T.C.
TRAKYA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
YÜKSEK LİSANS TEZİ



NİCEL BİRİKİMİN NİTEL DEĞİŞİME ETKİSİ: KÜTÜPHANE YÖNETİMİNDE VERİ MADENCİLİĞİ UYGULAMASI

ESRA TAŞKIN

DANIŞMAN
PROF. DR. FATMA LORCU

Edirne 2020

T.C.

TRAKYA ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

.....İŞLETME..... ANABİLİM DALI

.....İŞLETME.....PROGRAMI

YÜKSEK LİSANS TEZİ

.....Esra TAŞKIN..... tarafından hazırlanan .....Nicel Birikimin Nitel Değişime Etkisi : Kütüphane Yönetiminde Veri Madenciliği Uygulaması..... Konulu **Yüksek Lisans** tezinin Sınavı, Trakya Üniversitesi Lisansüstü Eğitim-Öğretim Yönetmeliği'nin 19-6 maddeleri uyarınca .....08.01.2020 Çarşamba..... günü saat .....16.00..... 'da yapılmış olup, yüksek lisans tezinin *.....Kabul edilmesine..... **OYBİRLİĞİ/~~OYÇOKLUĞU~~** ile karar verilmiştir.

JÜRİ ÜYELERİ	KANAAT	İMZA
Prof. Dr. Fatma LORCU (Danışman)	Kabul Edilmesine	
Doç. Dr. Emel YILDIZ	Kabul Edilmesine	
Dr. Öğr. Üyesi Yasemin BİLİR	Kabul Edilmesine	

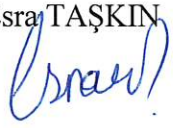
* Jüri üyelerinin, tezle ilgili kanaat açıklaması kısmında "Kabul Edilmesine/Reddine" seçeneklerinden birini tercih etmeleri gerekir.

TRAKYA ÜNİVERSİTESİ SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
DOĞRULUK BEYANI

Trakya Üniversitesi Sosyal Bilimler Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında, tüm verilerin bilimsel ve akademik kurallar çerçevesinde elde edildiğini, kullanılan verilerde tahrifat yapılmadığını, tezin akademik ve etik kurallara uygun olarak yazıldığını, kullanılan tüm literatür bilgilerinin bilimsel normlara uygun bir şekilde kaynak gösterilerek ilgili tezde yer aldığını ve bu tezin tamamı ya da herhangi bir bölümünün daha önceden Trakya Üniversitesi ya da farklı bir üniversitede tez çalışması olarak sunulmadığını beyan ederim.

08 / 01 / 2020

Esra TAŞKIN



Tezin Adı: Nicel Birikimin Nitel Değişime Etkisi: Kütüphane Yönetiminde Veri Madenciliği Uygulaması

Hazırlayan: Esra TAŞKIN

ÖZET

Bilişim teknolojilerinde yaşanan gelişmelerle birlikte rekabet koşullarının da değişmesiyle bilgi güç haline gelmiş ve bilgiye olan önem daha da artmıştır. Bununla birlikte, dünya ekonomisinden daha hızlı büyüyen veriyi doğru okuyup bilgiye dönüştürme süreci de zorlaşmıştır. Eğitim-öğretim ve bilimsel faaliyetleri destekleyen akademik kütüphanelerde de büyük miktarda kullanıcı ve materyal bilgileri depolanmaktadır. Bu depolanan verilerin miktarı arttıkça yönetimi de zorlaşmaktadır. Dağınık ve karmaşık yapıdaki büyük veri yığınlarının analizinde son yıllarda gelişme gösteren veri madenciliği analizi kullanılmaktadır. Veri madenciliği, veri ambarlarında depolanan büyük miktardaki yapılandırılmamış veri içerisinden faydalı bilgiye ulaşma işidir.

Bu araştırmada, Trakya Üniversitesi Merkez Kütüphanesi'nden ödünç yayın alan öğrencilerin davranışları incelenerek sonuçların, kütüphane yönetimindeki karar vericileri için yeni bir kaynak olması amaçlanmıştır. Bu amaç doğrultusunda, ilk bölümde, verilerin toplanıp kendi içinde anlamlandırılmasıyla başlayan ve devamında özetlenip çeşitli analizlerle bilgiye dönüşmesi süreci anlatılmış ve çevrimiçi kayıt sistemleri (Online Transaction Processing-OLTP) ve çevrimiçi analitik işleme (Online Analytical Processing-OLAP) sorgulamaları, karar destek sistemleri ve karar destek sistemlerinin gerçekleştirilebilmesi için bir alt yapı oluşturan veri ambarı kavramı açıklanmıştır. İkinci bölümde ise veri madenciliğinin tarihsel gelişimi, kullanılan yöntemler ve uygulama alanları anlatılmıştır. Son bölümde ise Trakya Üniversitesi Merkez Kütüphanesi'nden alınan kullanıcı ödünç yayın sirkülasyon verileri, RapidMiner programı kullanılarak analiz edilmiştir.

Yapılan analizler sonucunda, Trakya Üniversitesi Merkez Kütüphanesi'nden ödünç yayın alan öğrencilerin davranışları incelenerek akademik birim ve bölüm

bazlı tercih edilen yayınlar, ödünç alınan yayınların teslim durumu, yıl ve ay bazlı ödünç yayın alımı gibi istatistiksel sonuçların gösterilmesi sağlanmıştır. Çalışma sonuçlarından faydalanarak kütüphane kullanıcıları ve kütüphane yönetimi arasında önemli bir iletişim kanalı oluşturularak çalışmadan elde edilen bilgiyle yayın erişimini kolaylaştıracak uygulamalar geliştirilebilir.

Anahtar Kelimeler: Veri Madenciliği, Kümeleme Analizi, Birliktelik Kuralları Analizi, Kütüphane Verisi

Thesis Name: Qualitative Change Effect of Quantitative Accumulation: Data Mining Application in Library Management

Prepared by: Esra TAŞKIN

ABSTRACT

The developments in information technologies and together with changes in competition conditions, knowledge has become like power and the importance of knowledge has further increased. However, the data that is growing faster than the world economy has become a difficult right to analyze and the process of transformation into information. A large amount of user and material information is stored in also academic libraries that support education, training and scientific activities. As the amount of this stored data increases, its management becomes also difficult. Data mining analysis, which has been developing in recent years, is used in the analysis of dispersed and complex big data stacks. Data mining is an operation of accessing useful information from within a large amount of unstructured data stored in data warehouses.

In this research, the behaviors of students of work borrower from Trakya University Central Library were examined and the results were aimed to be a new source for decision-makers in library administration. In accordance with this purpose, in the first part, the process of data collection and interpretation, and the process of its transformation into information through various analyzes is explained and afterward, concepts of OLTP and OLAP queries, decision support systems and data warehouse are explained. In the second part, historical development of data mining, the methods used and application areas are explained.In the last part, borrow work circulation data received from Trakya University Central Library users were analyzed using the RapidMiner program.

As a result of the analyzes, the behaviors of students borrowed from Trakya University Central Library were examined and statistical results such as preferred units based on academic unit and department, delivery status of borrowed

publications, and borrowing on year and month basis were shown. By making use of the results of study, an important communication channel can be established between library users and library management and applications can be developed to facilitate publication access with the information obtained from the study.

Key Words: Data Mining, Cluster Analysis, Association Rules Analysis, Data of Library

ÖN SÖZ

Bu çalışmada, bilgi teknolojisinin doğal evriminin bir parçası olan ve hızla büyüyen veri madenciliği kavramı detaylı bir şekilde incelenmiştir. Veri madenciliği süreci, veri madenciliğinde kullanılan modeller, işletmeler ve kütüphanelerde veri madenciliği uygulama alanlarından bahsedilmiştir. Trakya Üniversitesi Merkez Kütüphanesi'nden elde edilen veriler üzerinde birliktelik ve kümeleme analizi uygulaması yapılmıştır. Çalışma sonuçları doğrultusunda öneriler sunulmuştur.

Araştırmamın her aşamasında değerli katkı ve eleştirileriyle yol gösteren saygıdeğer hocam Sayın Prof. Dr. Fatma LORCU'ya, araştırmamın uygulama aşamasında gerçekleşmesi için gerekli olan veri setinin sağlanmasında yardımcı olan Trakya Üniversitesi Merkez Kütüphanesi Daire Başkan Vekili Öğr. Gör. Bahattin DEMİRELLİ'ye teşekkür ederim.

Hayatım boyunca her konuda beni destekleyen ve hep yanımda olan annem Cahide TAŞKIN'a, babam İlyas TAŞKIN'a ve kardeşim Selin TAŞKIN'a, hayatımdaki değerini her gün daha iyi anladığım Hasan Yiğit KONAKLILAR'a, benden hiçbir zaman desteklerini esirgemeyen değerli arkadaşlarım Aslı ÇAVUŞ, Aytaç GÜLTEKİN, Kemal KUMRAL ve Merve ÇAKIR'a sonsuz teşekkür ederim.

Ocak, 2020 Esra TAŞKIN

İÇİNDEKİLER

ÖZET ..... **I**

ABSTRACT ..... **III**

ÖN SÖZ ..... **V**

İÇİNDEKİLER ..... **VI**

TABLOLAR LİSTESİ ..... **IX**

ŞEKİLLER LİSTESİ ..... **XI**

KISALTMALAR ..... **XIII**

GİRİŞ **1**

BİRİNCİ BÖLÜM

KARAR DESTEK SİSTEMLERİ ve VERİ AMBARLARI

1.1. Veri, Enformasyon, Bilgi ve Bilgelik Kavramları 5

1.2. Çevrimiçi Kayıt (OLTP) ve Çevrimiçi Analitik İşleme (OLAP) Sistemleri……..8

1.3. Karar Destek Sistemleri 10

1.4. Veri Ambarı 14

1.4.1. Veri Ambarı İşleyişi 17

1.4.2. ETL Süreçleri 18

1.5. Veri Ambarı ve Veri Madenciliği İlişkisi 19

İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

2.1. Veri Madenciliğinin Tanımı ve Gelişimi 21

2.2. Veri Madenciliği Süreci 25

2.3. Veri Madenciliği Modelleri 28

2.3.1. Tahmin Edici Modeller 29

2.3.1.1. Sınıflandırma 29

2.3.1.2. Regresyon Analizi 31

2.3.2. Tanımlayıcı Modeller 31

2.3.2.1. Kümeleme Analizi 32

2.3.2.2. Birliktelik Kuralları 33

2.4. Veri Madenciliğinde Karşılaşılan Sorunlar 35

2.5. Veri Madenciliğinin İşletmelerde Uygulama Alanları 36

2.6. Kütüphanelerde Veri Madenciliği ve Uygulama Örnekleri 40

ÜÇÜNCÜ BÖLÜM

KÜTÜPHANE VERİ MADENCİLİĞİ UYGULAMASI

3.1. Araştırmanın Amacı ve Önemi 44

3.2. Araştırmanın Evreni ve Örneklemi 45

3.3. Araştırmada Verilerin Elde Edilmesi ve Kullanılacak Program Seçimi 45

3.4. Veri Ön İşleme Süreci 47

3.5. Araştırma Bulguları 52

3.5.1. Kütüphane Kullanıcı İstatistikleri 52

3.5.2. Kütüphane Ödünç Yayın Sirkülasyon Verileri Üzerine Birliktelik Analizi 63

3.5.3. Kütüphane Kullanıcı Verileri Üzerine Kümeleme Analizi 69

SONUÇ ve ÖNERİLER 74

KAYNAKÇA 78

DİZİN 90

EKLER 90

TABLOLAR LİSTESİ

Tablo 1: OLTP ve OLAP Sistemleri Arasındaki Farklar..... 9

Tablo 2: Veri Madenciliğinin Tarihçesi..... 23

Tablo 3: İşletmelerde Veri Madenciliğinin Uygulama Alanları..... 37

Tablo 4: Veri Madenciliği Uygulama Sürecinde Kullanılan Tablolar..... 46

Tablo 5: Veri Önişleme İşlemleri Sonu Veri Madenciliğinde Kullanılacak Tabloların Durumu..... 51

Tablo 6: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Cinsiyete Göre Dağılımı..... 52

Tablo 7: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Eğitim Programına Göre Dağılımı..... 52

Tablo 8: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Üniversitedeki Akademik Birimlere Göre Dağılımı..... 53

Tablo 9: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Üniversitedeki Bölümlerine Göre Dağılımı (İlk ve Son Üç Bölüm)..... 56

Tablo 10: Kütüphaneden En Az Bir Yayın Ödünç Almış İktisadi ve İdari Bilimler Fakültesi Bölümlerinde Eğitim Gören Öğrencilerin Cinsiyete Göre Dağılımı..... 56

Tablo 11: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Yayınları Teslim Etme Durumuna Göre Dağılımı..... 58

Tablo 12: Kütüphaneden Ödünç Alınan Yayınların Cinsiyete Göre Teslim Durumu Dağılımı..... 58

Tablo 13: Kütüphaneden Ödünç Alınan Yayınların Eğitim Programına Göre Teslim Durumu Dağılımı..... 59

Tablo 14: Kütüphaneden Ödünç Alınan Yayınların Akademik Birimlere Göre Teslim Durumu Dağılımı 59

Tablo 15: Kütüphaneden Ödünç Alınan Yayınların Alındığı Aylara Göre Dağılımı 62

Tablo 16: Birliktelik Kuralı Tablosu..... 65

ŞEKİLLER LİSTESİ

Şekil 1: Bilgi Hiyerarşisi Piramidi 6

Şekil 2: Yönetim Düzeylerine Göre Karar Türleri 11

Şekil 3: Yıldız Şeması Tablosu 16

Şekil 4: Veri Ambarı İşleyişi 17

Şekil 5: ETL Süreci 18

Şekil 6: Veri Tabanlarında Bilgi Keşfi Sürecini Oluşturan Adımlar 26

Şekil 7: CRISP-DM Süreç Diyagramı 27

Şekil 8: Üye Tablosundaki Verilerin Ham Hali 47

Şekil 9: Ödünç Tablosundaki Verilerin Ham Hali 49

Şekil 10: Verilerin Ön İşlemden Geçmiş Hali 51

Şekil 11: Kütüphaneden Ödünç Alınan Yayınların Sınıflarına Göre Dağılımı (İlk On Yayın Sınıfı) 57

Şekil 12: Kütüphaneden Ödünç Alınan Yayınların Alındığı Yıllara Göre Dağılımı 63

Şekil 13: RapidMiner Üzerinde Oluşturulan Veri Madenciliği Süreci 64

Şekil 14: İktisadi ve İdari Bilimler Fakültesi Grubunda Oluşan Birliktelik Kuralının Grafik Olarak Gösterilmesi 68

Şekil 15: Kümeleme Analizi Süreci 70

Şekil 16: Kümeleme Analizi Sonucu Oluşan Kümeler 70

Şekil 17: Kümeleme Analizi Sonucu Kümelerdeki Cinsiyet Dağılımları ... 71

Şekil 18: Kümeleme Analizi Sonucu Oluşan Kümelerdeki Kullanıcı Eğitim Programı Dağılımı 71

Şekil 19: Kümeleme Analizi Sonucu Oluşan Birinci Kümedeki Kullanıcıların Bölüm Dağılımı 72

Şekil 20: Kümeleme Analizi Sonucu Oluşan İkinci Kümedeki Kullanıcıların Bölüm Dağılımı 72

Şekil 21: Kümeleme Analizi Sonucu Oluşan Kümelerdeki Kullanıcı Yayın Teslim Durumu Dağılımı 73

KISALTMALAR

CART: Veri Madenciliği Algoritması (Classification and Regression Trees)

CRISP-DM: Veri Madenciliği İçin Sektörler Arası Standart Süreç (CRoss-Industry Standard Process for Data Mining)

DDC: Dewey Onlu Sınıflama Sistemi (Dewey Decimal Classification)

EFT: Elektronik Fon Transferi

ENIAC: Elektronik Sayısal Entegreli Hesaplayıcı (Electronic Numerical Integrator and Computer)

ETL: Çıkarım, Dönüşüm, Yükleme (Extract, Transform, Load)

FAS: Finansal Erişim Araştırması (Financial Access Survey)

GRI: Veri Madenciliği Algoritması (Generalized Rule Induction)

KDS: Karar Destek Sistemleri

OLAP: Çevrimiçi Analitik İşleme (Online AnalyticalProcessing)

OLTP: Çevrimiçi Kayıt İşleme (Online TransactionProcessing)

URL: Tekbiçimli Kaynak Konumlayıcı (Uniform Resource Locator)

VTBK: Veri Tabanlarında Bilgi Keşfi

YALE: Başka Bir Öğrenme Ortamı (Yet Another Learning Environment)

GİRİŞ

Bilişim teknolojilerinde yaşanan gelişmelerle birlikte analogtan dijitale evrilmiş bir süreç yaşanmaktadır. Bu süreçte, dijital dünyada depolanan veri miktarı katlanarak artmaktadır. 2000'li yılların başında bir yılda üretilen veri, artık bir saatte üretilmektedir. Veri miktarındaki devam eden bu atışla birlikte hem bilimsel hem de toplumsal değeri olan ve önemli ölçüde ekonomik değer kaynağı olan verinin bilgi keşfi süreci zorlaşmıştır. Teknolojide yaşanan gelişmeler ve beraberinde getirdiği zorluklar, işletmelerin sürdürülebilir rekabet konusunda geçmişe göre daha fazla önem vermesine neden olmuştur. Örneğin; artan veri yığını üzerinde eski teknikler kullanılarak bilgiye ulaşılmaya çalışıldığında zaman almakta ve bilgi daha kullanılmadan eskimekteydi. Doğru verilerin zamanında işlenmesiyle ortaya çıkan bilgi işletmeler için değerli nitelikte olduğu için yeni nesil teknolojilere ve yöntemlere ihtiyaç duyulmuştur.

Bilgi keşfi sürecinde, teknolojinin etkin bir şekilde kullanılması işletmelere maliyet ve zaman tasarrufu, ürün/hizmet veya süreç iyileştirmeleri ve kısa sürede alınan doğru kararlarla iş dünyasındaki rakiplerine karşı sürdürülebilir rekabet üstünlüğü sağlamaktadır. Bu bağlamda, artan veri yığını arasında ilişki kurma ve anlamlı örüntüler çıkartabilmek için bilgi keşfinin ana unsuru, insanların ve bilgisayarların ortak bir çabası olan veri madenciliği teknolojisinden yararlanılmaktadır (Katardazics, 2011: 1)

İşletmelerde, veri madenciliği tekniklerinin etkin bir şekilde kullanılmasıyla kısa zamanda ve az maliyetle elde edilen bilgiyle işletmelerin anlık sorularına cevap alınabileceği gibi durumsal ve zamansal değişkenliklerin işin içine girmesiyle de oluşabilecek muhtemel sorulara da cevap alınabilmektedir. Örneğin; alışveriş yaparken kullanılan müşteri kartlarından satın alma alışkanlıkları ve tercihleri hakkında elde edilen verilerle oluşturulan kişiye özel kampanyalar, Google aramaları ve sosyal medya hesaplarındaki kapsamlı profil bilgileri ve beğenilerden yola çıkarak oluşturulan internet reklamları, örüntü tanıma yazılımı ile hastalık veya salgının

erken dönemde tespiti ve etkili tedavi süreci, coğrafi konum ve olağan kredi kartı faaliyetleri değerlendirilerek bu göstergelerin dışında kalan olağandışı bir işlemde ilgili banka tarafından işlemin yasallığını doğrulamak için gönderilen mesajlar veri madenciliğinin kullanıldığı alanlara birer örnektir.

Teknolojinin değişen doğası ve uygulamaların değişen karakteri ile dönüşen rekabet anlayışında veriye yeni anlamlar yükleyen veri madenciliği uygulamaları birçok alanda kullanıldığı gibi akademik kütüphanelerde de uygulanmaktadır. Bilişim teknolojilerindeki gelişmelerle beraber akademik toplumunun ihtiyaçlarını karşılayan üniversite kütüphanelerinde de dijitalleşme süreci başlamıştır (Çelik ve Uçak, 1993: 120).

Akademik kütüphanelerin veri yönetim sisteminde, kütüphane kullanıcıları tarafından günlük gerçekleştirilen yayın sorgulama, yayın ödünç alma ve iade etme gibi işlemler hakkında veriler üretilmekte ve depolanmaktadır. Üretilen verilerden kütüphane yönetimi tarafından çeşitli hizmet ve faaliyet raporları oluşturulmasına rağmen mevcut verileri dönüştürmek ve bu verilerde gizli örüntüleri keşfetmek için herhangi bir yöntemden yararlanılmamaktadır. Veri havuzu içindeki anlamlı bilgi kalıplarını bulmak veri madenciliği teknolojisi kullanılmaktadır. Kütüphane veri yönetimi sisteminde, veri madenciliği teknikleri uygulanarak yayın ödünç sirkülasyon verilerinden kullanıcıların ihtiyaçları ve yayınlar arasındaki ilişkilerin tahmini gibi yararlı bilgilere ulaşılabilmektedir.

Bu çalışmada, Trakya Üniversitesi Merkez Kütüphanesi kullanıcılarının ödünç yayın sirkülasyonu veri madenciliği ile incelenmiştir. Üç bölümden oluşan çalışmanın birinci bölümünde; bilgi hiyerarşisinin unsurları olan veri, enformasyon, bilgi ve bilgelik kavramları incelenmiştir. İlişkisel veri tabanları üzerinde yapılan OLTP ve OLAP sorgulamaları, karar destek sistemleri ve veri ambarı kavramı anlatılarak veri ambarının veri madenciliği ile ilişkisi üzerinde durulmuştur.

İkinci bölümde, bilgi keşfi sürecinde bilgiye giden yolda geçilmesi gereken noktalardan birisi olan veri madenciliği kavramı ve gelişim sürecinden bahsedilmiştir. Veri madenciliği ve modelleri, süreci ve karşılaşılan sorunlar

anlatılarak veri madenciliğinin işletmelerde ve kütüphanelerdeki uygulama alanları incelenmiştir.

Son bölümde ise araştırmanın amacı, önemi, evreni, örneklemi, verilerin toplanması ve uygulama süreci detaylı bir şekilde anlatılmıştır. Kullanılan veri madenciliği yöntemlerinin sonucunda kütüphane kullanıcılarının ödünç aldığı yayın sirkülasyon bilgilerinden birliktelik kuralları oluşturulmuş ve kütüphane kullanıcı bilgileri üzerinde kümeleme analizi uygulanarak kütüphane kullanıcılarının davranışları incelenmiştir.

BİRİNCİ BÖLÜM

KARAR DESTEK SİSTEMLERİ ve VERİ AMBARLARI

İnsanlık tarihi boyunca toplumlar, bilginin varlığı üzerine kurulmuş ve toplumsal yaşamda her zaman bilgiye ihtiyaç duyulmuştur. Bilgi, insanların neyi, nasıl yapacağı ve diğer karşılaşılan sorunların nasıl üstesinden gelme konusunda yaşamın beraberinde getirdiği sorulara karşılık cevaplar bulma arayışında toplumsal yaşama destek olmuş ve yol göstermiştir (http://www.phil.metu.edu.tr/). Bir başka deyişle insanlar, yaşadığı toplumda gerçekleşen olayları gözlemleyerek bu olayları anlamlandırmaya çalışmıştır.

Bilginin tarihsel gelişim süreci, MÖ 5. yüzyılda, sorduğu sorularla genelgeçer olan bilgiye ulaşmayı amaçlayan felsefe filozofu olan Sokrates ile başlamıştır (Güçlü ve Sotirofski, 2006: 352). Tarihsel gelişim sürecinde ilk aşamada bilgi; aletlere, süreçlere ve ürünlere uygulanarak Sanayi Devrimi'ni yaratmıştır. İkinci aşamada, 1880 yıllarında başlayıp 2. Dünya Savaşı'yla sona eren dönemde bilgi, çalışmalara uygulanmaya başlamış ve yeni bir boyut kazanmıştır. 2. Dünya Savaşı'ndan sonra ortaya çıkan son gelişim aşamasında ise bilgi, bilginin kendisine uygulanmaya ve üretim faktörü olarak kullanılmaya başlamıştır. Diğer bir deyişle bilgi, örgütlerin en önemli kaynağı haline gelerek doğal kaynak, emek ve sermaye faktörlerini bir kenara itmiştir (Drucker, 1993: 33-34).

18. yüzyılın sonu ve 19. yüzyılın başlarında James Watt'ın geliştirdiği buhar gücüne dayalı makine, Sanayi Devrimi'ni nasıl devinime geçirdiyse 2. Dünya Savaşı'ndan sonra bir grup bilim adamı tarafından geliştirilen ve ilk genel amaçlı bilgisayar olarak nitelendirilen elektronik sayısal entegreli hesaplayıcı (Electronic Numerical Integrator and Computer-ENIAC) da Bilgi Devrimi'nin tetikleyicisi olmuştur (Drucker, 2003: 14-15; Aydın, 2012: 183).

Dünya genelinde büyük etki yaratan küresel ekonominin doğuşu, bilgi toplumlarında ve teşebbüslerdeki değişim süreciyle birlikte geçmişe oranla birçok

kuruluşta veri üretimi artmıştır. Artan veri üretimiyle asgari düzeyde olan teknolojik yapı, veri saklama ve işleme yöntemleri konusunda yetersiz kalmıştır. Bu üç büyük değişim sürecinde, iş dünyası ve çevresindeki işletme ve yönetimlerde, karar destek uygulamalarında veri tabanı sistemlerinin ihtiyacı karşılamadığı görülmüştür. Yeni doğan ihtiyaçlar doğrultusunda, karar vericilerin veriye tek merkez üzerinden ve hızlı bir şekilde erişilmesine olanak sağlayan veri ambarı kavramı ortaya çıkmıştır (Karahoca ve Karahoca, 1998: 1; Özkan, 2013: 20).

Bu bölümde, çalışmanın temelini oluşturan veri kavramı ve bu kavram temelinde gelişen bilgi hiyerarşisinin unsurları olan veri, enformasyon, bilgi ve bilgelik kavramları incelenecektir. Ayrıca; OLTP, OLAP, veri ambarı ve karar destek sistemleri irdelenerek veri ambarı ve veri madenciliği arasındaki ilişkiden yine bu bölümde bahsedilecektir.

1.1. Veri, Enformasyon, Bilgi ve Bilgelik Kavramları

Dünyada bulunan çeşitli ve birçok sayıdaki öğenin, gözlemlenebilecek özelliklere sahip olduğu varsayılmaktadır. Veriler de bu gözlemlenebilir özelliklerin sembolik birer gösterimidir (Rowley, 2007: 170). Şeker (2013: 22), herhangi bir işlem görmemiş, gözlem veya ölçümle ulaşılan her türlü değeri, **veri** olarak nitelendirmektedir. Örneğin; işletmelerdeki personel sayısı ve çalışma saatleri, günlük ödünç alınan kitap sayısı birer veridir.

Veri tabanlarında depolanan veri; dağınık, biçimsiz ve karmaşık bir yapıdadır. Verinin işlenmemiş, ham haline, **yapılandırılmamış veri** denir. Dünya üzerindeki verilerin %80 veya %90'ının yapılandırılmamıştır. Yapılandırılmamış verilerin, bilgisayar tarafından okunabilen yapılandırılmış verilere dönüştürmek için işlenmesine ihtiyaç duyulmaktadır (Gürsakal, 2014: 35-56). Yapılandırılmamış verilerin güncel olması ve bu veriler üzerinden yapılan çıkarımların doğruluğu veriyi değerli kılan özelliktir.

1960'lı yıllarda, bilgisayar bilgisinden bir adım öne geçilip dönemin şartlarında ana üretkenlik etmeni olarak görülen enformasyon, verinin işlenmiş hali

olarak tanımlanabilir (Drucker, 2003: 45). **Enformasyon**, verilerin işlenerek ve ilişkilendirilerek işletmelerin karar verme aşamalarında kullanıldığı gibi bilginin oluşması için gerekli alt yapıyı sağlamaya da katkıda bulunmaktadır. Örneğin; bir pazarlama stratejilerini oluştururken kullanması vb.



Şekil 1: Bilgi Hiyerarşisi Piramidi

Kaynak: Veri, Enformasyon (Malumat), Bilgi ve Bilgelik (İrfan), https://ohantekten.com/wp-content/.../Veri_-Enformasyon-_Malumat__-Bilgi-ve.pdf, 6 (03.12.2018).

Bilgi, hem devletlerde hem de işletmelerde uyuşmazlık, rekabet ve mücadele durumunda ana kaynak olduğu için bu kaynağın doğru bir şekilde tanımlanması, ölçülmesi ve değerlenmesi son derece önemlidir. Şekil 1'de yer alan bilgi hiyerarşisi incelendiğinde, bilgiye ulaşmanın farklı aşamalardan geçtiği ve süreçteki aşamaların da kolay olmadığı görülmektedir. Ortaya çıkan durum, bilgi çağındaki önemli sorunlardan birisidir. Bilgi kaynağını; doğal kaynak, emek ve sermaye gibi diğer işletme kaynaklarından farklı kılan ve ölçülmesi zor olan özellikler şu şekilde sıralanabilir (Waltz, 1998: 49-50):

- Bilgi, bir varlığın (yer, açıklama, ölçüm gibi) veya bir sürecin (şifreleme işleme, ilişki gibi) şeklini alabildiği için soyut bir kavramdır.
- Bilgi, çoklu ve eşzamanlı kullanıma sahiptir. Örneğin; bir radyo vericisinin frekansı iletişim kurmak veya bozmak için kullanılabilir.
- Bilgi sınırsızdır; keşfedilebilir, yaratılabilir, dönüştürülebilir ve tekrarlanabilir ancak zamansaldır. Yeni bilgiler eyleme geçirilebilir bir değere sahiptir, eski bilgiler ise sadece geçmiş dönem bilgilerine sahip olabilir.

- Bilginin değeri veya faydası, yalnızca hacmiyle veya büyüklüğüyle alakalı değildir. Kullanılacak olan verinin potansiyeli, bu veriden oluşan enformasyonun içeriği ve enformasyondan elde edilecek bilginin de gerçek dünyadaki etkisi ve işlevselliğiyle ilişkilidir.

