sınıflama yapılmasını sağlayan, bilişim teknolojisinin insanlara sunduğu önemli bir kavramdır. Günümüzde tıp, finans, eğitim ve daha bir çok alanda veri madenciliğinin kullanımı giderek yaygınlaşmaktadır. Makine öğrenmesi, bilgisayarların geçmiş bilgilerden elde edilen tecrübelerden yararlanarak, gelecekteki olayları tahmin etmesine ve modelleme yapmasına imkan veren bir yapay zeka alanıdır. Veri madenciliği ve makine öğrenmesi sayesinde bilgiyi elde etme süreci klasik yöntemlere göre hızlı ve doğruluk oranı yüksektir.

Bu çalışmada; daha önceden yapılmış mesleki alan seçimlerinin başarı durumlarına göre, alan seçimi yapacak bireyin başarılı olup olamayacağı çeşitli kriterlere (cinsiyet, ekonomik durum, veli bilgileri, eğitim bilgileri, meslek alan tercih bilgileri vb.) göre tahmin edilmeye çalışılmaktadır. Bu tahmin yapılırken daha iyi bir sonuç vermesi bakımından, veri madenciliğinde Bayesyen sınıflandırma tekniğinden türetilmiş, bir Makine Öğrenmesi tekniği olan Naive Bayes sınıflandırma algoritması kullanılarak bir yazılım geliştirilmiştir. Yazılımda geliştirilen makine öğrenmesi algoritması, Türkiye'nin çeşitli bölgelerinde bulunan mesleki eğitim yapan kurumlarda bulunan öğrencilerden alınan veriler üzerinde uygulanarak, bireylerin hangi mesleki alanlarda başarılı olabilecekleri tahmin edilmeye çalışılmıştır. Bu yazılımın, temel eğitimden sonra mesleki eğitim sürecine girecek olan bireyler için yapılan mesleki alan yöneltme çalışmalarına yardımcı olarak bireyler için en uygun

mesleki alanın belirlenmesi ve bu çalışmalarda harcanan zaman ve emeğin azaltılarak ortaya çıkabilecek yanlış yönlendirmelerin de önüne geçilmesi amaçlanmıştır.



## 2. KONU İLE İLGİLİ YAPILMIŞ ÇALIŞMALAR

Mesleki alan seçimi ve makine öğrenmesi kavramlarına ilişkin önemin her geçen gün artmasıyla, bu konular hakkında yapılan çalışmaların arttığı gözlenmektedir. Bu bölümde, Yükseköğretim Kurulu Başkanlığı Yayın ve Dokümantasyon Dairesi Ulusal Tez Merkezi, kütüphaneler, yerli ve yabancı kaynakların taranmasıyla, konu hakkında ulaşılabilen bazı çalışmalar özetlenmiştir.

Yazıcı (2007), Eğitsel ve Mesleki Rehberlik adlı yaptığı çalışmada, meslek ve kariyer, meslek seçimini etkileyen faktörler, meslek seçiminde kullanılan araçlar, yurtdışında geliştirilen ve uygulanan modeller, meslek danışmanlığı ve Türk eğitim sisteminde mesleki yönlendirme ve meslek seçimi süreci konuları üzerinde durmuş ve Türkiye’de mesleki rehberlik hizmetlerine verilen önemin daha da artması gerektiği sonucuna varmıştır [7].

Yapılan diğer bir çalışmada, Çakır (2007) karar verme süreci aşamaları ve meslek seçimi, meslek seçimi ile ilgili bilim adamlarının kuramsal görüşlere yer vermiş ve mesleki rehberlik hizmetlerinde göz önünde tutulması gereken noktaların önemine işaret etmiştir [2].

Konuyla ilgili başka bir çalışma Sarıkaya ve Khorshid tarafından 2009 yılında yapılmıştır. Çalışma kapsamında üniversite öğrencilerinin yerleştikleri bölümleri seçme nedenleri araştırılmıştır. Yapılan araştırma sonucunda seçtikleri alandan memnun olmayan ve yaptıkları seçimi bilinçsiz yapan öğrencilerin olduğu tespit edilmiştir. Çalışmada mesleki alan seçiminin birçok faktöre dayandığı ve mesleki eğitime geçiş sürecinde bulunan öğrencilerin tercihlerine yardımcı olmak açısından rehberlik hizmetlerinin daha dikkatli yapılması gerektiği sonucuna varılmıştır [3].

Yıldız’ın 2006 yılında yaptığı “Web Destekli Çoklu Zeka Kuramına Dayalı Mesleki Yönlendirme” adlı çalışmada, ülkemizde ilköğretim kademesinde rehberlik faaliyetlerinin yeterli olmadığına işaret ederek, ortaöğretime geçiş sürecinde mesleki yönlendirme çalışmalarının önemine dikkat çekmek üzere çoklu zeka kuramına

dayalı, web tabanlı bir uzman sistem yazılımı geliştirmiştir. Geliştirilen yazılım ile birlikte ortaöğretime geçiş sürecindeki öğrencilere daha verimli ve etkili bir rehberlik ve yönlendirme yapılması amaçlanmış, rehberlik uygulama süreçlerinin kolaylaştırıldığı belirtilmiştir [8].

Hasebrook ve Nathusius (1997) tarafından yapılan çalışmada kariyer planı yapacak olan öğrencilere yol göstermede yardımcı olması açısından multimedya tabanlı bir yazılım geliştirilmiştir. Yazılımda öğrencilerle benzer özellikler taşıyan, geçmişte yapılan kariyer tercihlerinden elde edilen tecrübelerden faydalanılarak bireylerin doğru bir meslek seçim yapması amaçlanmış ve yöntemin doğru yönlendirmeler yapmakta başarılı olduğu sonucuna varılmıştır [9].

Ari ve Vatansever (2009) Bulanık Mantık Tabanlı Mesleki Yönlendirme adlı yaptıkları çalışmada geliştirdikleri yazılımda, öğrencilerin mesleki yeterliliklerin tespit edilerek hangi mesleklerde başarılı olabileceklerinin tahmini üzerinde çalışılmıştır. Mesleki ilgi ve yeteneklerin tespitinde, bulanık mantık kuralları ve bu türde gerçekleştirilmiş çalışmalar ile daha verimli sonuçlara ulaşılabileceği görülmüştür [10].

Uzun (2005) tarafından yapılan çalışmada, tıp alanında ameliyat sonrası hasta seyirlerinin takiplerinin bilgisayar sistemleri tarafından yapılması için, makine öğrenmesinde kullanılan sınıflandırma algoritmalarıyla bilgisayarların eğitilmesi amaçlanmış ve bu kapsamda makine öğrenmesi sınıflanma algoritmaları tıbbi veriler üzerinde uygulanmıştır. Yapılan uygulamalar sonucunda, bu alanda en iyi algoritmanın hangisi olduğu belirlenmeye çalışılmıştır [11].

Başkeleş (2006) yazılım mühendisliği alanında yazılım eforlarının makine öğrenmesi algoritmalarıyla önceden tahmin edilmesi üzerinde çalışmıştır. Yazılım eforlarının doğru tahmin edilmesinin, yazılım projelerinde maliyetleri düşüreceğine dikkat çekmiş ve yaptığı çalışma sonucunda kullanılan algoritmalarla yazılım eforu tahminlerinin yapılabileceği sonucuna varmıştır [12].

Amasyalı (2004) tarafından yapılan çalışmada yeni ilaç tasarımlarının makine öğrenmesi algoritmalarıyla bilgisayar destekli ortamda yapılması üzerinde durulmuş ve yeni makine öğrenmesi algoritmaları geliştirilmeye çalışılmıştır. Yapılan çalışma kapsamında, konuyla ilgili olarak makine öğrenmesinde bir çok yöntem geliştirilmiş ve bu yöntemlerin araştırmacı ve kullanıcılar için çok yararlı veri kümeleri ürettiği sonucuna varılmıştır [13].

Benzer bir çalışma ise Gökbay (2007) tarafından meme kanserinin erken teşhisinin makine öğrenmesi algoritmalarıyla gerçekleştirilebilmesi üzerine yapılmıştır. Çalışma kapsamında geliştirilen sistem mamografi görüntülerine uygulanmış ve önerilen metodun, radyologların mamografi görüntülemeye kullanabileceği alternatif bir yöntem olduğu sonucuna varılmıştır [14].

Makine öğrenmesi tekniklerinin finans alanında kullanılmasına yönelik bir çalışma ise Bahadır (2008) tarafından yapılan Bayes teoremi ve yapay sinir ağları modelleriyle borsa gelecek değer tahmini uygulaması isimli çalışmadır. Yapılan çalışmada Bayes teoremi temelli yapay sinir ağları ve uzman sistemler incelenmiş ve borsa hisse senedi değerlerinin gelecekteki değerlerinin tahmini yapılmaya çalışılmıştır [15].

Altunyaprak (2006) yaptığı çalışmada istenmeyen elektronik postaların belirlenmesi konusunu incelemiştir. Çalışma kapsamında Bayes teoremini kullanarak geliştirdiği yazılımı istenmeyen elektronik postalar ve normal postalardan oluşan mesajlar üzerine uygulayarak, istenmeyen postaların otomatik olarak filtrelenmesini amaçlamıştır. Bayes teoremi ile geliştirilen yazılım ile istenmeyen elektronik postaların filtrelenmesinde başarılı sonuçlar verdiği görülmüştür [16].

Kahraman (2009) yaptığı doktora çalışmasında, geleneksel öğretim sistemlerinde ortaya çıkan problemlerin çözümüne yönelik, Bayes teoremini kullanarak bir öğretim aracı geliştirmiştir. Geliştirilen araç sayesinde öğrencilerin geleneksel sistemlere kıyasla daha başarılı olduğu sonucuna varılmıştır [17].

### 3. MESLEKİ REHBERLİK VE MESLEK SEÇİMİ

Çalışmanın bu bölümünde meslek, mesleki rehberlik kavramları, mesleki rehberliğin gelişim süreci, ilgili kuramlar, meslek seçimi süreci ve mevcut rehberlik faaliyetleri konuları anlatılmıştır.

#### 3.1. Meslek

Meslek kavramı birçok kaynakta farklı şekillerde tanımlanmıştır. Türk Dil Kurumu'nun Büyük Sözlüğü'nde geçen tanıma göre meslek kavramı, *“Belli bir eğitim ile kazanılan sistemli bilgi ve becerilere dayalı, insanlara yararlı mal üretmek, hizmet vermek ve karşılığında para kazanmak için yapılan, kuralları belirlenmiş iş”* olarak tanımlanmaktadır (TDK, 2003).

Bir başka tanıma göre ise meslek, bireyin geçimini nasıl sağladığını, ne iş yaptığını, ne kadar kazandığını, nasıl bir çevrede oturduğunu ve kimlerle ilişki kurduğunu, kişisel bazı özelliklerinin ve kültürünün düzeyini gösteren, nispeten kişilik özelliklerine uygun iş, uğraş ve faaliyetlerden oluşan, resmen tanınmış ve grup ve üyelikle aidiyet özelliği sağlayan bir kavramdır [2].

Meslek genel anlamda bireylerin geçimlerini sağlaması yanı sıra, kişilerin toplumdaki unvanını ve yaşam tarzlarını belirleyen bir kavramdır. Çünkü insanların günlük yaşamlarının büyük bir bölümünü mesleği ile uğraşarak geçirmektedir. Bunun dışında kalan zamanlarını değerlendirirken, sosyal aktivitelerde amatör olarak yaptığı uğraşlarda(resim, heykel, bitki yetiştirme, müzik aleti çalma vb.) kişinin resmi olmasa da mesleği olarak görülebilir. Bu açıdan bakıldığında meslek, bireylerin yeteneklerini yansıtmaya imkan tanıyarak yaşamdan zevk almasını ve zamanını en iyi şekilde değerlendirerek kaliteli bir yaşam sürdürmesini sağlayan, aynı zamanda psikolojik doyum sağlayan bir kavramdır. Bütün bu nedenler ele alındığında, meslek seçimi insanların tüm yaşamını doğrudan etkileyen çok önemli bir karar olmaktadır.

### 3.2. Meslek Seçimi ve Mesleki Eğitim

Meslek seçimi bir kimsenin kendisine açık olan meslekleri çeşitli yönleri ile değerlendirip kendi ihtiyaçları ve beklentileri açısından istenilen yönleri çok, istenilmeyen yönleri az olan birine yönelmeye karar vermesidir [18]. Meslek seçen bir birey aynı zamanda hayatına nasıl bir yön çizeceğini de seçmektedir. Hatta meslek seçimi, bir kişinin kiminle evleneceğini, dünya görüşünü, günlük yaşam biçimini ve alışkanlıklarını belli bir biçime sokan etkilere de sahiptir [2]. Bireylerin hayatında bu kadar önemli olan meslek seçim kararı hatalı yapıldığında çok olumsuz sonuçlar doğuracağı açıktır. Sevdiği ve istediği işte çalışan biri mutlu olurken, ilgi, yetenek ve değerlerine uygun olmayan bir işte çalışmak zorunda kalan bir insan bilişsel bir çelişki yaşayacak ve işi onu hasta edecektir [1].

Kişinin mesleki alanını doğru seçmesi kadar, bu seçimin doğru zamanda yapılması da bir o kadar önemlidir. Mesleki yaşamına ve eğitimine erken yaşlarda başlayan bireylerin, daha geç yaşlarda başlayan bireylere göre daha deneyimli oldukları ve iş yaşamlarında mesleki açıdan daha hazır oldukları bir gerçektir. Mesleki eğitime üniversite çağlarında başlayan bireyler lise çağlarında başlayan bireylere göre mesleklerine adapte olmakta oldukça zorlandığı görülmektedir.

Günümüz ekonomik koşulları hem dünyada hem de ülkemizde bireylerin iş bulmalarını oldukça zorlaştırmaktadır. Artık iş bulmak için bir diploma sahibi olmak yetmemekte, işverenler eleman tercihlerinde, alacakları kişilerin deneyim ve tecrübelerine büyük önem vermekteler. Bu anlamda kişinin mesleki becerilerini geliştirmesi ve kendisini daha iyi yetiştirebilmesi bakımından mesleki eğitimin erken yaşlarda başlaması çok büyük bir önem kazanmaktadır.

Ülkemizde de son yıllarda mesleki eğitime verilen önem oldukça artmıştır. M.E.B. bünyesinde bulunan mesleki ve teknik ortaöğretim kurumları güncel modüler tabanlı eğitim sistemine geçmişlerdir. Bu kurumların fiziki alt yapıları ve donanım envanterleri her geçen gün daha da zenginleşmektedir. Bu kurumların iş dünyasına ara eleman yetiştirme konusunda önemli bir görevi de vardır. Üniversite sınavına

giriş koşullarının iyileştirilmesi ve bu tip okullardan mezun olan bireylerin daha erken çağlarda iş bulabilmeleri gibi etmenler son yıllarda mesleki ve teknik ortaöğretim kurumlarını cazip hale getirmiş ve bireylerin tercihlerini bu yönde etkilemiştir. Mesleki eğitime giren birey için en önemli karar ise hangi mesleki alanda eğitim alacağıdır. Bu kararı verme durumunda olan birey ya ilköğretim kademesinin son senesinde ya da lise birinci sınıfın sonunda bu seçimle karşı karşıya kalmaktadır.

Meslek seçimi yapılırken kişiyi etkileyen birçok etmen bulunmaktadır. Yazıcı(2007) bireyin yaşamında çok önemli bir aşama olan meslek seçim kararını etkileyen faktörleri şu maddelerle belirtmiştir [7];

- Yetenekler
- Okul Yaşantıları
- Aile
- Kişilik
- Öz kavramı (self-concept) ve Öz-yeterlik (self-efficacy)
- Değerler
- Cinsiyet
- İlgiler
- Çevresel Faktörler

Birey bu faktörlerin tümünün etkisinde kalarak, kendisi için doğru mesleki alanı seçmek için oldukça zorlu bir karar verme sürecinden geçmektedir. Genç yaşlarda bu kararı vermek durumunda olan bireyler mutlaka rehberlik hizmetlerinden yardım almalıdır.

Genç mesleğini seçebilmesi için kendini iyi tanımalı, kendi yeteneklerine ve özelliklerine en uygun meslekleri tanıyıp onlar arasından tercih yaparak karar verebilmelidir. Gencin sağlıklı bir seçim yapabilmesi için okullarda gelişmiş olan rehberlik hizmetlerinden yararlanmalı, ebeveynin ve sosyal çevrenin de olumlu katkısı sağlanmalıdır [1].

### 3.3. Mesleki Rehberlik

Mesleki rehberlik, rehberlik faaliyetlerinin bireyin mesleki gelişimi, eğitimi ve mesleki kararları yönünde yapılabilecek sistematik yardımları içerir [1].

Çakır (2007)'ın Super (1957)'den aktardığı tanıma göre mesleki rehberlik; bireyin kendi kişisel özellikleri ve meslekler hakkında bilgi sahibi yaparak, kendine uygun bir mesleği seçip bu alanda ilerlemesine yardım etme sürecidir [2].

Vurucu (2010) Özoğlu (1982)'ndan yaptığı aktarmada ise mesleki rehberliği, ferde mesleki tercihler yapmasında, bir meslek alanına yönelmesinde, bu meslek alanı içinde bir meslek seçmesinde ve mesleğe hazırlanmasında götürülecek etkinlikleri içeren yönlendirme çalışmaları olarak tanımlamaktadır [19].

Bir başka tanıma göre ise mesleki rehberlik, bir bireyin kendi nitelik ve şartlarını kavrayarak onlara en uygun bir mesleği seçmesi, bu meslek için hazırlanması ve bu mesleğe girerek orada başarı ile ilerlemesi oluşumuna yardım faaliyetleridir [20].