Veri ve enformasyon, bilgiyi meydana getiren yapıtaşlarıdır (Gülseçen, 2012: 20). Enformasyon kendiliğinden bilgiye dönüşmez. Enformasyon; bireyin onu algılaması, içselleştirmesi ve tecrübeleriyle şekillendirmesi sonucunda bilgeye dönüşür. Belirtilen bu özellikler, bilgiyi kişisel ve öznel kılmaktadır. Bilgi, farklı kaynaklardan sentezlenen enformasyonun teoriler, aksiyomlar veya yapılara yerleştirilmesidir (Baltzan, 2014: 10).

Genel olarak bakıldığında bilgi, enformasyondan; enformasyon ise veriden işlenerek üretilir. Bu dönüşüm sürecinde, ham halde olan veri, en değersiz ve çok işe yaramayan bir olgu gibi görülmektedir. Oysa veri, her türlü rekabet durumunda işletmeleri bir adım ileriye taşıyacak olan bilgiyi oluştururken kullanılan esas kaynaktır. Ayrıca, bu süreçteki en değerli ve en çok işe yarayan olgu bilgeliktir. **Bilgelik**, büyük resme bakıp olayı anlayabilme, altında yatan sebepleri görebilme ve yorumlayabilme yeteneğidir. Bilgelik, sübjektif bir yapısı olan ve doğru veya yanlış, iyi veya kötü gibi etik kavramların devreye girdiği süreçtir. Bilgelik, henüz bilgisayarlar tarafından ulaşılamamış bir seviye olarak görülmektedir (Ahsan ve Shah, 2006: 273).

Bilgi hiyerarşisi piramidinin amacı; piramidin en altında yer alan verinin; toplanması, işlenmesi ve ilişkilendirilmesiyle üretilen bilgiyi kullanacak olan kişilere bilge kararlar verdirebilmektir. Bu noktada, bilge kararlar verebilmek ve bilginin oluşum sürecinde esas kaynak olan verinin muhafaza edilebilmesi için veri depolama birimlerine ihtiyaç duyulmaktadır.

1.2. Çevrimiçi Kayıt (OLTP) ve Çevrimiçi Analitik İşleme (OLAP) Sistemleri

Bilişim teknolojilerinde yaşanan değişim ve gelişimden etkilenen işletme sayısı her geçen gün artmaktadır. Sürekli gelişen teknolojiyle beraber işletmelerin rekabet avantajı sağlayabilmeleri için bilgiye olan ihtiyaçlarını daha da artmıştır. Bilgiye sahip olanın güçlü olduğu bu dönemde, işletmeler verilerini saklayabilmek ve bu verilerden de karar verme sürecinde destek almak için veri depolama birimlerine ihtiyaç duymuşlardır.

Veri tabanı, veriyi merkezileştirerek kullanım amacına göre düzenleyen sistemlerdir. OLTP, veri tabanları üzerinden günlük hesaplamaların tutulduğu sistemlerdir. Bankacılık sektöründe gün içinde yapılan havale ve elektronik fon transferi-EFT işlemleri; eğitim sektöründe, öğrenci ders kayıtları ve sınav notları; kütüphane veri tabanlarından yapılan eser ayırtma veya alınan eserin iade tarihini uzatma işlemleri vb. gibi tipik sorgular OLTP'ye birer örnektir (Plattner, 2014: 18). OLTP sistemleri, olaylar ve işlemlerle ilgili verileri depoladığından sürekli yeni veri girişi yapılmaktadır. Bu yüzden OLTP sistemlerinde en güncel veriler bulunmaktadır (Gökmen, 2014: 28).

Geleneksel veri tabanları, birbirleriyle etkileşim halinde olan büyük miktardaki veriler üzerinde yapılan tam tablo taramaları, birleştirme ve bireysel ve esnek sorgulama gibi karmaşık işlemleri desteklememektedir. Daha üst veri yönetimi ve çok boyutlu analiz için OLAP sistemleri geliştirilmiştir.

OLAP, temel olarak son kullanıcılar tarafından gerçekleştirilen çeşitli etkinlikleri ifade etmektedir. Belirlenen amaçlara yönelik sorgulamaların yapıldığı platformdur. OLAP, veri tabanı yönetim sistemleri veya daha çok veri ambarı sistemlerindeki büyük veri setlerini modeller ve görselleştirir. Ayrıca veri tabanlarında daha karmaşık sorgulara cevap verebilmek için tasarlanmış sistemlerdir. Büyük veride yapılan işlemlerin tamamı OLAP üzerine kuruludur. Google'da yapılan aramalar bir OLAP sorgusudur. Yayıncılık sektöründe faaliyet gösteren bir işletmenin hizmet verdiği bölgelerde, farklı zaman dilimleri ve ürünler üzerinden

toplanmış ve gruplandırılmış satış rakamları OLAP'a örnek olarak verilebilir (Plattner, 2014: 17).

Tablo 1: OLTP ve OLAP Sistemleri Arasındaki Farklar

Özellik	OLTP	OLAP
Karakteristik	Operasyonel işleme	Bilgi işleme
Uyum	İşlem	Analiz
Kullanıcı	Memur, veri tabanı yöneticisi ve uzmanı	Profesyonel bilgi yöneticileri (müdür, yönetici, analist)
Fonksiyon	Günlük işlemler	Karar destek için uzun dönemli bilgi
Veri Tabanı Tasarımı	ER tabanlı, uygulama odaklı	Yıldız/Kar tanesi, karar odaklı
Veri	Güncel veriler	Geçmiş veriler
Özetleme	Basit, ayrıntılı	Özet, birleştirilmiş
Görüntüleme	Detaylı, yüzeysel ilişki	Özetlenmiş, çok boyutlu
İş Birim	Kısa, basit işlemler	Karmaşık sorgular
Erişim	Okuma/Yazma	Okuma
Odak	Veri girişleri	Bilgi çıktıları
İşleyiş	Birincil anahtar dizinleri	Birçok tarama
Erişilen Kaynak Sayısı	Binlerce	Milyonlarca
Kullanıcı Sayısı	Binlerce	Yüzlerce
Veri Tabanı Boyutu	Gigabaytlarca	>Terabayt
Öncelik	Yüksek performans ve kullanılabilirlik	Yüksek esneklik, son kullanıcı özerkliği
Ölçü	İşlem hacmi	Sorgu verimi, yanıt süresi

Kaynak: Chaudhuri, Surajit ve Dayal, Humeshwar, "An Overview of Data Warehousing and OLAP Technology", *ACM Sigmod Record*, Cilt: 26, Sayı: 1, Mayıs 1997, 65-74.

Tablo 1'de, OLTP ve OLAP sistemleri arasındaki farklar incelendiğinde; operasyonel veri tabanındaki veriler daha çok olmasına rağmen karar verme durumunda kullanılmak için yetersiz kalmaktadır. Karar desteği heterojen kaynaklardan gelen verilerin birleştirilmesini gerektirmektedir. Bu durum ise yüksek kaliteli ve temiz veriler sağlamaktadır. OLTP ve OLAP sistemleri farklı işlevler

sağladığından ve farklı türden veriler gerektirdiğinden, bu veri tabanlarının her birini farklı işlemlerde kullanmayı gerektirir (Han, Kamber ve Pei, 2001: 130).

1.3. Karar Destek Sistemleri

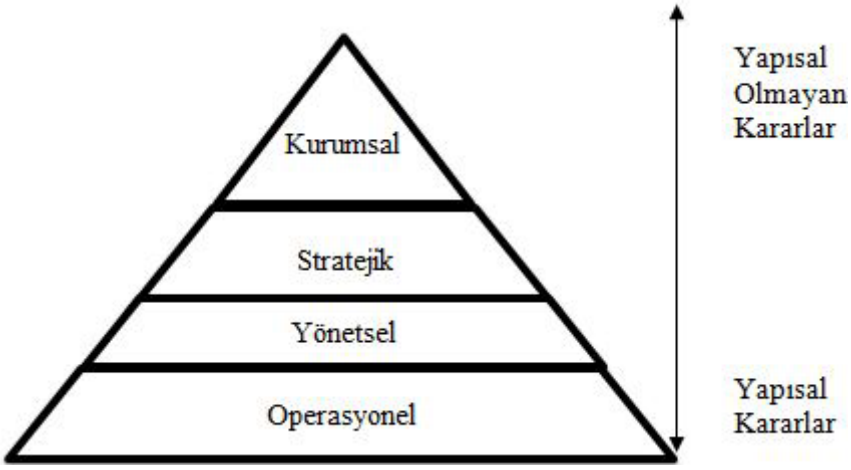
Geçmişten günümüze gelene kadar insanoğlu, hayatının her evresinde "karar verme" davranışı ile karşı karşıya kalmaktadır. Karar verme; günlük hayatta karşılaşılan problemler ve duyduğu istekler doğrultusunda kişinin kendi yararına olanı seçme eylemi olarak görülebilir.

Karar verme basit tanımıyla, mevcut seçenekler arasından birini seçme sürecidir (Köse, 2018: 6). Karar verme, belirlenen amaçlar ve sahip olunan bilgi çerçevesinde çeşitli seçenekler arasından yapılan bir tercihtir. Bu tercihler de zaman boyunca değişim göstermektedir (Sarr, 1974: 234). Zaman boyunca değişim gösteren bu kararlar, gün içerisinde hangi kıyafeti giyeceğini seçme olabileceği gibi bir fabrikada üretim sürecinde sınırlı kaynakların en uygun şekilde nasıl kullanılacağını belirlemek de olabilir. Mevcut bilginin yeterli olması, bilginin kalitesine, değişken sayısına ve karar verme aşamasında kullanılan modelin uygunluğuna bağlıdır ve bunlarda kararın kalitesini etkileyen unsurlardır. Buradaki önemli olan unsurlardan biri, karar vermede kullanılacak olan bilginin fazla olması değil amaca uygun olmasıdır (Demirci, 2012: 3).

Yönetimin farklı seviyelerinde alınan farklı karar tipleri bulunmaktadır. Bu karar yapıları üç ana nitelikte ele alınmaktadır. Bunlar; yapısal (structured), yarı-yapısal (semi-structured) ve yapısal olmayan (unstructured) kararlardır. **Yapısal kararlar**, programlanmış ya da tekrarlanan kararlar olarak da adlandırılmaktadır. Bu tip kararlar, yönetimin çevrim süreleri sonunda karşılaşılan sorunlar için tekrarlanan alışılagelen kararlardır. Mevcut üretimin değiştirilmesi kararı, satın alma ve üretim kararı, yenileme kararları yapısal kararlara örnek verilebilir. **Yarı-yapısal kararlar**, bazen önceden tanımlanmış bazen de yöneticinin kavramasına ve sezgisine dayanan kararlardır. Geliştirilecek ürün için üretim yeri ve işgücü planlanması bu karar türüne örnek verilebilir. **Yapısal olmayan kararlar**, programlanmamış kararlar olarak da bilinmektedir. Bu tip kararlar, belirli bir kurallar serisine bağlı olmayan ve gelişen

olayların özelliklerine göre şekillenen kararlardır. Örneğin; yeni ürün veya pazar kararları, tutundurma kararları ve finansal kararlar (Gökçen, Özkil, Yardımoğlu ve Peker, 2010: 2; Kuruüzüm, 1998: 9-10).

Yönetimin farklı karar düzeyleri ile farklı karar tipleri arasındaki ilişki, Şekil 2'de gösterilmektedir. Operasyonel yönetim seviyesi daha çok yapısal kararlarla kurumsal yönetim seviyesi ise yapısal olmayan kararlarla ilgilenmektedir.



Şekil 2: Yönetim Düzeylerine Göre Karar Türleri

Kaynak: Kuruüzüm, Ayşe, *Karar Destek Sistemlerinde Çok Amaçlı Yöntemler,* Akdeniz Üniversitesi Basımevi, 1998, 9.

Belirlenen amaçlar, doğan ihtiyaçlar, karşılaşılan sorunlar veya fırsatlar doğrultusunda verilen kararlar, karar vericileri istenen bir sonuca götürebileceği gibi çeşitli sorunlarla da karşı karşıya bırakabilmektedir. Bu sorunlar; karar vericinin değişen öncelikleri, durum karşısında oluşabilecek önyargısı, düşük olasılık tahmini gibi insani özelliklerinden kaynaklanabileceği gibi çevresel, sosyal ve ekonomik amaçların birbiriyle örtüşmemesiyle de ortaya çıkabilir (Demir ve Gümüşoğlu, 1994: 53). Bu noktada, oluşabilecek karmaşıklığı ortadan kaldırarak karar vericilerin daha etkin kararlar vermesinde onlara yardımcı olmak adına geliştirilen bilgisayar tabanlı bilgi sistemlerine **karar destek sistemleri-KDS** denilmektedir.

KDS kavramının ortaya çıkışı, bilgisayarların varoluşundan çok öncesine, yöneylem araştırması, davranışsal ve bilimsel yönetim teorileri ve istatistiksel süreç

kontrolünün ortaya çıktığı 1940'lı yıllara kadar dayandırılabilir (Singh, 2009: 581). Ancak KDS'nin ortaya çıkışı, Peter Keen ve Charles Stabell'in, Carneige Teknoloji Enstitüsü'nde örgütsel karar vermenin teorik dayanakları ve etkileşimli bilgisayar sistemleri üzerine yaptıkları teknik çalışmalar sonucunda olmuştur (Keen ve Scott-Morton, 1998). Scott-Morton (1971)'de yayımlanan kitabında da uygulamada model tabanlı KDS tanımlanmıştır.

KDS'nin ilk tanımlamalarında, yönetim ile ilgili yarı yapısal kararlarda karar vericileri desteklemeyi amaçlayan bir sistem olarak görüldüğü anlaşılmaktadır (Turban, Aronson ve Liang, 2007: 103). Little (2004: 1854), KDS'yi; yöneticiye karar vermesinde yardımcı olmak için veri ve kararların işlenebileceği model tabanlı süreçler kümesi olarak tanımlamaktadır. Bir başka tanıma göre KDS, yöneticilerin karar almada verimliliğini arttırmak yerine etkililiğini arttırmayı amaçlayan etkileşimli bilgisayar sistemleridir (Alagöz, Öge ve Koçyiğit, 2013: 34). KDS tanımının ve kapsamının, kullanılan amaç doğrultusunda yıllar geçtikçe farklılaştığını söyleyen Sol, Takkenberg ve VriesRobbé'un (1985), yıllara göre değişen KDS tanımlamaları şunlardır:

- 1970'lerin başında KDS, karar vermede yardımcı olan bilgisayar tabanlı sistemler,
- 1970'lerin ortalarında, karar vericilerin iyi yapılandırılmamış sorunlarını çözmek için veri tabanlarını ve modellerini kullanmalarında yardımcı olan etkileşimli bilgisayar tabanlı sistemler,
- 1970'lerin sonu ve 1980'lerin başlarında; yönetimsel ve mesleki faaliyetlerin etkinliğini arttırmak için uygun ve mevcut teknolojiyi kullanan sistemler,
- 1980'li yılların sonlarına doğru, uzman sistemler ve dokuman tabanlı sistemlerin ortaya çıkmasıyla akıllı iş istasyonlarının tasarımına yönelik kullanılan sistemlerdir (Sol, Takkenberg ve VriesRobbé, 1985: 1-2).

Etkileşimli bilgisayar destek sistemleriyle ilgili yürütülen teorik araştırmalar sonucunda ortaya çıkan KDS, 21. yüzyılda ticari işlemlerle ilgili verileri toplamak ve yöneticilerin işletme faaliyetlerini yürütürken bu verileri kullanarak kısa sürede doğru karar vermesini mümkün kılmaktadır (Averweg, 2009: 1753; Baysal, 2008: 4; Singh, 2009: 581).

Yapılan tanımlar incelendiğinde; KDS'nin ne olduğu hususunda ortak bir görüş olmadığı için standart nitelikleri ve yetenekleri konusunda net bir bilgiye de erişilememektedir. Turban ve Aronson (2001), temel KDS'nin nitelikleri ve yeteneklerini şu şekilde açıklamışlardır:

- KDS, bilgisayar yapay zekâsını ve insan yargısını bir araya getirerek, çoğunlukla yarı yapılandırılmış ve yapılandırılmamış durumlarda karar vericiler için destek sağlar.
- Üst düzey yöneticilerden bölüm yöneticilerine kadar tüm yönetim kademelerinde destek verir.
- Birey ve grup tabanlı karar desteği sağlar. Daha az yapılandırılmış sorunlarda, farklı bölümlerden ve örgütsel seviyelerden veya farklı kuruluşlardan bireylerin katıldığı gruplara destek verir.
- Birbirine bağımlı ve/veya birbirini izleyen kararlara destek sağlar. Kararlar bir kez, birkaç kez veya defalarca verilebilir.
- Karar verme sürecinin bilgi, tasarım, seçim ve uygulama aşamalarında destek sağlar.
- Çeşitli karar verme süreçleri ve tiplerinde destek sağlar.
- Zaman içinde değişen koşullara göre uyarlanabilir. KDS esnektir ve kullanıcılar temel öğeleri ekleyebilir, silebilir, birleştirebilir, değiştirebilir ve yeniden düzenleyebilir.
- Karar verme verimliliğinden (karar alma maliyeti) ziyade etkililiğinin (doğruluk, zamanındalık, kalite) iyileştirilmesine odaklanır.

- Karar verme sürecinin tüm adımlarında, karar verici, tam kontrole sahiptir. KDS, karar vericinin yerine geçmeyi değil, onu desteklemeyi amaçlar.
- Bilgi Sistemi uzmanlarının yardımıyla daha büyük sitemler oluşturulabilir. Veri ambarlarıyla birlikte OLAP yazılımı, kullanıcıların daha büyük ve karmaşık KDS süreçlerine olanak tanır.
- KDS'de modeller, farklı koşullar altında farklı stratejilerle karar verme durumlarının denemelerini yapar.
- Çeşitli veri kaynaklarına, formatlarına ve türlerine erişim sağlar.
- KDS, diğer uygulamalara entegre edilebilir, ağ ve Web teknolojileri kullanılarak dahili ve harici olarak dağıtılabilir (Turban ve Aronson, 2001: 108).

KDS'nin geniş bir uygulama alanı bulunmaktadır ve teorik olarak herhangi bir alanda geliştirilebilir. Örneğin; sağlık sektöründe klinik ve laboratuvar uygulamalarında, bankacılık ve finans sektöründe çeşitli risk uygulamalarında, askeri uygulamalar gibi birçok alanda KDS uygulamalarıyla karşılaşmak mümkündür (http://mis.sadievrenseker.com/). KDS uygulamalarının bir alt yapı oluşturularak gerçekleştirilebilmesi için veri ambarına ihtiyacı vardır.

1.4. Veri Ambarı

İşletmeler; veri tabanlarını kullanarak ödeme yapma, siparişleri işleme koyma ve müşterilere hizmet verme gibi temel işlemleri takip ederler. Bunun yanı sıra işletmeler, faaliyetlerini daha verimli yönetilmesine ve yöneticilerin ve çalışanların daha iyi kararlar almasına yardımcı olacak veri tabanlarına da ihtiyaç duyarlar. Örneğin; işletmeler, planlama stratejilerini oluştururken piyasada hangi ürününün daha çok satıldığını veya en kar getiren müşterisinin kim olduğunu bilmek ister. Özellikle hacim olarak çok büyük miktarlarda veriyle çalışan işletmeler, karar alma noktasında, tek merkezden veri yönetimi kontrolünü sağlamak isterler (Laudon ve Laudon, 2007: 162).

Veri depolama, firma yöneticilerine stratejik kararlar alma noktasında, verilerini sistematik olarak organize etme, anlama ve kullanma konusunda mimari yapı ve araçlar sağlamaktadır. Veri ambarı sistemleri, günümüzün rekabetçi ve hızla gelişen dünyasında değerli birer araçtır. Her sektörde rekabetin artmasıyla birçok işletme, veri depolamanın, müşteri ihtiyaçları hakkında daha fazla şey öğrenerek bilgi edinebilmenin bir yolu ve sahip olması gereken en son pazarlama silahı olduğunu düşünmektedir (Han, Kamber ve Pei, 2011: 126).

Yöneticilerin ve çalışanların örgütsel performansını iyileştiren ve daha iyi kararlar almasına yardımcı olmak için geliştirilen, veri tabanlarında saklanan verilerin kontrolünü sağlayan mimari modele, **veri ambarı** (data warehouse) denir. Yaygın olarak kullanılmakta olan veri ambarları, gün aşırı işlem hareketlerinin yer aldığı ve işletimsel düzeyde rutin işlemlerin gerçekleştirildiği veri tabanlarının birleştirilmiş ve işlemeye uygun hale getirilmiş türüdür.

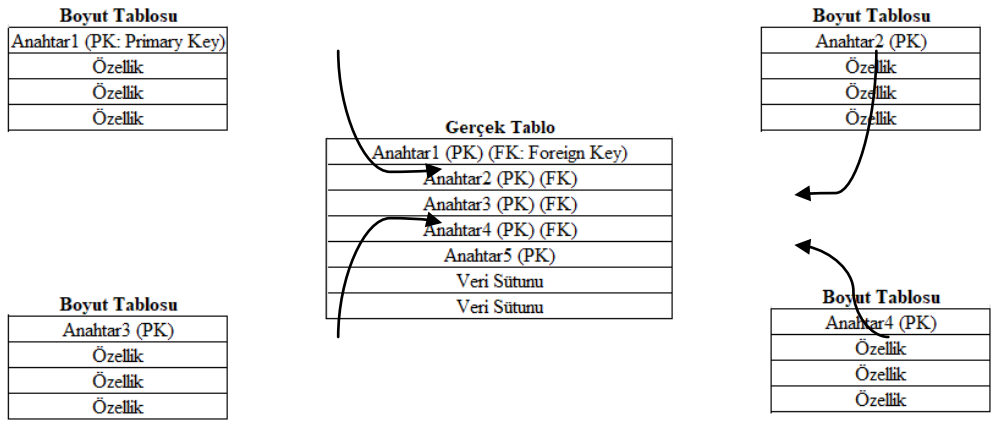
Veri ambarı; satış, stok ve maliyet gibi kurum içi verilerle, pazar raporları, yasal düzenleme şartları ve müşteri memnuniyeti gibi kurum dışı verilerin bütünleştirilip merkez havuzda birleştirilerek anlık sorgulamalar için hazırda bekletildiği bir ortamdır (Prytherch, 2005: 196). Bu sayede, işletmelerin belirlenen amaçlar doğrultusunda, ihtiyaç duyulan verilere hızlı bir şekilde erişimi sağlamaktadır.

Veri ambarı sisteminin mimarı olan William H. Inmon'a göre; veri ambarı, yönetimin kararlarını desteklemek için konu odaklı, bütünleşik, statik ve zamana bağlı bir veri koleksiyonudur (Inmon, 2002: 31). Konu odaklı, bütünleşik, statik ve zamana bağlı olması veri ambarının temel özellikleridir ve bu özellikler veri ambarını diğer depolama sistemlerinden ayırmaktadır.

Veri ambarının konu odaklı olması; işletmelerde gerçekleşen günlük işlem hareketlerinden ziyade müşteri, ürün veya satış odaklı özelleştirmelerle veri analizi ve modellemeler yaparak karar vericilere bilgi sağladığı anlamına gelmektedir. Veri ambarının bütünleşik olması; heterojen kaynaklardan sağlanan tutarsız verilerin veri ambarlarında birleştirilerek tek formatta tutulduğunu ifade etmektedir. Veri

ambarının statik olması; hata oluşmadığı sürece veri ambarındaki veriler üzerinde herhangi bir düzenleme yapılamayacağını, veri ambarlarında sadece okuma işlemi yapıldığını ve veri tabanlarından veri ambarlarına tek yönlü bir akışın olduğunu göstermektedir. Veri ambarının zamana bağlı olması ise karşılaştırma ve tahminlerde kullanılacak sorguların veri ambarlarında depolanan işletmenin geçmiş verileri üzerinden yapıldığını göstermektedir.

Veri ambarlarında çok boyutlu analiz sağlayan OLAP araçları, üç boyuttan oluşan küpler şeklinde düzenlenmiştir. Her bir boyut, bir dizi küme seviyesinden oluşmaktadır. Klasik OLAP işlemleri, bir boyut boyunca verinin toplanmasını veya birleştirilmesini, küpün belli kısımlarının seçilmesini ve verilerin çok boyutlu görünümüne dönüşümünü içermektedir (Vassiliadis ve Sellis, 1999: 64). Veri küpleri oluşturmak için en çok yıldız şeması kullanılmaktadır.



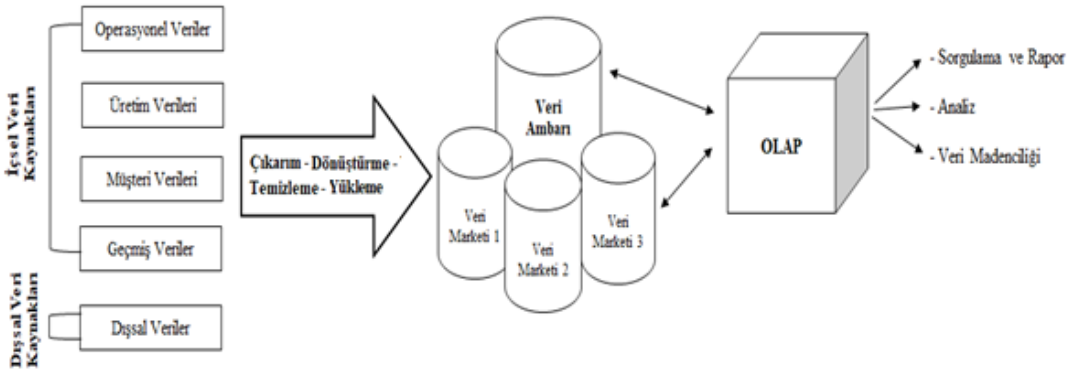
Şekil 3: Yıldız Şeması Tablosu

Kaynak: Şeker, Şadi E., "Veri Ambarı (Data Warehouse)", *YBS Ansiklopedisi*, Cilt: 2, Sayı: 4, Aralık 2015, 13.

Şekil 3'teki yıldız şeması, bir merkezi gerçek tablo ve birkaç denormalize edilmiş boyut tablolarından oluşmaktadır. OLAP'a ait ilgili ölçütler gerçek tabloda depolanır. Çok boyutlu modelin her bir boyutu için yapılan birleştirmeler ve özellikleri de boyut tablosunda saklanmaktadır.

1.4.1. Veri Ambarı İşleyişi

Veri ambarının temeli olan çok boyutlu veri modelinde işletmelerin mevcut verileri, farklı boyutlar ve bu boyutlara ait ölçüm değerlerine göre depolanmaktadır. Veri ambarının belli bir amaç için özelleştirilmiş ve boyut olarak küçültülmüş haline **veri marketi** (data mart) denilmektedir (Arslan ve Yılmaz, 2010: 77). Veri ambarı, birden fazla veri marketinin bir araya gelmesiyle oluşur. İşletmeler, veri marketlerindeki enformasyonu; müşteri satın alma modelleri oluştururken zaman dilimi ve bölgelere göre satış performanslarını karşılaştırarak üretim stratejilerini belirlerken ve kar payını arttıracak kaynaklar ararken kullanmaktadırlar. Veri ambarındaki enformasyon, karar vericiler tarafından özümsenerek birer işletme bilgisine dönüşmektedir (Han, Kamber ve Pei, 2011: 127).



Şekil 4: Veri Ambarı İşleyişi

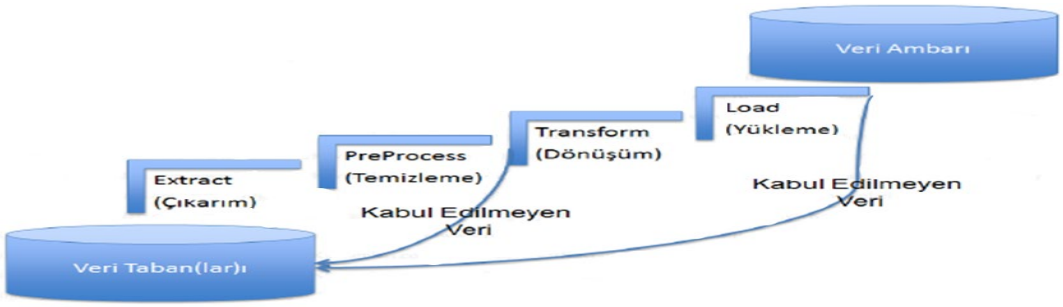
Kaynak: Laudon, Kenneth C. ve Laudon, Jane P., *Essentials of Business Information Systems,* 7. Baskı, Pearson Education, 2007, 163.

Şekil 3'te yer alan veri ambarındaki işleyiş sürecinde, işletme verileri, operasyonel, müşteri, üretim, geçmiş ve dış kaynaklı olmak üzere birçok temel işlemsel sistemden oluşmaktadır. Farklı veri tabanlarından gelen veriler temizleme, çıkarım, dönüştürme ve yükleme işlemleri uygulanarak standartlaştırılır ve veri ambarlarında depolanır. Veri ambarı veya veri marketlerinde saklanan veriler, OLAP sistemleri kullanılarak çok boyutlu analiz ve sorgulama sürecini takiben işletmelerin karar alma süreçlerinde kullanılabileceği bir forma dönüştürebilmek için çıkarım, dönüşüm ve yükleme (Extract, Transform, Load-ETL) süreçlerinden geçirilmektedir.

1.4.2. ETL Süreçleri

Veri ambarlama sürecinde işlenen veriler üç kısımda incelenmektedir. Bunlar; operasyonel, özet ve meta verilerdir. **Operasyonel veriler,** ana işlemlerde kullanılan verilerdir. **Özet veri,** soyutlama sürecinin ilk basamağı olup sürekli güncellenmeye ihtiyacı olan veri türüdür. Özet veri aynı zamanda veri tabanı şeması ile veriye özgü altyapıyı hazırlamaktadır. **Meta veri**, veri hakkında veri anlamına gelmektedir ve veri ambarının önemli bileşenlerinden birini oluşturup yönetim ve KDS için kullanılmaktadır (Biçen, 2002: 21-22).

Veri tabanlarındaki verilerin, veri ambarlarına taşınması esnasında bazı sorunlarla karşılaşılmaktadır. Örneğin; cinsiyet değişkeni bir veri tabanında "erkek veya kadın" şeklinde, bir diğer veri tabanında "E veya K", başka bir veri tabanında ise "bay veya bayan" girilmiş olabileceği gibi tekrar eden veriler ve boş bırakılan alanlar da olabilir. Bu gibi durumlar, taşınma sürecinde karışıklığa yol açacaktır. Bu sorunların önüne geçebilmek için ETL işlemleri uygulanmaktadır.



Şekil 5: ETL Süreci

Kaynak: Şeker, Şadi E., "Veri Ambarı (Data Warehouse)", *YBS Ansiklopedisi*, Cilt: 2, Sayı: 4, Aralık 2015, 9.

ETL, farklı veri tabanlarından verileri toplayan, ortak kurumsal bir tanım kümesi kullanarak verileri dönüştüren ve ardından veri ambarına yükleyen işlemdir (Baltzan, 2014: 228). ETL aynı zamanda analitik sistemin işlemsel sistemle senkronize edilmesidir (Plattner, 2014: 18). Şekil 5'te gösterilen ETL süreci üç aşamadan oluşmaktadır. Bu aşamalar, şu şekilde açıklanabilir:

- **Çıkarım (Extract):** Veri ambarlarına aktarılacak verilerin seçim sürecidir. Hangi verilerin seçileceğine belirleyen kriterler; doğruluk, bütünlük, zamanlılık, örnekleme ve sadeliktir. Çıkarım işlemi, veri tekrarlarını ve bütün verilerin depolanmasından doğacak olan kaynak, para ve zaman maliyetlerini önlemeyi sağlar.
- **Dönüşüm (Transform):** ETL işlemleri içinde en uzun süreli aşamadır. Farklı kaynaklar, formatlar ve biçimlerdeki veriler standardize edilir. Bu aşamadaki en önemli işlem, temizleme işlemidir. Yazım ve tarih hataları, eksik ve tekrarlanan veriler bu aşamada düzeltilmektedir.
- **Yükleme (Load):** Çıkarım ve dönüştürme işlemlerinden sonra verilerin, veri ambarlarına yüklenmesi aşamasıdır. Buradaki en önemli husus, yükleme aşamasında sistem performansının etkilenmemesidir (Şeker, 2015: 9; Ateş, 2008: 58-59).

ETL süreci, iş zekâsı çalışmalarında en çok zamanın harcandığı ve maliyetlerin yapıldığı aşamadır. Burns'e (2005) göre, toplam maliyetin %50'si ve zamanın da %70'i ETL sürecinde harcanmaktadır. Maliyet ve süre artışlarındaki nedenler; veri tanımlama ve çıkarma işlemlerinin karmaşıklığı, dönüştürme işleminde basit olmayan yöntemlerin kullanılması, veri kaynaklarının tutarsız oluşu ve verinin değişken yapısı ile ilişkilendirilebilir (Utley, 2008: 44). ETL sürecinde oluşabilecek herhangi bir hata, bu veriler üzerinden alınacak kararları etkileyeceğinden, süreç doğru bir şekilde planlanmalı ve tasarlanmalıdır.

1.5. Veri Ambarı ve Veri Madenciliği İlişkisi

Veri ambarlarında, farklı kaynaklardan gelen ve birtakım işlemden geçen ayrıntılı veriler, veri marketlerinde ise belli amaçlar için özelleştirilen veriler depolanmaktadır. Veri ambarındaki veriler, geçmişe ait ve ayrıntılı veriler kullanılarak bilinmeyenin arayışında olan veri madenciliği için çok önemli bir veri kaynağıdır. Bu sebeple veri ambarlarında bulunan ayrıntılı veriler, veri marketlerindeki özelleştirilmiş verilere göre bilgi keşfi sürecinde daha fazla tercih edilmektedir. Veri ambarındaki veriler; temizlenmiş, bütünleştirilmiş ve organize

edilmiştir. Aynı zamanda arşiv diye nitelendirilen geçmiş yılların verilerini de içermektedir. Veri ambarı, bilgi keşfi ve veri madenciliği sürecinde mükemmel bir kaynağı sağlamaktadır; ama bu süreçte dış kaynaklı diğer veriler de kullanılabilmektedir. Bu yüzden, veri ambarları, veri madenciliği sürecinde kullanılan tek kaynak olmamasına rağmen veri madenciliğinde, veri ambarlarının kullanılmasıyla süreç iyileşir ve başarı oranı artmaktadır. Tüm bunlar veri madenciliği sürecini başlatmak için ihtiyaç duyulan unsurlardır (Inmon, 2002: 53).

İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

Hesaplama işleminde birer kılavuz niteliğinde olan araçlar, tarihsel süreçte insanlığa her zaman yardımcı olmuştur. Parmaklar, eller, ayaklar, taşlar ve çubuklar kullanılarak yapılan hesaplama işlemleri yerlerini toplum arasında yürütülen alım satım etkinliklerinin de artmasıyla abaküse bırakmıştır. İlkel toplumlarda takas usulüyle yürütülen alım satım etkinlikleri, zaman içinde şekil değiştirerek, ticaret adı verilen kavramın temellerini oluşturmuştur. Ticaretin gelişmesi ve buna bağlı olarak satış işlemlerinin ivme kazanmasıyla insanlar, dört işlemden daha fazlasını yapabilen yeni sayı sayma sistemlerine ihtiyaç duymuşlar, 19. yüzyılda elektromekanik hesap makinesini takiben 20. yüzyılda bilgisayarların ortaya çıkmasıyla birlikte hesaplama işlemleri bugünkü pratikliğine ulaşmıştır.

Bu bölümde, veri madenciliğinin tanımı ve gelişim sürecinden bahsedilecektir. Veri madenciliği süreci ve modelleri, veri madenciliğinin işletmelerde ve kütüphanelerdeki uygulama alanları anlatılarak veri madenciliğinde karşılaşılan sorunlar üzerinde durulacaktır.

2.1. Veri Madenciliğinin Tanımı ve Gelişimi

İnternetin ortaya çıkışıyla birlikte büyük bir ivme kazanan dijitalleşme veya verilenme sürecinde, her gün üretilen veri miktarı katlanarak çoğalmaktadır. Örneğin; arama motorlarında yapılan her sorgulamada veri stokuna eklemeler yapılmaktadır. Yaklaşık 7,7 milyar kişi olan küresel nüfusta, 4 milyar internet kullanıcısı ve 3,7 milyar aktif sosyal medya kullanıcısı vardır. Günde yaklaşık olarak gönderilen e-posta sayısı 150 milyar, yazılan blog yazısı 3 milyon, atılan tweet sayısı 435 milyon, Google'da yapılan arama sayısı 4 milyardır (http://www.worldometers.info/). İnternette yapılan aramaların %77'sinin Google üzerinden gerçekleştirildiği varsayımıyla dünyada, günde yaklaşık 6 milyar kişi

internette arama yapmakta ve yapılan aramaların ise %60'ı mobil cihazlar kullanılarak gerçekleştirilmektedir (https://www.brandwatch.com/).

Dijital ortamda veri kaynakları ve hacminde yaşanan artışla beraber verinin, erişilebilirliği kolaylaşmış ve etki gücü artmıştır. Hızla çoğalan veri, pazarlamadan finansa, eğitimden sağlığa, insan ilişkilerinden iletişime hemen her alanda etkisini göstermektedir. İşletmeler de veri güdümlü artan rakipleriyle mücadele edebilmek için verileri, etkili bir şekilde kullanabilmek ve değişen rekabet koşullarında; daha iyi karar verme, iş becerisi ve yenilikçiliğe geçiş sürecinde rakiplerine karşı avantaj sağlayabilmek için yeni yöntem bulma arayışına yönelmişlerdir (Davenport ve Bean, 2018: 4).

Veri madenciliği; öngörülü bir yaklaşımla karar verme sürecinde, işletme yöneticilerine farklı bakış açıları kazandırmak ve stratejik rekabet üstünlüğü sağlamak amacıyla veri tabanlarındaki büyük miktarda veriyi kullanarak işletme için faydalı bilgiyi elde etmekte kullanılan bir işletme zekası yöntemidir (Çelik, Akçetin ve Gök, 2017: 36; Shmueli, Bruce, Yahav, Patel ve Lichtendahl Jr., 2018: 5).

Küresel bir araştırma ve danışma firması olan Gartner Group'a göre veri madenciliği; veri havuzlarında depolanan büyük miktardaki veriyi eleyerek anlamlı korelasyonları, örüntüleri ve eğilimleri keşfetme sürecidir. Veri madenciliği örüntü tanımlamanın yanı sıra istatistiksel ve matematiksel teknikleri de kullanmaktadır (https://www.gartner.com/).

Shumuli ve diğerleri (2018), veri madenciliğini; sayıların ötesine geçen işletme zekâsı yöntemlerini, tanımlayıcı teknikleri, raporlama ve iş kurallarına dayalı yöntemlerin bütünü olarak ifade etmektedir.

Bilgi keşfi sürecinin bir parçası olan veri madenciliğinin tarihçesi, Tablo 2'de dört dönem ve bu dönemde gerçekleşen olaylar çerçevesinde incelenmiştir. Bu dört dönem; istatistik, bilgisayar çağı, veri madenciliği ve günümüzdür.

Tablo 2: Veri Madenciliğinin Tarihçesi

	Olaylar	Yıl
İSTATİSTİK	Thomas Bayes Paper	1763
	Regresyon Analizi	1805
BİLGİSAYAR ÇAĞI	Evrensel Makine	1936
	Sinir Ağları	1943
	Evrimsel Hesaplama	1965
	Veri Tabanları	1970'ler
	Genetik Algoritma	1975
VERİ MADENCİLİĞİ	Veri Tabanlarından Bilgi Keşfi	1989
	Destek Vektör Makinesi	1992
	Veri Bilimi	2001
	Moneyball	2003
GÜNÜMÜZ	Büyük Veri	2015-…

Kaynak:https://dataconomy.com/2016/06/history-data-mining/?fbclid=IwAR2PF2ZC38P9bvwR65XnyrM6G3ywWSaUfp0k3tXr4PuOo6Jd5658VRl9k2g (23.10.2018).

Veri madenciliği sürecinde, büyük bir veri havuzu içinden çıkarımlarda bulunmak için istatistiksel bir model kullanılmaktadır. Tahmin edilen olasılıklara

dayanan karmaşık gerçekliklerin anlaşılmasına izin verdiği için veri madenciliğinin gelişim sürecinde Bayes Teoremi'nin yeri oldukça önemlidir. 1805 yılında, Adrien-Marie Legendre ve Carl Friedrich Gauss'ın Güneş yörüngesindeki gök cisimlerinin konumunu tahmin etmede kullandığı regresyon analizi, veri madenciliğinin kilit noktalarından birisi olarak kabul edilmektedir.

Büyük miktardaki verinin toplanmasını ve işlenmesini mümkün kılan ve bilgisayarının temellerini oluşturan Alan Turing'in evrensel makinesi, sinir ağındaki nöronların aktivitesine ilişkin kavramsal bir model oluşturulması, gerçek dünya problemlerinin çözümünde kullanılmak için geliştirilen evrimsel hesaplama, daha fazla verinin depolaması ve sorgulaması için geliştirilen veri tabanı yönetim sistemleri ve genetik algoritmalar üzerine yapılan çalışmalar veri madenciliğinin gelişimde etkili olmuştur.

1989 yılında, veri tabanlarında bilgi keşfi, veri içindeki potansiyel ve yararlı kalıpları tanımlamasıyla veri madenciliğindeki en önemli gelişim adımı olmuştur. 1990'lı yıllarda, perakende şirketleri ve finans toplulukları tarafından, müşteri tabanını genişletmek ve faiz oranı dalgalanmalarını, hisse senedi fiyatlarını ve müşteri talebini tahmin etmek için veri madenciliği kullanılmaya başlanmıştır. Verileri analiz eden ve sınıflandırma ve regresyon analizi için kullanılan kalıpları tanımlayan, destek vektör makineleri geliştirilmiş ve 2001 yılında veri biliminde yaşanan gelişmelerle birlikte 2003 yılında, Oakland Atletizm takımının oyuncu kalitelerini belirlemek için istatistiksel veri odaklı bir yaklaşım kullandığı Moneyball isimli kitap, büyük bir etki yaratmıştır.

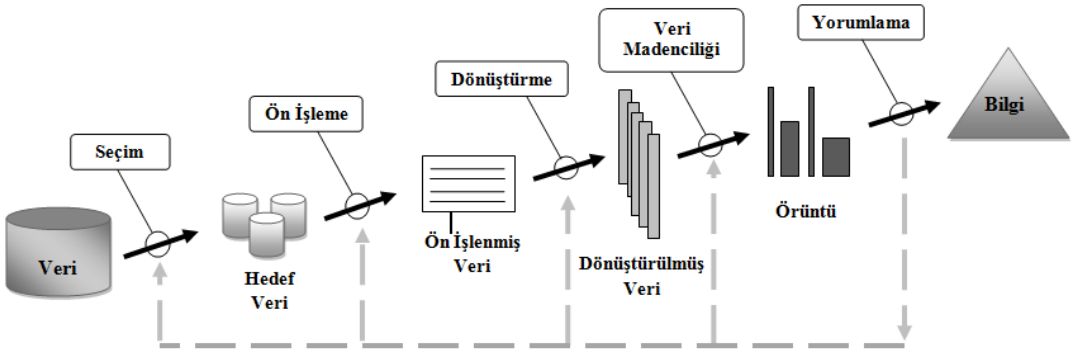
1770'lerden günümüze dek gelişen veri madenciliği, verinin olduğu her alanda kullanım imkânı bulmaktadır. Teknolojik ilerlemeler ve tüketicilerin artan veri kullanımıyla birlikte **büyük veri** olarak adlandırılmaya başlamış ve büyük miktardaki verinin etkin kullanılabilmesi için çeşitli veri madenciliği araçları geliştirilmiştir.

2.2. Veri Madenciliği Süreci

İnternet vasıtasıyla gerçekleştirilen her eylemle birlikte (online alışveriş, e-mail gönderme, sosyal medya paylaşımları vb.) dijital ortamda üretilen veri miktarındaki artış katlanarak çoğalmaktadır. Veri miktarındaki süregiden bu artış, işletmelerin faaliyet biçimlerinde köklü bir değişime neden olmuştur. İş dünyasında yaşanan bu değişimle birlikte veri, işletmeleri rekabet ortamında tutacak hayati ve stratejik bir unsur olarak görülmeye başlanmıştır (Marr, 2018: 9-10). İşletmelerin veriyi birer iş değerine dönüştürmesi ve faaliyet biçimlerini iyileştirmesi için birtakım kuram ve araçlara gereksinimi vardır (Marr, 2018: 30; Karagöz, 2007: 9).

Tarihsel olarak literatürde, veri içerisindeki faydalı yapıların ortaya çıkarılması olayına pek çok isim karşılık gelmiştir. Bunlar; veri madenciliği, bilgi çıkarımı, bilgi keşfi, bilgi harmanlama, veri arkeolojisi ve veri modelleme süreci (Fayyad, Piatetsky-Shapiro ve Smyth, 1996: 39). Veri tabanlarında bilgi keşfi-VTBK olarak da adlandırılan, **bilgi keşfi süreci**, veri tabanlarında tutulan büyük hacimli verilerin veri madenciliği teknikleri kullanılarak işlenmesidir. Diğer bir deyişle verilerdeki mevcut, eşi benzeri olmayan, potansiyel olarak kullanışlı ve anlaşılabilir örüntüleri tanımlamanın kolay olmayan bir süreci olarak tanımlanmaktadır (Cios, Pedrycz, Swiniarski ve Kurgan, 2007: 10). Çoğu kez birbirinin yerine kullanılan VTBK'nin ve veri madenciliği kavramları arasındaki fark; VTBK, veriden bilgi keşfi sürecinin tamamını kapsarken, veri madenciliğinin bu sürecin bir basamağı olmasıdır (Köktürk, Ankaralı ve Sümbüloğlu, 2009: 21).

Fayyad ve diğerlerine (1996) göre, VTBK sürecini oluşturan beş adım Şekil 6'da gösterilmektedir.

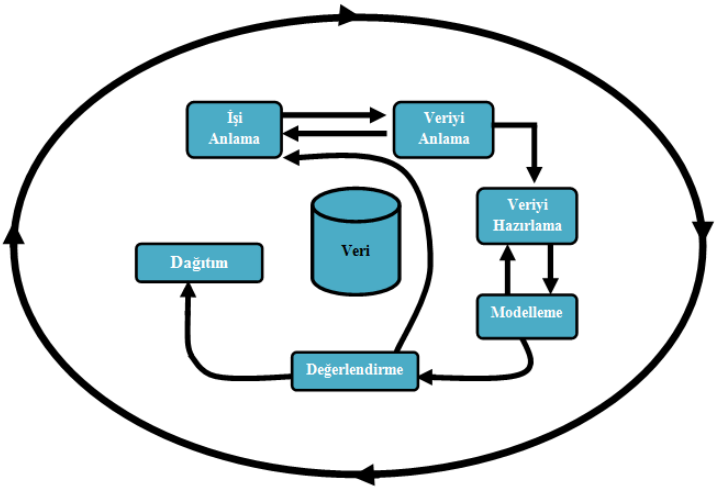


Şekil 6: Veri Tabanlarında Bilgi Keşfi Sürecini Oluşturan Adımlar

Kaynak: Fayyad Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic, "From Data Mining to Discovery Knowledge in Databases", *AI Magazine*, Cilt: 17, Sayı: 3, 1996, 41.

VTBK sürecine başlamadan; uygulama alanı, konu ile ilgili ön bilgi ve son kullanıcı hedefleri belirlenerek bir anlayış geliştirilmelidir. Buna takiben, sürecin ilk aşaması olan seçim adımına geçilmelidir. Seçim adımında, önemli olduğu öngörülen ve üzerinde keşif gerçekleştirilecek olan veri örnekleri seçilerek bir hedef veri seti oluşturulur. Ön işleme adımında, veri madenciliği uygulamasında kullanılacak olan hedef veri setindeki verilerin uygunluğunu ve kalitesini arttırmak için eksik verilerin tamamlanması, aykırı verilerin temizlenmesi ve tutarsız verilerin düzeltilmesi gereklidir. Dönüştürme adımında, farklı veri madenciliği metotlarında kullanılabilecek formatta veriler oluşturulur. Veri madenciliği aşamasında, sınırlanan amaçlar doğrultusunda seçilen veri madenciliği yöntemleriyle çeşitli örüntüler oluşturulur. Yorumlama aşamasında, veri madenciliği aşamasında çıkarılan örüntülerin yeterliliği, sınanıp yeterli bulunmadığı takdirde VTBK süreci tekrarlanmaktadır (Akpınar, 2017: 76-77; Fayyad vd., 1996: 42).

Şekil 6'da gösterilen veri tabanlarında bilgi keşfi sürecini oluşturan beş adım, VTBK ve veri madenciliği konusunda, verinin bilgiye dönüşümü noktasında geliştirilen birçok süreç bulunmaktadır. Bu süreç modellerden biri de 1996 yılı sonlarında, gelişmekte olan veri madenciliği pazarının dört büyük lideri olan Daimler-Benz, Integral Solutions Ltd. (ISL), NCR ve OHRA tarafından tasarlanan CRoss-Industry Standard Process for Data Mining'dir (CRISP-DM).



Şekil 7: CRISP-DM Süreç Diyagramı

Kaynak: Data Mining Process, http://crisp-dm.eu/reference-model/, (13.06.2019).

Veri madenciliği projesini yürütmek için kapsamlı bir yöntem ve süreç modeli olan CRISP-DM, veri madenciliği proje yaşam döngüsünü altı aşamada incelemektedir. Bunlar; işi anlama, veriyi anlama, veriyi hazırlama, modelleme, değerlendirme ve dağıtım aşamasıdır.

İşi anlama: CRISP-DM süreç modelinin ilk aşaması olan işi anlama, proje hedefi ve gereksinimlerini iş perspektifinden anlamaya ve daha sonra bu bilgiyi veri madenciliği problem tanımlama ve hedefe ulaşabilmek için tasarlanmış bir ön plana dönüştürmek için odaklanmaktadır.

Veriyi anlama: İşi anlama aşamasında belirlenen hedeflere paralel olarak ilgili verilerin toplandığı aşamadır. Veri kalitesi problemlerini tanımlamak, verileri ilişkin ilk izlenimleri keşfetmek ve veri içinde gizlenen enformasyonu ortaya çıkaracak hipotezler oluşturulmaktadır.

Veriyi hazırlama: Başlangıçtaki ham verileri, veri madenciliği modelleme araçlarında kullanılacak olan hedef veri setini dönüştürmek için yapılacak tüm faaliyetleri kapsamaktadır. Bunlar; tablo, kayıt ve öznitelik seçimi ile modelleme araçlarında kullanılması için verilerin dönüştürülmesi ve temizlenmesi işlemidir. Bu

işlemlerin önceden belirlenmiş bir sırası olmayıp birden çok defa tekrarlanabilmektedir.

Modelleme: Çeşitli modelleme teknikleri seçilir, test tasarımı ve modeller oluşturulur, modelleme teknikleri parametreleri en uygun değere ayarlanır, uygulanır ve modellerin değerlendirilmesi yapılmaktadır.

Değerlendirme: Bu aşamada, veri analisti tarafından oluşturulan modelin dağıtıma geçmeden önce modeli ayrıntılı bir şekilde değerlendirmek ve iş hedeflerine uygunluğundan emin olmak için modelin yapısını incelemek önemlidir. Proje lideri, bu aşamanın sonunda, veri madenciliği sonuçlarının nasıl kullanılacağına karar vermelidir. Bu aşamadaki kilit noktalar; sonuçların değerlendirilmesi, süreç incelemesi ve sonraki adımların belirlenmesidir.

Dağıtım: İşletme gereksinimlerine göre kolaylığı veya zorluğu değişebilen dağıtım aşamasında önemli nokta, süreç sonunda elde edilen bilginin, son kullanıcıların anlayabileceği şekilde düzenlenmesi ve sunulmasıdır. Dağıtım aşamasındaki kilit noktalar; dağıtım planı oluşturulmalı, oluşturulan dağıtım planı izlenmeli ve bakım işlemleri yapılmalı, nihai rapor hazırlanmalı ve proje gözden geçirilmelidir.

Şekil 7'de gösterilen altı aşamalı CRISP-DM veri madenciliği süreç modelinde; oklar, süreç aşamaları arasındaki en önemli ve en sık görülen bağlılıkları belirtirken dıştaki daire, veri madenciliğinin döngüsel niteliğini sembolize etmekte ve veri madenciliği sürecinde ve konuşlandırılmış çözümden öğrenilen derslerin yeni, genellikle daha odaklı iş sorunlarını tetikleyebileceğini göstermektedir.

2.3. Veri Madenciliği Modelleri

Uygulamada, veri madenciliğinin tahmin ve tanımlama olmak üzere iki ana hedefi vardır. Veri madenciliği; veri kümesinde tanımlı, konuya ilişkin değişkenler ve alanlar kullanılarak ilgili değişkenlerdeki bilinmeyen veya gelecekte alabileceği değerleri tahmin etmek ve karar vericiler tarafından yorumlanabilecek mevcut veri kümesini tanımlayan kalıpları bulmaya odaklanmaktadır (Kantardzic, 2011: 2).

Veri madenciliğinin ana hedefleri olan tahmin ve tanımlama, aynı zamanda veri madenciliği modellerini sınıflandırmada da kullanılmaktadır. Tahmin edici modellerde, sonuçları bilinen veri dizisinden hareket edilerek model geliştirilmekte ve bu model, sonuçları bilinmeyen bir veri kümesine uygulanarak sonuç değerleri tahmin edilmeye çalışılmaktadır. Tanımlayıcı modellerde ise karar vermede yol gösterici olarak kullanılabilecek veri dizisindeki saklı olan örüntüler ortaya çıkarılmaktadır (Çelik vd., 2017: 21).

2.3.1. Tahmin Edici Modeller

Tahmin edici modellerde, açıklayıcı değişken üzerinden hedef değişken için bir model oluşturulur. Tahmin edici modellerin amacı, hedef değişkenin tahmini ve gerçek değerleri arasında oluşabilecek hata payını en aza indiren bir model oluşturmaktır. Tahmin edici modeller; gelecekte pazarda oluşabilecek riskleri belirleyebilir, aynı zamanda dünyanın kaynak tüketimini de tahmin edilebilmektedir (Weiping ve Yuming, 2013: 158). Sınıflandırma ve regresyon analizi olmak üzere iki tahmin edici model vardır.

2.3.1.1. Sınıflandırma

Sınıflandırma, veri setindeki yeni bir nesnenin niteliklerini incelemek ve bu nesneyi önceden tanımlanmış bir sınıfa atamak veya ilgilenilen sayısal değişkenin değerini tahmin etmek için verilerdeki özellikleri kullanmaktadır (Saranya, Rathnavathy ve SureshBabu, 2014: 263). Bir başka ifade ile sınıflandırma, veri setindeki nesnelerin özniteliklerinin benzerliğine göre ayrıştırılmasıdır.

Bilgi keşfi uygulamalarının bir parçası olarak kullanılan sınıflandırma, veri madenciliği teknikleri arasında en yaygın kullanılanlardan biridir (Weiss ve Indurkhya, 1998: 34). Sınıflandırma tekniğinde girdi, veri kümesinde ayrık konumlanan ve farklı özniteliklerden oluşan nesnelerdir. Nesnelerin sahip olduğu öznitelikler kullanılarak sınıf niteliğini belirlemek için bir sınıflandırıcı model oluşturulur. Oluşturulan model kullanılarak da sınıf etiketi belirlenmemiş nesneler en iyi şekilde doğru sınıflara atanmaktadır.

Sınıfları doğrusal bir karar sınırı kullanarak mükemmel bir şekilde ayırmak mümkün değildir. Sınıflandırma analizinin performansı, örnek olaylarda yapılan hataların sayısı takip edilerek ölçülmektedir.

Sınıflandırma modeli, veri madenciliğinde iki önemli göreve hizmet etmektedir. İlk olarak, daha önce etiketlenmemiş nesneleri sınıflandırmak için öngörücü bir model olarak kullanılır. İyi bir sınıflandırma modeli, hızlı tepki gücüyle doğru tahminler sağlamalıdır. İkincisi, nesneleri farklı sınıflardan ayıran öznitelikleri tanımlamak için tanımlayıcı bir model görevi görür. Bu durum özellikle, sağlık sektöründe tıbbi teşhisler gibi kritik uygulamalarda önem arz etmektedir.

Büyük miktardaki veri kümesiyle çalışabilmesi ve kullanıcılar tarafından kolay yorumlanabilmesi sebebiyle en çok tercih edilen sınıflandırma yöntemleri; karar ağaçları ve K-En Yakın Komşu algoritmalarıdır. Karar ağaçları, çok aşamalı karar vermede kullanılan bir yöntemdir. Karar ağaçlarında amaç; çalışma örneğini mümkün olduğunca doğru bir şekilde sınıflandırması ve gizli kalmış örneklerin yüksek hassasiyetle sınıflandırılabilmeleri için çalışma örneğine genellemeler yapılabilmesidir (Safavian ve Landgrebe, 1991: 660). Bir dizi düğüm ve daldan oluşan bir karar ağacında, düğümler dallar yardımıyla alt düğümlere bağlanır. Düğümler, karar verme sürecinde ele alınan nitelikleri temsil eder ve farklı nitelik değerleri, farklı dallar oluşturur. Karar vermek işleminde kökten yapraklara kadar arama yapılır; karar verici, sınıfların belirlendiği yapraklardır (Weiping ve Yuming: 2013: 159). Karar ağaçları, daha fazla çalışma örneği olması durumunda güncelleme kolaylığının olması sebebiyle tercih edilmektedir. K-En Yakın Komşu, mesafeye dayalı sınıflandırma algoritmasıdır. Öznitelik değerlerine göre sınıflandırılan nesneler, kendisine en yakın komşuya atanır. Komşular yakınlaşıp uzaklaşmasıyla sınıflandırmaya etki gücünün belirlendiği K-En Yakın Komşu algoritması, komşuların etkilerinin mesafeye göre ağırlıklandırılmasıdır (Akpınar, 2017: 275).

Sınıflandırma algoritmalarının genel kullanım alanları arasında finansal piyasalardaki eğilimlerin sınıflandırılması, kredi başvuru değerlendirme süreci, hastalık teşhisi, sahtekârlık sınıflandırılması, e-posta başlık ve içeriklerinden elde edilen özelliklere göre spam filtreleme gibi uygulamalar yer almaktadır.

2.3.1.2. Regresyon Analizi

Regresyon analizi, veri madenciliği modelleri arasında en eski ve en bilinen istatistiksel tekniktir. Kavramsal olarak regresyon analizi, değişkenler arasındaki ilişkileri en az hatayla modellemek ve veri öğesini gerçek değerli tahmin değişkenine eşlemek için işlev bulmaya çalışan istatistiksel bir yöntemdir. Mevcut nicel veriler üzerinden uygun matematiksel formül geliştirilerek yeni verilerin gelecekteki davranışlarını tahmin etmek için kullanılmaktadır (Sahu, Shrma ve Gondhalakar, 2011: 115; Çelik vd., 2017: 81; Fayyad vd., 1996: 44).

Regresyon analizi; bağımsız değişkenin/değişkenlerin, bağımlı değişkeni etkileme derecesini belirlemek için kullanılan istatistiksel yöntemdir. Regresyon analizinde ilk adım, problemin tanımlanmasıyla başlar. Tanımlanan probleme ait bağımlı ve bağımsız değişkenler belirlenir. Analizin devamında, bağımsız değişken sayısına göre model belirlenir. Tek bağımsız değişkene sahip model basit regresyon ve iki ya da daha fazla bağımsız değişkenli model çoklu regresyon olarak adlandırılmaktadır. Toplanan verilerle, belirlenen modele ait çıktı değerleri tahmin edilmeye çalışılır.

Regresyon analizi; ürün/hizmet iyileştirmelerinde, gelecekte oluşabilecek fırsatları veya riskleri tahmin etmede, iş süreçlerini optimize etmede ve yönetim kararlarına destek olup yanlış kararların verilmesini engellemede kullanılabilir.

2.3.2. Tanımlayıcı Modeller

Tanımlayıcı modeller, hedef veri setindeki modelleri veya ilişkileri tanımlar ve verilerin temel özelliklerini açıklar. Tanımlayıcı modeller, geçmiş veri analizleri yardımıyla gelecekte neler olabileceğini açıklayan modellerdir. Başka bir ifadeyle tanımlayıcı modeller, geçmiş davranışlardan neler öğrenilebileceğine ve öğrenilenlerin gelecek sonuçları nasıl etkileyebileceğinin anlaşılmasına yardımcı olur. Tanımlayıcı modeller, raporlama ve izleme için verilerin özetlenmesi ve verilerin anlamlı bilgilere dönüştürülmesine odaklanır. Kümeleme analizi ve birliktelik kuralları olmak üzere iki tanımlayıcı model vardır. Kümeleme analizi, nesneleri özelliklerine dayanarak gruplara ayıran tanımlayıcı model türüdür.