Bu tanımlar ışığında mesleki rehberliğin, mesleki eğitim öncesinde bulunan bireyler için kendileri için en doğru mesleği seçmelerinde, mesleki eğitim sürecinde bulunan bireyler için ise meslek yaşamlarında başarılı bir şekilde ilerlemesini sağlayan önemli bir araç olduğu görülmektedir. Mesleki rehberlik hizmetlerinden faydalanmayan bireyler kendilerine uygun olmayan mesleki bir alanı seçebilir, doğru mesleki alanı seçse dahi seçtikleri alanda yanlış bir kariyer yaşamına sürüklenebilirler.

Mesleki rehberlik Milli Eğitim Bakanlığı'nın Mayıs 2001 tarih ve 2524 sayılı Tebliğler Dergisi'nde yayımladığı Milli Eğitim Bakanlığı Rehberlik Ve Psikolojik Danışma Hizmetleri Yönetmeliği'nin 8. maddesinde şu şekilde tanımlanır;

Eđitim s¼recinde her ¼đrenciye; mesleki tercih yapması, kendine uygun mesleđe yönelmesi, iş yaşamına ve mesleđe hazırlanması için gerekli rehberlik ve psikolojik danışma hizmetleri verilir.

Bu hizmetlerde ařađıdaki hususlar temel alınır:

- Hizmetler bir s¼reç olarak ele alınır, okul ¼ncesi eđitim ve ilköđretimin başlaması ile birlikte bu hizmetler verilir.
- Hizmetlerde ¼đrencinin içinde bulunduđu gelişim dönemi ve bireysel özellikleri dikkate alınır.
- ¼đrenciye ve velisine; ¼đrencinin özellikleri, iş dünyası, meslekler ve bunları edinme yollarına ilişkin güncel bilgiler sistemli olarak aktarılır.
- ¼đrenci, bir meslek alanı veya mesleđi seçme baskısı altında bırakılmaz [21].

### **3.3.1. Mesleki rehberliđin dünya ve ¼lkemizdeki tarihi gelişim süreci**

Mesleki rehberlik faaliyetlerini resmi olmasa da tarihin ilk dönemlerinde de görmek mümkündür. İlk çağlardan bu yana zamanın koşullarına uygun meslek alanlarında, her insan yetenekleri, ilgileri dođrultusunda uygun şekilde yönlendirildiğinde ve yetiştirildiğinde yaptığı işte başarılı olmaktadır. Tarihin önemli filozofları, bilim adamlarının yetiştirilmesinde dođru rehberlik faaliyetlerinin katkısı göz ardı edilemez. Bu kişilerinde yetiştirdiđi ¼đrenciler gene önemli çalışmalar yapmıştır. Tarihte yer alan Türk devletleri gerek medreseleri gerek Enderun adı verilen okullarıyla eđitime önem vermişler ve bu okullarda yetenekli ¼đrencilerin kabiliyetlerine uygun bir şekilde yetiştirilmesini sağlamışlardır. Osmanlılar döneminde Kanuni Sultan Süleyman'ın sarayda belli meslek alanlarında eleman yetiştirmek için çođunluđu gayrim¼slim olan ailelerin zeki, yetenekli çocuklarını Enderun denen bir okulda eđittiđini ve her birini özel yeteneklerine uygun bir meslekte istihdam ettiđi batılı kaynaklarca da belirtilmektedir [1].



Rehberlik, 1900'lü yılların başlarında ABD'de başlamış ve öncelikle mesleki rehberlik alanında gelişmeler görülmüştür. İlk faaliyetler 1895 yıllarında öğrencileri iş olanaklarından haberdar etmek kapsamında görülmüştür. Parsons 1908 yılında Boston'da mesleki büroyu kurması resmi olarak mesleki rehberlik faaliyetlerinin başlangıcı olarak kabul edilmektedir. 1910 yılında Parsons tarafından ilk mesleki rehberlik konferansı yapılmıştır. 1915'te öğrencilere, öğretmenlere ve velilere meslekler hakkında bilgi vermek üzere "Boston Mesleki Rehberlik Dairesi" kurulmuştur [2]. Daha sonraki yıllarda DOT (Dictionary of Occupational Titles) yayını ile meslekler sınıflandırılmıştır. Avrupa'da mesleki rehberlik faaliyetleri özellikle II. Dünya Savaşı sırasında ortaya çıkmış ve 1950'li yılların başından itibaren başta Danimarka, Fransa, Hollanda, İtalya, İspanya, İsviçre, Avusturya, Belçika, İngiltere ve diğer ülkelerde giderek artan bir şekilde devam etmiştir. Avrupa ülkeleri rehberlik faaliyetlerini başlangıçta okul dışı alanlarda uygulamış, daha sonra okullarda uygulamaya başlamıştır. Rehberlik faaliyetleri bu ülkelerde daha çok mesleki rehberlik faaliyetleri kapsamında görülmüştür.

Ülkemizde rehberlik faaliyetleri aynı Avrupa'da olduğu gibi II. Dünya Savaşı sırasında gelişmeye başlamıştır. Bu yıllarda ABD.'li uzmanlar ülkemizdeki eğitim sistemi üzerine yaptıkları araştırmalar sonucunda rehberlik faaliyetlerinin eksikliğini belirtmişler ve bunun üzerine bu ülkede yüksek lisans ve doktora seviyesinde eğitim almak üzere eğitimciler gönderilmiştir. İlk defa 1953-54 ders yılında Gazi Eğitim Enstitüsü Pedagojik ve Özel Eğitim Bölümlerinde rehberlik bağımsız bir ders olarak programda yer almıştır. Daha sonraki dönemlerde MEB şuralarında eğitsel ve mesleki rehberlik konularına yer verilmiştir. Üniversitelerde PDR alanında lisans ve lisansüstü eğitimlere başlanarak rehberlik alanında uzman kişiler yetiştirilmiştir.

### **3.3.2. Mesleki rehberliğin amaçları**

Mesleki rehberlikte temel amaç; insan yaşamının önemli bir alanını kapsayan meslek seçimi konusunda, bireyin uygun tercihler yapmasına ve seçtiği mesleğe uyum göstermesine katkı sağlamaktır. Mesleki rehberlik bir şemsiye kavram olarak düşünülmelidir. Bu şemsiye içinde; bilgi verme, tavsiyelerde bulunma, mesleki

danışma yapma, beceri eğitimi kazandırma ve mesleki değerlendirmeler yapma gibi farklı işlevler vardır [7].

Üre(2006) Özgüven(2000)'den yaptığı aktarmada mesleki rehberliğin amaçlarını şu maddeler ile açıklamıştır:

- Eğitimden ayrı düşünmeksizin öğrencilerin ilgi yetenek ve değerlerine uygun olan mesleklerin görevleri, nitelikleri, sorumlulukları ve sağlayacakları imkanlar hakkında doğru, gerçekçi ve objektif bilgi edinmelerine yardım etme.
- Öğrencilere kendi kişisel değer, ilgi ve yeteneklerini keşfedip anlayabilecekleri yaşantılar hazırlama ve sunma. Bu yolla gençlerin kendilerine uygun mesleklerle ilişkiye geçmelerine yardımcı olma.
- Kendi değer, yetenek ve ilgilerinin toplum gözündeki değerini anlamasına ve kendini uyarlamasına yardımcı olacak yaşantılar hazırlama.
- Girmeyi planladığı mesleki alanın gerçek sınırlarının farkına varmasını sağlama.
- Gençlerin farklı türdeki meslek ve faaliyetler hakkında detaylı bilgiler edinmesine yardımcı olma.
- Gençlerin çeşitli meslekleri eleştirel bir gözle incelemelerine ve bunları karşılaştırmalı olarak değerlendirebilmesine yardımcı olacak eğitim ortamları hazırlama.
- Mesleki karar aşamasında karşılaşacağı problemler için psikolojik danışmandan yardım alabilecek inisiyatifini kazanmasına yardım etme.
- Öğrencinin ihtiyaç duyduğu bütün bilgileri edinmesine yardım etme.
- Bir iş ya da mesleğe uygun hazırlık ve eğitim planlaması yapmasına yardımcı olma.
- İş ortamı ve çalışma grubu içindeki yerini anlamasına, grubun değerli bir üyesi olmasına yardım etme.
- Başarının çaba ile elde edilebileceği ve mesleki doyumun da işi bilinçli ve yeterli düzeyde yapmak ile sağlandığını anlamasına yardım etme.

- Zihinsel, bedensel ve ekonomik yönden özürü veya yetersiz olanların iş yaşamını gerçekleştirmelerine ve uyum sağlamlarına yardım etmek.

### 3.3.3. Mesleki rehberlikte özellik-faktör yaklaşımı

Meslek seçimini açıklamaya yönelik ilk kuramsal açıklamalar, Parsons'un "Bir Meslek Seçimi" (Choosing a Vocation, 1909) adlı kitabında açıkladığı ilke ve yöntemlere dayanan özellik-faktör (trait and factor) yaklaşımıdır. Bugün eğitim-öğretim kurumları içerisinde yürütülen mesleki rehberlik hizmetlerinde de bu kuramsal yaklaşımın görüşleri ağırlıklı olarak kullanılmaktadır [2]. Bu yaklaşım meslek seçimini bireyin psikolojik özellikleriyle mesleklerin gerekleri arasındaki uyumun bir fonksiyonu olarak ele almaktadır [22]. Bu kuramda mesleki rehberlik 3 aşamadan oluşmaktadır. Bu aşamalar sırasıyla; öğrenciyi tanıma, mesleklerin incelenmesi ve son olarak da bireyin kişisel özellikleri ile mesleklerin gerektirdiği özellikler arasında ilişki kurma karar verme aşamalarıdır. Öğrenciyi tanıma aşamasında bilgi toplarken dikkat edilmesi hususlar şöyle açıklanmaktadır [23, 24, 25];

#### 1. Bireyin Fiziki Özellikleri

- Cinsiyet, boy, yaş, ağırlık
- Bedensel özellikleri, özür durumu
- Sağlık durumu
- Görünüş, ses tonu

#### 2. Bireyin Kişilik Özellikleri

- Zeka düzeyi
- Özel yetenekleri
- İlgileri, değer yargıları
- Mesleki değer yargıları
- İhtiyaçları, motivasyonu, amaçları
- İletişim becerileri

#### 3. Akademik Özellikleri

- Okuldaki genel başarı derecesi, başarılı ve başarısız olduğu dersler

- Boş zaman hobileri
- Okuldaki kol faaliyetleri, sosyal ve kültürel etkinliklere katılma durumu
- Çalışma tarzı, alışkanlıkları, iş tecrübesi

#### 4.Sosyolojik Özellikleri

- Ailenin sosyo-ekonomik durumu
- İkamet edilen yer
- Ailenin ihtiyacı
- İşi etkilemesi mümkün yaşantı ve tecrübeler

Yapılan bu çalışmada öğrencilere ait veriler toplanırken Parsons'un geliştirdiği bu yaklaşımda önerilen özellikler dikkate alınmıştır. Öğrenci tanıma aşamasından sonra mevcut meslekler ile öğrencinin özellikleri değerlendirilerek, bireyin hangi mesleğe daha uygun olacağını belirlemek için karar verme aşamasına geçilmektedir.

#### 3.3.4. Mevcut mesleki yönlendirme teknikleri

Ülkemizde mesleki rehberlik hizmetleri, eğitim süreci içerisinde, rehber öğretmenlerin liderliğinde, öğretmen ve diğer ilgililerin yardımıyla gerçekleştirilmektedir [8]. Yönlendirmeler yapılırken anketler, meslek envanterleri, tutum ölçekleri, yetenek ve kişilik testleri yaygın olarak kullanılmaktadır. Günümüzde bilişim teknolojilerin sunduğu imkanlar ile bu araçlarla veri toplama işlemleri bilgisayar ortamında yapılmaya başlamıştır. Bu sayede hem zaman ve işgücünden büyük tasarruf edilmiş hem de verilerin daha doğru değerlendirilmesi sağlanmıştır. Ayrıca verilerin sağlıklı bir şekilde saklanması ve bu verilerden edinilen tecrübelerin, ileriye yönelik fikir vermesi açısından kullanılması sağlanmıştır. Ülkemizde halen yeterince rehber öğretmen ve mesleki yönlendirme uzmanı olmadığı düşünüldüğünde yönlendirme çalışmalarının bilgisayar destekli yapılması artık kaçınılmaz olmuştur. Ülkemizde bu tür çalışmalar, Kuzgun ve Sözüoğlu'nun ÖSYM'de başlattıkları ve BİLDEMER adı verilen uygulamalar ile başlamıştır [7]. Daha sonraki yıllarda bu tarz programların kullanımının ve geliştirilmesinin giderek yaygınlaştığı görülmektedir. Son zamanlarda MEB

tarafından, mesleki eğitime geçme sürecinde olan öğrencilerin yönlendirmelerine yardımcı olması bakımından MBS (Ulusal Mesleki Bilgi Sistemi) geliştirilmiştir.

## 4. VERİ MADENCİLİĞİ VE MAKİNE ÖĞRENMESİ

### 4.1. Veri Madenciliği

Veri madenciliği bilgi topluluklarından varılmak istenen hedef doğrultusunda anlamlı bilgilerin çıkarılmasını ve sınıflama yapılmasını sağlayan, bilişim teknolojisinin insanlara sunduğu önemli bir kavramdır.

Bilişim teknolojilerinin gelişmesine paralel olarak günlük yaşamdaki iş süreçleri, giderek artan bir şekilde elektronik ortama taşınmaktadır. Klasik veri yönetim işlemleri önemli bir zaman kaybına yol açarken aynı zamanda işgücü gereksinimine de ihtiyaç duyulmakta ve verilerin düzenli bir şekilde tutulmaması sonucu önemli bilgi kayıpları yaşanmaktaydı. Bunun yanı sıra verilerin yapılandırmadan saklanması, bu verilerden anlamlı bilgiler çıkarılmasını sağlayacak olan analiz işlemlerini de oldukça zorlaştırmaktadır. Tüm bu dezavantajlar ve her geçen gün daha da gelişen bilgisayar sistemlerinin sunduğu rahatlık, maliyetlerin ucuzlaması, bu sistemlere erişimin zaman ve mekan sınırı tanımadan sağlanması kurumları veri yönetim süreçlerini elektronik ortamda gerçekleştirmeye yöneltmiştir.

Bilgisayar sistemlerinin kullanımı ile verilerin saklanması ve yönetimi kurumlara büyük kolaylıklar sağlamıştır. Kurumlar açısından bilgisayar sistemlerinin verilerin saklanmasından daha çok verilerden anlamlı bilgiler elde edilmesinde kullanılması daha önemlidir. Çünkü veriler tek başlarına bir anlam ifade etmezler. Bir süpermarkette, müşterinin yaptığı bir alışverişte aldığı kalem sayısı o firma için yalnızca bir satış kayıdır yani veridir. Fakat o gün satılan toplam kalem sayısı, aylık ortalama kalem satış miktarı, kalem ile birlikte en çok hangi ürünün satıldığının belirlenmesi, gelecek ay bu ürünün ne kadar satılacağı vb. ifadeler ise anlamlı birer bilgidir. Süpermarket örneği için satılan kalem adedi verisi o ürünün satış ve reklam politikasını belirlerken bir anlam ifade etmez fakat bu ürünle birlikte en çok tercih edilen ürünün bilgisi ise ürüne yönelik reklam ve satış politikalarında oldukça belirleyicidir. Kalem ile birlikte en çok kalem ucu ürününün satıldığı bilgisi varsayılırsa bu durumda bu süpermarket bu iki ürün reyonunu yan yana koyabilir

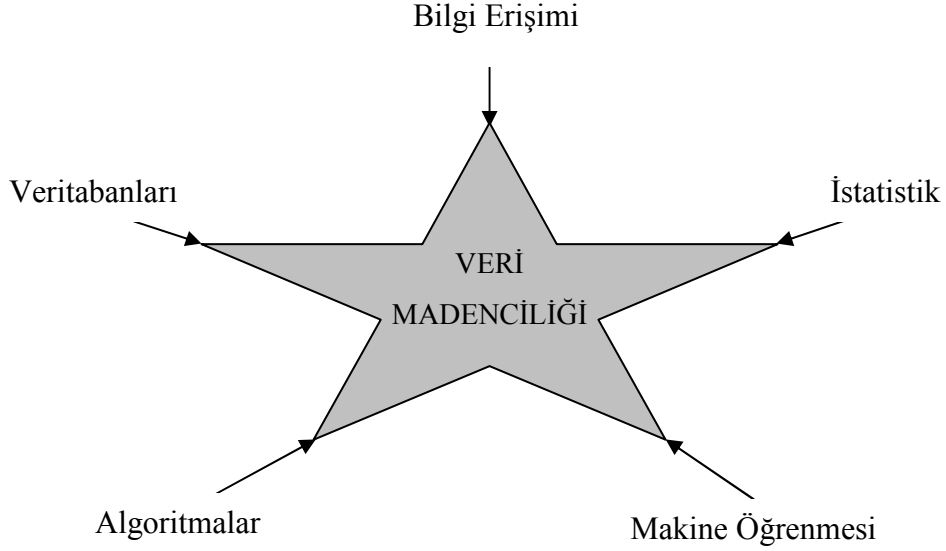
veya çeşitli promosyon çalışmalarında bu iki ürünü bir arada müşterilerine sunarak bu ürüne ait satış oranlarını arttırabilir. Verilen bu basit örnekteki, karar sürecinde veriler tek başlarına bir işe yaramaz iken ancak bilgiye dönüştürüldüklerinde bir anlam ifade etmektedirler.