Kümeleme analizinde amaç, aynı gruptaki nesnelerin birbiri ile benzer ve diğer gruptaki nesnelerden farklı olmasıdır. Birliktelik kurallarında ise farklı özellikler arasındaki ilişki ortaya çıkarılır.

2.3.2.1. Kümeleme Analizi

Veri madencilerinin sıklıkla kullandığı tanımlayıcı modellerden birisi olan kümeleme analizi; ayrışık veri yığınını, homojen kümelere ayıran çok değişkenli istatistiksel yöntemdir. Bir başka ifade ile kümeleme analizi; bir ilişkilendirme ölçütü kullanarak verileri benzer özniteliklerine göre gruplandırılmasıdır (Sahu vd., 2011: 115).

Kümeleme analizinde, her bir veri arasındaki uzaklıkların ya da benzerliklerin ölçülmesiyle kendi içinde homojen ve birbirleri arasında heterojen gruplar oluşmaktadır. Benzer özellikler gösteren veriler, koordinat sistemine göre birbirine yakın konumlanmaktadır. Değişkenleri benzer niteliklere göre gruplandırmada kullanılan uzaklık veya benzerlik ölçüleri, değişkenlerin ölçü birimine göre farklılık göstermektedir. Değişkenler oran ya da aralıklı ölçekle ölçülmüş ise Öklid uzaklık bağıntısından yararlanılmaktadır (Özel ve Ata, 2007: 222).

Kümeleme analizinin uygulanması için herhangi bir varsayımı sağlama ölçütü aranmamaktadır. Ancak, uzaklık değerlerinin normalliği yeterli kabul edilmektedir. Uzaklık değerleri, birim ve ölçeklerden etkilendiği için uzaklık değerlerini standartlaştırmak gerekir. Kümeleme analizinde birçok değişik benzerlik ölçüsü kullanılmakla birlikte en yaygın kullanılan benzerlik ölçüsü Öklid'in uzaklık ölçüsüdür (Kurtuluş, 2004: 409). Küme analizinde, seçilen örneklemin anakütleyi temsil edip etmediği ve değişkenler arasında çoklu bağlantı olup olmadığına dikkat edilmektedir.

Literatürde ve farklı yazılım ortamlarında geniş kümeleme algoritmaları bulunmaktadır. Çok boyutlu veri setlerinde bulunan çeşitli yapıların ortaya çıkarılmasında evrensel olarak uygulanabilir bir kümeleme yöntemi

bulunmamaktadır. Kullanıcının problemi ve ilgili veri tiplerini anlaması, uygun kümeleme yöntemini seçmede en iyi ölçüttür (Kantardzic, 2011: 252).

Kümeleme analizinde kullanılan teknikler, genel olarak hiyerarşik ve hiyerarşik olmayan (ayrıştırıcı) kümeleme teknikleri olmak üzere iki grupta incelenebilir. Hiyerarşik kümeleme yönteminin başlangıcında her bir birim küme olarak alınır. Sürecin devamında ise benzer özniteliklere sahip birimler aynı gruba atılarak adım adım birleştirilir veya bütünden adım adım ayrılır (Akpınar, 2017: 376). Hiyerarşik olmayan kümeleme yöntemi ise araştırmacının anlamlı bir küme sayısı belirlediği noktada tercih edilmektedir (Özel ve Ata, 2007: 223).

Kümeleme analizinin yaygın olarak kullanıldığı uygulama alanlarını şu şekilde örneklemek mümkündür:

- Tıp alanında, Çelik (2013) Türkiye'deki 81 ilin sağlık yapılarının belirlenmesinde ve Haberal ve Oğul (2014) kanser hastalığının teşhisinde,
- İşletme alanında; Akoğul ve Tuna (2016) ürünlerin etkin bir şekilde pazarlanması için yapılan pazar araştırmalarında, Hosseini, Maleki ve Gholamian (2010) müşterilerin başka bir işletmeye geçme ihtimalinin incelenmesinde, Özcan (2014) dolandırıcılıkların önlenmesinde ve Alsayat ve El-Sayed (2016) sosyal ağ analizi çalışmalarında,
- Enformatik alanında, Chen ve Gopalakrishnan (1998) konuşma tanıma ve Şen ve Yazıcı (2017) belirli bir zaman ve bölgede, benzer türlerde işlenen suçların nedenlerinin araştırılması vb.

2.3.2.2. Birliktelik Kuralları

Birliktelik kuralları, verilerin özelliklerini güçlü bir şekilde ilişkilendiren modelleri tanımlamak için kullanılmaktadır. Literatürde "pazar sepet analizi" olarak adlandırılan birliktelik kuralları, büyük hacimli işlem verileri arasından olayların birlikte olma olasılığını ortaya çıkarır. Birliktelik kurallarında amaç, büyük hacimli

veriler arasından en ilginç ilişkiyi çıkarmak için etkili bir model oluşturmaktır (https://docs.oracle.com/).

Birliktelik kurallarının ilk kullanımı perakendecilik sektöründe olmuştur. Perakendecilik sektöründe gerçekleşen işlem verilerinin analiziyle ürün yerleşimi ve satış artımına yönelik kullanılmıştır. Örneğin; Amerikan perakende mağaza zinciri Wal-Mart'ın yaptığı araştırmaya göre bebek bezi ve bira arasında güçlü bir ilişki vardır. Yapılan analizler, Cuma günü saat 17:00 ile 19:00 saatleri arasında bebek bezi almaya gelen müşterilerin çoğunun bira da aldığını göstermektedir. Wal-Mart yetkilileri bu sonuca dayanarak bebek bezi ve bira reyonlarını yan yana getirmiş ve bebek bezi alıp bira almayan müşterilerin bile bira aldığı yapılan analizler sonucunda ortaya çıkmıştır. Yine Wal-Mart'ın geliştirdiği The Scan & Go isimli, konum belirleme özelliği içeren, otomatik ödeme uygulamasıyla müşterilerin mağaza içindeki satın alma davranışları analiz edilerek alışveriş esnasında müşteriye özel kuponlar sunulmaktadır (Center for Media Justice, Color of Change ve Sum of Us, 2013: 19).

Birliktelik kuralları farklı şekillerde kullanılabilir. Örneğin; müşteri kitap satın aldığında çevrimiçi mağaza, ilgili diğer kitapları önerebilir. Markette, sık sık birlikte alınan ekmek ve süt gibi ilişkili ürünler yan yana raflara yerleştirilerek müşterilerin alışveriş süresi kısaltılabilir veya ilişkili ürünler birbirinden en uzak noktadaki raflara yerleştirilip, ara raflara da diğer ilişkili ürünler yerleştirilerek bu ürünlerin alımı teşvik edilebilir (Silberschatz, Korth ve Sudarshan, 2011: 904).

Birliktelik kurallarında en önemli ilişkileri tanımlamak için destek ve güven ölçütleri kullanılır. Destek ölçütü, bir işlem setinin tüm işlemlerde ne sıklıkta gerçekleştiğini gösterir. Güven ölçütü, Y ürününü alan müşterinin hangi olasılıkla X ürününü de alacağını gösterir. Güven ölçütü, X'in Y'ye koşullu olasılığının tahminini sağlar. İyi bir kural, yüksek güvenilirlik ve güçlü bir destekle elde edilir.

Birliktelik kurallarında en çok kullanılan Apriori algoritmasıdır. Apriori algoritması, en sık kullanılan öğe kümelerini bulmak için kapsamlı bir arama stratejisi ile önceden belirlenmiş bir toplama yöntemini kullanır ve daha sonra

yinelemeli olarak öğe kümelerinin güven düzeyine göre tek tek kurallar oluşturulur (Danping ve Jin, 2011: 143).

2.4. Veri Madenciliğinde Karşılaşılan Sorunlar

Dijital veri toplama ve depolama teknolojisindeki ilerleme, veri tabanlarında gelişmelere yol açmıştır. Veri analizinde kullanılan mevcut teknikler, veri tabanlarının çok yüksek boyutlara ulaşmasıyla yetersiz kalmıştır. Örneğin; süpermarket işlem verileri, kredi kartı kullanım kaydı, telefon görüşmeleri ayrıntıları ve devlet istatistikleri gibi eskiden daha çok insan çabasıyla yürütülen dâhili işlemler, büyük veri tabanlarında yaşanan gelişmelerle yerini, insan müdahalesini minimuma indiren harici işlemlere bırakmıştır.

Verinin olduğu her alanda kullanım imkânı bulan veri madenciliği, gelişmiş veri analizi teknikleriyle karar vericilere, bilinmeyen ve beklenmeyen sonuçları göstererek karar alma noktasında kolaylık sağlamaktadır. Veri madenciliğinin kullanıcılarına sağladığı avantajların karşılaşılan güçlükler de bulunmaktadır. Veri madenciliği uygulamalarında karşılaşılan sorunları şu şekilde özetlemek mümkündür:

Bunlardan ilki eksik veri ile ilgili güçlüktür. Eksik veri; araştırmanın evreni oluşturulurken veri toplama aşamasında, verilerin öznitelik değerlerinin boş bırakılması durumunu ifade etmektedir. **Eksik veri**; bir veya birden fazla veri öznitelik değerinin, kullanıcı kaynaklı veri giriş hatası veya soruların cevap verici tarafından bilerek boş bırakılması gibi nedenlerle veri dizesinde eksik kalmasıdır (Akpınar, 2017: 111). Veri dizisinde eksik veri içeren kayıtların silinmesi, mevcut veriler göz önünde bulundurularak en uygun değerlerin belirlenip eksik verilerin yerine girilmesi veya eksik veri içeren değişkenin ortalaması alınarak eksik verilerin yerine yazılması gibi yöntemlerle hata toleransını azaltmak mümkündür.

Atık veri; veri setinde analize uygun olmayan gereksiz veya önemsiz özelliklerin var olma durumudur. Veri seti oluşturmak için normalleştirilmemiş ilişkisel tabloları birleştirmek, kullanıcının bilmediği gereksiz özelliklere yol açarak atık veri sorununu oluşturur. Atık veri sorunu, özellik seçimi algoritması ile

giderilebilmektedir. Özellik seçimi, hedef kavramları tanımlamak için gerekli ve yeterli görülen küçük bir özellik alt kümesi oluşturma işlemidir. Özellik seçimi arama alanını daraltarak hem sınıflandırma işlemini hızlandırmakta hem de sınıflandırma kalitesini arttırmaktadır (Deogun, Raghavan ve Sever, 1997: 17).

Dinamik veri; yeni veriler ekleyerek veya var olan verileri değiştirerek veri tabanlarının sürekli güncellenme durumunu ifade etmektedir. Bu nedenle, bilgi keşfi için kullanılan veri madenciliği araçları zamanla geliştirilmelidir. Başka bir ifadeyle veri tabanlarındaki her veri değişiminde, elde edilen bilgilerin de aşamalı olarak güncellenmesi gerekmektedir (Cios, Pedrycz ve Swiniarski, 2012: 6).

Gürültülü (noisy) veri; veri tabanlarında veri öznitelik değerlerindeki yanlış ölçümler, veri girişi esnasında yapılan hatalar veya öznel yargılara dayanan veri girişlerinden kaynaklanmaktadır. Bu tür sistematik olmayan hatalar gürültü olarak adlandırılmaktadır. Veri özniteliklerinin yanlış girilmesiyle oluşan gürültülü veriler, hatalı analiz sonucuna sebep olmaktadır. Gürültülü veri sorunu; regresyon, paketleme veya kümeleme yöntemleri kullanılarak çözümlenebilmektedir (Quinlan, 1986: 92-96).

2.5. Veri Madenciliğinin İşletmelerde Uygulama Alanları

İşletme faaliyetlerinde eskiden doğal kaynak, sermaye ve emek üçlüsünü yönetmek yeterli olurken küreselleşme ve teknolojide yaşanan gelişmelerle birlikte yönetilecek yeni bir kavram ortaya çıkmıştır. Veri, bu işletme kaynakları gibi yönetilecek önemli bir organizasyonel kaynak olarak kabul edilmiştir (Martin, DeHayes, Hoffer ve Perkins, 1991: 500).

Büyük, karmaşık ve bilgi açısından zengin veri setlerini anlama ihtiyacı hemen hemen tüm işletme, bilim ve mühendislik alanlarının ortak noktasıdır. İş dünyasında, kurumsal veriler ve müşteri verileri stratejik bir varlık olarak kabul edilmektedir. Bu verilerde saklı olan, yararlı bilgileri çıkarma ve bu bilgilere göre hareket etme yeteneği rekabetin yoğun yaşandığı dünyamızda giderek daha önemli hale gelmektedir. Verilerden bilgi edinmek için bilgisayar tabanlı bir yöntem

uygulama süreci olan veri madenciliği yoğun rekabet ortamında işletmelere büyük bir avantaj sağlamaktadır (Katardazics, 2011: 1).

Beklenmeyen ilişkileri bulan ve verileri, veri sahibi için hem anlaşılır hem de yararlı olan yeni yollarla özetleyen veri madenciliği, işletme disiplininde çeşitli çalışma alanlarında tercih edilmektedir (Hand, Mannila ve Smyth, 2001: 1; Çelik vd., 2017: 56). Tablo 3'te, veri madenciliğinin işletme disiplinindeki kullanım alanları özetlenmektedir.

Tablo 3: İşletmelerde Veri Madenciliğinin Uygulama Alanları

Uygulama Alanları	Uygulama	Veri Madenciliği Modelleri	Açıklama
Sağlık	Sağlık hizmetlerinin sunumu	Karar Ağacı	Geçmiş veriler üzerinden hasta semptomlarına göre kanser, kalp krizi gibi çeşitli hastalıkların ön tanısı, risk seviyelerinin belirlenmesi ve ortaya çıkabilecek hataların önüne geçilmesi
	Sağlık politikalarına yön verilmesi	Birliktelik Kuralları	Sağlık politikalarına ve sağlık kurumları yönetimine etki edebilecek tüm değişkenlerin değerlendirilerek yeni yol haritaları oluşturulması
	Sağlık kurumlarının yönetimi	Kümeleme Analizi	
		Sinir Ağları	
Borsa	Risk tahmini	Karar Ağaçları	Döviz kuru bilgileriyle işlem gören hisse senetlerinin risk seviyesinin belirlenmesi
	Genel piyasa değerlendirmesi	Genetik Algoritma	Hisse saptama ve hisse alım-satım işlemlerinin tutarlılığı
		Sinir Ağları	

Tablo 3 (Devam)

Telekomünikasyon	Müşteri kaybı analizi	Kümeleme Analizi	Müşteri kararlarını etkileyen unsurların belirlenerek olası müşteri kayıplarının saptanması
	Kalite iyileştirme yönetimi		Değişen müşteri ihtiyaçlarının analiz edilerek ürün veya hizmet kalitesinde sürekli iyileştirilmeler yapılması
	Şebeke alt yapı yönetimi	Birliktelik Analizi	Anlık ağ performansı ve veri trafiği ölçümleriyle hat yoğunluk tahmini yapılarak olası sorunların önüne geçilmesi
	Benzerlik konumlandırma		Birlikte satın alınan ürünlerin belirlenip görsel olarak birleştirilmesi
Pazarlama ve Perakendecilik	Çapraz satış	Birliktelik Kuralları	Müşterilerin genel satın alma davranışlarına göre tamamlayıcı veya ilişkili ürünler önerilmesi
	Pazar sepet analizi		Müşteri satın alma davranışları analiz edilerek ilişkili ürünlerin market içi konumlandırılması, fiyatlandırılması, reklam kampanya planlama süreçlerinin hazırlanmasında
	Müşteri ilişkileri yönetimi		Müşteri sadakatini değerlendirerek yeni pazarlama stratejileri oluşturulması
Bankacılık ve Sigortacılık	Müşteri ilişkileri yönetimi	Karar Ağaçları	Müşterilerin bir ürün veya hizmetin olası alternatiflerine karşı değerini belirleyerek geliri en üst düzeye çıkarmak için planlamalar yapılması
	Sahtekarlık tespiti		Kredi taleplerinin değerlendirilmesi
		Kümeleme Analizi	Kredi kartıyla yapılan tüketim harcamalarına göre müşteri profilinin belirlenmesi
			Geçmiş veriler üzerinden kural bazlı senaryolar tanımlanarak olası dolandırıcılıklara karşı hızlı müdahale edilmesi

Kaynak: Olson, David L. ve Delen, Dursun, *"Advanced Data Mining Techniques"*, Springer, 2008, 7; Ekinci, Tuğba, Veri Madenciliği ve Telekomünikasyon Sektöründe Bir Uygulama, (Yüksek Lisans Tezi, Celal Bayar Üniversitesi, Sosyal Bilimler Enstitüsü), 2009 Manisa.

Tablo 3, veri madenciliğinin işletmelerde uygulandığı alanları, uygulama amacını ve verilerin analizinde kullanılan veri madenciliği yöntemlerinin bir

genellemesi olup alan yazını incelendiğinde, veri madenciliğinin işletmelerdeki kullanımına ilişkin örnekler aşağıda sıralanmıştır.

Akıncı (2019), Financial Access Survey (FAS) veri tabanından elde edilen sekiz farklı ile Orta Doğu ve Kuzey Afrika ülkelerindeki finansal erişim kümeleme analiziyle açıklanmaya çalışılmıştır.

Zhuang, Yao ve Liu (2018), banka müşterilerinin davranışlarını modelleyerek, vadeli mevduat katılımını tahmin etmek ve banka pazarlamasının etkinliğini ve doğruluğunu arttırmak için veri madenciliği tekniklerini kullanmışlardır.

Gültekin (2017), otomobil sektöründe yer alan ikinci el araçların fiyat karşılaştırmasında veri madenciliği modellerinden yapay sinir ağları ve doğrusal regresyon kullanarak ikinci el araçların fiyat tahmini amaçlanmıştır. Yapay sinir ağları ile yapılan tahminlerdeki sapma miktarının, doğrusal regresyon tahminlerinden daha iyi olduğu ve doğru verilerle yapılan tahminlerde sapma miktarının azaldığı sonucuna ulaşılmıştır.

2011-2014 yıllarında, Borsa İstanbul'da sınaî endekste kesintisiz işlem gören işletmelerin yıllık faaliyet ve bağımsız denetim raporları incelenerek olumlu görüş dışındaki denetim görüşlerinin tahmini amaçlanmıştır. Veri madenciliği algoritmalarında CART, GRI ve C5.0 kullanılarak olumlu görüş dışındaki denetim görüşünün tahmininde en etkili değişkenin "önceki denetim görüşü" olduğu sonucuna ulaşılmıştır (Yaşar, 2016: 81-109).

Kaliforniya Eyalet Üniversitesi'nde öğrenim gören öğrencilerin verileri kullanılarak, üniversitenin mezun verme oranı yükseltmek için okuldan erken ayrılma eğilimi olan öğrencileri önceden tahmin eden bir model tasarlanmıştır (Xiangyan, 2015).

Akyel ve Seçkin (2012); örnek olaylar üzerinden belirli nitelikler dikkate alınarak çalışan personel ya da yeni işe başvuran adayların hileye başvurma eğilimleri K-En Yakın Komşu algoritması kullanılarak ölçülmüştür.

Telekomünikasyon sektöründe faaliyet gösteren bir işletmenin verileriyle müşteri kaybı tahmini için veri madenciliği teknolojisini kullanan bir karar destek sistemi oluşturulan çalışmada, müşteri kaybı tahmini için müşterilerin davranışlarını önceden tahmin eden bir model tasarlanmıştır (Jadhav ve Pawar, 2011).

Yanbin (2006); elektronik pazarlardaki satıcıları geri bildirim puanlarına göre "yeni veya deneyimli" olarak sınıflandırarak, etkili ve verimli elektronik pazarlar oluşturmak için gerekli olan satıcı davranışı ve etkili satış stratejileri veri madenciliği algoritmalarıyla açıklanmaya çalışılmıştır.

Sigorta sektöründe yapılan çalışmada Smith, Willis ve Brooks (2000), müşteri taleplerini öngörebilmek ve poliçe sahiplerinin, poliçe yenileme veya sonlandırma işlem tahmini için veri madenciliği algoritmalarından Yapay Sinir Ağları ve Karar Ağaçları kullanılmıştır.

Birmingham Hastanesi'nin 1996 yılındaki hasta verileri kullanılarak yapılan çalışmada, veri madenciliği yöntemleri hastane enfeksiyon kontrolünün sağlanmasında kullanılmıştır. Analiz sonucunda, hastalarda enfeksiyon oluşumunda "Pseudomonasaeruginosa" bakterisinin etkili olduğu sonucuna ulaşılmıştır (Brosette, Sprague, Hardin, Waites, Jones ve Moser, 1998: 373-381).

İşletmelerde veri madenciliği tekniklerinin etkin bir şekilde kullanılmasıyla kısa zamanda ve az maliyetle bilgi elde edilerek işletmelerdeki anlık sorulara cevap alınabilmektedir. Teknolojinin değişen doğasına ayak uydurmaya çalışan kütüphane sistemlerinde de dijitalleşme sürecinde veri madenciliği yöntemleri kullanılmaktadır.

2.6. Kütüphanelerde Veri Madenciliği ve Uygulama Örnekleri

Toplumun bilgiye olan gereksinimini karşılamayı amaçlayan kütüphaneler; birçok uygarlık ve iletişim biçiminden geçerek varlıklarını devam ettiren ve kullanıcılarının yararlanabilmesi için her türlü bilgi kaynağını bünyesinde barındıran ve bu kaynaklara erişimini sağlayan en önemli kurumlardandır (Çakın, 1986: 9-10).

Toplumun bilgiye ulaşma ve bilgiyi yayma çabasının artması, teknolojinin değişen doğasının en önemli kazanım ve karakteristiklerinden birisidir.

Çeşitli disiplinlerde uygulama alanı bularak bilgiye ulaşma ve yayma noktasında güçlü ve etkili teknikler sunan veri madenciliği, potansiyelleri ele alarak farklı kullanıcı ihtiyaçlarını karşılamaktadır (Shieh, 2010: 709). Dijitalleşme sürecinde olan kütüphanelerde de katlanarak çoğalan veri yığınlarını geleneksel yöntemler kullanarak işlemenin, yönetmenin ve arşivlemenin zorlaşmasıyla kütüphanelerde veri madenciliği tekniklerini kullanılmaya başlanmıştır.

Kütüphanelerde oluşan büyük boyutlardaki veri yığınlarını veri madenciliği yöntemleri kullanarak incelemek "bibliomining" olarak adlandırılmaktadır. 2003 yılında Scott Nicholson'ın kütüphane veri madenciliği üzerine yaptığı çalışmalarını "bibliomining" olarak adlandırmasıyla alan yazınına yeni bir terim kazandırılmıştır (Uçan, 2010: 21).

Bibliomining olarak adlandırılan kütüphaneler için veri madenciliği, kütüphane sistemlerinden davranış temelli örüntü kalıplarını çıkarmak için kullanılan veri madenciliği teknikleri olarak tanımlanmaktadır. (Shieh, 2010: 710). Bir başka deyişle bibliomining, kütüphane yönetimindeki karar vericilere yardımcı olmak veya yapılan faaliyetleri gerekçelendirmek amacıyla istatistik ve örüntü tanıma araçlarının kütüphane sisteminde bulunan ilgili büyük miktardaki veriye uygulanmasıdır (Nicholson, 2003: 146).

Bibliomining süreci; hedef tanılama, veri setlerinin toplanması ve ön işlenmesi, verideki bilginin keşfi, sonuçların değerlendirilmesi ve edinilen bilginin pratikte uygulanması gibi birkaç adımdan oluşmaktadır. Bibliomining sürecinin ilk adımında, kütüphanedeki belirli bir problem veya karar vermeyi gerektiren bir durum ilgili alan yöneticileri tarafından belirlenir. Hedef alan belirlendikten sonra ikinci adımda, dâhili ve harici veri kaynaklarından veriler toplanır ve ETL araçları kullanılarak bu veriler temizlenip, öznitelikler anonimleştirilerek veri ambarı oluşturulur. Üçüncü adımda, veri madenciliği uzmanları tarafından en uygun değerler

elde edilene kadar farklı veri madenciliği teknikleri seçilir ve uygulanır. Son adımda ise analiz sonuçları kütüphane karar vericileri tarafından değerlendirilerek uygunluğu durumunda uygulamaya konur (Shieh, 2010: 710; Nicholson, 2003: 147-148).

Bibliomining sürecinin amacı, kütüphane kullanıcı ilişkilerinin daha etkin bir şekilde yönetilmesini ve kütüphane kullanıcılarının davranış kalıplarını elde ederek kütüphanelerin sunduğu hizmetlerin kalitesinin iyileştirilmesini sağlamaktır. Bu sebeple bibliomining, kütüphane yönetiminin karar alma süreçlerini desteklemek için önemli bir araç haline gelmektedir (Hajek ve Stejskal, 2012: 339).

Huancheng, Tingting ve Rocha (2018), Çin akademik kütüphanelerinde veri madenciliği uygulamalarına ilişkin literatür analizi yapılmıştır. 2008-2017 yılları arasında yayınlanan, veri madenciliği ve akademik kütüphane anahtar kelimelerini içeren eserler taranmıştır. Birliktelik kuralları ve kümeleme analizi kullanılarak yayınlanan eserlerdeki etkin noktalar ve gelecekteki araştırma konularının eğilimi analiz edilmiştir.

Demiral, Soba ve Armutlu (2017), 2007-2017 yılları arasında, Uşak Üniversitesi Merkez Kütüphanesi verileri kullanılarak öğrencilerin kitap alma davranışları incelenmiştir. İktisadi ve İdari Bilimler Fakültesi öğrencilerinin verileri üzerinde ayırıcı hiyerarşik kümeleme tekniği (DIANA) kullanılmış ve on bir küme elde edilmiştir. Kümelerin ödünç aldığı kitaplar incelendiğinde branşlarıyla ilgili kitapları daha fazla aldıkları ve yıllara göre ödünç kitap alımlarında dalgalanmalar yaşandığı belirlenmiştir.

Kütüphane kullanıcılarının ödünç alma verileri üzerinde birliktelik kuralları analizi uygulanan çalışmada; ödünç alınan kitaplar arasındaki ilişkilerden, kütüphane kullanıcılarının ihtiyaçları tahmin edilmeye çalışılmıştır (Song ve Wei, 2011: 1525-1528).

Uçan (2010); Akdeniz Üniversitesi Merkez Kütüphanesi'nde, kütüphane kullanıcı ve eser bilgileri yardımıyla, kullanıcı ödünç alma verileri Apriori ve TwoStep algoritmaları uygulanarak irdelenmiştir. Akdeniz Üniversitesi Merkez Kütüphanesi'nin uygulamış olduğu kitap sınıflandırmalarına göre kullanıcıların

ödünç aldığı kitap gruplarından hangileri arasında ilişki olduğu Apriori algoritmasıyla ortaya koyulmuştur. Birliktelik analizinin sonucuna göre en yüksek güven oranına sahip Türk Edebiyatı eserlerinin en fazla ilişkisel yapıya sahip olduğu belirlenmiştir. Türk Edebiyatı eserlerine ilgi gösteren kullanıcıların, Almanca Flemenkçe İskandinav Edebiyatı ve Fransızca İtalyanca İspanyolca Edebiyatı eserlerini de birlikte ödünç aldığı sonucuna ulaşılmıştır. Kütüphane kullanıcılarının ödünç sirkülasyon verileri üzerinden bölüm, cinsiyet, grup ve toplam ödünç yayın sayısına göre kullanıcılar, TwoStep algoritması kullanılarak kümelenmiştir. TwoStep analizi sonucunda dört küme elde edilmiş ve kümelerde cinsiyet unsurunun baskın özellik olduğu keşfedilmiştir.

İş akışlarına uyum sağlayan hesaplanabilir çevrimiçi sınıflandırma hizmeti geliştirmek amacıyla dijital kütüphane sistemleri üzerinde veri ve metin madenciliği algoritmaları kullanılmıştır (Sanderson ve Watry, 2007: 73-79).

Battioui (2007), Louisville Üniversitesi'ndeki Ekstrom Kütüphanesi'nin veri tabanında bulunan dijital kaynaklara erişim ile ilgili URL bilgileri ile diğer kümeler arasındaki ilişkiler tanımlanmaya çalışılmıştır. SAS Enterprise Miner programı kullanılarak yapılan Path analizinde, kullanıcıların kütüphane web sitesinde oturum açtıktan sonra yaptığı aramalarla kurulan bağlantıların kaç kez oluştuğu ve bağlantılar arasındaki ilişki tanımlanmıştır.

Veri madenciliği tekniklerinden Karar Ağaçları, Yapay Sinir Ağları, Diskriminant ve Lojistik Regresyon analizi kullanarak dijital kütüphane ortamında otomatik koleksiyon geliştirmek için öngörücü bir model oluşturulmaya çalışmıştır (Nicholson, 2003: 146-156).

Gebze Yüksek Teknoloji Enstitüsü Kütüphanesi web sitesinin günlük verilerine dayanarak K-En Yakın Komşu algoritmasını kullanarak kütüphane kullanıcılarına ait erişim örüntüleri oluşturulmaya çalışılmıştır. Kütüphane kullanıcıları erişim yapılan dosyanın niteliğine göre iyi ve kötü kullanıcı olarak tanımlanmıştır. Kullanıcıların web sitesi erişimlerine göre kişiselleştirilmiş hizmetler sunulmaya çalışılmıştır (Takçı ve Soğukpınar, 2002: 12-26).

ÜÇÜNCÜ BÖLÜM

KÜTÜPHANE VERİ MADENCİLİĞİ UYGULAMASI

Bu bölümde, Trakya Üniversitesi Merkez Kütüphanesi veri tabanından alınan kütüphane verilerinin veri madenciliği ile analizi anlatılmıştır. Araştırmanın amacı, önemi, evreni, örneklemi ve uygulama süreci detaylı bir şekilde ele alınmıştır.

3.1. Araştırmanın Amacı ve Önemi

Rekabet koşullarının yoğun olduğu dönemde, geniş veri derlemini (koleksiyon) eleyerek işletme için yararlı bilgiye ulaşma yolunu açan ve kurumsal verileri avantaja dönüştüren veri madenciliğinin önemi gün geçtikçe artmaktadır. İnsanların ve bilgisayarların ortak bir çabası olan veri madenciliği (Katardazics, 2011: 1), Türkiye'de de yeni bir süreç olarak uygulanmaktadır.

Bilgi keşfinin bir parçası olan veri madenciliği, verinin toplandığı ve depolandığı tüm alanlarda sektör farkı gözetmeksizin uygulama olanağı bulduğu gibi kütüphanelerde de kullanılmaktadır. Teknolojinin gelişmesiyle kütüphane kavramı da gelenekselden dijitale çevrilmiştir.

Dijitalleşme sürecinde olan kütüphanelerde, katlanarak çoğalan veri yığınlarını geleneksel yöntemler kullanarak işlemek, yönetmek ve arşivlemek zorlaşmıştır. Bu durum, kütüphanelerde veri madenciliği tekniklerini kullanmayı zorunlu kılmıştır.

Araştırmada; bilgi yönetimi çağında, dijitalleşme sürecindeki kütüphanelerde oluşan veri akınlarından anlamlı bilgiler çıkartılması yani bilginin keşfi sağlanmaya çalışılmıştır. Trakya Üniversitesi Merkez Kütüphanesi otomasyon sistemi üzerinde yapılan çalışmalar ile sonuçların, kütüphane yönetimindeki karar vericileri için yeni bir kaynak olması amaçlanmış, bu konuda yeni ve stratejik kararlar alma aşamasında destek sağlayan bir model önerisi ortaya konulmuştur. Kütüphane veri tabanında veri madenciliği tekniklerinin uygulanma sürecinin henüz yeni olmasından dolayı düzenli olmayan veriler üzerinde temizleme işleme

uygulanmış ve çeşitli modelleme teknikleri kullanılmıştır. Çalışmada özellikle, yayın ödünç sirkülasyon verileri üzerine odaklanarak kütüphane kullanıcılarının davranışları incelenmiştir.

3.2. Araştırmanın Evreni ve Örneklemi

Bu çalışmada, orta büyüklükte bir veri tabanına sahip Trakya Üniversitesi Merkez Kütüphanesi'nden elde edilen veriler kullanılmıştır. Kütüphanenin dijitalleşme süreci 1998 yılından itibaren başlamış ve Türkiye'de geniş bir kullanıcı kitlesine sahip olan YORDAM kütüphane otomasyon sistemi yazılımı ile sağlanmıştır. YORDAM yazılımı; kitap, süreli (dergi ve gazete) ve gri (standart, patent, tez vb.) yayınlar ve her tür kayıt dışı belge (CD, CD-ROM, videokaset, harita, fotoğraf vb.) giriş verilerini, kütüphane içi hizmet ve kullanıcı bilgilerini kayıt altına alabilmektedir.

Bu çalışmanın araştırma evrenini, 2003-2018 eğitim öğretim yılında Trakya Üniversitesi Merkez Kütüphanesi'nde araştırma yapan öğrencilerin kitap ödünç alma davranışlarını içeren veri seti oluşturmaktadır. Araştırma, bir veri madenciliği çalışması olduğu için veriden örneklem alınmamış, tüm veriler veri madenciliği algoritmaları tarafından değerlendirilmiştir.

3.3. Araştırmada Verilerin Elde Edilmesi ve Kullanılacak Program Seçimi

Trakya Üniversitesi Merkez Kütüphanesi'nin 15 yıllık verileri, YORDAM kütüphane otomasyon sistemi yazılımından Microsoft Office Excel ortamına aktarılması ile elde edilmiştir. Araştırmanın uygulama aşamasında Tablo 4'te belirtilen tabloların kullanılması uygun görülmüştür. Tablo 4'te belirtilen ve veri madenciliği çalışmasında kullanılacak olan tablolar, herhangi bir işlem uygulanmamış haliyle yansıtılmıştır.

Tablo 4: Veri Madenciliği Uygulama Sürecinde Kullanılan Tablolar

Tablo Adı	Alan Sayısı	Kayıt Sayısı
Üye	17	30.075
Ödünç	11	269.041

Üye tablosu içerisinde kütüphaneden en az bir yayın almış olan öğrenci, akademik ve idari personelin kayıtları bulunmaktadır. Ödünç tablosu ise kütüphane kullanıcılarının yayın alım bilgilerini içermektedir. Ödünç tablosu, kullanıcıların üye numaralarına göre aldığı yayınların sınıfını, ödünç tarihi ve zamanı, iade tarihi ve zamanı ve iade zamanının aşımı sonucunda alınan ceza kayıtlarını içermektedir.

Veri temizleme işlemleri, Microsoft Access ve Excel 2007 programlarında çeşitli sorgulamalar yapılarak gerçekleştirilmiştir. Veri temizleme işlemlerinin ardından ilişkili üye ve ödünç tabloları Microsoft Access ortamında birleştirilerek veri madenciliği sürecinde kullanılmak üzere hazır hale getirilmiştir.

RapidMiner, genel bir veri madenciliği sürecinin tasarımını ve belgelenmesini destekleyen bir sistemdir. Programlama bilgisine ihtiyaç duymadan sürükle-bırak-tıkla işlevleriyle sağladığı kullanım kolaylığı, veri ve sonuçlarının görselleştirilmesi ve diğer veri madenciliği programlarına kıyasla daha fazla özellik içermesi nedeniyle bu araştırmada, RapidMiner 8.2 (Educational Licence) programı kullanılması tercih edilmiştir.

Almanya'daki Dortmund Teknik Üniversitesi Yapay Zeka Birimi'ndeki bilim insanları tarafından makine öğrenimi, veri madenciliği, metin madenciliği, yapay zeka ve doğal dil işleme konuları hakkında derinlemesine bilgi sağlamak amacıyla bir veri madenciliği yazılım projesine başlanmıştır. "Yet Another Learning Environment-YALE" adlı bu proje, daha sonra geliştirilerek RapidMiner olarak yeniden adlandırılmıştır. RapidMiner; açık kaynak kodlu oluşu, esnek ve genişletilebilir yapısı ve desteklediği özel uzantılı dosyalar sayesinde bilimsel, endüstriyel ve ticari uygulamalar için ideal bir çözüm aracı olmuştur.

3.4. Veri Ön İşleme Süreci

Trakya Üniversitesi Merkez Kütüphanesi veri tabanında, çok büyük miktarda bir veri kümesi olmamasına rağmen karmaşık ve hatalı kayıtlara sahiptir. Bu hatalı kayıtların önemli bir kısmı, kütüphane veri tabanına üye, demirbaş veya ödünç yayın gibi veri girişlerinin farklı kütüphane çalışanları tarafından yapılmasından kaynaklanmaktadır. Veri kümesinin, veri madenciliği programlarında kullanılmak üzere elverişli olmayışı sebebiyle oluşan hatalı kayıtlar veri tabanından çıkarılmış ya da düzeltilmiştir.

Üye ve ödünç kayıtlarının olduğu tablolardan veri madenciliğinde kullanılması düşünülmeyen alanlar çıkartılmıştır. Üye tablosunda üye adı, soyadı, doğum tarihi, yaşı gibi kişisel bilgiler gizlilik kuralları çerçevesinde veri ambarına dâhil edilmemiştir.

Üye No	Cinsiyet	Tabiyeti	Durumu	Grup	Fakülte	Bölüm
10902030xx	Erkek	T.C.	Aktif	Öğrenci	Mühendislik Fakültesi	Bilgisayar Mühendisliği
10982271xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	TARİH ÖĞETMENLİĞİ(İ.Ö.)(TEZSİZ)
10969200xx	Erkek	T.C.	Aktif	Öğrenci	Edirne Sosyal Bilimler Mes	Ağırlama Hizmetleri
10901020xx	Erkek	T.C.	Aktif	Öğrenci	Fen-Edebiyat Fakültesi	Fizik
20801053xx	Erkek	T.C.	Pasif	Öğrenci	Fen-Edebiyat Fakültesi	Türk Dili ve Edebiyatı (İ.Ö.)
10982271xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	TARİH ÖĞETMENLİĞİ(İ.Ö)(TEZSİZ)
10862030xx	Kadın	T.C.	Aktif	Öğrenci	Havsa Meslek Yüksekokulu	
10981051xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Fen Bilimleri Enstitüsü	Bilgisayar Mühendisliği
10982081xx	Kadın	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	Kamu Yönetimi
20705028xx	Kadın	T.C.	Aktif	Öğrenci	Eğitim Fakültesi	İngilizce Öğretmenliği(İ.Ö.)
P0004xx	Kadın	T.C.	Aktif	İdari Personel	T.Ü.Rektörlüğü	İdari Mali İşler Daire Başkanlığı
20824023xx	Kadın	T.C.	Aktif	Öğrenci	Uygulamalı Bilimler Yüksek	Turizm İşletmeciliği ve Otelcilik (İ.Ö.)
10901080xx	Erkek	T.C.	Aktif	Öğrenci	Fen-Edebiyat Fakültesi	Sanat Tarihi
10982121xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	Uluslararası İlişkiler
10804010xx	Erkek	T.C.	Pasif	Öğrenci	Tıp Fakültesi	Tıp
A0012xx	Erkek	T.C.	Aktif	Akademik	T.Ü.Rektörlüğü	
20705013xx	Kadın	T.C.	Aktif	Öğrenci	Eğitim Fakültesi	Sınıf Öğretmenliği (İ.Ö.)
10862030xx	Erkek	T.C.	Aktif	Öğrenci	Havsa Meslek Yüksekokulu	Bankacılık
10705018xx	Kadın	T.C.	Aktif	Öğrenci	Eğitim Fakültesi	Sınıf Öğretmenliği
20707023xx	Erkek	T.C.	Aktif	Öğrenci	İktisadi ve İdari Bilimler Fa	İŞLETME
10121020xx	Kadın	T.C.	Aktif	Öğrenci	Edirne Sağlık Yüksekokulu	Ebelik
10926040xx	Erkek	T.C.	Aktif	Öğrenci	Keşan Yusuf Çapraz Uygulal	İşletme Bilgi Yönetimi
10905080xx	Kadın	T.C.	Aktif	Öğrenci	Eğitim Fakültesi	Sosyal Bilgiler Öğretmenliği
10982061xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	İktisat
10982231xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	İşletme (İ.Ö.)Tezsiz
A0012xx	Erkek	T.C.	Aktif	Akademik	Tıp Fakültesi	Genel Cerrahi
10982181xx	Erkek	T.C.	Aktif	Y.L. Öğrencisi	Sosyal Bilimler Enstitüsü	TARİH ÖĞETMENLİĞİ(TEZSİZ)
P0004xx	Erkek	T.C.	Aktif	İdari Personel	T.Ü.Rektörlüğü	

Şekil 8: Üye Tablosundaki Verilerin Ham Hali

Şekil 8'deki üye tablosunda, **Üye No** alanı birincil anahtar olarak belirlenmiş ve öğrenci, akademik ve idari personel numaraları ve kimlik bilgileri gizlenerek veri madenciliği uygulama aşamasında kullanılmamıştır. **Üye No** alanında, akademik personel numarası "A0000xx" ve idari personel numarası

“P0000xx” gibi yedi karakter ve öğrenci numarası “19904011xx” gibi on karakter olma zorunluluğuna uymayan veriler ayıklanarak tablodan çıkartılmıştır.

Cinsiyet alanında, kütüphane çalışanları tarafından, tüzel kişi olarak kodlanmış kayıtlar çıkartılarak düzeltilmiştir. Tabiiyeti alanına girilen kayıtlar analize dâhil edilmeyeceği için çıkartılmıştır. Üyelik durumunu gösteren **Durumu** alanı, aktif ve pasif olarak kodlanmıştır. Pasif, üyelerin mezun olma veya çeşitli sebeplerle üniversite ile olan ilişiğinin kesildiği anlamına gelmektedir. Üye tablosundaki pasif üyelerin sayısı, toplam kayıt sayısının neredeyse yarısına eşit olduğu için çıkarılmamış ve analize dâhil edilmiştir.

Grup alanında üyeler; **Öğrenci Y.L. Öğrencisi** ve **Doktora**, **İdari** ve **Akademik Personel** ve **Vatandaş** olarak kodlanmıştır. Uygulamada, yalnızca öğrencilerin kütüphane sirkülasyon davranışlarının incelenmesine karar verilmiş ve grup alanındaki vatandaş, idari ve akademik personel kayıtları veri madenciliği aşamasına dahil edilmemiştir. Grup alanında Doktora kayıtları **Doktora Öğrencisi**, Y.L. Öğrencisi kayıtları **Yüksek Lisans Öğrencisi** olarak değiştirilmiştir. Üyelerin bağlı olduğu fakülte alanında, meslek yüksekokulu olarak kayıtlı öğrenciler, **Ön Lisans Öğrencisi** olarak değiştirildikten sonra geri kalan öğrenci grubunun ismi **Lisans Öğrencisi** olarak düzenlemiştir. Grup alanın ismi **Eğitim Programı** olarak değiştirilmiştir.

Fakülte ve bölüm alanlarında kütüphane çalışanları tarafından hatalı girilen kullanıcı bilgileri mevcuttur. Örneğin; İktisadi ve İdari Bilimler Fakültesi’nde İşletme bölümü için İşletme, İŞLETME, İşletme (İ.Ö.) gibi birden fazla kodlamalar yapılmıştır. Bu gibi hatalı kayıtlar düzeltilip ve örgün öğretim ile ikinci öğretim kayıtları da tek bir bölüm altında birleştirilerek toplanmıştır. Fakülte alanında, Kırklareli Üniversitesi ve Tekirdağ Namık Kemal Üniversitesi olarak kodlanan 535 kayıt silinerek uygulamaya dâhil edilmemiştir. Fen-Edebiyat Fakültesi de **Fen Fakültesi** ve **Edebiyat Fakültesi** olarak ayrılıp gruplandırılmıştır. Tıp Fakültesi’nde bölüm, branş ve dal seçimi mezun olduktan sonra yapıldığı için Tıp Fakültesi’nde öğrenim gören lisans öğrencilerinin bölüm alanına **Tıp** kodlaması yapılmıştır. Yabancı Diller Fakültesi’ne kayıtlı öğrenciler, Mütercim Tercümanlık bölümünün

hazırlık aşamasındaki öğrenciler olup bu fakültedeki öğrencilerin bölüm kodlaması **Mütercim Tercümanlık (Hazırlık)** olarak değiştirilmiştir. Eczacılık ve İlahiyat Fakültesi'ne kayıtlı öğrencilerin bölüm alanına kütüphane çalışanları tarafından **İlahiyat** ve **Eczacılık** kodlaması yapılmış ve bu kodlamalara dokunulmamıştır. Fakülte alanında, **meslek yüksekokulu** olarak kayıtlı öğrencilerin bölüm alanlarına öğrenim gördükleri bölümün programları ve **enstitü** adı altında kayıtlı yüksek lisans ve doktora öğrencilerinin bölüm alanlarına öğrenim görülen bilim dalı şeklinde kodlama yapılmıştır. Bölüm karışıklığını ortadan kaldırmak amacıyla ön lisans öğrencilerinin bölüm alanına, programın bağlı olduğu bölümün adı, yüksek lisans ve doktora öğrencilerinin ise bilim dalının bağlı olduğu anabilim dalının adı olarak değiştirilmiştir. Fakülte alanı ismi de **Akademik Birimler** olarak değiştirilmiştir.

Demirbaş	Sınıflama	İadeTarihi	İade Zamanı	Ödunç Tarihi	Ödünç Zamanı	Üye Numarası	Eser Adı
0024830	664.06/sin	07.01.2003	16:48:07	07.01.2003	00:20:28	A0003xx	Food process design and evaluation / Bakesh K. Singh / A [6
0004053	150/MCK p.	18.01.2003	10:40:31	04.01.2003	00:20:30	A0000xx	Psychology / Wilbert James McKeachie, Chorlotte Lackner D
0004112	150/Joh p.	18.01.2003	10:40:46	04.01.2003	00:20:30	A0000xx	Psychology: A Problem-Solving Approach / Donald M. Johnso
0000899	370.956/ATA	21.01.2003	14:52:04	09.01.2003	00:20:31	10243090xx	Atatürk ve Eğitim : Fotograflar-Hatıralar-Belgeler-Kronoloji /
0000072	512.8/BAY s.	23.01.2003	11:33:51	07.01.2003	00:20:32	10001030xx	Soyut cebir ve sayılar teorisi / Mustafa Bayraktar / A [512.8
0022111	516/ste	23.01.2003	11:34:03	07.01.2003	00:20:32	10001030xx	Calculus ve analitik geometri Cilt:2 / Sherman K. Stein, Antho
0009743	658.4/GAR d.IV	23.01.2003	13:13:15	18.01.2003	00:20:32	10107018xx	Deneyimlerim IV / Üzeyir Garih / A [658.4/GAR d.IV] 05 Okun
0021514	658.314/tos	23.01.2003	13:13:25	18.01.2003	00:20:32	10107018xx	İşletme Yönetimi / Kemal Tosun / A [658.314/tos] 05 DEPO
0021057	659.2/seç	24.01.2003	11:22:15	14.01.2003	00:20:32	10147130xx	Halkla ilişkiler ortam ve araçları / Haluk Gürgen; Ed. Kazım S
0006368	658/İŞL i.	24.01.2003	11:22:21	14.01.2003	00:20:32	10147130xx	İşletme Yönetiminde Güncel Konular-1 / A [658/İŞL i.] 05 Oku
0024636	382/ata	25.01.2003	12:12:33	18.01.2003	00:20:32	A0000xx	Dış ticaret işlemleri ve muhasebesi / Ümit Ataman, Haluk Sur
0024716	332.7/ber	25.01.2003	12:12:49	18.01.2003	00:20:32	A0000xx	Bankacılıkta pazara yönelik kredi yönetimi / Niyazi Berk / A
0024751	332.153/tun	25.01.2003	12:13:00	18.01.2003	00:20:32	A0000xx	Ticari bankalarda optimal kaynak yönetimi:teorik ve analitil
0021490	658.8/paz	28.01.2003	11:32:14	14.01.2003	00:20:33	10147130xx	Pazarlama yönetimi seminer notları / A [658.8/paz] 05 DEPO
0003867	597.09561/MAT t.	29.01.2003	15:14:39	06.01.2003	00:20:33	A0000xx	Türkiye deniz balıkları atlası / Savaş Mater, Savaş Mater, Oğ
0021879	617.89/ann 110(1-6)	31.01.2003	12:16:41	30.01.2003	00:20:34	A0000xx	Annals of otology rhinology and laryngology / Ed. Brian F. M
0021878	617.89/ann 110(7-12)	31.01.2003	12:17:08	30.01.2003	00:20:34	A0000xx	Annals of otology rhinology and laryngology / Ed. Brian F. M
0026554	813/tür	31.01.2003	15:20:23	28.01.2003	00:20:34	10201040xx	Güven : roman / Vedat Türkali / A [813/tür] 05 Okuma Salon
0026362	891.734/gor	31.01.2003	15:20:33	28.01.2003	00:20:34	10201040xx	Benim üniversitelerim / Maksim Gorki, çeviren Mehmet Zorlu
0015317	616.2/tex	03.02.2003	11:50:52	03.02.2003	00:20:34	A0004xx	Text book of pulmonary diseases sixth edition / Gerald L. Ba
0024390	616.2/tex/c.1	03.02.2003	11:50:59	03.02.2003	00:20:35	A0004xx	Textbook of respiratory medicine / Edited by John F. Murray,
0014900	001.19561/tur	03.02.2003	16:12:28	21.01.2003	00:20:35	10243090xx	Bilgelerin Yolunda / Sadık Tural; Ed. Elmas Kılıç / A [001.195
0022398	005.1/çak	03.02.2003	16:12:36	21.01.2003	00:20:35	10243090xx	Bilgisayar 2 : Excel paket programı bilgisayrlı muhasebe / İb
0026245	617.96/roi	03.02.2003	16:22:14	18.01.2003	00:20:35	A0000xx	Klinik anestezi esasları / Michael F. Roizen, Lee A. Fleisher ; ç
0025319	617.96/kay	03.02.2003	16:22:24	18.01.2003	00:20:35	A0000xx	Klinik anestezi / Zeynep Kayhan / A [617.96/kay] 05 Okuma
0016097	828.1/mit 1.	04.02.2003	13:32:48	24.01.2003	00:20:35	10143090xx	Rüzgâr Gibi Geçti / Margaret Mitchell, çeviren Fuat İstanbull
0025460	828.33/koo	04.02.2003	13:32:55	24.01.2003	00:20:35	10143090xx	Kurbanlar=phantoms / Dean R. Koontz, çeviren Mehmet Harn

Şekil 9: Ödünç Tablosundaki Verilerin Ham Hali

Şekil 9, kütüphane kullanıcılarının yayın sirkülasyon bilgilerini içermektedir. Demirbaş ve üye numaraları bilgilerinin yanında ödünç alınan yayının adı, sınıflaması, ödünç tarihi ve zamanı ile iade tarihi ve zamanı bilgilerini içermektedir. Sınıflama alanındaki verileri anlamlandırmak amacıyla kitap kodlarının hangi alana karşılık geldiğinin belirlenmesi için Trakya Üniversitesi Kütüphane ve Dokümantasyon Daire Başkanlığı internet sayfasındaki Dewey Onlu Sınıflama Sistemi (Dewey Decimal Classification-DDC) şemasından yararlanılmıştır.

DDC sisteminde on tane temel konu grubu bulunmaktadır. Bu konu gruplarının sayı aralıkları şu şekilde düzenlenmiştir:

- 000-099 Genel Konular
- 100-199 Felsefe ve Psikoloji
- 200-299 Din
- 300-399 Toplum Bilimleri
- 400-499 Dil ve Dil Bilim
- 500-599 Doğa Bilimleri ve Matematik
- 600-699 Teknoloji (Uygulamalı Bilimler)
- 700-799 Sanat (Güzel Sanatlar)
- 800-899 Edebiyat (Retorik)
- 900-999 Coğrafya ve Tarih

Ödünç tablosunun sınıflama alanı da DDC sistemindeki on temel konu grubunun içinde yer alan alt konu gruplarına göre yeniden düzenlenmiştir.

Trakya Üniversitesi Merkez Kütüphane Yönergesinde belirtilen hususlar çerçevesinde ödünç yayın verme süresi ön lisans ve lisans öğrencileri için on beş gün, yüksek lisans ve doktora öğrencileri için ise otuz gündür. Şekil 9'daki ödünç tablosunda, ödünç alınan yayının zamanında veya gecikmeli teslim edildiğini görmek için ilk adımda, iade tarihinden ödünç tarihi matematiksel fark işlemi yapılarak oluşturulan **İade Süresi** alanına aktarılmıştır. İkinci adımda, **İade Durumu** alanı oluşturulmuştur. Son adımda ise iade süresi, Trakya Üniversitesi Merkez Kütüphane Yönergesinde belirtilen ödünç verme süre sınırlarının altında olan kayıtlara **Zamanında Teslim**, bu süre sınırının aşıldığı kayıtlar ise **Gecikmeli Teslim** olarak iade durumu alanında kodlanmıştır. Eser adı, ödünç ve iade zamanı, iade süresi alanları da veri madenciliği aşamasına dâhil edilmemiştir.

Tablo 5: Veri Önişleme İşlemleri Sonu Veri Madenciliğinde Kullanılacak Tabloların Durumu

Tablo Adı	Alan Sayısı	Kayıt Sayısı
Üye	5	23.227
Ödünç	6	209.769

Üye ve ödünç tabloları üzerinde yapılan veri önişleme işlemleri Microsoft Excel ve Access sorguları kullanılarak gerçekleştirilmiştir. Ancak RapidMiner yazılımı aracılığıyla veri ambarı oluşturulurken ihtiyaca göre veri dönüştürme işlemleri veri madenciliğinin tüm aşamalarında devam etmektedir. Tablo 5'te veri önişleme işlemleri sonucunda veri madenciliğinde kullanılacak olan nihai tabloların durumu gösterilmektedir.

uye_no	cinsiyet	egitim_programi	akademik_birim	bolum	yayin_sinifi	teslim_durumu
21316513xx	Erkek	Lisans Öğrencisi	Eğitim Fakültesi	Yabancı Diller Eğitimi	Mühendislik Bilimi	Gecikmeli teslim
21316513xx	Kadın	Lisans Öğrencisi	İktisadi ve İdari Bilimler Fakültesi	Çalışma Ekonomisi ve Endüstri İlişkileri	Mühendislik Bilimi	Gecikmeli teslim
21002013xx	Erkek	Lisans Öğrencisi	Mühendislik Fakültesi	Makine Mühendisliği	Mühendislik Bilimi	Zamanında teslim
11768250xx	Erkek	Lisans Öğrencisi	Tıp Fakültesi	Tıp	Mühendislik Bilimi	Zamanında teslim
11768250xx	Erkek	Lisans Öğrencisi	Tıp Fakültesi	Tıp	Mühendislik Bilimi	Zamanında teslim
11716040xx	Erkek	Lisans Öğrencisi	Mühendislik Fakültesi	Elektrik-Elektronik Mühendisliği	Mühendislik Bilimi	Gecikmeli teslim
11716040xx	Erkek	Lisans Öğrencisi	Mühendislik Fakültesi	Elektrik-Elektronik Mühendisliği	Mühendislik Bilimi	Gecikmeli teslim
11716040xx	Kadın	Lisans Öğrencisi	Mühendislik Fakültesi	Makine Mühendisliği	Mühendislik Bilimi	Zamanında teslim
11716040xx	Kadın	Ön Lisans Öğrencisi	Şehit Ressam Hasan Rıza Güzel Sanatlar Meslek Yüksekokulu	Mimarlık ve Şehir Planlama	Mühendislik Bilimi	Gecikmeli teslim
11716040xx	Kadın	Ön Lisans Öğrencisi	Şehit Ressam Hasan Rıza Güzel Sanatlar Meslek Yüksekokulu	Mimarlık ve Şehir Planlama	Mühendislik Bilimi	Gecikmeli teslim
11716040xx	Kadın	Lisans Öğrencisi	Tıp Fakültesi	Tıp	Mühendislik Bilimi	Gecikmeli teslim
11616040xx	Erkek	Lisans Öğrencisi	Mühendislik Fakültesi	Elektrik-Elektronik Mühendisliği	Mühendislik Bilimi	Gecikmeli teslim
11616040xx	Erkek	Lisans Öğrencisi	Mühendislik Fakültesi	Elektrik-Elektronik Mühendisliği	Mühendislik Bilimi	Gecikmeli teslim
11616040xx	Erkek	Lisans Öğrencisi	Mühendislik Fakültesi	Elektrik-Elektronik Mühendisliği	Mühendislik Bilimi	Gecikmeli teslim

Şekil 10: Verilerin Ön İşlemden Geçmiş Hali

Şekil 10'da, üye ve ödünç tablolarının yukarıda belirtilen ön işlemlerden geçirilerek ilişkili tablolar Microsoft Office Access veri tabanında birleştirilmiştir. Veri setindeki veriler, Microsoft Office Excel ortamında .csv formatında kaydedilerek veri madenciliği aşamasında kullanılmak üzere RapidMiner'a aktarılacaktır.

3.5. Araştırma Bulguları

Bu bölümde kütüphane kullanıcılarının ödünç sirkülasyon verileri üzerinde yapılan çeşitli kullanıcı tanımlayıcı ve kümeleme analizine ilişkin bulgular yer almaktadır.

3.5.1. Kütüphane Kullanıcı İstatistikleri

Çalışmada ilk olarak kütüphaneden en az bir yayın ödünç almış kullanıcılara ait veriler, IBM SPSS Statistics 24 programı üzerinden frekans tekniği kullanılarak incelenmiştir. Bunlardan ilki kullanıcılara ait cinsiyet dağılımdır.

Tablo 6: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Cinsiyete Göre Dağılımı

Değişken	Grup	Frekans	%
Cinsiyet	Kadın	14.313	61,6
	Erkek	8.914	38,4
	Toplam	23.227	100,0

Tablo 6'dakütüphaneden en az bir yayın ödünç almış kullanıcıların cinsiyete göre dağılımı incelendiğinde %61,6'sı kadın ve %38,4'ü erkek olduğu görülmektedir.

Tablo 7: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Eğitim Programına Göre Dağılımı

Değişken	Grup	Frekans	%
Eğitim Programı	Lisans Öğrencisi	18.597	80,1
	Yüksek Lisans	1.645	7,1
	Ön Lisans	2.842	12,2
	Doktora	143	0,6
	Toplam	23.227	100,0

Tablo 7'de, en az bir yayın almış kullanıcıların eğitim programına göre dağılımı incelendiğinde; en çok kütüphaneden ödünç yayın alan grubun lisans öğrencileri (%80,1) olduğu açık bir şekilde görülmektedir. Geri kalan kütüphane kullanıcılarının %12,2'si ön lisans, %7,1'i yüksek lisans ve %0,6'sı ise doktora öğrencisidir.

Tablo 8: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Üniversitedeki Akademik Birimlere Göre Dağılımı

	Değişken	Frekans	%
Akademik Birimler	**İktisadi ve İdari Bilimler Fakültesi**	4.293	18,48
	Eğitim Fakültesi	3.305	14,23
	Edebiyat Fakültesi	2.721	11,71
	Sağlık Bilimleri Fakültesi	1.942	8,36
	Mühendislik Fakültesi	1.446	6,23
	Sosyal Bilimler Enstitüsü	1.342	5,78
	Tıp Fakültesi	1.303	5,61
	Sağlık Hizmetleri Meslek YO	1.009	4,34
	Fen Fakültesi	983	4,23
	Kırkpınar Beden Eğitimi ve Spor YO	686	2,95
	Edirne Sosyal Bilimler Meslek YO	667	2,87

Tablo 8 (Devam)

	Değişken	Frekans	%
Akademik Birimler	Uygulamalı Bilimler Yüksekokulu	626	2,7
	Mimarlık Fakültesi	532	2,29
	Edirne Teknik Bilimler Meslek YO	488	2,1
	Havsa Meslek Yüksekokulu	337	1,45
	Fen Bilimleri Enstitüsü	277	1,19
	Eczacılık Fakültesi	209	0,9
	Şehit Ressam Hasan Rıza Güzel Sanatlar MYO	206	0,89
	Sağlık Bilimleri Enstitüsü	169	0,73
	Diş Hekimliği Fakültesi	147	0,63
	İlahiyat Fakültesi	141	0,61

Tablo 8 (Devam)

	Değişken	Frekans	%
Akademik Birimler	**Güzel Sanatlar Fakültesi**	108	0,46
	Arda Meslek Yüksekokulu	104	0,45
	Keşan Yusuf Çapraz Uygulamalı Bilimler YO	77	0,33
	Devlet Konservatuvarı	49	0,21
	Yabancı Diller Yüksekokulu	25	0,11
	Uzunköprü Meslek Yüksekokulu	22	0,09
	İpsala Meslek Yüksekokulu	5	0,02
	Keşan Hakkı Yörük Sağlık Yüksekokulu	4	0,02
	Keşan Meslek Yüksekokulu	2	0,01
	Tunca Meslek Yüksekokulu	2	0,01
	Toplam	23.227	100,00

Tablo 8'de, kütüphaneden en az bir yayın almış kullanıcıların, Trakya Üniversitesi'nde kayıtlı oldukları akademik birimlere göre dağılımı incelenmiştir. İlk beş sırada yer alan akademik birimler; İktisadi ve İdari Bilimler Fakültesi (%18,48), Eğitim Fakültesi (%14,23), Edebiyat Fakültesi (%11,71), Sağlık Bilimleri Fakültesi (%8,36) ve Mühendislik Fakültesi (%6,23) tüm akademik birimler içerisinde %59,01'lik kısmı oluşturmaktadır.