Günümüz bilgi toplumunda her kurum kendi alanında rekabet edebilmesi için iş süreçlerinde elde ettiği verilerden mutlaka anlamlı bilgiler çıkartmalıdır. Kurumlar veriler üzerinde yaptıkları analiz çalışmaları sonucu elde ettiği anlamlı bilgiler ışığında gelecekteki politikalarını belirlemektedirler. Bu bilgiler sayesinde belirlenen politikalar, kurumların ileriye yönelik düzeltici ve iyileştirici adımlar atmasına olanak tanımaktadır. Gelecek planlamasını yapabilen kurumlar maliyet, zaman, işgücü anlamında da kazançlı olmaktadır.

Veritabanı temelli bilgisayar sistemlerinin yaygınlaşması ile birlikte her türlü veri elektronik ortama aktarılmakta ve bunun sonucunda büyük veri yığınları ortaya çıkmaktadır. Veri tabanlarında saklanan veri, bir dağa benzetilirse, bu veri dağı tek başına değersizdir ve kullanıcı için çok fazla bir anlam ifade etmez. Ancak bu veri dağı, belirli bir amaç doğrultusunda sistematik olarak işlenir ve analiz edilirse, değersiz görülen veri yığımında, amaca yönelik sorulara cevap verebilecek çok değerli bilgilere ulaşılabilir [26]. Bu büyük veri yığınlarından anlamlı bilgilerin nasıl elde edileceği konusunda ise veri madenciliği kavramı ortaya çıkmaktadır.

Veri madenciliği büyük veri yığınlarından anlamlı ve gelecekteki süreçlerde karar vermeye yardımcı olacak nitelikte bilgilerin elde edilme sürecidir. Bir başka tanıma göre ise veri madenciliği; veri ambarlarındaki tutulan, çok çeşitli ve çok miktarda veriye dayanarak daha önce keşfedilmemiş bilgileri ortaya çıkarmak, bunları karar verme ve eylem planını gerçekleştirmek için kullanma sürecidir [6]. Veri madenciliği sürecinde bu bilgiler ortaya çıkartılırken veritabanı teknolojileri, istatistiksel yöntemler, algoritmalar, makine öğrenmesi teknikleri ve yapay zeka alanlarından faydalanılır.

Veri madenciliğinin tarihsel gelişiminde yararlanılan konular Şekil 4.1’de gösterilmiştir:



Şekil 4.1. Veri madenciliğinin tarihsel perspektifi [27]

Veri madenciliği ham veriler üzerinden anlamlı bilgi elde etme ile ilgilidir. Bu anlamda bakıldığında istatistik uygulamalarına benzemektedir. Ancak klasik istatistiksel uygulamalar yeterince düzenlenmiş ve çoğunlukla özet veriler üzerinde çalıştırılır. Veri madenciliğinde ise milyonlarca ve hatta milyarlarca veri ve çok daha fazla değişken ile ilgilenilir [28]. Veri madenciliği süreci sonucunda elde edilen bilgiler karar destek sistemlerinde kullanılmaktadır.

#### 4.1.1. Veri madenciliği ve veritabanı sistemlerinin tarihi gelişimi

Veri madenciliği insan ve bilgisayarların ortak çabası ile ortaya çıkmıştır. İnsanlar problemleri tarif ederler, veritabanlarını tasarlarlar ve amaçları belirler. Bilgisayarlar ise veriler arasında bu amaçlara uygun verileri eleterek desenleri ortaya çıkartırlar [29].

Veri madenciliği tarihinin başlangıcı 1700’lü yılların sonunda Bayes teoreminin ortaya çıkışı ile başladığı söylenebilir. 1900’lü yılların başında istatistik bilimindeki



gelişmeler regresyon analizinin geliştirilmesi ile devam etmiştir. 1950'li yılların başında yapay sinir ağları, 1960'lı yılların başlarında makine öğrenmesi kavramı ile yapay zeka alanında ilerlemeler sırasıyla karar ağaçları, kümeleme yöntemleri ile devam etmiş ve veri madenciliğinin gelişiminde önemli bir rol oynamıştır. 1960'ların sonunda veritabanı teknolojisi ile elektronik ortamda sayısal verilerin kaydedilmesi ve yönetilmesi süreci başlamıştır.

Bilgi sistemlerinin ilk yıllarında basit sıralı dosyalar ile veri yönetim süreçleri gerçekleştirilmiş, ardından veritabanı kavramının ortaya çıkması ile beraber ileri uygulamalar gerçekleştirilmeye başlanmıştır. Veritabanları, türlerine göre(sayısal, metinsel, tarihsel, mantıksal vb.) ve yüksek miktarlarda veri saklanmasına imkan tanırken, geliştirilen SQL ve benzeri sorgulama dilleri ile veriler üzerinde istatistiksel hesaplama işlemlerinin yapılmasını sağlamıştır. İlişkisel veritabanı sistemleri ile veri yönetim şekli ve kapsamı ilerlemiş, yazılım teknolojilerindeki gelişmeler ile birlikte bankacılık, otomasyon sistemleri gibi yüksek seviyeli uygulamalar ortaya çıkarılmıştır. 1970 ve 1980'li yıllarda ise ilişkisel veri modeli, genetik algoritmalar, karar ağaçları algoritmaları geliştirilmiştir. 1990'lı yılların başında internet teknolojilerinin yaygınlaşması ile birlikte uygulamalar ağ üzerine taşınarak sistemler daha çok kullanıcıya ulaşmış ve arama motorları geliştirilmiştir. Ağ üzerinde bulunan uygulamalar, XML kavramının ortaya çıkması ile birlikte ortak veritabanı ve bilgi kullanımı sürecine geçmişlerdir.

Çevrimiçi hareket işleme (OLTP) yöntemleri ile kurumlar günlük iş akışları ve raporlama uygulamalarını bu sistemler üzerinde gerçekleştirmeye başlamışlardır. Veritabanı ve yazılım teknolojilerinin gelişmesi ve daha yaygın olarak kullanılması ile birlikte her geçen gün depolanan veri miktarı artmış ve büyük veri ambarları ortaya çıkmıştır. Bu büyük veri ambarlarından karar destek sistemlerinde kullanılmak üzere bilgi elde etme sürecinde ise klasik OLTP sistemler yerine analiz çözümleri sunan OLAP sistemleri kullanılarak ileri veri madenciliği uygulamalarına geçilmiştir. Veri sayısı çok olunca, bazı özel analiz algoritmalarının geliştirilmesi gerekmiş, ayrıca verinin saklandığı ortamların da örneğin veri ambarı biçiminde yeniden düzenlenmesini gerekli kılmıştır [28]. Hızla artan veri kayıtları (GB/saat),

otomatik istasyonlar, uydu ve uzaktan algılama sistemleri, teleskopla uzay taramaları, gen teknolojisindeki gelişmeler, bilimsel hesaplamalar, benzetimler, modeller, veri madenciliğini zorunlu kılmıştır [6].

#### **4.1.2. Veri madenciliği uygulama alanları**

Günümüzde kurumların iş süreçlerini elektronik ortama taşınması sonucu oluşan ve giderek artan veri miktarı ve bu verilerin analiz edilerek karar destek sistemlerinde kullanılma ihtiyacı ile birlikte, veri madenciliği birçok alanda yaygın olarak kullanılmaktadır. Bu alanlar aşağıdaki gibi özetlenmiştir [30, 28, 31, 32]:

##### **Pazarlama**

- Müşterilerin satın alma örüntülerinin belirlenmesi
- Müşterilerin demografik özellikleri arasındaki bağlantıların bulunması
- Posta kampanyalarında cevap verme oranının artırılması
- Mevcut müşterilerin elde tutulması, yeni müşterilerin kazanılması
- Pazar sepeti analizi
- Müşteri ilişkileri yönetimi
- Müşteri değerlendirme
- Satış tahmini

##### **Bankacılık**

- Farklı finansal göstergeler arasında gizli korelasyonların bulunması
- Kredi kartı dolandırıcılıklarının tespiti
- Kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi
- Kredi taleplerinin değerlendirilmesi

##### **Sigortacılık**

- Yeni poliçe talep edecek müşterilerin tahmin edilmesi
- Sigorta dolandırıcılıklarının tespiti
- Riskli müşteri örüntülerinin belirlenmesi

##### **Elektronik Ticaret**

- Saldırıların çözümlenmesi
- e-CRM uygulamalarının yönetimi

- Web sayfalarına yapılan ziyaretlerin çözümlenmesi

### **Savunma Sistemleri**

- Suç ve suçlularla ilgili örüntülerin saptanması
- Terörist ve düşman eylemlerinin modellenmesi ve kestirimi
- Uçak kazalarında hataların saptanması ve önlemlerin alınması
- Adli verilerin analizi ve suç unsurlarının öngörülmesi

### **Borsa**

- Hisse senedi fiyat tahmini
- Genel piyasa analizleri
- Alım-satım stratejilerinin optimizasyonu

### **Telekomünikasyon**

- Kalite ve iyileştirme analizleri
- Abonelik tespitleri
- Hatların yoğunluk tahminleri

### **Sağlık ve İlaç**

- Test sonuçlarının tahmini
- Ürün geliştirme
- Tıbbi teşhis
- Tedavi sürecinin belirlenmesi

### **Endüstri**

- Kalite kontrol analizleri
- Lojistik
- Üretim süreçlerinin optimizasyonu

### **Eğitim**

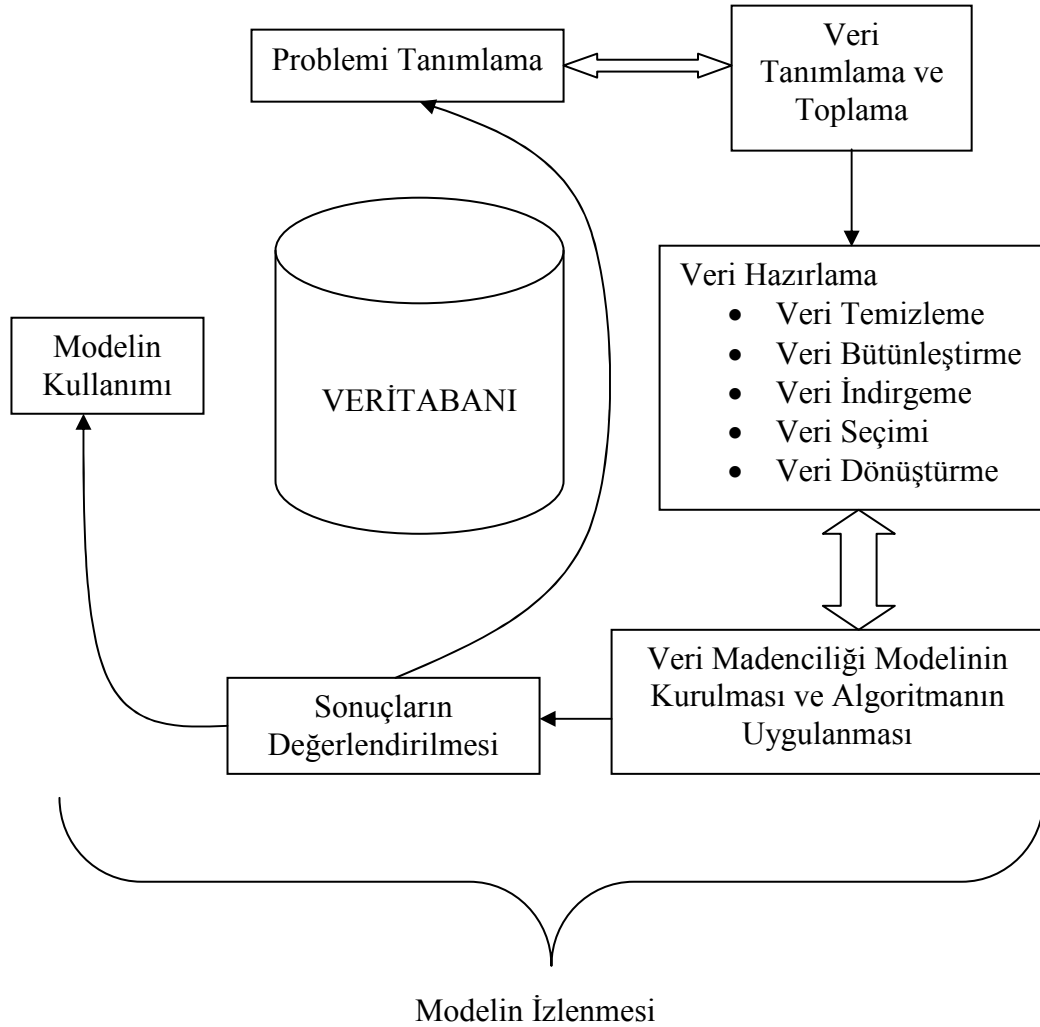
- Ölçme ve değerlendirme çalışmaları
- Mesleki rehberlik faaliyetleri
- Sınav başarısı analizleri

### **İnternet**

- Arama motoru optimizasyonu
- Kelime anlamı çıkarma
- Önemsiz elektronik postaların belirlenmesi

### 4.1.3. Veri madenciliği uygulama süreci

Veri madenciliği uygulamaları bir kez yapıldıktan sonra hemen istenilen sonuçlara ulaşılan yöntemler değildir. Veri madenciliği uygulamalarında veritabanlarından bilgi keşfi yapılırken, birbirini izleyen adımlardan oluşan ve tekrar eden bir süreç gerektirmektedir. Veri madenciliği sürekli tekrarlanabilen dinamik bir süreçtir. Sürecin bazı aşamalarında elde edilen sonuçlara göre önceki aşamaların tekrar düzenlenmesi gerekebilir. Şekil 4.2’de veri madenciliği süreci gösterilmiştir.



Şekil 4.2. Veri madenciliği uygulama süreci

Veri madenciliği süreci sonucunda hedeflenen bilgiye ulaşmak için başlangıçta bu sürecin aşamaları doğru planlanmalıdır. Bu süreçte çıkabilecek sorunlar problemler veri madenciliği algoritmaları ve ürünleri tarafından mutlaka ele alınmalıdır [27].

Dunham(2002) veri madenciliğinin uygulanmasında yer alan problemleri şöyle sıralamıştır:

- İnsan etkileşimi
- Çok değişkenlilik
- Sapan değer
- Sonuçların yorumlanması
- Sonuçların görüntülenmesi
- Büyük veri setleri
- Yüksek boyutluluk
- Multimedya veri
- Kayıp veri
- Konu dışı veri
- Gürültülü veri
- Dinamik veri
- Entegrasyon
- Uygulama [27].

### Problemin Tanımlanması

Veri madenciliği uygulama sürecinin ilk adımı yapılacak olan işin yani problemin tanımlanmasıdır. Bu aşamada ihtiyaçlar ve süreç sonunda elde edilecek bilginin hangi amaçla kullanılacağı belirlenmektedir. Ayrıca veri madenciliği uygulanırken kullanılacak yazılım, donanım ve iş planına da karar verilmektedir. Süreçte ortaya çıkabilecek riskler tespit edilmelidir. İlk aşama sürecin devamındaki aşamaları belirleyeceğinden bu bölümde çok dikkatli planlanma yapılmalı ve hedefler ortaya açıkça konulmalıdır. Ayrıca yanlış tahminlerde katlanılacak olan maliyetlere ve doğru tahminlerde kazanılacak faydalara ilişkin tahminlere de bu aşamada yer verilmelidir [30].

### Veri Tanımlama ve Toplama

Problemi tanımladıktan ve anladıktan sonraki adım, veriyi toplamak, veri kalitesinin yeterli olup olmadığını ve iş hedeflerine uygun olup olmadığını belirlemektir. Veri yapısı daha iyi anlaşıldıkça problem tanımı değiştirilebilir veya yeniden yapılabilir [33]. Veriler toplanırken kurumun kendi verileri dışında belirlenen amaca uygun olarak başka kuruluşların veritabanlarından da faydalanılabilir.

### Veri Hazırlama

Tanımlanan probleme uygun toplanan verilerin uygulanacak olan veri madenciliği modeline uygun hale getirildiği aşamadır. Modelleme aşamasının sağlıklı sonuç vermesi bu aşamada hazırlanan verilere bağlıdır. Veri madenciliği sürecinde verilerin modele bağlı olarak yeniden düzenlenme ihtiyacı doğarsa veri hazırlama aşaması tekrarlanabilmektedir.

Veri hazırlama sürecinde yapılan işlemler aşağıdaki gibi sıralanmıştır.

- Veri Temizleme
- Veri Bütünleştirme
- Veri İndirgeme
- Veri Seçimi
- Veri Dönüştürme

### *Veri Temizleme*

Farklı kaynaklardan çeşitli yollarla toplanarak elde edilen veriler tam veya istenildiği gibi olmayabilir. Kullanıcılardan alınan veriler eksik, tutarsız veya aşırı uçlarda girilmiş olabilir. Veritabanlarında yer alan bu tür verilere gürültülü veriler denir. Yüzlük sistemde girilmesi gereken bir notun boşluk sistemde girilmesi, bireylerin mezuniyet durumlarını boş bırakması, yaşını 324 girmesi gürültülü verilere örnek olarak verilebilir. Kirli veriler veri madenciliğinde karmaşıklığa ve sonuçların

güvensiz olmasına neden olmaktadır [32]. Gürültülü ya da eksik veriler modellemeye geçilmeden önce düzeltilmelidir. Verilerin düzeltilmesinde kullanılabilecek teknikler aşağıdaki gibi sıralanmıştır [28].

- Eksik değer içeren kayıtlar veri kümesinden atılabilir.
- Kayıp değerlerin yerine genel bir sabit kullanılabilir. Bütün kayıp değerler için aynı sabit kullanılabilir. Örneğin “bilinmiyor” değeri bu eksik veri yerine kullanılabilir. Ancak bütün değişkenlerde kayıp değerler yerine aynı sabit değer kullanımı sorun yaratacaktır.
- Değişkenin tüm verileri kullanılarak ortalaması hesaplanır ve eksik değer yerine bu değer kullanılabilir.
- Değişkenin tüm verileri yerine, sadece bir sınıfa ait örneklerin değişken ortalaması hesaplanarak eksik değer yerine kullanılabilir.
- Verilere uygun bir tahmin yapılarak, örneğin regresyon ya da karar ağacı modeli kurularak eksik değer tahmin edilebilir ve eksik değer yerine kullanılabilir.