Tablo 9: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Üniversitedeki Bölümlerine Göre Dağılımı (İlk ve Son Üç Bölüm)

Değişken	Grup	Frekans	%
Bölüm	Tıp	1.303	5,61
	Temel Eğitim	1.043	4,49
	Tıbbi Hizmetler ve Teknikler	1.009	4,34
	Çocuk Bakımı ve Gençlik Hizmetleri	1	0,00
	Felsefe ve Din Bilimleri A.B.D.	1	0,00
	İmmünoloji A.B.D.	1	0,00

Tablo 9'da, kütüphaneden en az bir yayın almış kullanıcıların kayıtlı oldukları bölümlerden ilk ve son üçü gösterilmektedir. Kütüphaneden en çok yayın alan ilk üç bölüme bakıldığında ilk sırada %5,61 oranla Tıp bölümü olduğu görülmektedir. Son üç bölüme bakıldığında ise Çocuk Bakımı ve Gençlik Hizmetleri, Felsefe ve Din Bilimleri A.B.D. ve İmmünoloji A.B.D. bölümlerinden birer öğrencinin kütüphaneden ödünç yayın aldığı görülmektedir.

Çalışmada, her bir akademik birim bölümlerinde eğitim göre öğrencilerin cinsiyete göre dağılımları EK1'de verilmiş olup İktisadi ve İdari Bilimler Fakültesi bölümlerinde eğitim gören öğrencilerin cinsiyete göre dağılımları incelenmiştir.

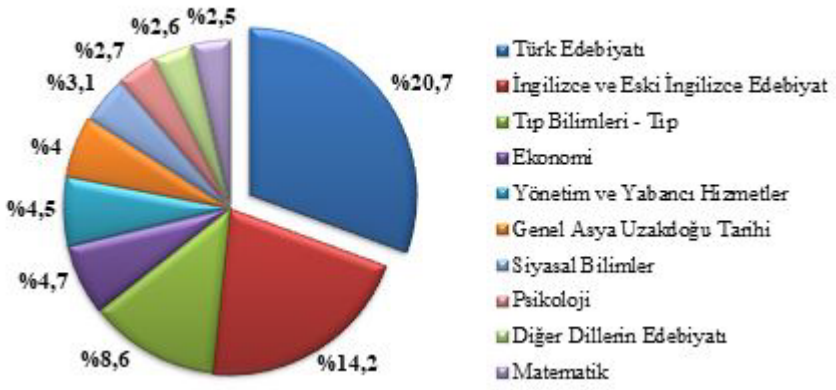
Tablo 10: Kütüphaneden En Az Bir Yayın Ödünç Almış İktisadi ve İdari Bilimler Fakültesi Bölümlerinde Eğitim Gören Öğrencilerin Cinsiyete Göre Dağılımı

Bölüm	Cinsiyet	Frekans	%
Çalışma Ekonomisi ve Endüstri İlişkileri	**Erkek**	118	2,75
	Kadın	284	6,62
Ekonometri	**Erkek**	167	3,89
	Kadın	321	7,48
İktisat	**Erkek**	346	8,06
	Kadın	410	9,55
İşletme	**Erkek**	310	7,22
	Kadın	375	8,74

Tablo 10 (Devam)

Maliye	**Erkek**	247	5,75
	Kadın	335	7,8
Siyaset Bilimi ve Kamu Yönetimi	**Erkek**	327	7,62
	Kadın	387	9,01
Uluslararası İlişkiler	**Erkek**	294	6,85
	Kadın	372	8,67
	Toplam	4.293	100,00

Tablo 10'da, kütüphaneden en az bir yayın ödünç almış İktisadi ve İdari Bilimler Fakültesi bölümlerinde öğrenim gören öğrencilerin cinsiyet dağılımlarına bakıldığında, kadın oranının daha fazla olduğu görülmektedir. Çalışma Ekonomisi ve Endüstri İlişkileri bölümünün %6,62'si, Ekonometri bölümünün %7,48'i, İktisat bölümünün %9,55'i, İşletme bölümünün %8,74'ü, Maliye bölümünün %7,8'i, Siyaset Bilimi ve Kamu Yönetimi bölümünün %9,01'i ve Uluslararası İlişkiler bölümünün %8,67'si kadınlardan oluşmaktadır.



Şekil 11: Kütüphaneden Ödünç Alınan Yayınların Sınıflarına Göre Dağılımı (İlk On Yayın Sınıfı)

Şekil 11 incelendiğinde, Türk Edebiyatı sınıfı yayınlarının ödünç alınan yayınlar arasında 43.323 ile en çok alınan ödünç yayın olduğu görülmektedir. Bu sayı ödünç alınan tüm eserler arasında %20,7'e denk gelmektedir. Sırasıyla İngilizce ve Eski İngilizce Edebiyatı ve Tıp Bilimleri-Tıp sınıflarındaki yayınlar 29.790 ve

18.010 yayınla kütüphaneden en çok ödünç alınan 2. ve 3. yayın grubudur. Tüm ödünç alınan yayınlar içerisinde İngilizce ve Eski İngilizce Edebiyatı yayınları %14,2 ve Tıp Bilimleri-Tıp yayınları %8,6 oranında bulunmaktadır. Ekonomi, Yönetim ve Yabancı Hizmetler, Genel Asya Uzakdoğu Tarihi, Siyasal Bilimler, Psikoloji, Diğer Dillerin Edebiyatı ve Matematik yayın gruplarının ödünç alınan tüm yayınlar içerisindeki yüzdeleri sırasıyla 4,7, 4,5, 4, 3,1, 2,7, 2,6 ve 2,5'tir. En çok ödünç alınan on yayın sınıfı, ödünç alınan tüm eserlerin %67,6'sını oluşturmaktadır.

Tablo 11: Kütüphaneden En Az Bir Yayın Ödünç Almış Kullanıcıların Yayınları Teslim Etme Durumuna Göre Dağılımı

Değişken	Grup	Frekans	%
Teslim Durumu	**Gecikmeli Teslim**	106.719	50,9
	Zamanında Teslim	103.050	49,1
	Toplam	209.769	100,0

Tablo 11 incelendiğinde, kütüphaneden en az bir yayın ödünç almış kullanıcıların %50,9'unun aldıkları yayınları kütüphaneye gecikmeli teslim ettiği görülmektedir. Ödünç alınan yayınları zamanında teslim edenlerin ise oranı %49,1'dir.

Tablo 12: Kütüphaneden Ödünç Alınan Yayınların Cinsiyete Göre Teslim Durumu Dağılımı

Cinsiyet	Teslim Durumu	Frekans	%
Erkek	Gecikmeli Teslim	37.298	17,78
	Zamanında Teslim	34.014	16,21
Kadın	Gecikmeli Teslim	69.421	33,09
	Zamanında Teslim	69.036	32,91
	Toplam	209.769	100,00

Tablo 12'de kütüphaneden ödünç yayın alan kadınların %33,09'u ve erkeklerin %17,78'i yayınları iade tarihinden sonra teslim ettiği görülmektedir. Kütüphaneden alınan ödünç yayınları zamanında teslim eden kullanıcıların %16,21'i erkek, %32,91'i kadındır.

Tablo 13: Kütüphaneden Ödünç Alınan Yayınların Eğitim Programına Göre Teslim Durumu Dağılımı

Eğitim Programı	Teslim Durumu	Frekans	%
Doktora	Gecikmeli Teslim	876	0,42
	Zamanında Teslim	1.016	0,48
Lisans	Gecikmeli Teslim	90.761	43,27
	Zamanında Teslim	84.948	40,50
Ön Lisans	Gecikmeli Teslim	8.336	3,97
	Zamanında Teslim	7.722	3,68
Yüksek Lisans	Gecikmeli Teslim	6.746	3,22
	Zamanında Teslim	9.364	4,46
	Toplam	209.769	100,00

Kütüphaneden ödünç alınan kitapların eğitim programına göre teslim durumu dağılımının incelendiği Tablo 13'te, doktora öğrencilerinin %0,48'i ve yüksek lisans öğrencilerinin de %4,46'sı ödünç aldığı yayınları zamanında teslim etmiştir. Lisans öğrencilerinin %43,27'si ve ön lisans öğrencilerinin %3,97'si ise aldıkları yayınları iade tarihi geçtikten sonra teslim etmişlerdir.

Tablo 14: Kütüphaneden Ödünç Alınan Yayınların Akademik Birimlere Göre Teslim Durumu Dağılımı

Akademik Birimler	Durum	Frekans	%
Arda MYO	Gecikmeli Teslim	264	50,3
	Zamanında Teslim	261	49,7
Devlet Konservatuvarı	Gecikmeli Teslim	235	45,9
	Zamanında Teslim	277	54,1
Diş Hekimliği Fakültesi	Gecikmeli Teslim	548	54,1
	Zamanında Teslim	465	45,9
Eczacılık Fakültesi	Gecikmeli Teslim	933	55,3
	Zamanında Teslim	753	44,7
Edebiyat Fakültesi	Gecikmeli Teslim	13.286	50,1
	Zamanında Teslim	13.254	49,9
Edirne Sosyal Bilimler MYO	Gecikmeli Teslim	1.756	53,8
	Zamanında Teslim	1.505	46,2

Tablo 14 (Devam)

Akademik Birimler	Durum	Frekans	%
Edirne Teknik Bilimler Meslek Yüksekokulu	Gecikmeli Teslim	1.331	51,1
	Zamanında Teslim	1.272	48,9
Eğitim Fakültesi	Gecikmeli Teslim	13.904	48,6
	Zamanında Teslim	14.703	51,4
Fen Bilimleri Enstitüsü	Gecikmeli Teslim	1.062	43,8
	Zamanında Teslim	1.363	56,2
Fen Fakültesi	Gecikmeli Teslim	5.230	55,5
	Zamanında Teslim	4.189	44,5
Güzel Sanatlar Fakültesi	Gecikmeli Teslim	356	62,7
	Zamanında Teslim	212	37,3
Havsa Meslek Yüksekokulu	Gecikmeli Teslim	984	64,3
	Zamanında Teslim	547	35,7
İktisadi ve İdari Bilimler Fakültesi	Gecikmeli Teslim	21.494	51,1
	Zamanında Teslim	20.553	48,9
İlahiyat Fakültesi	Gecikmeli Teslim	437	49,2
	Zamanında Teslim	451	50,8
İpsala MYO	Gecikmeli Teslim	12	75,0
	Zamanında Teslim	4	25,0
Keşan Hakkı Yörük Sağlık YO	Gecikmeli Teslim	10	71,4
	Zamanında Teslim	4	28,6
Keşan MYO	Zamanında Teslim	4	100,0
Keşan Yusuf Çapraz Uygulamalı Bilimler YO	Gecikmeli Teslim	146	44,0
	Zamanında Teslim	186	56,0
Kırkpınar Beden Eğitimi ve Spor YO	Gecikmeli Teslim	2.520	55,5
	Zamanında Teslim	2.021	44,5
Mimarlık Fakültesi	Gecikmeli Teslim	2.175	56,8
	Zamanında Teslim	1.654	43,2

Tablo 14 (Devam)

Akademik Birimler	Durum	Frekans	%
Mühendislik Fakültesi	Gecikmeli Teslim	6.360	54,7
	Zamanında Teslim	5.275	45,3
Sağlık Bilimleri Enstitüsü	Gecikmeli Teslim	631	42,5
	Zamanında Teslim	855	57,5
Sağlık Bilimleri Fakültesi	Gecikmeli Teslim	10.734	47,0
	Zamanında Teslim	12.082	53,0
Sağlık Hizmetleri Meslek Yüksekokulu	Gecikmeli Teslim	3.335	48,5
	Zamanında Teslim	3.535	51,5
Sosyal Bilimler Enstitüsü	Gecikmeli Teslim	5.929	42,1
	Zamanında Teslim	8.162	57,9
Şehit Ressam Hasan Rıza Güzel Sanatlar MYO	Gecikmeli Teslim	569	50,2
	Zamanında Teslim	564	49,8
Tıp Fakültesi	Gecikmeli Teslim	10.053	60,9
	Zamanında Teslim	6.442	39,1
Tunca MYO	Gecikmeli Teslim	1	25,0
	Zamanında Teslim	3	75,0
Uygulamalı Bilimler YO	Gecikmeli Teslim	2.248	49,0
	Zamanında Teslim	2.336	51,0
Uzunköprü MYO	Gecikmeli Teslim	84	75,7
	Zamanında Teslim	27	24,3
Yabancı Diller YO	Gecikmeli Teslim	92	50,3
	Zamanında Teslim	91	49,7
	Toplam	209.769	100,0

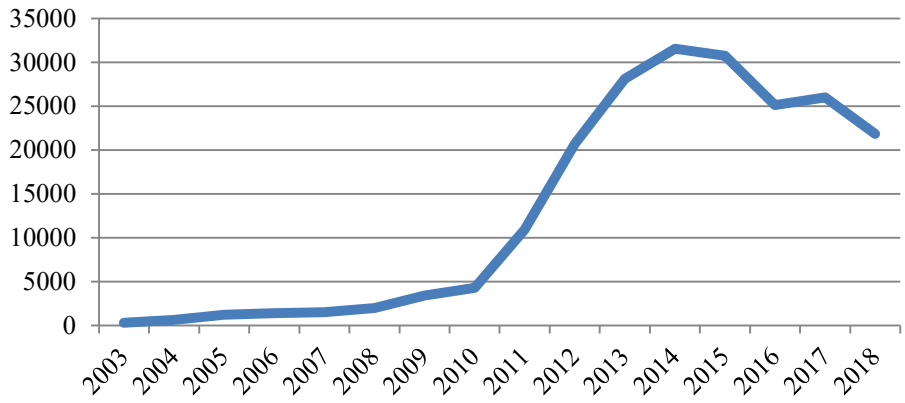
Kütüphaneden ödünç alınan yayınların akademik birimlere göre teslim durumunu gösteren Tablo 14'te, Keşan Meslek Yüksekokulu'nda öğrenim görüp ödünç yayın alan dört öğrencinin de aldığı yayınları zamanında teslim ettiği görülmektedir. Fen Bilimleri Enstitüsü'nün %56,2'si, Sağlık Bilimleri Enstitüsü'nün %57,5'i ve Sosyal Bilimler Enstitüsü'nün %57,9'u ödünç alınan yayınları zamanında

teslim etmiştir. İktisadi ve İdari Bilimler Fakültesi'nde öğrenim gören öğrencilerin %51,1'i kütüphaneden ödünç alınan yayınları iade tarihinden sonra teslim ettikleri görülmektedir.

Tablo 15: Kütüphaneden Ödünç Alınan Yayınların Alındığı Aylara Göre Dağılımı

Değişken	Grup	Frekans	%
Yayın Ödünç Alma Ayı	**Ekim**	34.743	16,563
	Aralık	29.960	14,282
	Kasım	26.313	12,544
	Mart	25.074	11,953
	Şubat	22.778	10,859
	Nisan	20.795	9,913
	Eylül	15.773	7,519
	Mayıs	14.091	6,717
	Ocak	12.262	5,845
	Haziran	3.878	1,849
	Temmuz	2.240	1,068
	Ağustos	1.862	0,888
	Toplam	209.769	100,000

Kütüphane kullanıcılarının yayın ödünç alma sirkülasyonu aylara göre dağılımını gösteren Tablo 15'te, en çok yayının ödünç alındığı ilk üç ay sırasıyla %16,563 oranla Ekim ayı, %14,282 oranla Aralık ayı ve %12,544 ile Kasım ayıdır. En az yayının ödünç alındığı aylar ise %1,85 oranla Haziran ayı, %1,07 oranla Temmuz ayı ve %0,89 oranla Ağustos ayıdır. Haziran, Temmuz ve Ağustos aylarında yayın ödünç alma yüzdesinin diğer aylara göre düşük olmasının nedeni öğrencilerin akademik yaz tatilinde olmasındandır.

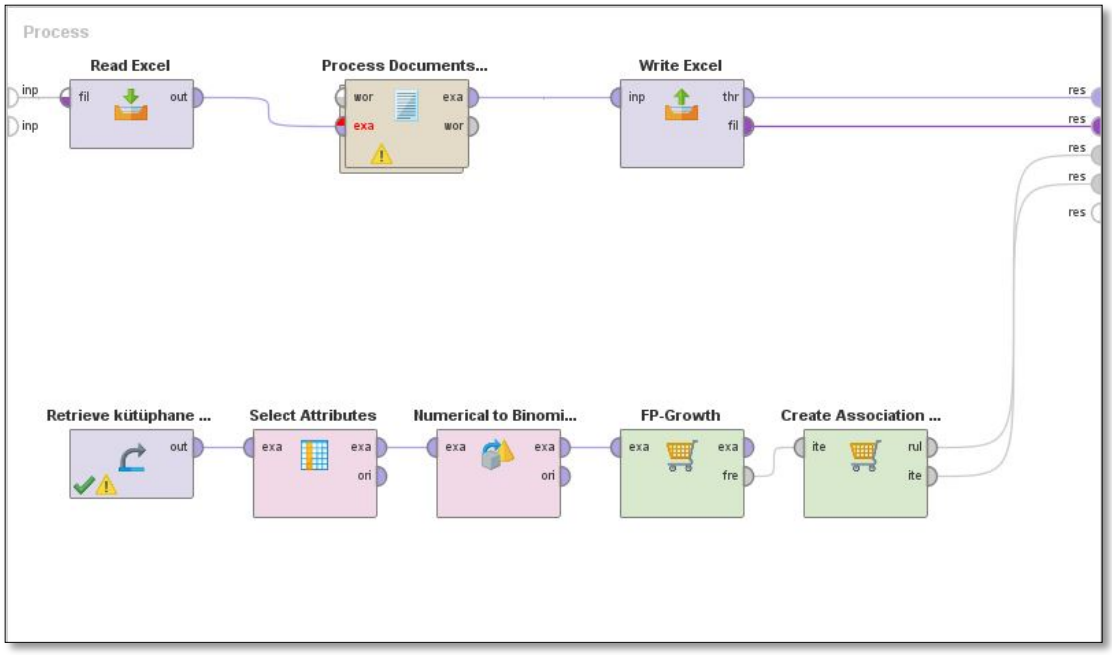


Şekil 12: Kütüphaneden Ödünç Alınan Yayınların Alındığı Yıllara Göre Dağılımı

Şekil 12 incelendiğinde, en çok yayın alınan yılın 31.544 yayınla 2014 yılı ve en az yayın alınan yılın ise 323 yayınla 2003 yılında olduğu belirlenmiştir. Trakya Üniversitesi'ne kayıtlı öğrenci sayısındaki artışla birlikte 2003 ve 2014 yılları arasında kütüphaneden ödünç alınan kitapların sayısında devamlı bir artış olduğu açık bir şekilde görülmektedir. Kütüphane tarafından sağlanabilen olanakların artması ve kütüphane kullanıcılarının elektronik kaynaklara yönelmesi etkisiyle 2014 yılından sonrasında kullanıcıların yayın ödünç alım sayısında azalma olduğunu söylemek mümkündür.

3.5.2. Kütüphane Ödünç Yayın Sirkülasyon Verileri Üzerine Birliktelik Analizi

Bu çalışmada, kütüphaneden en az bir yayın ödünç almış kullanıcıların bilgileri ile ödünç yayın sirkülasyon verileri arasında var olabilecek birliktelik kurallarının saptaması yapılmıştır. Ön işlemin ardından Excel ortamında toplanan veri setindeki 209.769 verideki kategorik alanlar, sayısal alanlara dönüştürülerek birliktelik analizine uygun hale getirilmiştir. Buradaki amaç, kütüphane içindeki ödünç yayın sirkülasyon bilgilerine göre kullanıcıların en çok hangi yayınları aldığı ve tercih ettiği yayın sınıflarını belirlemektir.



Şekil 13: RapidMiner Üzerinde Oluşturulan Veri Madenciliği Süreci

Çalışmada, birliktelik analizinde uygulanan veri madenciliği süreci ve işlem bağlantıları Şekil 13'te gösterilmiştir. Şekil 13, tek süreç oluşturularak iki adımda gerçekleşen veri madenciliği sürecini temsil etmektedir. Veri madenciliği sürecinin ilk adımında, kütüphane ödünç yayın veri seti **Read Excel** operatörüyle sürece eklenmiş, kategorik yapıda olan veri seti **Process Documents from Data** operatörü kullanılarak sayısal yapıya dönüştürülmüş ve dönüştürülen veriler **Write Excel** operatörüyle ilgili Excel dosyasına aktarılmıştır. İkinci adımda ise dönüştürülen veri seti, RapidMiner veri ambarına yüklenerek öznitelik seçimi yapılmış ve numerical yapıda olan veriler analiz için binominal yapılara dönüştürülerek birliktelik kuralı oluşturma işlemleri uygulanmıştır. FP-Growth algoritması kullanılarak birliktelik kuralları oluşturulmuştur.

FP-Growth algoritmasında veri tabanı iki defa taranmaktadır. İlk taramada, algoritma içinde belirlenen eşik değeri baz alınarak tüm öğelerin destek değerleri hesaplanır. Belirlenen destek eşik değerine eşit ve bu değerden büyük olan öğeler F listesine eklenerek destek değerleri büyükten küçüğü sıralanır. İkinci taramada ise oluşturulan kurallardaki öğe destek değerleri sıralanıp sıkıştırılarak ağaç yapısı oluşturulmaktadır. Bu süreç içerisinde yer alan bir öğenin ağaçta olmaması

durumunda o öğe için yeni bir düğüm oluşturulur ve destek değeri 1 olarak hesaplanır. Eğer o öğe daha önce ağaçta yer alıyorsa o düğümün destek değeri 1 arttırılır (Özdoğan, 2010: 27).

Kütüphane kullanıcılarının verileri ve yayın sınıflandırmaları arasında uygulanan FP-Growth algoritması için minimum destek değeri %0,01, minimum frekans değeri 10, minimum güven değeri %0,1 ve maksimum öncül öğe sayısı 2 olarak belirlenmiştir. Kütüphane ödünç yayın veri seti işlemsel kayıtları içerdiğinden değerler düşük tutulmuştur. Belirlenen kurallar, 2003-2018 yıllarında Trakya Üniversitesi Merkez Kütüphanesi'nde gerçekleşen 209.769 ödünç yayın verisinde uygulanmıştır. Birliktelik analizi sonucunda 57 kural ve 103 işlemsel kayıt ortaya çıkmıştır. Analiz sonucu oluşturulan kurallar Tablo 16'da gösterilmiştir. Kütüphane kullanıcılarının eğitim gördüğü akademik birim ve ödünç aldığı yayın sınıfları arasındaki ilişki hesaplanmıştır. Birliktelik kurallarında en çok kural oluşturan yayın sınıfı grupları **Türk Edebiyatı** ve **İngilizce ve Eski İngilizce Edebiyatı**'dır. Akademik birimler içerisinde en yoğun kural oluşturan grup ise **İktisadi ve İdari Bilimler Fakültesi**'dir.

Tablo 16: Birliktelik Kuralı Tablosu

Birliktelik Kuralları
[İNGİLİZCE VE ESKİ İNGIİLİZCE EDEBİYAT] --> [EDEBİYAT FAKÜLTESİ] (confidence: 0.107)
[GENEL ASYA UZAKDOĞU TARİHİ] --> [EĞİTİM FAKÜLTESİ] (confidence: 0.109)
[İKTİSADİ VE İDARİ BİLİMLER FAKÜLTESİ] --> [EKONOMİ] (confidence: 0.109)
[SOSYAL BILIMLER ENSTITÜSÜ] --> [EKONOMI] (confidence: 0.118)
[TIP FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.119)
[EDEBIYAT FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.120)
[FEN FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.125)
[TIP FAKÜLTESI] --> [TÜRK EDEBIYATI] (confidence: 0.126)

Tablo 16 (Devam)

[TÜRK EDEBIYATI] --> [SAĞLIK BILIMLERI FAKÜLTESI] (confidence: 0.129)
[İKTISADI VE İDARI BILIMLER FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.130)
[GENEL ASYA UZAKDOĞU TARIHI] --> [SOSYAL BILIMLER ENSTITÜSÜ] (confidence: 0.141)
[MÜHENDISLIK FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.142)
[FEN FAKÜLTESI] --> [TÜRK EDEBIYATI] (confidence: 0.150)
[SOSYAL BILIMLER ENSTITÜSÜ] --> [YÖNETIM VE YARDIMCI HIZMETLER] (confidence: 0.151)
[MATEMATIK] --> [MÜHENDISLIK FAKÜLTESI] (confidence: 0.156)
[INGILIZCE VE ESKI İNGILIZCE EDEBIYAT] --> [SAĞLIK BILIMLERI FAKÜLTESI] (confidence
[GENEL ASYA UZAKDOĞU TARIHI] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.159)
[MÜHENDISLIK FAKÜLTESI] --> [TÜRK EDEBIYATI] (confidence: 0.168)
[EKONOMI] --> [SOSYAL BILIMLER ENSTITÜSÜ] (confidence: 0.170)
[DIĞER DILLERIN EDEBIYATI] --> [EĞITIM FAKÜLTESI] (confidence: 0.170)
[İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] --> [EĞITIM FAKÜLTESI] (confidence: 0.170)
[TÜRK EDEBIYATI] --> [EĞITIM FAKÜLTESI] (confidence: 0.173)
[EĞITIM FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.177)
[TÜRK EDEBIYATI] --> [EDEBIYAT FAKÜLTESI] (confidence: 0.180)
[TÜRK EDEBIYATI] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.181)
[İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.184)
[SAĞLIK BILIMLERI FAKÜLTESI] --> [TIP BILIMLERI - TIP] (confidence: 0.185)
[DIĞER DILLERIN EDEBIYATI] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.186)
[İKTISADI VE İDARI BILIMLER FAKÜLTESI] --> [TÜRK EDEBIYATI] (confidence: 0.186)
[FEN FAKÜLTESI] --> [MATEMATIK] (confidence: 0.186)
[PSIKOLOJI] --> [EĞITIM FAKÜLTESI] (confidence: 0.189)
[SIYASAL BILIMLER] --> [SOSYAL BILIMLER ENSTITÜSÜ] (confidence: 0.190)

Tablo 16 (Devam)

[UYGULAMALI BILIMLER YÜKSEKOKULU] --> [TÜRK EDEBIYATI] (confidence: 0.195)
[PSIKOLOJI] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.195)
[MATEMATIK] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.200)
[SAĞLIK BILIMLERI FAKÜLTESI] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.204)
[SAĞLIK HIZMETLERI MESLEK YÜKSEKOKULU] --> [İNGILIZCE VE ESKI İNGILIZCE EDEBIYAT] (confidence: 0.216)
[YÖNETİM VE YARDIMCI HİZMETLER] --> [SOSYAL BİLİMLER ENSTİTÜSÜ] (confidence: 0.226)
[TIP BİLİMLERİ - TIP] --> [SAĞLIK BİLİMLERİ FAKÜLTESİ] (confidence: 0.234)
[SAĞLIK BİLİMLERİ FAKÜLTESİ] --> [TÜRK EDEBİYATI] (confidence: 0.244)
[KIRKPINAR BEDEN EĞİTİMİ VE SPOR YÜKSEK OKULU] --> [TÜRK EDEBİYATI] (confidence: 0.246)
[SAĞLIK HIZMETLERI MESLEK YÜKSEKOKULU] --> [TÜRK EDEBİYATI] (confidence: 0.249)
[EĞİTİM FAKÜLTESİ] --> [TÜRK EDEBİYATI] (confidence: 0.262)
[EDIRNE SOSYAL BILIMLER MESLEK YÜKSEKOKULU] --> [TÜRK EDEBIYATI] (confidence: 0.270)
[EDEBIYAT FAKÜLTESI] --> [TÜRK EDEBIYATI] (confidence: 0.294)
[GENEL ASYA UZAKDOĞU TARIHI] --> [EDEBIYAT FAKÜLTESI] (confidence: 0.313)
[MATEMATIK] --> [FEN FAKÜLTESI] (confidence: 0.333)
[TOPLUM BILIMLERI] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.370)
[YÖNETIM VE YARDIMCI HIZMETLER] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.389)
[TIP BILIMLERI - TIP] --> [TIP FAKÜLTESI] (confidence: 0.404)
[SIYASAL BILIMLER] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.434)
[TÜRK DILI] --> [EDEBIYAT FAKÜLTESI] (confidence: 0.435)
[TIP FAKÜLTESI] --> [TIP BILIMLERI - TIP] (confidence: 0.441)
[EKONOMI] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.466)

Tablo 16 (Devam)

[MÜHENDISLIK VE İLGILI İŞLER] --> [MÜHENDISLIK FAKÜLTESI] (confidence: 0.479)
[EĞITIM] --> [EĞITIM FAKÜLTESI] (confidence: 0.492)
[HUKUK] --> [İKTISADI VE İDARI BILIMLER FAKÜLTESI] (confidence: 0.511)

Şekil 14'te İktisadi ve İdari Bilimler Fakültesi grubunda oluşan birliktelik kuralı grafik halinde gösterilmiştir. Şekil 14 incelendiğinde, İktisadi ve İdari Bilimler Fakültesi grubunda oluşan kuralların destek ve güven değerleriyle hangi yayın sınıfları arasında ilişkinin kurulduğu belirtilmektedir. Kural 57'de, destek değeri 0,009 ve güven değeri 0,511'dir. Hukuk sınıfında yayınların %51'i İktisadi ve İdari Bilimler Fakültesi'nde öğrenim gören kullanıcılar tarafından ödünç alındığı görülmektedir. Kural 57'deki destek değerinin %1 olması analize dahil edilen tüm yayın sınıfları arasında Hukuk sınıfı yayınlarının oranını temsil etmektedir.