### *Veri Bütünleştirme*

Toplanan veriler farklı veri kaynaklarında, veritabanlarında ve değişik kayıt ortamlarında ise analiz işlemleri için verilerin tek bir noktaya toplanması gerekmektedir. Veri bütünleştirme, farklı ortamlardan elde edilen verileri tek bir veri tabanına toplama işlemlerinin tümüdür [31].

### *Veri İndirgeme*

Veri madenciliği uygulanacak veri seti büyük miktarlarda veri içeriyorsa, yöntemin uygulanmasını kolaylaştırmak için veri indirgeme çalışması yapılabilir. Eğer eldeki değişkenlerden sonuca etkisi olmayacaklar var ise çözümleme aşamasında dikkate alınmayabilir. Örneğin kayıtları tanımlamada kullanılan kimlik numarası, okul numarası, kayıt tarihi, isim vb. bilgiler model için hazırlanan veri kümesinden çıkartılabilir. Veri indirgeme çalışmaları çözümleme çalışmalarında harcanacak olan

işgücü ve zamanı önemli ölçüde kısıltacaktır. Veri indirgeme yapılırken veri küpü oluşturma, boyut indirgeme, veri sıkıştırma, örnekleme ve genelleme teknikleri kullanılabilir [34].

### *Veri Seçimi*

Veri madenciliğinde kullanılacak olan model için veri setinden veritabanı sorgulama teknikleri kullanılarak veri seçimi yapılır. Örneğin tahmin edici bir model için, bu adım bağımlı ve bağımsız değişkenlerin ve modelin eğitiminde kullanılacak veri kümesinin seçilmesi anlamını taşımaktadır. Verilerin görselleştirilmesine olanak sağlayan grafik araçlar ve bunların sunduğu ilişkiler, bağımsız değişkenlerin seçilmesinde önemli yararlar sağlayabilir [30].

### *Veri Dönüştürme*

Veri madenciliği algoritmalarının bazıları belli türdeki verilerle çalışırken başka tür verilerle çalışmamaktadır [31]. Değişkenlerin ortalama ve varyansları birbirinden önemli ölçüde farklı olduğu takdirde büyük ortalama ve varyansa sahip değişkenlerin diğerleri üzerinde baskısı daha fazla olur ve onların rollerini önemli ölçüde azaltır. Ayrıca değişkenlerin sahip olduğu çok büyük ve çok küçük değerlerde çözümlerinin sağlıklı biçimde yapılmasını engeller [28]. Bu nedenlerden dolayı model için seçim yapıldıktan sonra, verilerin istatistiksel analizlerinin yapılarak çıkan sonuçlara göre gerekiyorsa dönüştürülmelidir.

Veri dönüştürme işlemi yapılırken çeşitli teknikler kullanılır. Verileri 0 ve 1 arasında değerlere dönüştürmek için Min-Max Normalleştirme adı verilen yöntem kullanılır. A değişkeni için min max normalleştirme aşağıdaki formül kullanılarak uygulanır.

$$A' = \frac{A - A_{\min}}{A_{\max} - A_{\min}} \quad (4.1)$$



Bu formülde A gözlenen,  $A_{\min}$  en küçük gözlenen,  $A_{\max}$  en büyük gözlenen ve  $A'$  ise dönüştürme sonucunda elde edilen değeri temsil etmektedir.

Normalleştirme 0-1 aralığından farklı bir aralıkta yapılacaksa, uygulanması gereken formül aşağıdaki biçimde değiştirilmelidir [35].

$$A' = \frac{A - A_{\min}}{A_{\max} - A_{\min}} (\text{hedef}_{\text{maks}} - \text{hedef}_{\text{min}}) + \text{hedef}_{\text{min}} \quad (4.2)$$

Dönüştürme yapılırken kullanılan bir diğer yöntem ise Z-score Standartlaştırmadır. Bu yöntem, verilerin ortalaması ve standart hatası göz önüne alınarak yeni değerlere dönüştürülmesi esasına dayanmaktadır [28]. B değişkeni için Z-score Standartlaştırma yönteminin uygulanması aşağıdaki formüldeki gibidir.

$$B' = \frac{B - \bar{B}}{\sigma_B} \quad (4.3)$$

Bu formülde B gözlenen,  $\bar{B}$  gözlenen değerlerin aritmetik ortalaması ve  $\sigma_B$  ise gözlenen değerlerin standart sapmasını temsil etmektedir.

### Veri Madenciliği Modelinin Kurulması ve Algoritmanın Uygulanması

Veriler hazırlandıktan sonra veri madenciliği modelinin kurularak uygulandığı yani analizin yapıldığı aşamadır. Veri madenciliği modelleri genel olarak tahmin edici ve tanımlayıcı modeller olmak üzere ikiye ayrılır.

Veri madenciliği modellerini gördükleri işlemlere göre,

- Sınıflama ve regresyon
- Kümeleme
- Birliktelik kuralları ve ardışık zamanlı örüntüler

olmak üzere üç ana başlık altında incelemek mümkündür. Sınıflama ve regresyon modelleri tahmin edici, kümeleme, birliktelik kuralları ve ardışık zamanlı örüntü modelleri tanımlayıcı modellerdir [30].

### Sonuçların Değerlendirilmesi

Veri madenciliği modeli uygulanması ile elde edilen sonuçlar değerlendirilerek kurulan modelin kullanılmaya geçilip geçilmeyeceğine karar verilir. Sonuçların başlangıçta belirlenen hedeflere uygun olmadığı görülürse problem tanımlama aşamasına dönülebilir.

### **4.2. Makine Öğrenmesi**

Günümüzde bilgisayarlar iş gücü olarak insanların yüklerini oldukça hafifletmişlerdir fakat beyin gücü olarak hala istenilen seviyeye ulaşamamışlardır. Yapay zeka teknolojileri bilgisayarların bu alanda gelişmelerini sağlamaya çalışmaktadırlar. Makine öğrenmesi ile geçmiş veriler dayanarak bilgisayara tecrübe kazandırmak yoluyla olayların analizi ve bunların gelecekte alabileceği durumları hakkında tahminleri yaptırılmaktadır. Bu sayede karar verme sürecinde bilgisayarların rolü daha da artacaktır.

Makine öğrenmesi, bilgisayarların geçmiş bilgilerden elde edilen tecrübelerden yararlanarak, gelecekteki olayları tahmin etmesine ve modelleme yapmasına imkan veren bir yapay zeka alanıdır. Bilgisayarın bir olay ile ilgili bilgileri ve tecrübeleri öğrenerek gelecekte oluşacak benzeri olaylar hakkında kararlar verebilmesi ve problemlere çözümler üretebilmesi olarak da tanımlanabilir. Bir başka tanıma göre ise makine öğrenmesi, E deneyimi ve buna uyan bazı T görev sınıflarından, P performansı ile öğrenmesi söylenen bir bilgisayar programıdır. Eğer T görevlerindeki performansı P ile ölçülebiliyorsa, E deneyimi ile geliştirilebilir. Bu yapıya aşağıdaki örnek verilmiştir.

- Görev T: Dama oynamak
- Performans ölçütü P: Rakiplerle oynanan oyunlardaki kazanma yüzdesi
- Antrenman deneyimi E: Kendisi ile alıştırmaya oyunları oynaması [36].

Makine öğrenmesi ile bilgisayar yazılımlarının, olayları daha önceki örneklerden edinilmiş tecrübelerin kullanılarak öğrenmesi sağlanır. Öğrenecek olan bilgisayar sistemleri önce bir örnek almakta ve bu örnekten bazı bilgileri öğrenmektedir. Daha sonra ikinci örneğe bakarak biraz daha bilgi edinmektedir. Bu işlemi öğrenilecek olay ile ilgili genellemeler yapılmaktadır. Bu olaya tecrübelerden öğrenmenin bir yolu olarak bakmak mümkündür [37].

Makine öğrenmesi ve veri madenciliği arasında doğrudan bir ilişki vardır. Makine öğrenmesi tekniklerinin büyük veritabanlarına uygulanması veri madenciliğidir [38]. Makine öğrenmesi veri madenciliği sürecinde uygulama aşamasında yer almaktadır. Bu aşamada seçilen bir makine öğrenmesi tekniği veri seti üzerine uygulanarak sonuçlar elde edilir. Makine öğrenmesi yalnızca veriler üzerinde uygulanan bir teknik değil, bir yapay zeka alanıdır. Veri madenciliği çıkan bilgi ve bu bilgilerin değerlendirilmesi ile uğraşmaktadır. Makine öğrenmesi ise bu bilgiyi çıkarmayı sağlayan teknikler ve bu teknikleri kullanan bilgisayar programlarının kendilerini geliştirmesi ile ilgilenir. İki konu arasındaki en büyük fark; makine öğrenmesi, öğrenme metotlarını geliştirerek, tahminleri ya da tanımları en iyi şekilde, yüksek performans ile nasıl çıkarabileceği ile ilgilenirken, veri madenciliğinin ortaya çıkan bilgi ile ilgilenmesidir [39].

#### **4.2.1. Makine öğrenmesi uygulama alanları**

Makine öğrenmesi, eğitim, finans ve meteoroloji gibi birçok alanda geçmişte bulunan veriler üzerinde uygulanarak gelecekte olabilecek durumların tahminleri ve verilerin sınıflandırmaları yapılabilmektedir. Bu tahminlerin en önemli özelliği, gerçekleşmesi uzun sürebilen ve gerçekleştikten sonra geri dönülmez sonuçlar üretebilen olayların önceden öngörülmesidir. Veri madenciliğine paralel olarak makine öğrenmesi teknikleri birçok alanda uygulanmaktadır. Aşağıda makine öğrenmesinin uygulandığı alanlar özetlenmiştir.

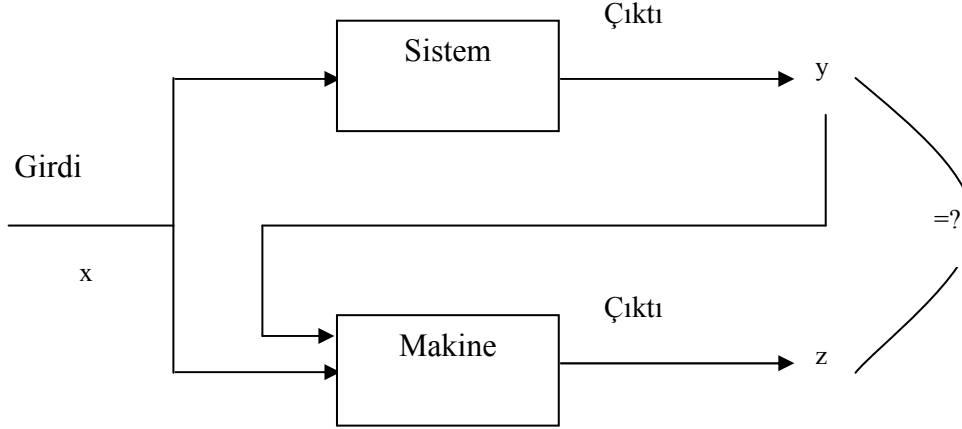
- Astronomik verilerin otomatik sınıflandırılması
- Tıpta hastalıkların semptomlarla tespiti

- Veri Madenciliği ve Analizi(Ekonomi,Finans ve Pazarlama)
- Enerji Sistemleri
- Uçuş eğitimlerinde strateji geliştirme
- Çimento dayanıklılık tahmini
- Doğal dil işleme ve öğretme
- Türkçe sözcük anlamı açıklama
- Karar ağaç yapılarının modellenmesi
- Tıbbi verilerin sınıflandırılması
- Yazılım efor büyüklüğünün tahmin edilmesi
- Kanser türlerinin sınıflandırılması ve teşhisi
- İlaç tasarımı
- Metinlerin otomatik olarak sınıflandırılması
- Kredi riski hesaplaması
- Kredi kartı çalınmasının tespiti
- Üretim hatalarının tespiti
- Teknik direktörler için sporcu performansı analizi
- Haber ve web aramalarının özelleştirilmesi
- Sürücüden bağımsız araba kullanımı

#### **4.2.2. Makine öğrenmesi modeli**

Makine öğrenmesinde amaç bilgisayar programlarının kendisine öğretilen tecrübeler yoluyla gelecek durumları tahmin etmesini sağlamaktır. Makine öğrenmesinde tahmin temeline dayanan tekniklerde, veriler eğitim ve test veri seti olmak üzere iki bölüme ayrılmaktadır. Programlar eğitim veri seti ile tecrübe edinerek, test verileri üzerinde tahmin yapmaya çalışmaktadırlar. Bu tahminlerden en iyi sonucu elde etmek için eğitim veri setinin geniş tutulması gerekmektedir. Çünkü makineler ne kadar çok öğrenirse o kadar tecrübe edinecek ve iyi sonuçlar üretebileceklerdir.

Uzun(2005), Forsyth(1989)'dan yaptığı aktarmada, makine öğrenmesi modelinin çalışma şeması Şekil 4.3'de gösterilmiştir.



Şekil 4.3. Makine öğrenmesi şeması [11]

Şekil 4.3’de yer alan modelin çalışması aşağıdaki gibi belirtilmiştir.

1.  $(x;y)$ 'nin bir kümesi alınır, burada  $x$  bir girdi vektörü ve  $y$  uygun bir çıktıdır.
2.  $y=f(x)$  fonksiyonu, önceden bildirilen bir modelin oluşumudur;
  - Modelin kalite ölçümüne, bir kriter tanımlanır
  - Modelin kullanılacağı bir eğitim kümesi oluşturulur
  - Modelin kullanacağı geçerli bir test kümesi oluşturulur [11].

#### 4.2.3. Makine öğrenmesi teknikleri

Makine öğrenmesi teknikleri uygulanırken birçok algoritma kullanılmaktadır. Bu bölümde sınıflandırma, kümeleme, birliktelik kuralları tekniklerinde yer alan makine öğrenmesi algoritmaları anlatılmıştır.

##### Sınıflandırma

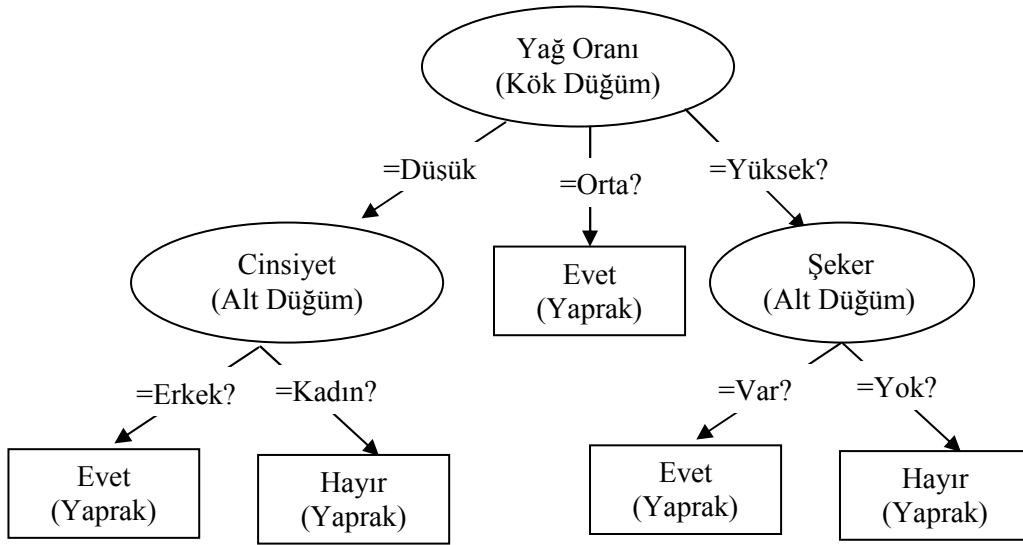
Sınıflandırma tekniklerinin amacı, verileri ortak özelliklerine göre istenilen yapıda gruplara ayırmaktır. Sınıflandırma ile birlikte sınıfları ya da ait oldukları grupları bilinmeyen veriler belirlenmektedir. Sınıflandırma tekniği birçok alanda kullanılmaktadır. Sağlık alanında hastaların kanser risklerinin belirlenmesi, finans alanında müşterilerin kredi ödeme durumlarının belirlenmesi, inşaat malzemelerinin

dayanıklılık durumlarının belirlenmesi, eğitimde öğrenci başarı durumlarının tespiti, hava durumu tahmini, ilaçların sınıflandırılmaları gibi örnekler olarak verilebilir.

Sınıflandırma bir öğrenme algoritmasına dayanır. Öğrenmenin amacı bir sınıflandırma modelinin yaratılmasıdır [28]. Sınıflandırma algoritmaları uygulanırken verilerin bir kısmı eğitim verisi olarak seçilir ve algoritma sınıfı bilinmeyen test verisinin özellikleri ile eğitim veri kümesinin üzerinde çalıştırılarak test verisinin hangi sınıfa ait olduğu belirlenmeye çalışılır. Makine öğrenmesinde sınıflandırma yapılırken karar ağaçları, regresyon ağaçları, destek vektör makinesi ve istatistiksel sınıflandırma yöntemleri kullanılmaktadır. Her bir yöntem için farklı algoritmalar geliştirilmiştir. Sınıflandırma yöntemleri tahmin edici modellerdir [40]. Bu modellerin yanı sıra yapay zeka tekniklerinden genetik algoritmalar ve yapay sinir ağları da sınıflandırma yapılırken kullanılmaktadır.