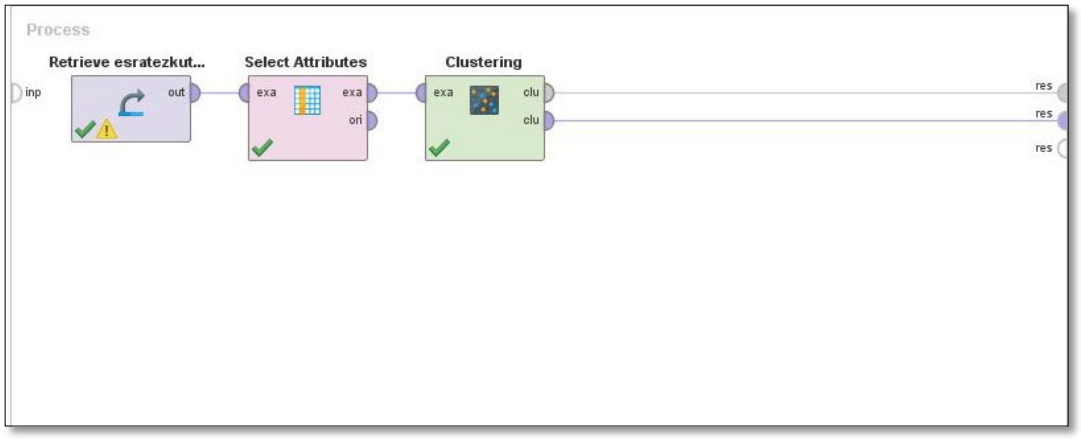
Şekil 14: İktisadi ve İdari Bilimler Fakültesi Grubunda Oluşan Birliktelik Kuralının Grafik Olarak Gösterilmesi

Birliktelik analizinde, kütüphane kullanıcılarının eğitim gördüğü akademik birim ve kütüphaneden ödünç aldığı yayınlar arasındaki ilişki incelenmiş ve bir sonraki aşamada, kütüphane kullanıcılarının özelliklerine göre kümeleme analizi yapılmıştır.

3.5.3. Kütüphane Kullanıcı Verileri Üzerine Kümeleme Analizi

Bu çalışmada, kütüphane kullanıcılarının bilgileri ile kümeleme analizi yapılmıştır. Kümeleme analizi, aynı gruptaki nesnelerin diğer gruptaki nesnelere kıyasla daha benzer şekilde gruplandırılması işlemidir. Benzerlik, kütüphane kullanıcılarının yaşı, ödünç aldığı toplam yayın sayısı gibi niceliksel veya cinsiyeti, eğitim bilgileri gibi niteliksel bazda incelenmektedir. Bu tip değişkenler, kategorik olarak adlandırılmaktadır. *K*-Ortalamalar, bu tip değişkenler için en çok kullanılan hiyerarşik olmayan kümeleme algoritmasıdır (Rezankova ve Everitt, 2009: 216). *K*-Ortalamalar algoritmasında ilk adımda, her küme için merkez noktası olarak kullanılacak olan k adet nesne rastgele seçilir. İkinci adımda, her bir nesne mesafeye bağlı olarak en yakın kümeye atanır. Öklid mesafesi en yaygın kullanılan uzaklık hesaplama yöntemidir. Son adımda ise kümelerin merkez noktalarının konumlarını optimize etmek için tekrar tekrar hesaplamalar yapılır.

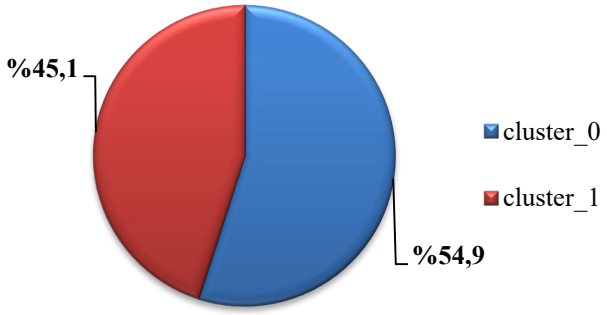
Çalışmada, ön işlemin ardından Excel ortamında toplanan veri setindeki 23.227 verideki kullanıcıların; cinsiyet, eğitim programı, öğrenim gördüğü bölüm ve yayın teslim durumu gibi kategorik bilgileri kümeleme analizinde kullanılmıştır. Buradaki amaç, kütüphane kullanıcılarının bilgilerine göre yapılan kümeleme analizi sonucunda, kullanıcı davranışları tespit edilerek kullanıcıların kütüphane içi kullanım olanaklarının genişletilebilmesi adına kütüphane yöneticileri tarafından çalışmalar yapılması yönündedir.



Şekil 15: Kümeleme Analizi Süreci

Çalışmada, kümeleme analizinde uygulanan veri madenciliği süreci ve işlem bağlantıları Şekil 15'te gösterilmiştir. Şekil 15 incelendiğinde, RapidMiner veri ambarına yüklenen kütüphane kullanıcı veri seti **Read Excel** operatörüyle sürece eklenmiş, **Select Attributes** operatörüyle kümeleme analizinde kullanılacak öznitelikler belirlendikten sonra *K*-Ortalamalar algoritması kullanılarak kümeleme analizi yapılmıştır.

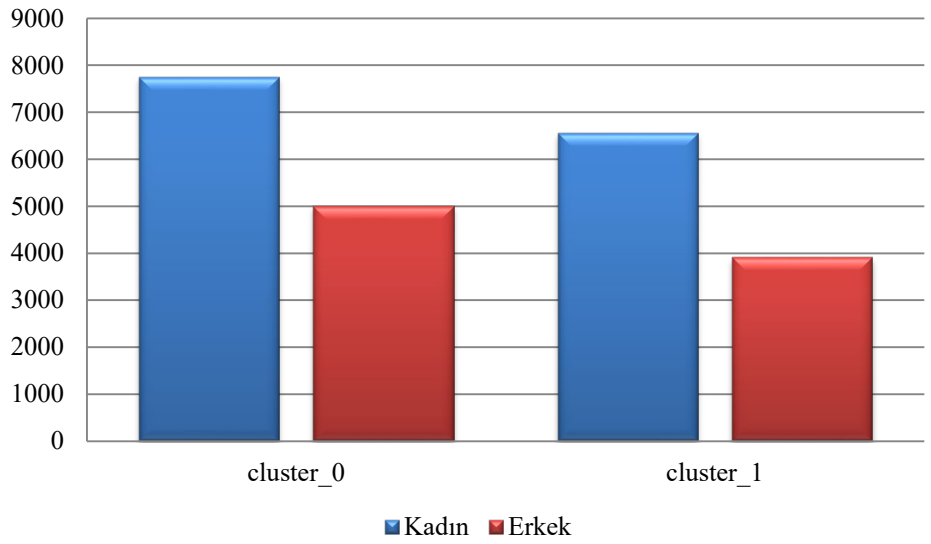
Kütüphane kullanıcılarının verileri üzerinde uygulanan *K*-Ortalamalar algoritmasında en yakın komşu kümeleri bulmak için kullanılacak ölçü türü **Nominal Measures** ve **Nominal Distance** olarak belirlenmiştir.



Şekil 16: Kümeleme Analizi Sonucu Oluşan Kümeler

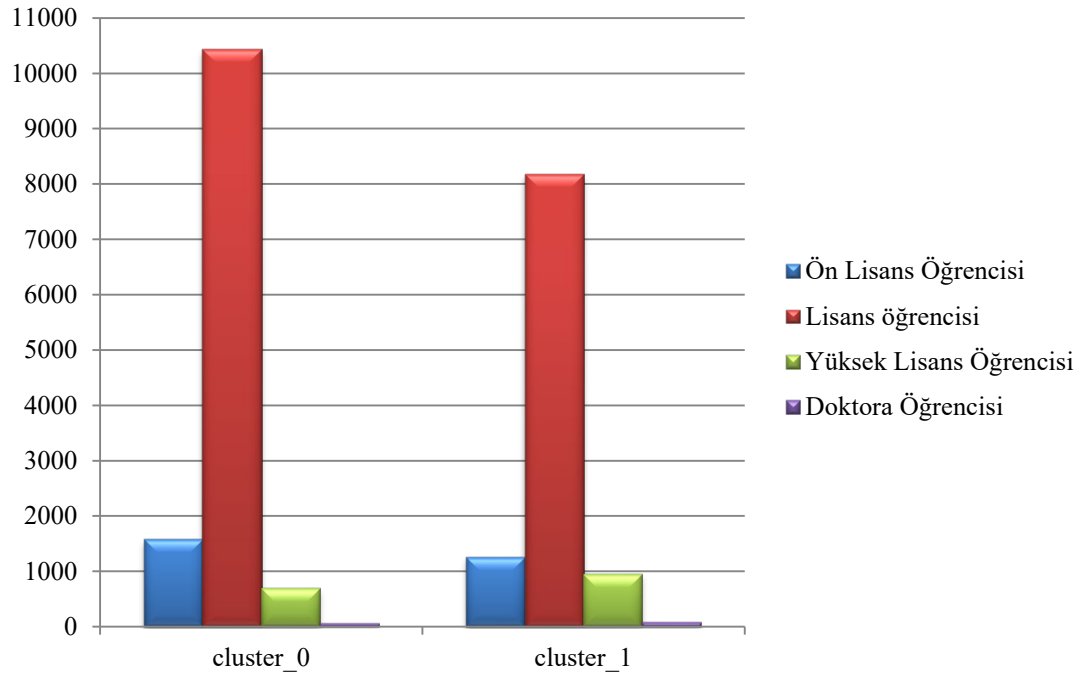
Şekil 16'da kullanıcı verileri üzerinde *K*-Ortalamalar algoritması kullanılarak yapılan kümeleme analizi sonucunda iki küme belirlenmiştir.

RapidMiner programında küme numaralandırılması 0'dan başlamaktadır. Birinci kümede 12.761 ve ikinci kümede 10.466 kayıt bulunmaktadır.



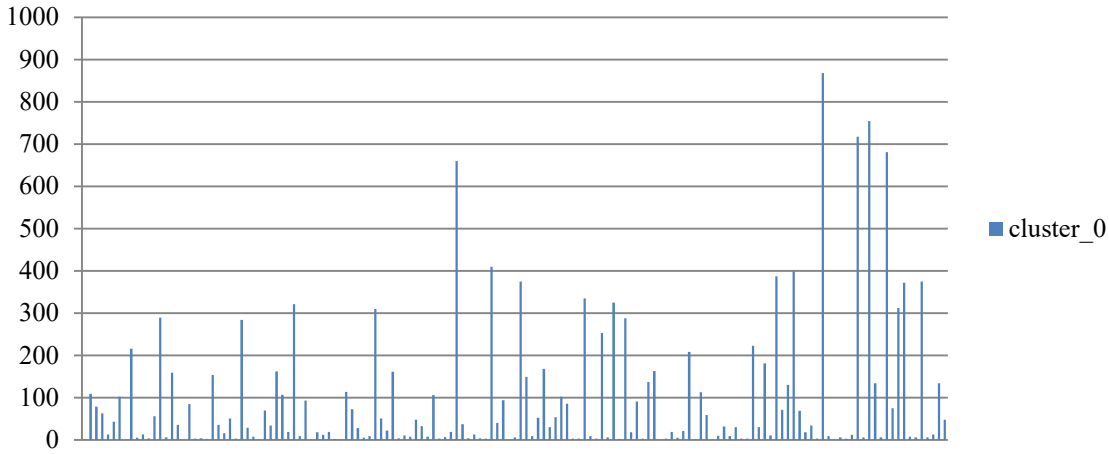
Şekil 17: Kümeleme Analizi Sonucu Kümelerdeki Cinsiyet Dağılımları

Kümeleme analizi sonucu kümelerdeki cinsiyet dağılımlarının gösterildiği Şekil 17'de, birinci kümenin %60,7'si kadın ve %39,3'ü erkek kullanıcı; ikinci kümenin ise %62,7'si kadın ve %37,3'ü erkek kullanıcılardan oluşmaktadır.



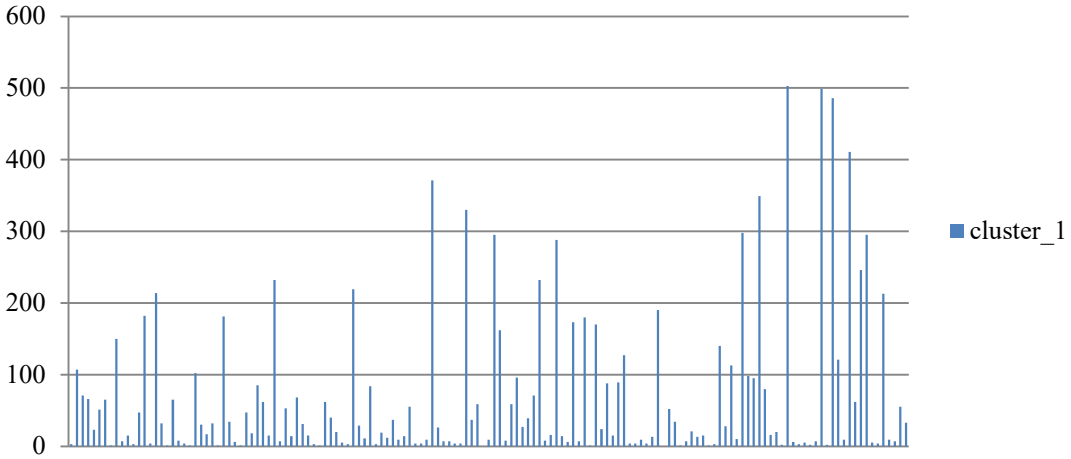
Şekil 18: Kümeleme Analizi Sonucu Oluşan Kümelerdeki Kullanıcı Eğitim Programı Dağılımı

Şekil 18'de kümeleme analizi sonucu oluşan kümelerdeki kullanıcı eğitim programı dağılımı incelendiğinde, her iki kümede de lisans öğrencilerinin baskın olduğu görülmektedir.



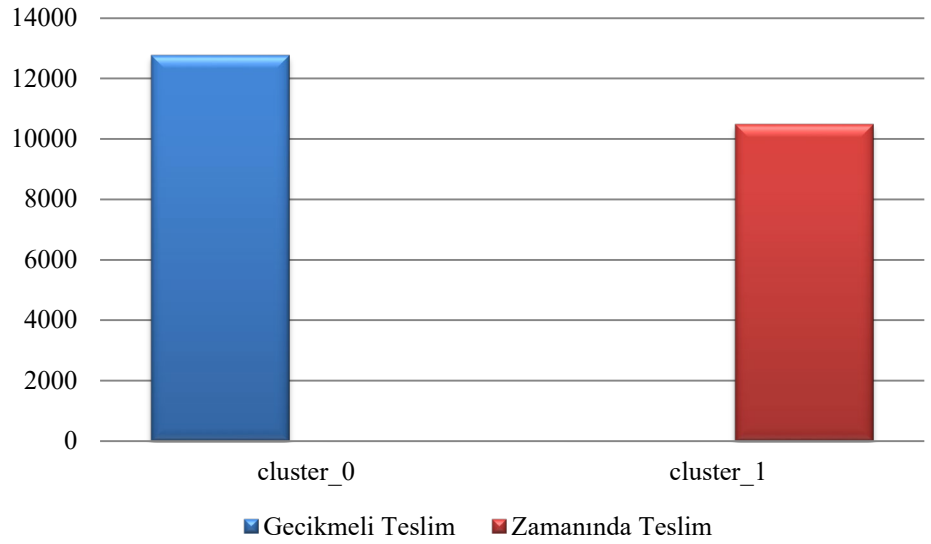
Şekil 19: Kümeleme Analizi Sonucu Oluşan Birinci Kümedeki Kullanıcıların Bölüm Dağılımı

Şekil 19'da kümeleme analizi sonucu oluşan birinci kümedeki kullanıcıların bölüm dağılımı verilmiştir. Birinci kümede en çok yayın alan 817 kayıtla Tıp bölümü kullanıcıları ve en az yayın alan bölümler ise 1'er kayıtla Alman Dili ve Edebiyatı A.B.D., Biyoistatistik ve Tıbbi Bilişim A.B.D., Gıda Mühendisliği A.B.D., Histoloji ve Embriyoloji A.B.D., Temel Eczacılık Bilimleri A.B.D. ve İmmünoloji A.B.D. kullanıcılarıdır.



Şekil 20: Kümeleme Analizi Sonucu Oluşan İkinci Kümedeki Kullanıcıların Bölüm Dağılımı

Şekil 20'de kümeleme analizi sonucu oluşan ikinci kümedeki kullanıcıların bölüm dağılımı verilmiştir. İkinci kümede en çok yayın alan 503 kayıtla Temel Eğitim bölümü kullanıcıları ve en az yayın alan bölümler ise 1'er kayıtla Balkan Dilleri ve Edebiyatları A.B.D., Bilgisayar Teknolojisi ve Bilişim Sistemleri, Biyoistatistik ve Tıbbi Bilişim A.B.D., Büro Hizmetleri ve Sekreterlik, Felsefe ve Din Bilimleri A.B.D., Matematik ve Fen Bilimleri Eğitimi A.B.D., Peyzaj Mimarlığı A.B.D., Resim A.S.D. ve Çocuk Bakımı ve Gençlik Hizmetleri kullanıcılarıdır.



Şekil 21: Kümeleme Analizi Sonucu Oluşan Kümelerdeki Kullanıcı Yayın Teslim Durumu Dağılımı

Şekil 21'de kümeleme analizi sonucu kümelerdeki kullanıcı yayın teslim durumu dağılımı incelendiğinde, birinci kümede sadece gecikmeli teslimlerin; ikinci kümede ise sadece zamanında teslim edenlerin olduğu görülmektedir. Belirgin bir şekilde ayrıldığı görülen teslim durumunun kümelerde baskın bir özellik olduğu ortaya konmuştur.

SONUÇ ve ÖNERİLER

Bilgi verilere, içgörülere ve deneyime dayanan eylem potansiyeli; işletmelerin elinde bulunan tüm insani, teknolojik ve örgütsel kaynakları birbirine bağlayan, hayatta kalma, performans ve rekabet avantajını temsil eden kritik bağlantıdır. Uygulamada, veri ve bilgi kaynaklarının etkin kullanımı karar vericiler için önemlidir. Dijitalleşme sürecinde yaşanan radikal ve aralıklı değişimler; toplumların, kuruluşların ve bireylerin yaşadığı ve çalıştığı ortamları, mevcut küresel iş akışı ve uygulamalarını doğrudan etkilemektedir. Karar verme sürecinde etkili olan bilginin yönetimi de yaşanan değişimlerden etkilenmektedir. Bu noktada bilgi yönetimindeki değişim karşısında işletmelerin gelişimlerini etkileyebilecek unsurları ortaya çıkararak iş süreçlerindeki yerleşik rutinlerini sürekli değerlendirmesi gereklidir (Malhotra, 2003: 66-67). Sürecin iyi yönetilebilmesi için veri madenciliği tekniği oldukça önemli bir yer teşkil etmektedir. Veri madenciliği işletmelerden diğer kuruluşlara taşınmıştır (Nicholson, 2006: 785).

Dünyada hızla gelişen ama Türkiye'de henüz gelişme aşamasında olan dijitalleşme sürecinde kütüphane alanında da yeterli derecede yaygınlaşmamıştır (Ateş, 2015: 561). Çalışmada; bilgi yönetimi çağında, dijitalleşme sürecindeki kütüphanelerde oluşan veri akınlarından anlamlı bilgiler çıkartılarak bilgi keşfi sağlanmaya çalışılmış ve kütüphane yönetimindeki karar vericilere yeni bir bakış açısı sunması amaçlanmıştır.

Çalışmanın ilk bölümünde, bilgi hiyerarşisinin unsurları, OLTP ve OLAP sorgulamaları, karar destek sistemleri, veri ambarı ve veri madenciliği ile olan ilişkisi; ikinci bölümde, veri madenciliği kavramının gelişim süreci, modelleri, aşamaları ve karşılaşılan sorunlar anlatılarak veri madenciliğinin işletmelerde ve kütüphanelerdeki uygulama alanları incelenmiştir. Üçüncü bölümde ise Trakya Üniversitesi Merkez Kütüphanesi sunucularında bulunan 2003-2018 yılları arasındaki kütüphane kullanıcılarının ödünç aldığı yayın sirkülasyon bilgilerinden birliktelik kuralları oluşturulmuş ve kümeleme analizi ile de kütüphane kullanıcılarının davranışları incelenmiştir.

Uygulamaya başlamadan önce verilerin depolanacağı ve veri madenciliği sürecinin uygulanacağı programlar belirlenmiştir. Çalışmada, sadece öğrenci kullanıcılarının davranışları incelenmek istendiği için akademik ve idari personel kütüphane kullanıcıları analize dahil edilmemiştir. Eksik ve hatalı veriler Microsoft Access sorguları kullanılarak temizlenmiştir. Veri ön işleme işlemleri, veri madenciliği sürecinin en çok zaman alan bölümü olup ihtiyaca göre veri madenciliği sürecinin tüm aşamalarında devam etmektedir. Sonraki aşamada ise RapidMiner yazılımı aracılığıyla veri ambarı oluşturulmuştur.

Veri madenciliği modellemeleri oluşturulmadan önce kütüphane kullanıcılarına ait tanımlayıcı bulgular IBM SPSS 24 programı üzerinde frekans tekniği kullanılarak tespit edilmiştir. Kütüphane kullanıcılarının cinsiyet dağılımı incelendiğinde %61,6'sının kadın ve %38,4'ünün erkek olduğu tespit edilmiştir. Kullanıcıların eğitim programına göre dağılımına bakıldığında, en çok kütüphaneden ödünç yayın alan grubun %80,1 ile lisans öğrencisi olduğu görülmektedir. Üniversitedeki akademik birimlere göre dağılımında sırasıyla İktisadi ve İdari Bilimler Fakültesi, Eğitim Fakültesi, Edebiyat Fakültesi ilk üçte yer almaktadır. İlk on akademik birim arasında, sosyal alanlarda eğitim veren kurumların daha fazla olduğu görülmektedir. Kullanıcıların bölüm bazlı dağılımlarına bakıldığında sırasıyla ilk üç bölüm Tıp, Temel Eğitim, Tıbbi Hizmetler ve Teknikler; son üç bölüm Çocuk Bakımı ve Gençlik Hizmetleri, Felsefe ve Din Bilimleri A.B.D., İmmünoloji A.B.D. olduğu sonucuna ulaşılmıştır. İktisadi ve İdari Bilimler Fakültesi bölümlerinde öğrenim gören öğrencilerin cinsiyet dağılımları incelendiğinde, kadın oranının her bölümde daha fazla olduğu sonucuna ulaşılmıştır. Kullanıcıların kütüphaneden ödünç aldığı yayınların sınıflarına göre dağılımlarında, en çok ödünç alınan yayın grubu %43,323 ile Türk Edebiyatı olmuştur. Ödünç alınan yayınların %50,9'unun kütüphaneye geç teslim edildiği, %49,1'inin ise zamanında teslim edildiği sonucuna ulaşılmıştır. Geç teslim eden kullanıcıların %17,78'inin erkek, %33,09'unun ise kadın kullanıcılar olduğu belirlenmiştir. Geç teslim eden kullanıcıların %0,42'si Doktora, %3,22'si Yüksek Lisans, %3,97'si Ön Lisans ve %43,27'si Lisans öğrencisidir. Keşan Meslek Yüksekokulu'nda öğrenim gören dört öğrencinin de aldığı yayınları zamanında teslim ettiği belirlenmiştir. Kütüphaneden ay bazında en

çok ödünç alınan ilk üç ay sırasıyla Ekim, Aralık ve Kasım aylarıdır. 2003 ve 2014 yılları arasında kullanıcıların ödünç yayın sayısında devamlı bir artış görülmüştür.

Çalışmada kullanılan veri madenciliği modelleri ikiye ayrılmaktadır. İlk adımda, kütüphaneden ödünç yayın alan kullanıcıların bilgileri ile yayın sirkülasyon verileri arasında oluşabilecek birliktelik kuralarının tespiti yapılmıştır. Kütüphane ödünç yayın sirkülasyon bilgilerine göre kullanıcıların eğitim gördüğü akademik birim ve ödünç aldığı yayın sınıfları arasındaki ilişki FP-Growth algoritması kullanılarak incelenmiş ve birliktelik kuralları oluşturulmuştur. Belirlenen ölçüm değerleri sonucunda 57 kural ve 103 işlemsel kayıt ortaya çıkmıştır. Birliktelik kuralları içerisinde en çok kural oluşturan yayın sınıfı grupları Türk Edebiyatı ve İngilizce ve Eski İngilizce Edebiyatı'dır. İktisadi ve İdari Bilimler Fakültesi ise akademik birimler içerisinde en yoğun kuralı oluşturan grup olduğu belirlenmiştir. İngilizce ve Eski İngilizce Edebiyat yayın sınıfının tüm yayın sınıfları arasında ödünç alınma oranı %3'tür. Bu yayın grubunu ödünç alan öğrencilerin %20'si İktisadi ve İdari Bilimler Fakültesi'nde öğrenim görmektedir. Akademik birimler içerisinde her bölümün çoğunlukla kendi alanlarıyla ilgili kitaplar tercih ettiği görülmektedir.

İkinci adımda ise kütüphane kullanıcılarının bilgileri ile kümeleme analizi yapılmıştır. Yapılan kümeleme modellemesi ise kütüphane kullanıcılarının cinsiyet, öğrenim gördüğü bölüm, eğitim programları ve yayın teslim durumları baz alınarak analize dahil edilmiştir. K-Ortalamalar algoritması kullanılarak elde edilen analiz sonucunda kullanıcıları iki kümeye ayırmıştır. Birinci kümede, 12.761 kayıt; ikinci kümede ise 10.466 kayıt bulunmaktadır. Birinci kümenin %60,7'si kadın ve %39,3'ü erkek; ikinci kümenin %62,7'si kadın ve %37,3'ü erkek kullanıcılardan oluşmaktadır. Küme bazlı eğitim programı dağılımında hem birinci hem de ikinci kümede lisans öğrencilerinin baskın konumda olduğu belirlenmiştir. Birinci kümedeki kullanıcıların bölüm dağılımında ilk sırada Tıp bölümü, ikinci kümede ise Temel Eğitim bölümü yer almaktadır. Kümelerdeki ödünç alınan yayınların teslim durumu dağılımı incelendiğinde birinci kümede sadece gecikmeli teslimlerin, ikinci kümede ise sadece zamanında teslimlerin olduğu görülmektedir. Belirgin bir şekilde

ayrıldığı görülen teslim durumunun kümelemede baskın bir özellik olduğu belirlenmiştir.

Çalışmada elde edilen bulgularla;

- Kütüphaneye alınması planlanan yayınlar kullanıcı gruplarının tercihlerine göre belirlenebilir.
- Kütüphane kullanıcılarının elektronik kaynaklara yönelmesi ile kitap ödünç alımlarında meydana gelen düşüşler kütüphane yönetimindeki karar vericiler tarafından değerlendirilmelidir.
- Kullanıcıların tercih ettiği yayınlar daha ayrıntılı incelenerek bir yayın öneri sistemi oluşturulabilir. Böylelikle kütüphane bünyesinde hiç ödünç alınmamış veya diğer yayınlara göre daha az ödünç alınmış yayınlar değerlendirilebilir.
- Kütüphane kullanıcılarının yayın ödünç alma davranışına göre kullanıcıların genel not ortalamasına sağladığı katkı ölçülebilir.
- Yayın sınıflarına göre kullanıcılar tarafından en çok ödünç alınan yayın grupları ve bu yayınların teslim durumları belirlenerek kütüphane bünyesinde yeterli miktarda yayın olup olmadığı tespit edilebilir.

Çalışmada, Trakya Üniversitesi Merkez Kütüphanesi'nden 2003-2018 yılları arasındaki kütüphane ödünç yayın sirkülasyon ve kullanıcı verileri toplanmıştır. Zaman periyodunun uzun tutulması, farklı veri setleri eklenmesi ve farklı veri madenciliği yöntemlerinin kullanılması durumunda farklı sonuçlara ulaşılabileceği göz önünde bulundurulmalıdır.

KAYNAKÇA

Ahsan, Syed ve Shah, Abad, "Data, information, knowledge, wisdom: A doubly linked chain", *The Proceedings of the 2006 International Conference on Information Knowledge Engineering*, 2006, 270-278.

Akıncı, Şerife, "Veri Madenciliğinde Kümeleme Analizi ve Bankacılık Sektöründe Bir Uygulama: Ortadoğu ve Kuzey Afrika Ülkelerinde Finansal Erişim", https://www.researchgate.net/publication/333786706_VERI_MADENCILIGINDE_KUMELEME_ANALIZI_VE_BANKACILIK_SEKTORUNDE_BIR_UYGULAMA, (29.08.2019).

Akoğul, Serkan ve Tuna, Elif, "Kümeleme ve Çok Boyutlu Ölçekleme Analizleri ile Endüstriyel Pazar Bölümlendirmesi ve Etkili Ürünlerin Belirlenmesi", *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, Cilt: 25, Sayı: 1, 2016, 29-42.

Aksoy, Barış, İşletmelerin Finansal Başarısızlık Tahmininde Veri Madenciliği Yöntemlerinin Karşılaştırılması: BIST'te Bir Uygulama, (Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme ABD, Doktora Tezi), Kayseri 2018.

Akyel, Nermin ve Seçkin, Keziban, "K-en Yakın Komşuluk Algoritmasının Hile Denetiminde Kullanımı", *Muhasebe ve Vergi Uygulamaları Dergisi (MUVU),* Cilt: 5, Sayı: 1, 2012, 21-39.

Alagöz, Ali, Öge, Serdar ve Koçyiğit, Nezahat, "Muhasebe Bilgi Sistemi ve Karar Destek Sistemleri İlişkisinin Yönetsel Karar Alma Faaliyetlerine Etkisi", *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi,* Sayı: 30, 2013, 27-40.