Karar ağaçları içlerinde verilerin hangi dala yönlendireceğini belirleyen karar düğümlerinden ve bu dalların uçlarında gelen verinin hangi sınıfta olduğunu söyleyen sınıf etiketlerini içeren yapraklardan oluşan hiyerarşik bir yapıdır [13]. Karar ağaçları modeli kural tabanlı öğrenme olarak da ifade edilmektedir. Bu model kurulurken başlangıçta bir kök düğüm belirlenir ve her bir düğümde seçilen niteliğin durumlarına göre verilen karar doğrultusunda alt düğümler oluşturulur. Her bir düğüm tek bir nitelik durumuna düşürüldüğünde, bu durum yaprak olarak adlandırılır ve o düğüm sonucunda bir sınıf belirlenmiş olur. Her düğümün sonunda sınıf belirleninceye kadar bu işlem özyinelemeli olarak devam eder.

Makine öğrenmesinde karar ağacı modeli oluşturmak için, ID3, C4.5, CHAID, CRT, QUEST gibi algoritmalar geliştirilmiştir. Bu algoritmalar, kök düğümden sonra alt düğümlerin oluşturulmasında belirlenecek kriterler açısından farklılıklar göstermektedirler. Şekil 4.4'de örnek bir karar ağacı modeli gösterilmiştir.



Şekil 4.4. Karar ağacı modeli örneği

Regresyon ağaçları ile sınıflandırma tekniği karar ağaçları modelinde olduğu gibi bir kök düğümden alt düğümler ve yapraklar elde edilir. Fakat bu teknikte her bir düğüm sol ve sağ olmak üzere yalnızca iki alt düğüme bölünerek ilerlemektedir. CART, Twoing, Gini bu teknik için geliştirilmiş algoritmalarıdır.

Destek vektör makinesi (SVM) yöntemi sınıflandırmayı bir doğrusal ya da doğrusal olmayan bir fonksiyon yardımıyla yerine getirir. Destek vektör makinesi yöntemi, veriyi birbirinden ayırmak için en uygun fonksiyonun tahmin edilmesi esasına dayanır [28]. SVM yönteminde sınıfları birbirinden ayıran özel düz bir çizginin bulunması amaçlanır. Sınıflandırma yapılırken bu düz çizginin birden çok çizilme ihtimali vardır. SVM bu çizgilerden her iki sınıfa en uzak olanı bulur ve hataya tolerans en fazla hale gelir. Eğitim verileri ile sınır çizgisi bulunduktan sonra test verileri sınırın hangi tarafında kaldıklarına göre sınıflandırılırlar [41].

İstatistiksel sınıflandırma modelleri koşula bağlı olasılık teorisini temel alarak verileri üzerine tahmin yapmaktadır. Bu modelde en çok kullanılan yöntem Bayes teoremidir. Sonraki bölümlerde bu teorem ve Naive Bayes algoritması detaylı olarak anlatılacaktır.

## Kümeleme

Kümeleme analizi, X veri matrisinde yer alan ve doğal gruplamaları bilinmeyen birimleri, değişkenleri ya da birim ve değişkenleri birbirleri ile benzer olan alt kümelere(grup, sınıf) ayırmaya yardımcı olan yöntemler topluluğudur [42].

Kümeleme analizi birçok alanda kullanılmaktadır. Tıp, biyoloji, psikoloji, sosyoloji, arkeoloji gibi belirsizlik koşullarının ve karmaşık oluşumların bulunduğu bilim alanlarında ise daha yoğun olarak yararlanılan bir yöntemdir [42]. Veri madenciliği uygulamalarında pazarlama faaliyetlerinde sıkça kullanılır. Örneğin bir mamulden farklı beklentilerine göre müşterileri kümelere ayırabilir [28]. Kümeleme yöntemleri denetimsiz öğrenme temeline dayanan tanımlayıcı modellerdir [40].

Kümeleme yöntemleri uygulanırken öncelikle uygun veri matrisi elde edilir. Ardından bu veri matrisi üzerinde uzaklıklar hesaplanır. Bu uzaklıklar hesaplanırken Öklid, Manhattan, Minskowski, Pearson, Korelasyon, Binary Öklid uzaklığı gibi istatistiksel yöntemler kullanılmaktadır. Uzaklık ölçüleri hesaplandıktan sonra verilere kümeleme algoritması uygulanmaktadır. Makine öğrenmesinde kümeleme yöntemleri için en yakın komşu, en uzak komşu ve K Means (Ortalamalar) algoritmaları geliştirilmiştir. En basit ve iyi çalışan en yakın komşu algoritmasıdır [39]. En yakın komşu algoritmasında kümeler arası uzaklık birbirine en yakın olan kümelerin uzaklığı olarak kabul edilirken, en uzak komşu algoritmasında ise uzaklık birbirine en uzak olan kümelerin uzaklığı kabul edilmektedir. En yakın komşu algoritmasında hedef fonksiyon ayrık veya reel değerli olabilir [36]. K Means algoritması yönteminde, başlangıçta belli sayıdaki küme için toplam ortalama hatayı minimize etmek amaçlanır [28].

## Birliktelik Kuralları

Birliktelik kuralı, geçmiş verilerin analiz edilerek bu veriler içindeki birliktelik davranışlarının tespiti ile geleceğe yönelik çalışmalar yapılmasını destekleyen bir yaklaşımdır [43].



Birliktelik kuralları tekniğinde verilerin birbirileri ile olan potansiyel ilişkilerinin ortaya çıkarılması amaçlanmaktadır. Sepet analizlerinde, müşterilerin hangi ürünleri birlikte satın alma eğiliminde olduklarını belirlerken birliktelik kuralları yönteminden faydalanılmaktadır. Bu eğilim belirlenerek firmalar satış ve reklam politikalarını belirleyebilmektedirler. Birliktelik kuralları için Apriori, Sampling algoritmaları geliştirilmiştir.

### 4.3. Bayes Teoremi

Bayes teoremi, İngiliz matematikçi Thomas Bayes tarafından ortaya konulmuştur. Bayes teoremi, Thomas Bayes'in yaşamından sonra geliştirilerek tıp, ekonomi, istatistik, arkeoloji, hukuk, atmosfer bilimi, ölçme değerlendirme, fizik, genetik gibi birçok alanda uygulanmıştır [44]. Bayes teoremi koşullu olasılık temeline dayanmaktadır. Bir B olayının gerçekleşmesi A olayına bağlı ise bu durum koşullu olasılık ile açıklanmaktadır.  $A \cap B \neq \{\}$  olmak üzere B olayının gerçekleşme olasılığı aşağıdaki formüldeki gibi ifade edilmektedir.

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)} \quad (4.4)$$

Burada  $P(B|A)$  sonrasal olasılığı,  $P(B)$  önsel olasılığı,  $P(A|B)$  koşullu olasılığı ve  $P(A)$  ise A olayının bağımsız olasılığını temsil etmektedir. Üyelik sınıfı bilinmeyen bir  $X=\{x_1, \dots, x_n\}$  veri setinde her bir  $x_i$  özelliğinin,  $C_1$  ve  $C_2$  gibi sınıflardan hangisine ait olduğunun olasılığını hesaplayan Bayes teoremi aşağıdaki şekilde ifade edilmektedir;

$$P(C_1 | x_i) = \frac{P(x_i | C_1)P(C_1)}{P(x_i | C_1)P(C_1) + P(x_i | C_2)P(C_2)} \quad (4.5)$$

Burada  $C_1$  ve  $C_2$  olarak gösterilen iki ayrı hipotezin, başka bir deyişle iki ayrı sınıfın olduğu kabul edilmektedir.  $P(C_1 | x_i)$   $x_i$ 'nin  $C_1$  sınıfında olma olasılığını ifade

etmektedir.  $P(x_i)$ ,  $x_i$  değerinin veritabanındaki bulunma sıklığı/sayıdır. Aynı şekilde  $x_i$ 'nin  $C_2$  sınıfında olma olasılığı  $P(C_2 | x_i)$  şöyle hesaplanmaktadır.

$$P(C_2 | x_i) = \frac{P(x_i | C_2)P(C_2)}{P(x_i | C_1)P(C_1) + P(x_i | C_2)P(C_2)} \quad (4.6)$$

Bayes teoremine göre,  $X$  değerinin hangi sınıfa ait olacağına ait karar verilirken aşağıdaki gibi bir hipotez kurulur.

$$P(C | X) = \begin{cases} \frac{P(X | C_1)}{P(X | C_2)} > 1 & C_1 \\ \text{Aksi Halde} & C_2 \end{cases} \quad (4.7)$$

Bu hipoteze göre,  $X$  değerinin  $C_1$  sınıfında olmasına ait hesaplanan ihtimal değeri,  $C_2$  sınıfında olması ihtimalinden daha büyük ise  $X$  değeri  $C_1$  sınıfına aittir, küçük ise  $C_2$  sınıfına aittir şeklinde ifade edilir. Naive Bayes'de Artımlı (Incremental) olarak tabir edilen online bir öğrenme durumu vardır; her bir talim örneği artımlı olarak bir hipotezin doğru olma olasılığını artırır veya azaltır [16]. Bayes teoreminde eğer  $m$  adet sınıf olduğu düşünülürse bu durumda teorem aşağıdaki gibi ifade edilmektedir [35].

$$P(x_i) = \sum_{j=1}^m P(x_i | C_j)P(C_j) \quad (4.8)$$

#### 4.3.1. Bayes sınıflandırma tekniği ve Naive Bayes algoritması

Bayes sınıflandırma tekniği, elde var olan, hali hazırda sınıflanmış verileri kullanarak yeni bir verinin mevcut sınıflardan herhangi birine girme olasılığını hesaplayan bir yöntemdir [35]. Sınıflandırma yöntemleri değişik sonuçlar üreten modeller olarak tanımlanmıştır. Bu yöntemlerin hepsinde test veri örneğinin, eğitim veri setindeki örneklere göre analiz edilerek sınıflandırılması bulunmaktadır [45].

$X=\{x_1, \dots, x_m\}$  gibi m sayıda kritere sahip bir test verisinin C sınıfına ait olma olasılığı aşağıdaki formülle hesaplanmaktadır [36].

$$P(X | C) = \prod_{i=1}^m P(x_i | C) \quad (4.9)$$

Her bir C sınıfı için X test verisinin ihtimal hesabı yapıldıktan sonra, X'in hangi C sınıfına ait olduğu en büyük sonrasal sınıflandırma yöntemi  $C_{MAP}$ (Maximum A Posteriori Classification= MAP) ile bulunur [28].

$$C_{MAP} = \arg \max \prod_{i=1}^m P(x_i | C) \quad (4.10)$$

Temel olarak verilerin olumlu ve olumsuz olarak iki durumu göz önünde bulundurularak çoklu kriterlere göre sınıflandırma yapılmaktadır. Bayes sınıflandırıcılarında, bir örneğin sınıfını belirleyen özellikler bu örneğin sınıfının belirlenmesinde birbirlerinden bağımsız olarak etkili olmaktadır [17]. Sınıflandırma yapılırken sınıflanmaya girecek olan veri test veri seti olarak ele alınır. Sisteme daha önceden girilen veriler ise eğitim veri setidir. Eğitim veri setinde her bir örnek yapılacak olan sınıflandırma sonucuna göre evet / hayır gibi iki duruma sahiptir. Test verisi sahip olduğu özellikler ile eğitim veri setine göre hangi sınıfa gireceği Naive Bayes algoritması yardımı ile belirlenmektedir. M adet kritere sahip bir test verisinin, toplam N adet veriden oluşan bir eğitim veri setine göre, olumlu( $C_{olumlu}$ ) ve olumsuz( $C_{olumsuz}$ ) sınıfa girme ihtimalleri aşağıdaki gibi hesaplanmaktadır.

$$P(\text{TEST} | C_{olumlu}) = (P/N) \times (M_{p1}/P) \times (M_{p2}/P) \times \dots \times (M_{pm}/P) \quad (4.11)$$

$$P(\text{TEST} | C_{olumsuz}) = (Q/N) \times (M_{q1}/Q) \times (M_{q2}/Q) \times \dots \times (M_{qm}/Q) \quad (4.12)$$

Test verisinin hangi sınıfa ait olduğu ise MAP yöntemi ile aşağıdaki gibi bulunur.

$$C_{MAP} = \operatorname{argmax} (C_{olumlu}, C_{olumsuz}) \quad (4.13)$$

Bu formüllerde M kriter sayısını, N toplam eğitim örnek sayısını, P olumlu örnek sayısını, Q olumsuz örnek sayısını,  $M_p$  test kriterine ait değer olumlu eğitim örnekleri ile eşleşen sayısını,  $M_q$  test kriterine ait değer olumsuz eğitim örnekleri ile eşleşen sayısını temsil etmektedir.  $P(\text{TEST} | C_{olumlu})$  test verisinin olumlu sınıfta olma ihtimalini ve  $P(\text{TEST} | C_{olumsuz})$  ise test verisinin olumsuz sınıfta olma ihtimalini belirtmektedir.

Algoritmada, test veri setinin olumlu sınıfa girme ihtimali, test verisinin her bir kriterine ait değerinin eğitim veri setinde olumlu sınıfa giren veri setlerinin değerleri ile aynı olanlarının sayısının, toplam veri seti sayısına bölünerek elde edilen ihtimal hesaplarının çarpımları sonucu belirlenmektedir. Her bir kriterine ait olumlu (p) ihtimal hesabı, test verisinde bulunan kriterine ait değer eğitim verisindeki tüm örnekler içerisinde bu değere eşit ve olumlu sınıfa giren örneklerin sayısının tüm olumlu örnek sayısına bölünmesi şeklinde bulunur. Olumsuz (q) ihtimal hesabı yapılırken de test verisinde bulunan kriterine ait değer eğitim verisindeki örnekler içerisinde bu değere eşit olan ve olumsuz sınıfa giren örneklerin sayısı tüm olumsuz örnek sayısına bölünür. Son olarak test verisinin olumlu sınıfa girme ihtimali; tüm kriterlerin olumlu ihtimallerinin birbirleriyle ve de olumlu örnek sayısının toplam örnek sayısına bölünmesi sonucu ortaya çıkan değer ile çarpılarak bulunur. Olumsuz sınıfa girme ihtimali hesaplanırken de tüm kriterlerin olumsuz ihtimallerin birbirleriyle ve olumsuz örnek sayısının toplam örnek sayısına bölünmesi sonucu ortaya çıkan değer ile çarpılması işlemi yapılır. Olumlu ihtimal ve olumsuz ihtimal hesaplanarak eğer olumlu ihtimal olumsuz ihtimalden daha büyük ise test verisi için olumlu sınıfa gireceği tahmini yapılmaktadır.

Çizelge 4.1.'de 14 gün boyunca izlenen hava, ısı, nem ve rüzgar kriterlerine ait değerlere göre tenis maçının yapıp yapılmadığı gözlenmiştir. Bu verilere dayanarak bir sonraki günde havanın güneşli, sıcaklığın serin, nemin yüksek ve rüzgârın

kuvvetli olma durumunda tenis maçının yapılıp yapılamayacağı aşağıdaki gibi tahmin edilmektedir.

Çizelge 4.1. Eğitim veri seti

Gün	Hava	Isı	Nem	Rüzgar	Oyun
1	Güneşli	Sıcak	Yüksek	Zayıf	Hayır
2	Güneşli	Sıcak	Yüksek	Kuvvetli	Hayır
3	Bulutlu	Sıcak	Yüksek	Zayıf	Evet
4	Yağmurlu	Ilık	Yüksek	Zayıf	Evet
5	Yağmurlu	Serin	Normal	Zayıf	Evet
6	Yağmurlu	Serin	Normal	Kuvvetli	Hayır
7	Bulutlu	Serin	Normal	Kuvvetli	Evet
8	Güneşli	Ilık	Yüksek	Zayıf	Hayır
9	Güneşli	Serin	Normal	Zayıf	Evet
10	Yağmurlu	Ilık	Normal	Zayıf	Evet
11	Güneşli	Ilık	Normal	Kuvvetli	Evet
12	Bulutlu	Ilık	Yüksek	Kuvvetli	Evet
13	Bulutlu	Sıcak	Normal	Zayıf	Evet
14	Yağmurlu	Ilık	Yüksek	Kuvvetli	Hayır

Çizelge 4.2. Test verisine ait kriterlerin frekansları

Sınıf / Kriter	Güneşli	Serin	Yüksek	Kuvvetli	Oyun
Evet	2	3	3	3	9
Hayır	3	1	4	3	5

Bu durumda p ve q ihtimal oranları şu şekilde hesaplanır;

$$P(\text{TEST} | C_{\text{oyunanır}}) = (9/14) \times (2/9) \times (3/9) \times (3/9) \times (3/9) = 0,0053$$

$$P(\text{TEST} | C_{\text{oynamaz}}) = (5/14) \times (3/5) \times (1/5) \times (4/5) \times (3/5) = 0,0206$$

$C_{\text{oyunanır}} = 0,0053 < C_{\text{oynamaz}} = 0,0206$  olduğundan dolayı, bir sonraki günde tenis maçının oynanmayacağı tahmini yapılır.

Test verisinin hangi sınıfa ait olduğu tahmin edilirken, test verisine ait bir kriter değerinin eğitim veri setine ait değerlerden hiçbiri ile eşleşmemesi durumunda, sıfır sorunu ortaya çıkmaktadır. İhtimal hesabı yapılırken her bir test veri kriterinin, eğitim veri setindeki eşleşen kriter değer sayıları ile çarpımları yapıldığından en az bir kriterin 0 çıkması sonucu 0 yapmaya yeterli olmaktadır. Çizelge 4.1.'de yer alan eğitim setinde 6.gün için havanın serin yerine sıcak olduğunu varsayalım. Bu durumda bir sonraki gün için tahmin yapılırken Çizelge 4.2.'de yer alan frekans sayıları aşağıdaki gibi olacaktır.

Çizelge 4.3. Sıfır sorunu için test verisine ait kriterlerin frekansları

Sınıf / Kriter	Güneşli	Serin	Yüksek	Kuvvetli	Oyun
Evet	2	4	3	3	9
Hayır	3	0	4	3	5

Sonraki günde tenis maçının oynanmama olasılığı  $P(\text{TEST} \mid C_{\text{Oynanmaz}})$  hesaplandığında sonuç 0 çıkacaktır. Bu sorun her bir kriter değerine bir u sayısı ilave edilerek çözülebilmektedir. Her bir  $M_{q_1}/Q$  oranına u sayısı ilave edilmesi ile  $\frac{M_{q_1} + up}{Q + u}$  oranı elde edilir. Burada u ilave edilecek sayı, p ise 0 sorununa neden olan kriterin farklı değer sayısıdır. Bu durumda oynanmama olasılığı aşağıdaki şekilde hesaplanmaktadır.

$$\frac{3+1/3}{3+1} \times \frac{0+1/3}{3+1} \times \frac{4+1/3}{3+1} \times \frac{3+1/3}{3+1} \quad (4.14)$$

Bu yöntemde her bir kriterin payına, ilave edilen sayının yani 1'in 0 sorununa neden olan kriterin farklı değer sayısına oranı, paydasına ise 1 sayısı eklenmektedir. 0 sorununa neden olan ısı kriterinin serin, ılık ve sıcak olmak üzere toplam 3 farklı değere sahip olduğundan her bir kriterin payına 1/3 oranı eklenmiştir. Her bir kriter ilave edilerek sıfır sorununun çözüm yöntemi Fransız matematikçi Pierre Laplace'ın ismiyle Laplace dönüşümü olarak ifade edilmektedir [45]. Makine

öğrenmesi tekniklerinde sistem ne kadar çok eğitilir ise o kadar iyi sonuç alınmaktadır. Eğitim verisinin geniş tutulması, kriterlerin sahip olduğu farklı değer sayısını arttırarak, test verilerinin değerlendirilmesinde ortaya çıkacak bu tarz sorunlarında önüne geçecektir.

Naive Bayes’de bir kritere ait veriler sayısal biçimde ise bu durumda o kritere ait hesaplama aşağıdaki standart olasılık yoğunluk fonksiyonu kullanılarak yapılmaktadır [45]. Burada  $\mu$  ortalama,  $\sigma$  standart sapma ve  $x$  ihtimali hesaplanacak olan kriter değeridir.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.15)$$

## **5. MATERİYAL VE METOD**

Bu bölümde, çalışma kapsamında geliştirilen mesleki yönlendirme işleminin yapılabilmesini mümkün kılan sistemin, genel özelliklerinden ve süreçte kullanılan araçlardan bahsedilecektir.

### **5.1. Mesleki Yönlendirmede Kullanılan Kriterlerin ve Algoritmanın Seçimi**

Yapılan çalışmada yönlendirme sürecinde bireylere ait 31 kriter, özellik-faktör yaklaşımında önerilen özellikler temel alınarak genel, veli, akademik başarı ve tercih bilgileri olmak üzere 4 ana başlık altında belirlenmiştir. Uygulamada kullanılan veritabanı bu kriterlere uygun olarak tasarlanmıştır.

Bayes sınıflandırma tekniğinde, test verisinin eğitim verilerine göre sınıfı belirlenirken, seçilen kriterlerin birbirlerinden bağımsız olarak bu sınıflamayı etkilemesi ve konu ile ilgili yapılan çalışmaların incelenmesi sonucunda, sınıflandırma algoritmasının yüksek doğruluk ile kesin ve hızlı sonuçlar üretmesi, çevrimiçi olarak sürekli eğitilebilir olması nedeniyle, çalışmada Naive Bayes algoritması tercih edilmiştir.

### **5.2. Verilerin Elde Edilme Süreci**

Uygulamada kullanılacak olan veriler, Türkiye'deki farklı mesleki türde eğitim veren okullarda bulunan, mezun olmuş ve öğrenime devam eden meslek lisesi öğrencilerinden hazırlanan anket formunun okullarda ve internet üzerinden uygulanması yöntemi ile toplanmıştır. İnternet üzerinde yapılan anket formu uygulaması için PHP programlama dili ve bu dille uyumlu olan MYSQL veritabanı kullanılmıştır. MYSQL veritabanında toplanan veriler, uygulamada kullanılan MSSQL 2005 veritabanına oluşturulan SQL transfer sorgularıyla aktarılmıştır. Veriler elde edildikten sonra kayıp ve gürültülü veriler üzerinde veri madenciliği teknikleri kullanılarak düzeltimler yapılmıştır.



### **5.3. Geliştirilen Mesleki Yönlendirme Yazılımı**

Bu bölümde tez çalışması kapsamında geliştirilen mesleki yönlendirme yazılımı için kullanılan araçlar, veritabanı tasarımı ve program ara yüzleri anlatılacaktır.

#### **5.3.1. Yazılımın geliştirilmesinde kullanılan araçlar**

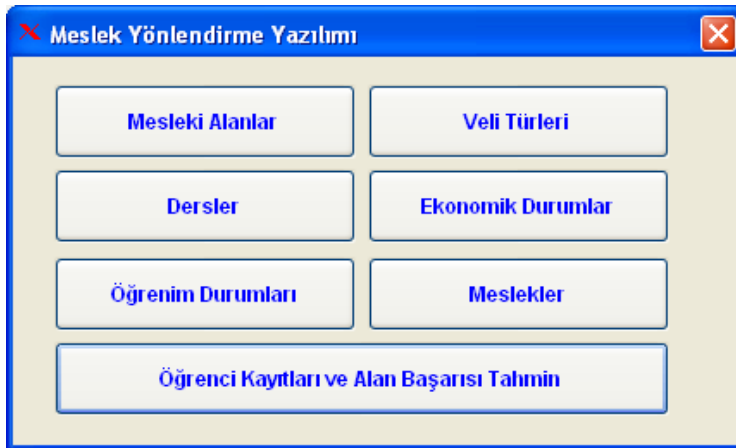
Yazılım, Microsoft Visual Studio 2008 uygulama geliştirme platformu üzerinde nesne yönelimli bir programlama dili olan C#.Net kullanılarak geliştirilmiştir. Veritabanı tasarımı için, ilişkisel veritabanı modeline uygun olan MSSQL 2005 veritabanı motoru kullanılmıştır. Veritabanı tasarımı yapılırken Microsoft SQL Server Management Studio ara yüz programından yararlanılmıştır.

#### **5.3.2. Veritabanı tasarımı**

Veriler genel, veli, mesleki tercih, kültür dersleri ve mesleki alan dersleri olmak üzere 5 ana tabloda toplanmıştır. Bu tablolara kaynaklık etmek ve ilişkide bulunmak üzere bölüm, cinsiyet, ders, ekonomik durum, meslek, mezuniyet durum ve veli tür tabloları da tasarlanmıştır. Tablolar arası ilişkiler ile algoritmanın uygulanmasında kullanılan eğitim ve test view nesnelere SQL sorgulama dili kullanılarak oluşturulmuştur.

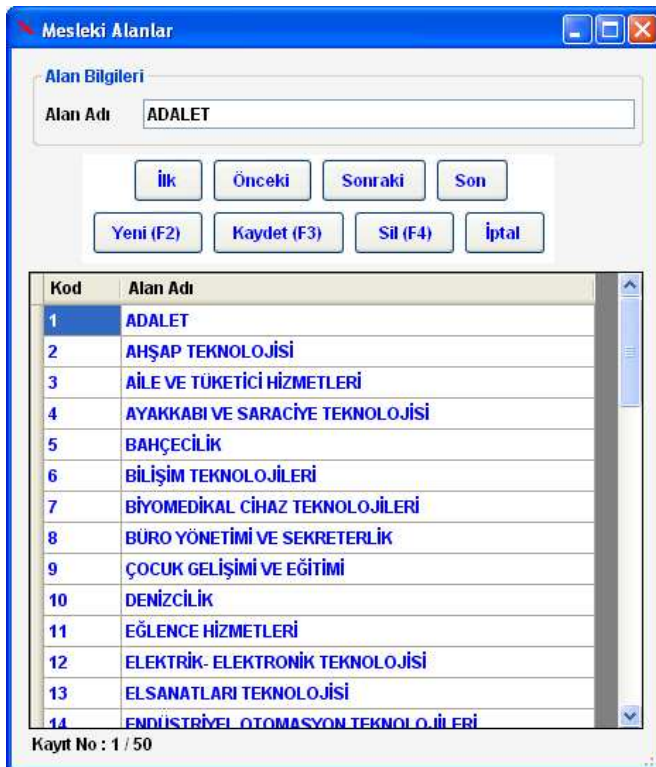
#### **5.3.3. Yazılım ara yüz tanıtımı**

Yazılımda, veritabanında yer alan ana ve alt tablolara ait verilerin düzenlendiği formlar, eğitim ve test verilerinin girildiği form ve başarı tahmininin yaptırıldığı form ile bu formların hepsine ulaşım sağlanan bir ana form bulunmaktadır. Yazılım ilk çalıştırıldığında kullanıcıların karşısına Şekil 5.1.'de yer alan ana form gelmektedir.



Şekil 5.1. Ana form

Mesleki alanlar, veli türleri, dersler, ekonomik durumlar, öğrenim durumları ve meslekler formları ile algoritmada kullanılan kriterlere kaynak olan tablolardaki veriler düzenlenmektedir.



Şekil 5.2. Mesleki alan düzenleme formu

**Veli Türleri**

Veli Tür Bilgileri

Tür Adı: Anne

İlk Önceki Sonraki Son

Yeni (F2) Kaydet (F3) Sil (F4) İptal

Kod	Veli Türü
1	Anne
2	Baba
3	Kardeş
4	Diğer

Kayıt No : 1 / 4

Şekil 5.3. Veli türleri düzenleme formu

**Dersler**

Ders Bilgileri

Ders Adı: BEDEN EĞİTİMİ

İlk Önceki Sonraki Son

Yeni (F2) Kaydet (F3) Sil (F4) İptal

Kod	Ders Adı
1	BEDEN EĞİTİMİ
2	BİLGİ VE İLETİŞİM TEKNOLOJİSİ
3	BİYOLOJİ
4	COĞRAFYA
5	DİL VE ANLATIM
6	DİN KÜLTÜRÜ VE AHLAK BİLGİSİ
7	FİZİK
8	KİMYA
9	MATEMATİK
10	SAĞLIK BİLGİSİ
11	TARİH
12	TÜRK EDEBİYATI
13	YABANCI DİL
14	SEÇMELİ DERS

Kayıt No : 1 / 14

Şekil 5.4. Ders kayıtları düzenleme formu

**Ekonomik Durumlar**

Ekonomik Durum Bilgileri

Durum Adı Çok Kötü

İlk Önceki Sonraki Son

Yeni (F2) Kaydet (F3) Sil (F4) İptal

Kod	Ekonomik Durum
1	Çok Kötü
2	Düşük
3	Orta
4	İyi
5	Çok İyi

Kayıt No : 1 / 5

Şekil 5.5. Ekonomik durum kayıtları düzenleme formu

**Öğrenim Durumları**

Öğrenim Durum Bilgileri

Öğrenim Durumu Okuma Yazma Bilmiyor

İlk Önceki Sonraki Son

Yeni (F2) Kaydet (F3) Sil (F4) İptal

Kod	Öğrenim Durumu
1	Okuma Yazma Bilmiyor
2	İlk Okul
3	Orta Okul
4	Lise
5	Meslek Yüksek Okulu(2 Yıllık)
6	Üniversite
7	Yüksek Lisans
8	Doktora

Kayıt No : 1 / 8

Şekil 5.6. Öğrenim durum kayıtları düzenleme formu

**Meslekler**

Meslek Bilgileri

Meslek Adı

İlk Önceki Sonraki Son

Yeni (F2) Kaydet (F3) Sil (F4) İptal

Kod	Meslek
1	Esnaf
2	İşçi
3	Memur
4	Öğretmen
5	Askeri Personel
6	Yargı Mensubu
7	Doktor
8	Polis
9	Teknik Personel
10	Zannıkar
11	Ev Hanımı
12	Diğer
14	Emekli

Kayıt No : 1 / 13

Şekil 5.7. Meslek kayıtları düzenleme formu

Öğrenci kayıtları ve alan başarısı tahmin formunda ise öğrenci bilgileri, genel, veli, alan tercihleri, kültür ve meslek ders bilgileri olmak üzere 5 alt bölümde düzenlenmektedir.

**Öğrenci Kayıtları**

Genel Bilgiler Veli Bilgileri Alan Tercihleri Kültür Ders Bilgileri Mesleki Ders Bilgileri

Öğrenci Ara

Eğitim Veri Setini Göster Test Veri Setini Göster Öğrenci Kodu

**Kişi Bilgileri**

Cinsiyeti Kız Boyu 65 Kilosu 60

Doğum Yılı 1992 Kardeş Sayısı 6

Sınıfı 12 Ekonomik Durumu Orta

Bölümü BİLİŞİM TEKNOLOJİLERİ  Kendisine Ait Çalışma Odası Var  Bir İşte Çalışıyor

Diploma Notu 4,07 OKS Puanı 0  Özel Eğitim Kapsamında Kaynaştırma Öğrencisi

İlk Önceki Sonraki Son Yeni (F2) Kaydet (F3) Sil (F4) İptal Çıkış

OGRENCI_KOD	CINSIYET	DOGUM_YIL	SINIF	BOLUM	DIPLOMA_NOT	OKS
1	Kız	1992	12	BİLİŞİM TE...	4,07	0
2	Erkek	1991	12	BİLİŞİM TE...	4,15	193
4	Kız	1992	12	BİLİŞİM TE...	4,45	0
6	Kız	1992	12	BİLİŞİM TE...	0,00	0
7	Kız	1992	12	BİLİŞİM TE...	0,00	0

Kayıt Sayısı: 104

BİLİŞİM TEKNOLOJİLERİ Bölümü İçin BAŞARI TAHMİNİ HESAPLA

Tahmin Kriterlerini Seç

Şekil 5.8. Öğrenci kayıtları ve alan başarısı tahmin formu

Bu formda öğrenci bilgilerinin girişi ve başarı tahmininin yapılması işlemleri sonraki bölümde anlatılacaktır.

## 6. MESLEKİ ALAN SEÇİMLERİNİN MAKİNE ÖĞRENMESİ ALGORİTMASI KULLANILARAK BELİRLENMESİ

Bu bölümde, geliştirilen yazılımda mesleki yönlendirme yaptırılırken algoritmanın uygulanma aşamaları ve bu süreçte yazılımın nasıl kullanıldığından bahsedilecektir.

### 6.1. Mesleki Yönlendirme İçin Naive Bayes Algoritmasının Uygulanması

Mesleki alan başarısı tahmini yaptırılırken, eğitim ve test veri olarak iki adet view nesnesi kullanılmaktadır. Bu view nesnelere ile veritabanında yer alan tablolarda bulunan kriterler tek bir tabloda birleştirilmiştir. TEST\_VERI ve EGITIM\_VERI isimli bu view nesnelere ile Naive Bayes algoritmasının uygulanacağı test ve eğitim veri setleri oluşturulmaktadır. Bu view nesnelere sayesinde mevcut verilerin bozulmadan veri madenciliği kurallarına göre uygun bir şekilde yapılandırılması sağlanarak ve farklı alt tablolarda bulunan öğrenci özellikleri tek bir tablo altında birleştirilmiştir. Ayrıca birleştirme işlemi yapılırken, algoritmada yapılacak hesaplamalara uygun olacak şekilde veriler standart hale getirilmektedir. Sisteme yeni eklenen öğrenci, özellikleri ile mevcut veriler ile karşılaştırılarak bu alanda başarılı olma durumu tahmini yaptırılmaktadır. Bir veri setinin eğitim verisi mi test verisi mi olduğu ise mesleki alan başarı ortalamasının bulunması ile kontrol edilmektedir. TEST\_VERI view nesnesi oluşturulurken de bu kriter göz önüne alınmıştır.

Tahmin yaptırılırken ilk olarak başarı tahmini yapılmak istenen bölüme göre veriler, eğitim verileri ve tahmin yapılmak istenen öğrencinin kod değerine göre oluşturulan test veri setleri DataRow nesnelere aktarılmaktadır. İkinci aşamada ise seçilen mesleki alanda bulunan örneklerin başarılı ve başarısız olanların toplam sayıları  $p$  ve  $q$  olarak iki değişkende hesaplatılmaktadır. Üçüncü aşamada her bir kriterin başarılı ve başarısız ihtimal durumu için frekans sayıları hesaplanmaktadır. Dördüncü aşamada, başarılı ve başarısız ihtimal hesaplamasında 0 sorununa yol açabilecek kriterler belirlenmektedir. Eğer seçilen test verisine ait bir kritere ait değer, eğitim veri setinde yer alan kritere ait değerlerden hiç biri ile eşleşmiyor ise bu durumda o

kriter için Laplace dönüşümünde kullanılmak üzere farklı özellik değerleri sayısı belirlenmektedir. Bu belirlenen değerler hesaplamada her bir kriterin pay ve paydasına ilave edilmektedir. Son aşamada ise, başarılı ve başarısız olma ihtimalleri kriterlerin elde edilen frekansları kullanılarak hesaplanmaktadır. Sonuç olarak, eğer başarılı olma oranı başarısız olma oranından daha büyük ise tahmin başarılı, küçük ise tahmin başarısız olarak kullanıcıya yansıtılmaktadır.

## 6.2. Mesleki Yönlendirme Yazılımının Kullanılması

Yazılımda, mesleki yönlendirme işlemleri için Şekil 6.1’de yer alan öğrenci kayıtları ve alan başarısı tahmin formu kullanılmaktadır. Bu formda, öğrencilerin kayıtları düzenlenmekte ve başarı tahminleri yapılmaktadır. Formda öğrencilere ait bilgiler genel, veli, alan tercihleri, kültür dersi bilgileri ve meslek dersi bilgileri olmak üzere 5 bölümde tutulmaktadır. Sisteme yeni bir öğrenci eklenirken genel bilgiler sekmesi altında bulunan yeni kayıt butonuna basılarak öğrenci ile ilgili genel bilgilerin girilmesi sağlanmaktadır. Şekil 6.1.’deki gibi öğrenciye ait genel bilgiler girildikten sonra kaydet butonuna basılarak öğrencinin sisteme eklenmesi sağlanmaktadır.

OGRENCI_KOD	CINSİYET	DOGUM_YIL	SINIF	BOLUM	DIPLOMA_NOT	OKS
1	Kız	1992	12	BİLİŞİM TE...	4,07	0
2	Erkek	1991	12	BİLİŞİM TE...	4,15	193
4	Kız	1992	12	BİLİŞİM TE...	4,45	0
6	Kız	1992	12	BİLİŞİM TE...	0,00	0
7	Kız	1992	12	BİLİŞİM TE...	0,00	0

Şekil 6.1. Öğrenci genel bilgileri sekmesi



Şekil 6.1.'deki gibi öğrenciye ait cinsiyet, doğum yılı, sınıf, boy, kilo, başarı notu, ekonomik durum vb. genel bilgiler girildikten sonra kaydet butonuna basılarak öğrencinin sisteme eklenmesi sağlanmaktadır. Öğrenci sisteme eklendikten sonra, anne, baba ve velisine ait bilgilerin girişi için Şekil 6.2.'de yer alan veli bilgileri sekmesi kullanılmaktadır. Bu sekmede anne, baba ve veliye ait meslek, öğrenim durumu, birlikte yaşama durumları ve hayatta olup olmadıklarına ait bilgiler kaydedilmektedir.

Şekil 6.2. Öğrenci veli bilgileri sekmesi

Şekil 6.3.'de gösterilen alan tercihleri sekmesinde, öğrencinin öğrenim görmek istediği mesleki alanlar için yaptığı tercihlerin kaydı yapılmaktadır.

**Öğrenci Kayıtları**

Genel Bilgiler Veli Bilgileri **Alan Tercihleri** Kültür Ders Bilgileri Mesleki Ders Bilgileri

Kod: 2

Tercih Bilgileri

Tercih Edilen Bölüm: BİLİŞİM TEKNOLOJİLERİ

Tercih Sırası: 1

Yeni Kaydet Sil İptal

TERCIH_KOD	OGRENCI_KOD	TERCIH_SIRA	BOLUM_KOD
6	2	1	BİLİŞİM TEKNOLOJİLERİ
7	2	2	TEKSTİL TEKNOLOJİSİ
8	2	3	EL SANATLARI TEKNOLOJİSİ
9	2	4	ÇOCUK GELİŞİMİ VE EĞİTİMİ

Şekil 6.3. Öğrenci alan tercihleri sekmesi

**Öğrenci Kayıtları**

Genel Bilgiler Veli Bilgileri Alan Tercihleri **Kültür Ders Bilgileri** Mesleki Ders Bilgileri

Kod: 2

Kültür Ders Bilgileri

Ders Adı: BEDEN EĞİTİMİ Notu: 5

İlk Önceki Sonraki Son Kaydet İptal

Kod	Ders	Puan
15	BEDEN EĞİTİMİ	5
16	BİLGİ VE İLETİŞİM TEKNOLOJİSİ	4
17	BİYOLOJİ	2
18	COĞRAFYA	3
19	DİL VE ANLATIM	3
20	DİN KÜLTÜRÜ VE AHLAK BİLGİSİ	4
21	FİZİK	2
22	KİMYA	3
23	MATEMATİK	2
24	SAĞLIK BİLGİSİ	3
25	TARİH	4
26	TÜRK EDEBİYATI	4
27	YABANCI DİL	5
28	SEÇMELİ DERS	4

Şekil 6.4. Öğrenci kültür ders bilgileri sekmesi

Şekil 6.4.'de yer alan kültür dersi bilgileri sekmesinde, öğrencinin mesleki alanda eğitim görmeden önce görmüş olduğu temel derslere ait not bilgilerinin girişi yapılmaktadır.

The screenshot shows the 'Öğrenci Kayıtları' application window. The 'Mesleki Ders Bilgileri' tab is active. The form shows 'Ders Adı' as 'temel elektronik', 'Saati' as 3, and 'Notu' as 3. Below the form are buttons for 'İlk', 'Önceki', 'Sonraki', 'Son', 'Yeni', 'Kaydet', 'Sil', and 'İptal'. A table below the form lists course records with columns 'Kod', 'Ders Adı', 'Saati', and 'Puanı'.

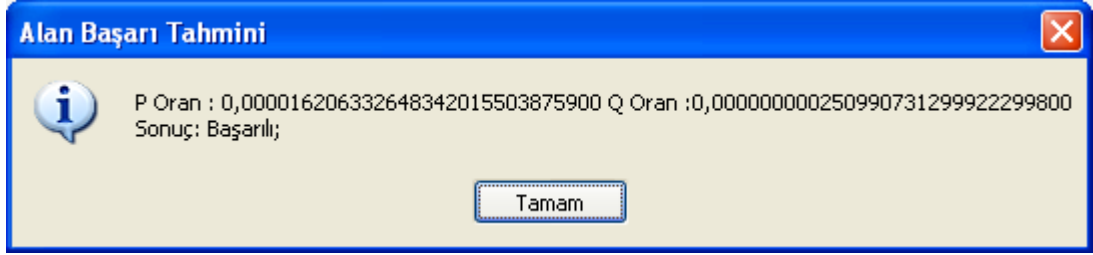
Kod	Ders Adı	Saati	Puanı
21	temel elektronik	3	3
22	grafik animasyon	4	5
23	veri tabanı	8	2
24	sistem bakım onarım	3	3
25	mesleki gelişim	1	5
26	web tasarımı	9	3
27	paket programlama	4	3
28	internet	4	3
29	teknik resim	2	4

Not Ortalaması: 3,11

Şekil 6.5. Öğrenci mesleki ders bilgileri sekmesi

Şekil 6.5.'de gösterilen sekmede öğrencinin seçmiş olduğu mesleki alanda almış olduğu derslere ait saat ve not bilgileri girilmektedir. Burada girilen bilgiler ile öğrencinin alanındaki başarı ortalaması hesaplanmaktadır. Mesleki ders bilgileri yalnızca mesleki eğitim sürecinde olan ve mezun olmuş öğrenciler için doldurulmaktadır. Mesleki alan seçimi sürecinde bulunan öğrenciler için bu sekmede bilgi girişi yapılmamaktadır. Sistemdeki öğrencilerin mesleki başarı ortalaması bulunuyor ise eğitim veri setine dahil olmaktadır.

Öğrenciye ait tüm bilgiler girildikten sonra Şekil 6.1.'de gösterilen öğrenci genel bilgileri sekmesinde, mesleki alan başarı tahmini yapılacak öğrenci kaydı ve başarı tahmini yapılacak olan mesleki alan seçildikten sonra başarı tahmini hesaplanarak, Şekil 6.6.'da gösterilen mesaj formu ile kullanıcıya sonuçlar gösterilmektedir.



Şekil 6.6. Alan başarı tahmin sonuç mesajı

Yazılımda, tahminde kullanılan kriterlerin seçimi kullanıcılar tarafından yapılabilmektedir. Başarı tahmini yapılacak öğrenci ve bölüm seçildikten sonra kullanıcılar tahmin kriterlerini seç onay kutusunu işaretler ise bu durumda Şekil 6.7.'de yer alan form açılarak kullanıcıların, tahmin hesaplamasında kullanılacak olan kriterleri seçmeleri sağlanmaktadır. Bu durumda oluşturulan eğitim ve test veri setleri seçilen kriterlere göre oluşturularak başarı tahmini yaptırılmaktadır.



Şekil 6.7. Başarı tahmini kriter seçim formu

## 7. SONUÇ VE ÖNERİLER

Meslek seçimi insanların hayatında verdikleri en önemli kararlardan bir tanesidir. Kişiler yaptıkları meslek seçimi ile hayatlarının en uzun evresi olan çalışma döneminin nasıl olacağına karar vermektedirler. Bireyin kendisine uygun olmayan, başarılı olamayacağı bir mesleği seçmesi hayatı boyunca mutsuz olmasına neden olmaktadır. Her geçen gün artan nüfus ve güçleşen ekonomik koşullar bireylerin iş bulmalarını giderek güçleştirmektedir. Günümüzde artık iyi bir okuldan mezun olmak iş sahibi olmak için yeterli olmamakta, bireylerin meslekleri ile ilgili edindikleri tecrübeler işverenler için belirleyici bir unsur olmaktadır. Bireyler meslek yaşamlarına ne kadar erken atılırlarsa, alanlarında sahip olacakları tecrübe o oranda artacaktır. Bunun içinde bireylerin mesleklerine erken yaşlarda karar vermeleri gerekmektedir.

Ülkemizde ve dünyada mesleki ve teknik eğitime verilen önem giderek artmaktadır. Bireylerin meslek hayatlarına erken atılmaları için lise düzeyinde bu eğitime başlamaları önemli bir fırsattır. Bu seviyede, mesleklerini seçmek durumunda olan bireylerin, doğru seçim yapmaları için mutlaka rehberlik hizmetlerinden faydalanmaları gerekmektedir. Mevcut mesleki yönlendirme faaliyetlerinde, bireylerin seçecekleri mesleki alanda başarılı olup olamayacakları önceden ön görülememektedir. Bireyler ilgi ve isteklerine uygun olan mesleki alana yöneltilseler de, eğitim sürecinde başarısız sonuçlara rastlanmaktadır. Eğitim sürecinde başarısız olan bireyler, telafisi çok da mümkün olmayan bir zaman kaybı yaşamakta ve de hayal ettiği mesleki alandan da uzaklaşmaktadırlar.

Bu çalışmada mevcut mesleki yönlendirme faaliyetlerinde dikkate alınmayan unsurların tercih sürecine dahil edilerek, bireylerin yalnızca istedikleri değil, başarılı olabilecekleri mesleği de seçmesi amaçlanmıştır. Bireylerin, seçmek istedikleri mesleki alanda daha önceden eğitim görmüş bireylerin gösterdikleri başarı durumuna göre, o alanda ne kadar başarılı olabileceği tahmin edilmeye çalışılarak mesleki yönlendirme sürecine katkı yapılmaya çalışılmıştır.

Klasik yönlendirme faaliyetlerinin meydana getirdiği insan hatası, bilgiyi elde etme süreci, maliyet ve zaman kaybı sorunlarına da çözüm getirmek adına bir yazılım geliştirilmiştir. Geliştirilen yazılımın, mesleki alan seçiminde karar verme sürecinde olan öğrenci ve mesleki rehberlik hizmetini yürüten kişilere yol gösterici olması hedeflenmiştir.

Yazılım geliştirilirken geçmiş tecrübelerle dayanarak tahmin modellemesi yapan ve bir yapay zeka alanı olan makine öğrenmesi teknikleri incelenmiştir. Yapılan incelemeler sonucunda problemin çözümü için, yüksek doğrulukta sonuçlar üretmesi, seçilen kriterlerin birbirlerinden bağımsız olarak sınıflandırmayı etkilemesi ve de çevrimiçi sürekli eğitilebilir olmasından dolayı Bayes sınıflandırma tekniği kullanılmıştır.

Tercih sürecinde olan bireylere ait 4 ana başlıkta 31 adet kriter, özellik faktör yaklaşımına göre seçilmiştir. Seçilen kriterler doğrultusunda veriler, Türkiye'nin farklı mesleki eğitim veren kurumlarından mezun olmuş ve halen öğrenim gören öğrencilerden toplanmıştır. Veriler toplandıktan sonra veri madenciliği yöntemleri uygulanarak toplanan verilerin algoritmanın uygulanması için hazır hale getirilmiştir.

Sistemdeki veriler eğitim ve test verisi olmak üzere temelde iki grupta değerlendirilmiştir. Yazılım, 5 farklı mesleki alanda toplam 108 öğrenci üzerinde uygulanmıştır. Bu öğrencilerden mesleki eğitim sürecinde olan 93 tanesi eğitim veri seti olarak kullanılmıştır. Test veri setinde yer alan 15 öğrenci için farklı alanlarda başarı tahmini yaptırılmıştır. Yapılan tahmin sonucunda öğrencilerin daha önceden yapmış oldukları mevcut mesleki alan seçimleri ve sistemin bu tercihler için yaptığı başarı tahmini sonuçları Çizelge 7.1'de verilmiştir.

Çizelge 7.1. Test verileri için yapılan başarı tahmini sonuçları

No	Seçilen Mevcut Mesleki Alan	Sistem Tarafından Yapılan Başarı Tahmini
1	Bilişim Teknolojileri	Başarılı
2	Bilişim Teknolojileri	Başarılı
3	Bilişim Teknolojileri	Başarısız
4	Muhasebe ve Finansman	Başarılı
5	Muhasebe ve Finansman	Başarısız
6	Bilişim Teknolojileri	Başarısız
7	Pazarlama ve Perakende	Başarısız
8	Pazarlama ve Perakende	Başarısız
9	Pazarlama ve Perakende	Başarılı
10	Pazarlama ve Perakende	Başarılı
11	Bilişim Teknolojileri	Başarılı
12	Bilişim Teknolojileri	Başarılı
13	Eİ Sanatları Teknolojisi	Başarılı
14	Eİ Sanatları Teknolojisi	Başarılı
15	Çocuk Gelişimi ve Eğitimi	Başarılı

Sistemin ne kadar çok eğitilirse o kadar iyi sonuçlar verdiği görülmüştür. Yazılımda tahmin yaptırılırken, kullanıcılara istedikleri kriterleri seçme imkanı da verilmiştir. Yazılımın işlevsel olması ve kolay kullanılabilirliği mesleki yönlendirme faaliyetlerini, klasik yöntemlere kıyasla kolaylaştırdığı görülmüştür.

Öğrenciler için belirlenen kriterlerin, ülkemizdeki tüm ilk ve orta öğretim öğrencilerinin bilgilerinin bulunduğu e-okul sistemi ile de uyumlu olması, ilerleyen zamanlarda sistemin tüm ülke çapında kullanılmasına da olanak vermektedir. Yazılımın daha geniş kitlelere ulaşması için, gelecekte web ortamına aktarılması ileride yapılacak çalışmalar arasında planlanmaktadır.

Yapılan bu çalışmanın, mesleki eğitim sürecine girecek olan tüm bireylere ve bu süreçte yönlendirme faaliyetinde bulunan öğretmenlere, karar verme sürecinde önemli bir yarar sağlayacağı düşünülmektedir.

## KAYNAKLAR

1. Üre, Ö., “Mesleki rehberlik”, Psikolojik Danışma ve Rehberlik, Deniz, M. Z., Erözkan, A., *Tekağaç Eylül Yayıncılık*, Ankara, 183-203 (2006).
2. Çakır, M. A., “Mesleki rehberlik”, Psikolojik Danışma ve Rehberlik, Kaya, A., *Anı Yayıncılık*, Ankara, 137-167 (2007).
3. Sarıkaya, T., Khorshid, L., "Üniversite öğrencilerinin meslek seçimini etkileyen etmenlerin incelenmesi: üniversite öğrencilerinin meslek seçimi", *Türk Eğitim Bilimleri Dergisi*, 7(2): 393-423 (2009).
4. Telman, N. ve Ünsal, P., “Çalışan Memnuniyeti”, *Epsilon Yayıncılık*, İstanbul (2004).
5. İnternet: KobiFinans “İşinden Memnun Olmayan Çalışan, Krizin Bitmesini Bekliyor” [http://www.kobifinans.com.tr/tr/bilgi\\_merkezi/020706/23810](http://www.kobifinans.com.tr/tr/bilgi_merkezi/020706/23810) (2010).
6. Dener, M., Dörterler, M., Orman, A., “Açık kaynak kodlu veri madenciliği programları: WEKA’da örnek uygulama”, *Akademik Bilişim’09 - XI. Akademik Bilişim Konferansı*, Şanlıurfa, 787-796 (2009).
7. Yazıcı, H., “Eğitsel ve mesleki rehberlik”, Psikolojik Danışma ve Rehberlik, Can, G., *Pegam A Yayıncılık*, Ankara, 77-116 (2007).
8. Yıldız, T.T., “Web destekli çoklu zeka kuramına dayalı mesleki yönlendirme”, Yüksek Lisans Tezi, *Gazi Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 46-48, 7 (2005).
9. Hasebrook, J.P. and Nathusius, W., “An expert advisor for vocational guidance”, *Journal of Artificial Intelligence in Education*, 8(1): 21-41 (1997).
10. Ari, E. ve Vatansaver, F., “Bulanık mantık tabanlı mesleki yönlendirme”, 5. *Uluslar arası İleri Teknolojiler Sempozyumu (İATS’09)*, Karabük, 788-791 (2009).
11. Uzun, Y., “Tıbbi veriler üzerinde makine öğrenme algoritmalar ve bulanık mantık ile kurallar öğrenme”, Yüksek Lisans Tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya, 1-3,12 (2005).
12. Başkeleş, B., “Yazılım eforunun makine öğrenme yöntemleri kullanarak tahmini”, Yüksek Lisans Tezi, *Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, 1-73 (2006).



13. Amasyalı, M.F., “Yeni makine öğrenmesi metotları ve ilaç tasarımına uygulamaları”, Doktora Tezi, *Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, 204-213,24 (2008).
14. Gökbay, Z.İ., “Meme kanserinde otomatik öğrenme teknikleri”, Yüksek Lisans Tezi, *Bahçeşehir Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, 1-34 (2007).
15. Bahadır, İ., “Bayes teoremi ve yapay sinir ağları modelleriyle borsa gelecek değer tahmini uygulaması”, Yüksek Lisans Tezi, *TOBB Ekonomi ve Teknoloji Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 1-78 (2008).
16. Altunyaprak, C. “Bayes yöntemi kullanarak istenmeyen elektronik postaların filtrelenmesi”, Yüksek Lisans Tezi, *Muğla Üniversitesi Fen Bilimleri Enstitüsü, Muğla*, 1-54,29 (2006).
17. Kahraman, H.T., “Web-tabanlı uyarlanır zeki öğretim sistemi tasarımı ve uygulaması”, Doktora Tezi, *Gazi Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 131-136,7 (2009).
18. Yeşilyaprak, B., Güngör, A., Kurç, G., “Eğitsel ve Mesleki Rehberlik”, *Varan Matbaacılık*, Ankara (1998).
19. Vurucu, F.,”Meslek lisesi öğrencilerinin meslek seçimi yeterliliği ve meslek seçimini etkileyen faktörler”, Yüksek Lisans Tezi, *Yeditepe Üniversitesi Sosyal Bilimleri Enstitüsü*, İstanbul, 33 (2010).
20. Güvenç, M.,”Okullardaki rehberlik faaliyetlerinin yürütülmesinde karşılaşılan güçlükler”, Yüksek Lisans Tezi, *Niğde Üniversitesi Sosyal Bilimleri Enstitüsü*, Niğde, 10 (2001).
21. Milli Eğitim Bakanlığı, “Milli Eğitim Bakanlığı Rehberlik Ve Psikolojik Danışma Hizmetleri Yönetmeliği”, *Tebliğler Dergisi*, (2001).
22. Osipow, H.S., Fitzgerald, F.L., “Theories of Career Development(4th.ed.)”, *MA:Ally and Bacon*, Boston (1996).
23. Kuzgun, Y., “Meslek Danışmanlığı Kuramlar Uygulamalar”, *Nobel Yayın Dağıtım*, Ankara (2000).
24. Özgüven, İ.E., “Çağdaş Eğitimde Psikolojik Danışma ve Rehberlik”, *Pdrem Yayınları*, Ankara (1999).
25. Kepçeoğlu, M., “Psikolojik Danışma ve Rehberlik”, *Özerler Matbaası*, Ankara (1996).
26. Özkes, S., “Veri madenciliği modelleri ve uygulama alanları”, *İstanbul Ticaret Üniversitesi Dergisi*, 3: 65-82 (2003).

27. Dunham, M.H., "Data Mining Introductory and Advanced Topics", *Pearson Education*, New Jersey (2002).
28. Özkan, Y., "Veri Madenciliği Yöntemleri", *Papatya Yayıncılık Eğitim*, İstanbul (2008).
29. Weiss, S.M. and Indurkha, N., "Predictive Data Mining: A Practical Guide", *Morgan Kaufmann Publications*, USA (1998).
30. Akpınar, H., "Veri tabanlarında bilgi keşfi ve veri madenciliği", *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 29(1): 1-22 (2000).
31. Bozkır, A.S., "OLAP ve veri madenciliği teknolojilerinden yararlanılarak web tabanlı bir karar destek sisteminin geliştirilmesi", Yüksek Lisans Tezi, *Hacettepe Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 20,25,28 (2009).
32. Akın, Y.K., "Veri madenciliğinde kümeleme algoritmaları ve kümeleme analizi", Doktora Tezi, *Marmara Üniversitesi Sosyal Bilimler Enstitüsü*, İstanbul, 35-37,59 (2008).
33. Kocabaş, F.M., "Veri madenciliği süreci ve gerçek bir veri seti üzerinde uygulanması", Yüksek Lisans Tezi, *Hacettepe Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 15 (2010).
34. Han, J. and Kamber, M., "Data Mining: Concept and Techniques", *Morgan Kaufmann Publications*, USA (2001).
35. Silahtaroglu, G., "Kavram ve Algoritmalarıyla Temel Veri Madenciliği", *Papatya Yayıncılık Eğitim*, İstanbul (2008).
36. Mitchell, M.T., "Machine Learning", *McGraw-Hill*, USA (1997).
37. Öztemel, E., "Yapay Sinir Ağları", *Papatya Yayıncılık Eğitim*, İstanbul (2003).
38. Alpaydın, E., "Introduction to Machine Learning", *The MIT Press*, London (2004).
39. Dalyan, T., "Makine öğrenmesinde 1R algoritması ve ikinci kuralın (2R) oluşturulması", Yüksek Lisans Tezi, *Kocaeli Üniversitesi Fen Bilimleri Enstitüsü*, Kocaeli, 14,17 (2006).
40. Albayrak, A.S., Yılmaz, K.Ş., "Veri madenciliği: karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama", Süleyman *Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(1):31-52 (2009).

41. İnternet: Yıldız Teknik Üniversitesi “Makine Öğrenmesine Giriş (M.Fatih Amasyalı Ders Notları)” <http://www.ce.yildiz.edu.tr/mygetfile.php?id=868> (2010).
42. Özdamar, K., “Paket Programlar İle İstatistiksel Veri Analizi -2 (Çok Değişkenli Analizler)”, *Kaan Kitapevi*, Eskişehir(2004).
43. Özçakır, C.Ö., Çamurcu, A.Y., “Birliktelik kuralı yöntemi için bir veri madenciliği yazılımı tasarımı ve uygulaması”, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 6(12):21-37 (2007).
44. Murat, N., “Model seçiminde bayesci yaklaşımların kullanımı”, Yüksek Lisans Tezi, *Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü*, Samsun, 2-3 (2007).
45. Witten, I.H, Frank, E., “Practical Machine Learning Tools and Techniques Second Edition”, *Morgan Kaufmann Publications*, USA (2005).
46. Bülbül, H.İ., Ünsal, Ö. “Determination of vocational fields with machine learning algorithm”, *The Ninth International Conference on Machine Learning and Applications (ICMLA 2010)*, *IEEE Computer Society*, Washington D.C., 710-713 (2010).

**EKLER**

EK-1. İnternet üzerinde yapılan anket formu

## MESLEK LİSELERİNDE BÖLÜM SEÇİM BAŞARILARININ TESPİT EDİLMESİ

Sayın Katılımcı, bu araştırma, yüksek lisans tez çalışması için yürütülmektedir. Araştırmadan elde edilen veriler, araştırmanın amacı olan bilimsel çalışma dışında kullanılmayacağı gibi, herhangi bir kişi veya kuruluşa da verilmeyecektir. Araştırmanın amacı, ülkemizde çok önemli bir konuma sahip olan meslek liselerinde bölüm seçimlerinin ne kadar isabetli olduğunu belirlemek ve bu seçimlerin ne tür değişkenlerden etkilendiğini ortaya koymaktır. Anket 4 bölümden oluşmaktadır. Birinci bölümde genel bilgilerinize ilişkin sorular, ikinci bölümde aile-veli bilgilerinize ilişkin sorular, üçüncü bölümde mesleki alan tercihlerinize yönelik sorular, dördüncü bölümde ise ders notlarınıza yönelik sorular yer almaktadır. Anketi yanıtlamanız yaklaşık 10 dakikanızı alacaktır. Her sorudaki duruma ilişkin gerçek bilgilerinizi girmeniz doğru meslek seçimlerine yönlendirmede önemli bir yer tutacaktır. Araştırmaya katkılarınızdan dolayı teşekkür ederim.

Özkan ÜNSAL

### 1-)Kişisel Bilgiler

**Cinsiyetiniz**  Erkek  Kız

**Doğum Yılı**

**Sınıfınız**

**Bölümünüz**

**İlköğretim Diploma Notunuz**  (5'lik sistemde yazınız)

**OKS Puanınız**

**Boyunuz(cm)**

**Kilonuz(kg)**

**Kardeş Sayınız**

**Ekonomik Durumunuz**

Kendime Ait Bir Çalışma Odam Var

Bir İşte Çalışıyorum

Özel Eğitim Kapsamında Kaynaştırma Öğrencisiyim

EK-1. (Devam) İnternet üzerinde yapılan anket formu

2-)Aile - Veli Bilgileri

**A)Veli Bilgileri**

Veliniz

Mesleği

Mezuniyeti

**B)Baba Bilgileri**

Mesleği

Mezuniyeti

Birlikte - Ayrı  Birlikte Yaşıyorum  Ayrı Yaşıyorum

Sağ - Ölü  Sağ  Ölü

**C)Anne Bilgileri**

Mesleği

Mezuniyeti

Birlikte - Ayrı  Birlikte Yaşıyorum  Ayrı Yaşıyorum

Sağ - Ölü  Sağ  Ölü

3-)Bölüm Tercih Bilgileri

**Bu Bölümde Yapmış Olduğunuz Bölüm Tercihlerini Sırası İle Seçiniz.**

1. Tercihiniz

2. Tercihiniz

3. Tercihiniz

4. Tercihiniz

5. Tercihiniz

EK-1. (Devam) İnternet üzerinde yapılan anket formu

4-)Ders Notu Bilgileri

**A)9.Sınıf Not Bilgileri (Karne notlarınızı 5'lik sistemde yazınız)**

<b>Ders Adı</b>	<b>Yıl Sonu Karne Notu</b>
<b>BEDEN EĞİTİMİ</b>	<input type="text"/>
<b>BİLGİ VE İLETİŞİM TEKNOLOJİSİ</b>	<input type="text"/>
<b>BİYOLOJİ</b>	<input type="text"/>
<b>COĞRAFYA</b>	<input type="text"/>
<b>DİL VE ANLATIM</b>	<input type="text"/>
<b>DİN KÜLTÜRÜ VE AHLÂK BİLGİSİ</b>	<input type="text"/>
<b>FİZİK</b>	<input type="text"/>
<b>KİMYA</b>	<input type="text"/>
<b>MATEMATİK</b>	<input type="text"/>
<b>SAĞLIK BİLGİSİ</b>	<input type="text"/>
<b>TARİH</b>	<input type="text"/>
<b>TÜRK EDEBİYATI</b>	<input type="text"/>
<b>YABANCI DİL</b>	<input type="text"/>
<b>SEÇMELİ DERS</b>	<input type="text"/>





## EK-2. Eğitim ve test veri setlerinin oluşturulmasında kullanılan t-sql komutları

```

SELECT  o.BOLUM, CASE WHEN o.BOLUM_ORTALAMA >= 3 THEN 'TRUE' ELSE 'FALSE' END AS
BASARILI, o.CINSIYET,
        CASE WHEN KARDES_SAYI > 2 THEN 'ÇOK' ELSE 'NORMAL' END AS KARDES_SAYI,
o.EKONOMIK_DURUM_KOD, o.ODA_DURUM,
        o.CALISMA_DURUM, CASE WHEN o.DIPLOMA_NOT >= 2.50 THEN 'TRUE' ELSE 'FALSE'
END AS DIPLOMA_NOT, v.VELI_TUR_KOD, v.VELI_MESLEK,
        v.VELI_MEZUNİYET, v.BABA_MESLEK, v.BABA_BERABER, v.BABA_HAYATTA,
v.ANNE_MESLEK, v.ANNE_MEZUNİYET, v.ANNE_BERABER,
        v.ANNE_HAYATTA,
        (SELECT TOP (1) BOLUM_KOD
FROM      dbo.TERCIH
WHERE     (OGRENCI_KOD = o.OGRENCI_KOD)) AS TERCIH,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT
WHERE     (DERS_KOD = 1) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS1,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_13
WHERE     (DERS_KOD = 2) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS2,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_12
WHERE     (DERS_KOD = 3) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS3,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_11
WHERE     (DERS_KOD = 4) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS4,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_10
WHERE     (DERS_KOD = 5) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS5,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_9
WHERE     (DERS_KOD = 6) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS6,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_8
WHERE     (DERS_KOD = 7) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS7,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_7
WHERE     (DERS_KOD = 8) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS8,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_6
WHERE     (DERS_KOD = 9) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS9,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_5
WHERE     (DERS_KOD = 10) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS10,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_4
WHERE     (DERS_KOD = 11) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS11,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_3
WHERE     (DERS_KOD = 12) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS12,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_2
WHERE     (DERS_KOD = 13) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS13,
(SELECT  DURUM
FROM      dbo.KULTUR_NOT AS KULTUR_NOT_1
WHERE     (DERS_KOD = 14) AND (OGRENCI_KOD = o.OGRENCI_KOD)) AS DERS14
FROM      dbo.OGRENCI_GENEL AS o INNER JOIN
        dbo.VELI AS v ON o.OGRENCI_KOD = v.OGRENCI_KOD
WHERE     (o.BOLUM_ORTALAMA > 0)

```

### EK-3. Algoritmanın uygulanmasında kullanılan C#.NET dili kodları

```

DataRow[] eğitim = veri.kayıtsac("SELECT * FROM EGITIM_VERI WHERE BOLUM=" +
comboBox3.SelectedValue.ToString());
DataRow[] test = veri.kayıtsac("SELECT * FROM TEST_VERI WHERE OGRENCI_KOD=" +
txtKod.Text);

int i = eğitim.Length;

//Seçilen Mesleki Alan İçin Toplam Başarılı ve Başarısız Örneklerin Hesaplanması
int p = 0, q = 0;
for (int j = 0; j < i; j++)
{
    if (eğitim[j]["BASARILI"].ToString() == "TRUE")
        p++;
    else
        q++;
}

decimal[, ] durum = new decimal[31,2];

for (int j = 0; j < i; j++)
{
    for (int k = 0; k < 31; k++)
    {
        if ((eğitim[j][k + 2].ToString() == test[0][k].ToString()) && (eğitim[j]["BASARILI"].ToString() ==
"TRUE"))
            durum[k, 0]++;
        if ((eğitim[j][k + 2].ToString() == test[0][k].ToString()) && (eğitim[j]["BASARILI"].ToString() ==
"FALSE"))
            durum[k, 1]++;
    }
}

DataTable tablo = veri.tabloyükle("SELECT * FROM EGITIM_VERI");

int ekp = 0, ekq = 0;
decimal ekppay = 0, ekqpay = 0;

for (int k = 0; k < 31; k++)
{
    if (durum[k, 0] == 0)
    {
        string alan_ad=tablo.Columns[k+2].ToString();
        DataRow[] alan_sınıflar = veri.kayıtsac("SELECT DISTINCT " + alan_ad + " FROM
EGITIM_VERI WHERE BOLUM=" + comboBox3.SelectedValue.ToString());
        ekp++;
        ekppay += (decimal)1 / alan_sınıflar.Length;
    }

    if (durum[k, 1] == 0)
    {
        string alan_ad = tablo.Columns[k + 2].ToString();
        DataRow[] alan_sınıflar = veri.kayıtsac("SELECT DISTINCT " + alan_ad + " FROM
EGITIM_VERI WHERE BOLUM=" + comboBox3.SelectedValue.ToString());
        ekq++;
        ekqpay += (decimal)1 / alan_sınıflar.Length;
    }
}

```

### EK-3. (Devam) Algoritmanın uygulanmasında kullanılan C#.NET dili kodları

```
//Hesaplamaya Girecek Her Bir Kriter İçin Olumlu ve Olumsuz İhtimal Hesaplaması

decimal[,] ihtimal = new decimal[2, 31];

for (int j = 0; j < i; j++)
{
    for (int k = 0; k < 31; k++)

        if ((eğitim[j][k+2].ToString() == test[0][k].ToString()) && (eğitim[j]["BASARILI"].ToString()
== "TRUE"))
            ihtimal[0, k]++;
        else if ((eğitim[j][k + 2].ToString() == test[0][k].ToString()) &&
(eğitim[j]["BASARILI"].ToString() == "FALSE"))
            ihtimal[1, k]++;
    }

    decimal poran = 0, qoran = 0;

    poran = (decimal)p / i;
    qoran = (decimal)q / i;

    //p ve q oranlarının hesaplanması
    for (int a = 0; a < 31; a++)
    {
        poran *= (decimal)(ihtimal[0, a]+ekppay) / (p+ekp);
        qoran *= (decimal)(ihtimal[1, a]+ekqpay) / (q+ekq);
    }

    if (poran > qoran)
        MessageBox.Show("P Oran : " + string.Format("{0:F30}", poran) + " Q Oran : " +
string.Format("{0:F30}", qoran) + "\nSonuç: Başarılı;", "Alan Başarı
Tahmini", MessageBoxButtons.OK, MessageBoxIcon.Information);
    else
        MessageBox.Show("P Oran : " + string.Format("{0:F30}", poran) + " Q Oran : " + string.Format("{0:F30}",
qoran) + "\nSonuç: Başarısız;", "Alan Başarı Tahmini", MessageBoxButtons.OK, MessageBoxIcon.Information);
}
```

## ÖZGEÇMİŞ

### Kişisel Bilgiler

Soyadı, adı : ÜNSAL, Özkan  
Uyruğu : T.C.  
Doğum tarihi ve yeri : 19.05.1986 Ankara  
Medeni hali : Bekar  
Telefon : 0 (312) 252 44 85  
E-mail : ozkanunsal@gmail.com

### Eğitim

Derece	Eğitim Birimi	Mezuniyet tarihi
Lisans	Gazi Üniversitesi/ Bilgisayar Eğitimi	2008
Lise	Ankara Atatürk Lisesi	2004

### İş Deneyimi

Yıl	Yer	Görev
2008-2011	Milli Eğitim Bakanlığı	Bilgisayar Öğretmeni

### Yabancı Dil

İngilizce

### Yayınlar

1. Bülbül, H.İ., Ünsal, Ö. "Determination of vocational fields with machine learning algorithm", The Ninth International Conference on Machine Learning and Applications (ICMLA 2010), IEEE Computer Society, Washington D.C., 710-713 (2010).