Alsayat, Ahmed, ve El-Sayed, Hoda, "Social Media Analysis Using Optimized K-Means Clustering", *2016 IEEE 14th International Conference on Software Engineering Research, Management and Applications (SERA),* 2016, 61-66.

Arslan, Volkan ve Yılmaz, Güray, "Karar Destek Sistemlerinin Kullanılması İçin Uygun Bir Model Geliştirilmesi", *Havacılık ve Uzay Teknolojileri Dergisi*, Cilt: 4, Sayı:4, Temmuz 2010, 75-82.

Ateş, Hilal, *Karar Vermede İş Zekâsının Önemi: Tekstil Sektöründe Bir Araştırma,* (Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme ABD, Yüksek Lisans Tezi), İzmir 2008.

Ateş K., Esmeray, "Dijitalleştirme Süreci ve Dijital Kütüphane Uygulamalarına Bir Bakış", *International Periodical fort he Languages, Literature and History of Turkish or Turkic,* Cilt: 10, Sayı: 10, 2015, 561-574.

Averweg, Udo R., "Historical Overview of Decision Support Systems (DSS)", *In Encyclopedia of Information Science and Technology IGI Global*, 2. Baskı, 2009, 1753.

Aydın, İsmail, "Bilişim Sektörü ve Türkiye'nin Sektördeki Potansiyeli", *International Journal of New Trends in Arts, Sports &ScienceEducation,* Cilt: 1, Sayı: 1, 2012, 180-200.

Baltzan, Paige, *Business Driven Information Systems,* McGraw-Hill, 4. Baskı, New York 2014.

Battioui, Chakib, "Data Mining Techniques to Analyze A Library Database", *SAS Institute Inc.Paper,* 2007, 076-31.

Baysal, Ahmet C., *Bayi Değerlendirmesi İçin Veri Madenciliği Uygulaması,* (İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İşletme Mühendisliği ABD, Yüksek Lisans Tezi), İstanbul 2008.

Biçen, Pelin, *Veri Madenciliği: Sınıflandırma ve Tahmin Yöntemlerini Kullanarak Bir Uygulama,* (Yıldız Teknik Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme ABD, Yüksek Lisans Tezi), İstanbul 2002.

Brossette, Stephen E., Sprague, Alan P., Hardin, Michael J., Waites, Ken B., Jones, Warren T. ve Moser, Stephen A., "Association Rules and Data Mining in

Hospital Infection Control and Public Health Surveillance", *Journal of the American Medical Informatics Association,* Cilt: 5, Sayı: 4, Haziran 1998, 373-381.

Burns, Michael, "Business Intelligence Survey", *CA Magazine*, Cilt: 138, Sayı: 5, 2005, 18.

Center for Media Justice, Color of Change ve Sum of Us**, "**Consumers, Big Data, and Online Tracking in the Retail Industry A Case Study Of Walmart", November 2013, (https://docplayer.net/819235-Consumers-big-data-and-online-tracking-in-the-retail-industry-a-case-study-of-walmart.html), (27.10.2019).

Chaudhuri, Surajit ve Dayal, Humeshwar, "An Overview of Data Warehousing and OLAP Technology", *ACM Sigmod Record,* Cilt: 26, Sayı: 1, Mayıs 1997, 65-74.

Chen, Scott S. ve Gopalakrishnan, Ponani S., "Clustering Via The Bayesian Information Criterion With Applications in Speech Recognition", *In Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98,* Cilt: 2, 1998, 645-648.

Cios, Krzysztof J., Pedrycz, Witold ve Swiniarski, Roman W., *Data Mining Methods for Knowledge Discovery*, Springer Science & Business Media, New York 2012.

Cios, Krzysztof J., Pedrycz, Witold, Swiniarski, Roman W. ve Kurgan, Lukasz A., *Data Mining: A Knowledge Discovery Approach,* Springer Science & Business Media, New York 2007.

Çakın, İrfan, "Kütüphanenin Toplumsal Konumu ve İşlevleri", *Türk Kütüphaneciliği,* Cilt: 35, Sayı: 1, 1986, 8-16.

Çelik, Şenol, "Kümeleme Analizi İle Sağlık Göstergelerine Göre Türkiye'deki İllerin Sınıflandırılması", *Doğuş Üniversitesi Dergisi,* Cilt: 14, Sayı: 2, 2013, 175-194.

Çelik, Ufuk, Akçetin, Eyüp ve Gök, Murat, *RapidMiner ile Uygulamalı Veri Madenciliği,* Pusula 20 Teknoloji ve Yayıncılık A.Ş., İstanbul 2007.

Çelik, Ahmet ve Uçak, Nazan, "Üniversite Kütüphaneleri Üzerine", *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi,* Cilt: 10, Sayı: 2, Aralık 1993, 115-121.

Danping, Zhang ve Jin, Deng, "The Data Mining of the Human Resources Data Warehouse in University Based on Association Rule", *Journal of Computers*, Cilt: 6, Sayı: 1, Ocak 2011, 139-146.

Data Mining, https://www.gartner.com/it-glossary/data-mining/, (13.07.2019).

Davenport, Thomas H. ve Bean, Randy, "Big Data Executive Survey 2018", *New Vantage Partners LLC*, 2018, 1-18.

Demir, Hulusi M. ve Gümüşoğlu, Şevkinaz, *Üretim Yönetimi (İşlemler Yönetimi),* 4. Baskı, Beta Yayınevi, İstanbul 1994.

Demiral, Gülten, Soba, Mustafa ve Armutlu, Şükran, "Kütüphane Veri Tabanında Veri Madenciliği: Uşak Üniversitesi Örneği", *Bartın Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,* Cilt: 8, Sayı: 16, 2017, 241-264.

Demirci, Nazan, *Karar Destek Sistemlerinin Bir Durum Çalışmasına Uygulanması,* (Trakya Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği ABD, Yüksek Lisans Tezi), Edirne 2012.

Deogun, Jitender S., Raghavan, Vijay V. ve Sever, Hayri, "Data Mining: Research Trends, Challenges, and Applications", *Roughs Sets and Data Mining: Analysis of Imprecise Data*, 1997, 9-45.

Descriptive Data MiningModels,https://docs.oracle.com/cd/B14117_01/datamine.101/b10698/4descrip.htm, (31.08.2019).

Drucker, Peter F., *Geleceğin Toplumunda Yönetim,* Çev. Doç. Dr. Mehmet Zaman, Hayat Yayınları, İstanbul, 2003.

Drucker, Peter F., *Kapitalist Ötesi Toplum*, Çev. Belkıs Çorakçı, İnkılap Kitapevi, İstanbul, 1993.

Fayyad Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic, "From Data Mining to Discovery Knowledge in Databases", *AI Magazine*, Cilt: 17, Sayı: 3, 1996, 37-54.

Gökçen, Hadi, Özkil, Altan, Yardımoğlu, Hülya ve Peker, Deniz, "Kamuda Karar Destek Sistemlerinin Kullanımı ve Bir Model Önerisi", *TBD Kamu-BİB Kamu Bilişim Platformu XII,* 2010, 1-12.

Gökmen, Şenol, Müşteri İlişkileri Yönetiminde Bir Araç Olarak Veri Madenciliği ve Perakende Sektöründe Bir Uygulama, (Yıldı Teknik Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme ABD, Yüksek Lisans Tezi), İstanbul 2014.

Güçlü, Nezahat ve Sotirofski, Kseanela, "Bilgi Yönetimi", *Türk Eğitim Bilimleri Dergisi*, Cilt: 4, Sayı: 4, Güz 2006, 351-373.

Gülseçen, Sevinç, *Bilgi ve Bilginin Yönetimi,* Papatya Yayıncılık, İstanbul 2012.

Gültekin, Sait U., *Veri Madenciliği: Yapay Sinir Ağı ve Doğrusal Regresyon Yöntemleri İle Fiyat Tahmini,* (Pamukkale Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme ABD, Doktora Tezi), Denizli 2017.

Gürsakal, Necmi, *Büyük Veri,* Dora Yayıncılık, 2. Baskı, Bursa 2014.

Haberal, İsmail, Oğul, Hasan, "Prostat Kanseri İlintili MikroRNA Kümelerinin Tespiti" *22nd Signal Processing and Communications Applications Conference (SIU 2014),* Trabzon 2014.

Hajek, Petr ve Stejskal, Jan, "Analysis of User Behavior in a Public Library Using Bibliomining", *Advances in Environment, Computational Chemistry and Bioscience,* 2012, 339-344.

Han, Jiawei, Kamber, Micheline ve Pei, Jian, *Data Mining Concepts and Techniques*, Morgan Kaufmann, 3. Baskı, Massachusetts 2011.

Hand, David, Mannila, Heikki ve Smyth, Padhraic, *Principle of Data Mining,* MIT Press, Massachusetts 2001.

Hosseini, Seyed M.S., Maleki, Anahita ve Gholamian, Mohammad R., "Cluster Analysis Using Data Mining Approach to Develop CRM Methodology to Assess The Customer Loyalty", *Expert Systems with Applications*, Cilt: 37, Sayı: 7, 2010, 5259-5264.

http://www.worldometers.info/, (08.01.2019).

Huancheng, Liu, Tingting, Wu ve Rocha, Alvaro, "An Analysis of Research Trends on Data Mining in Chinese Academic Libraries", *Journal of Grid Computing,* Cilt: 1, Sayı: 11, 2018, 591-601.

Inmon, William H., *Building The Data Warehouse*, 3. Baskı, Wiley Computer Publishing, New Jersey 2002.

Jadhav, Rahul J ve Pawar, Usharani T., "Chur Prediction in Telecommunation Using Data Mining Technology", *International Journal of Advanced Computer Science and Applications,* Cilt: 2, No: 2, 2011, 11-16.

Kantardzic, Mehmed, *Data Mining: Concepts, Models, Methods, and Algorithms,* John Wiley & Sons, 2. Baskı, New York 2011.

Karahoca, Dilek ve Karahoca, Adem, *İşletmeciler, Mühendisler ve Yöneticiler İçin Yönetim Bilişim Sistemleri ve Uygulamaları,* Beta Yayınları, İstanbul 1998.

Köktürk, Füruzan, Ankaralı, Handan, Sümbüloğlu, Vildan, "Veri Madenciliği Yöntemlerine Bakış", *Türkiye Klinikleri Journal of Biostatistics,* Cilt: 1, Sayı: 1, 2009, 20-25.

Köse, Ertuğrul, *Karar Verme Yetkinliği ve Yönetsel Performans,* Nobel Bilimsel Eserler Yayınevi, Ankara 2018.

Kurtuluş, Kemal, *Pazarlama Araştırmaları,* 7. Baskı, Literatür Yayınevi, İstanbul 2004.

Kuruüzüm, Ayşe, *Karar Destek Sistemlerinde Çok Amaçlı Yöntemler,* Akdeniz Üniversitesi Basımevi, Antalya 1998.

Laudon, Kenneth C. ve Laudon, Jane P., *Essentials of Business Information Systems,* Pearson Education, 7. Baskı, Londra 2007.

Little, John D. C., "Models and Managers: The Concept of a Decision Calculus", *Management Science,* Cilt: 50, Sayı: 12, 2004, 1854-1860.

Malhotra, Yogesh, "Is Knowledge the Ultimate Competitive Advantage?", *Business Management Asia,* Cilt: 3, Sayı: 4, 2003, 66-69.

Marr, Bernard, How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#49f000e660ba, (08.01.2019).

Martin, Wainright E., DeHayes, Daniel W., Hoffer, Jeffrey A. ve Perkins, William C., *Managing Information Technology: What Managers Need to Know,* Macmillan Publishing Company, 1. Baskı, Londra 1991.

Nicholson, Scott, "Bibliomining for Automated Collection Development in a Digital Library Setting: Using Data Mining to Discover Web-Based Scholarly Research Works", *Journal of the American Society for Information Science and Technology,* Cilt: 54, Sayı: 12, 2003, 1081-1090.

Nicholson, Scott, "The Basis for Bibliomining: Frameworks for Bringing Together Usage-Based Data Mining and Bibliometrics Through Data Warehousing in Digital Library Services", *Information Processing & Management,* Cilt: 42, Sayı: 3, 2006, 785-804.

Nicholson, Scott, "The Bibliomining Process: Data Warehousing and Data Mining for Library Decision-Making", *Information Technology and Libraries,* Cilt: 22, Sayı: 4, 2003, 146-156.

Özcan, Canan, Veri Madenciliğinin Güvenlik Uygulama Alanları ve Veri Madenciliği İle Sahtekarlık Analizi, (İstanbul Bilgi Üniversitesi, Sosyal Bilimler Enstitüsü, Bilişim ve Teknoloji Hukuku Yüksek Lisans) İstanbul 2014.

Özdoğan Ö., *Gülistan, Öbek Bilgisayarlarda Paralel FP-Growth Gerçekleştirimi*, (TOBB Ekonomi ve Teknoloji Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği ABD, Yüksek Lisans Tezi), Ankara 2010.

Özel, Gamze ve Ata, Nihal, "Eksik Gözlem Değerlerine Sahip OECD Ülkelerinin Bebek Sağlığı İle İlgili Analizinde Yerine Koyma Yöntemlerinin Kullanılması", *Doğuş Üniversitesi Dergisi*, Cilt: 8, Sayı: 2, 2007, 218-233.

Özkan, Yalçın, *Veri Madenciliği Yöntemleri,* Papatya Yayıncılık Eğitim, 2. Baskı, İstanbul 2013.

Plattner, Hasso, *A Course in In-Memory Data Management: The Inner Mechanics of In-Memory Databases,* Springer, 2. Baskı, New York 2014.

Prytherch, Ray, *Harrod's Librarians' Glossary and Reference Book,* Ashgate Publishing Limited, 10. Baskı, Farnham 2005.

Quinlan, J. Ross, “Induction of Decision Trees”, *Machine Learning,* Cilt: 1, Sayı: 81, 1986, 81-106.

Rezankova, Hana ve Everitt, Brian, “Cluster Analysis and Categorical Data”, *Statistika,* Cilt: 89, Sayı: 2, 2009, 216-232.

Rowley, Jennifer, “The Wisdom Hierarchy: Representations of the DIKW Hierarchy”, *Journal of Information Science,* Cilt: 33, Sayı: 2, 2007, 163-180.

Safavian, S. Rasoul ve Landgrece, David, "A Survey of Desicion Tree Claasifier Methodology", *IEEE Transactions on Systems, Man, and Cybernetics,* Cilt: 21, Sayı: 3, 1991, 660-674.

Sahu, Hemlata, Shrma, Shalini ve Gondhalakar, Seema, “A Brief Overview on Data Mining Survey”, *International Journal of Computer Technology and Electronics Engineering (IJCTEE),* Cilt: 1, Sayı: 3, 2011, 114-121.

Sanderson, Robert ve Watry, Paul, “Integrating Data and Text Mining Processes for Digital Library Applications”, *In Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries,* Haziran 2007, 73-79.

Santos, Ricardo J., Vieira, Marco ve Bernardino, Jorge, *“XSX: Lightweight Encryption for Data Warehousing Environments: Proceedings of yhe 18th International Big Data Analytics and Knowledge Discovery Conference”,* Springer, Eylül 2016, 281-295.

Saranya, M. K., Rathnavathy, Ramanathan ve SureshBabu, G. N. K., “A Study of Predictive Data Mining Techniques”, *International Journal of Computer Science and Information Technology Research,* Cilt: 2, Sayı: 2, Nisan – Haziran 2014, 258-264.

Sarr, M. D, “Karar Süreci ve Bilişimin Katkısı”, Çev. Engin Okyay, *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, Cilt: 3, Sayı: 2, Kasım 1974, 224-236.

Scott-Morton, Michael S., *Management Decision Systems: Computer-Based Support for Decision Making*, Harvard University Press, Massachusetts 1971.

Shieh, Jiann-Cherng, "The Integration System for Librarians' Bibliomining", *The Electronic Library,* Cilt: 28, Sayı: 5, 2010, 709-721.

Shmueli, Galit, Bruce, Peter C., Yahav, Inbal, Patel, Nitin R ve Lichtendahl Jr., Kenneth C., *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R,* John Wiley & Sons, New Jersey 2018.

Silberschatz, Abraham, Korth, Henry F. ve Sudarshan, S., *Database System Concepts,* McGraw – Hill, 6. Baskı, New York 2011.

Singh, Shio K., *Database Systems Concepts, Design and Applications,* Prentice Hall, New Jersey 2009.

Smith, Kate A., Willis, Robert J. ve Brooks, Malcolm, "An Analysis of Customer Retention and Insurance Claim Patterns Using Data Mining: A Case Study", *Journal of the Operational Research Society*, Cilt: 51, Sayı: 5, 2000, 532-541.

Smith, Kit, 121 Amazing Social Media Statistics and Facts, https://www.brandwatch.com/blog/amazing-social-media-statistics-and-facts/#section-10, (08.01.2019).

Sol, Henk G., Takkenberg, Cees A. Th. ve De VriesRobbé, Pieter F., *Expert Systems and Artificial Intelligence in Decision Support Systems: Proceedings of the Second Mini Euroconference*, Springer Science & Business Media, Berlin 1985.

Song, Yunlong ve Wei, Ran, "Research on Application of Data Mining Based on Fp-Growth Algorithm for Digital Library", *Second International Conference on Mechanic Automationand Control Engineering*, Temmuz 2011, 1525-1528.

Şeker, Şadi E., "Veri Ambarı (Data Warehouse)", *YBS Ansiklopedisi*, Cilt: 2, Sayı: 4, Aralık 2015, 9-17.

Şeker, Şadi E., Karar Destek Sitemleri, http://mis.sadievrenseker.com/2014/02/karar-destek-sistemleri-kds-decision-support-systems-dss/?fbclid=IwAR0zlMEXuLGNmNlGFUGu9OeAFIPdHCoXMTrR_SaONXrtHf_XmRHC4JClSyo, (07.01.2019).

Şen, Hülya ve Yazıcı, Kübra, "Türkiye'deki İllerin Cinsel Suçlar Açısından İncelenmesi", *Alphanumeric Journal*, Cilt: 5, Sayı: 2, 2017, 293-308.

Takçı, Hidayet ve Soğukpınar, İbrahim, "Kütüphane Kullanıcılarının Erişim Örüntülerinin Keşfi", *Bilgi Dünyası,* Cilt: 3, Sayı: 1, 2002, 12-26.

Turban, Efraim, Aronson, Jay E. ve Liang, Ting-Peng, *Decision Support Systems and Intelligent Systems,* PrenticeHall,7. Baskı, New Jersey 2007.

Uçan, Ömer, *Dijital Kütüphanelerde Veri Madenciliği Uygulamaları: Akdeniz Üniversitesi Merkez Kütüphanesi Örneği,* (Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme ABD, Yüksek Lisans Tezi), Antalya 2010.

Utley, Craig,*Business Intelligence with Microsoft Office Performance Point Server 2007,* McGraw-Hill, New York 2008.

Vassiliadis, Panos ve Sellis, Timos, "A Survey on Logical Models for OLAP Databases", *ACM Sigmod Record,* Cilt: 28, Sayı: 4, 1999, 64-69.

Veri, Enformasyon (Malumat), Bilgi ve Bilgelik (İrfan), https://ohantekten.com/wp-content/.../Veri_-Enformasyon-_Malumat__-Bilgi-ve.pdf, (03.12.2018).

Waltz, Edward, *Information Warfare: Principles and Operations*, Artech House, Massachusetts 1998.

Weiping, Fang ve Yuming, Wang, "The Development of Data Mining", *International Journal of Business and Social Science,* Cilt: 4, Sayı: 16, Aralık 2013, 158 - 159.

Weiss, Sholom M. ve Indurkhya, Nitin, *Predictive Data Mining: A Practical Guide*, Morgan Kaufmann Publishers, Massachusetts 1998.

Xiangyan, Tian, *Predicting Student Dropout: An Application of Data Mining Techniques on Get Data,* (California State University, Computer Science, PhD Thesis), Los Angeles 2015.

Yanbin, Tu, *Analysis of Selling Strategies in Electronic Marketplaces,* (University of Connecticut, Business Administration, PhD Thesis), Connecticut 2006.

Yaşar, Alparslan, "Olumlu Görüş Dışındaki Denetim Görüşlerinin Veri Madenciliği Yöntemleriyle Tahminine İlişkin Karar ve Birliktelik Kuralları", *Mali Çözüm Dergisi,* Cilt: 26, Sayı: 133, 2016, 81-109.

Yurdadoğ U., Berin, Aktaş, Ziya, Güvenç, Bozkurt, İnam, Ahmet ve Kazancı, Metin, Bilgi Toplumu ve Türkiye, http://www.phil.metu.edu.tr/ahmet-inam/bilgi-toplumu.htm, (29.11.2018).

Zhuang, Quntao R., Yao, Yaw W. ve Liu, O., "Application of Data Mining in Term Deposit Marketing", *Proceedings of the International Multi Conference of Engineers and Computer Scientists,* Hong Kong 2018, 1-4.

DİZİN

EKLER

EK- 1. Kütüphaneden En Az Bir Yayın Ödünç Almış İktisadi ve İdari Bilimler Fakültesi Bölümlerinde Eğitim Gören Öğrencilerin Cinsiyete Göre Dağılımı

Bölüm	Cinsiyet	Frekans	%
Alman Dili ve Edebiyatı A.B.D.	**Erkek**	2	0,009
	Kadın	2	0,009
Antrenörlük Eğitimi	**Erkek**	185	0,796
	Kadın	109	0,469
Arkeoloji	**Erkek**	156	0,672
	Kadın	142	0,611
Arkeoloji A.B.D.	**Erkek**	23	0,099
	Kadın	13	0,056
Balkan Çalışmaları A.B.D.	**Erkek**	42	0,181
	Kadın	43	0,185
Balkan Dilleri ve Edebiyatları	**Erkek**	52	0,224
	Kadın	102	0,439
Balkan Dilleri ve Edebiyatları A.B.D.	**Erkek**	1	0,004
Bankacılık	**Erkek**	126	0,542
	Kadın	216	0,930
Bankacılık A.B.D.	**Erkek**	7	0,030
	Kadın	5	0,022
Bankacılık ve Sigortacılık	**Erkek**	8	0,034
	Kadın	13	0,056
Beden Eğitimi ve Spor A.B.D.	**Erkek**	5	0,022
	Kadın	4	0,017
Beden Eğitimi ve Spor Öğretmenliği	**Erkek**	64	0,276
	Kadın	56	0,241
Beslenme ve Diyetetik	**Erkek**	59	0,254
	Kadın	289	1,244
Beslenme ve Diyetetik A.B.D.	**Kadın**	6	0,026
Bilgisayar Mühendisliği	**Erkek**	316	1,360
	Kadın	159	0,685
Bilgisayar Teknolojileri	**Erkek**	43	0,185
	Kadın	36	0,155
Bilgisayar Teknolojisi ve Bilişim Sistemleri	**Erkek**	4	0,017

Bilgisayar ve Öğretim Teknolojileri Eğitimi	**Erkek**	86	0,370
	Kadın	85	0,366
Bilgisayar ve Öğretim Teknolojileri Eğitimi A.B.D.	**Erkek**	11	0,047
	Kadın	3	0,013
Biyofizik A.B.D.	**Erkek**	4	0,017
	Kadın	4	0,017
Biyoistatistik ve Tıbbi Bilişim A.B.D.	**Kadın**	2	0,009
Biyoloji	**Erkek**	45	0,194
	Kadın	154	0,663
Biyoloji A.B.D.	**Erkek**	19	0,082
	Kadın	36	0,155
Biyoteknoloji ve Genetik A.B.D.	**Erkek**	9	0,039
	Kadın	16	0,069
Büro Hizmetleri ve Sekreterlik	**Erkek**	19	0,082
	Kadın	54	0,232
Çalışma Ekonomisi ve Endüstri İlişkileri	**Erkek**	118	0,508
	Kadın	284	1,223
Çalışma Ekonomisi ve Endüstri İlişkileri A.B.D.	**Erkek**	18	0,077
	Kadın	29	0,125
Çalışma İktisadı A.B.D.	**Erkek**	6	0,026
	Kadın	8	0,034
Çocuk Bakımı ve Gençlik Hizmetleri	**Kadın**	1	0,004
Dış Ticaret	**Erkek**	34	0,146
	Kadın	70	0,301
Ebelik	**Kadın**	34	0,146
Eczacılık	**Erkek**	47	0,202
	Kadın	162	0,697
Eğitim Bilimleri	**Erkek**	44	0,189
	Kadın	107	0,461
Eğitim Bilimleri A.B.D.	**Erkek**	6	0,026
	Kadın	19	0,082
Ekonometri	**Erkek**	167	0,719
	Kadın	321	1,382
Ekonometri A.B.D.	**Erkek**	2	0,009
	Kadın	9	0,039
El Sanatları	**Erkek**	11	0,047
	Kadın	93	0,400
Elektrik ve Enerji	**Erkek**	39	0,168
Elektrik-Elektronik Mühendisliği	**Erkek**	120	0,517
	Kadın	18	0,077

Elektronik ve Otomasyon	**Erkek**	49	0,211
	Kadın	12	0,052
Engelli Çalışmaları A.B.D.	**Erkek**	7	0,030
	Kadın	19	0,082
Ermeni Dili ve Edebiyatı	**Erkek**	3	0,013
	Kadın	2	0,009
Felsefe ve Din Bilimleri A.B.D.	**Kadın**	1	0,004
Finans Bankacılık ve Sigortacılık	**Erkek**	36	0,155
	Kadın	114	0,491
Fizik	**Erkek**	70	0,301
	Kadın	73	0,314
Fizik A.B.D.	**Erkek**	22	0,095
	Kadın	28	0,121
Fizik Tedavi ve Rehabilitasyon A.B.D.	**Erkek**	2	0,009
	Kadın	5	0,022
Fizyoloji A.B.D.	**Erkek**	2	0,009
	Kadın	9	0,039
Fizyoterapi ve Rehabilitasyon	**Erkek**	170	0,732
	Kadın	310	1,335
Genetik ve Biyomühendislik	**Erkek**	14	0,060
	Kadın	51	0,220
Gıda İşleme	**Erkek**	19	0,082
	Kadın	22	0,095
Gıda Mühendisliği	**Erkek**	39	0,168
	Kadın	161	0,693
Gıda Mühendisliği A.B.D.	**Kadın**	4	0,017
Girişimcilik A.B.D.	**Erkek**	16	0,069
	Kadın	11	0,047
Görsel Kültür A.B.D.	**Erkek**	9	0,039
	Kadın	8	0,034
Görsel, İşitsel Teknikler ve Medya Yapımcılığı	**Erkek**	61	0,263
	Kadın	48	0,207
Grafik Tasarımı	**Erkek**	8	0,034
	Kadın	33	0,142
Gümrük İşletme	**Erkek**	15	0,065
	Kadın	8	0,034
Güzel Sanatlar Eğitimi	**Erkek**	45	0,194
	Kadın	106	0,456
Güzel Sanatlar Eğitimi A.B.D.	**Erkek**	4	0,017
	Kadın	3	0,013

Halk Sağlığı A.B.D.	**Erkek**	1	0,004
	Kadın	7	0,030
Hazırlık (Mütercim-Tercümanlık)	**Erkek**	6	0,026
	Kadın	19	0,082
Hemşirelik	**Erkek**	95	0,409
	Kadın	660	2,842
Hemşirelik A.B.D.	**Erkek**	3	0,013
	Kadın	37	0,159
Hesaplamalı Bilimler A.B.D.	**Erkek**	7	0,030
	Kadın	4	0,017
Heykel	**Erkek**	10	0,043
	Kadın	13	0,056
Heykel A.S.D.	**Kadın**	4	0,017
Histoloji ve Embriyoloji A.B.D.	**Erkek**	2	0,009
	Kadın	3	0,013
İktisat	**Erkek**	346	1,490
	Kadın	410	1,765
İktisat A.B.D.	**Erkek**	36	0,155
	Kadın	40	0,172
İlahiyat	**Erkek**	47	0,202
	Kadın	94	0,405
İmmünoloji A.B.D.	**Erkek**	1	0,004
İnsan ve Toplum Araştırmaları A.B.D.	**Erkek**	5	0,022
	Kadın	6	0,026
İşletme	**Erkek**	310	1,335
	Kadın	375	1,615
İşletme A.B.D.	**Erkek**	146	0,629
	Kadın	149	0,641
İşletme Bilgi Yönetimi	**Erkek**	12	0,052
	Kadın	9	0,039
Kamu Yönetimi A.B.D.	**Erkek**	38	0,164
	Kadın	52	0,224
Kimya	**Erkek**	84	0,362
	Kadın	168	0,723
Kimya A.B.D.	**Erkek**	18	0,077
	Kadın	30	0,129
Kimya ve Kimyasal İşleme Teknolojileri	**Erkek**	11	0,047
	Kadın	54	0,232
Klinik Bilimler	**Erkek**	45	0,194
	Kadın	102	0,439

Makine Mühendisliği	**Erkek**	482	2,075
	Kadın	86	0,370
Makine Mühendisliği A.B.D	**Erkek**	11	0,047
	Kadın	3	0,013
Makine ve Metal Teknolojileri	**Erkek**	25	0,108
	Kadın	3	0,013
Maliye	**Erkek**	247	1,063
	Kadın	335	1,442
Maliye A.B.D.	**Erkek**	12	0,052
	Kadın	9	0,039
Malzeme ve Malzeme İşleme Teknolojileri	**Erkek**	13	0,056
	Kadın	3	0,013
Matematik	**Erkek**	136	0,586
	Kadın	253	1,089
Matematik A.B.D.	**Erkek**	5	0,022
	Kadın	6	0,026
Matematik ve Fen Bilimleri Eğitimi	**Erkek**	60	0,258
	Kadın	325	1,399
Matematik ve Fen Bilimleri Eğitimi A.B.D.	**Erkek**	3	0,013
	Kadın	2	0,009
Mimarlık	**Erkek**	166	0,715
	Kadın	288	1,240
Mimarlık A.B.D.	**Erkek**	21	0,090
	Kadın	18	0,077
Mimarlık ve Şehir Planlama	**Erkek**	77	0,332
	Kadın	91	0,392
Motorlu Taşıtlar ve Ulaştırma Teknolojileri	**Erkek**	20	0,086
	Kadın	3	0,013
Muhasebe ve Vergi	**Erkek**	58	0,250
	Kadın	137	0,590
Mütercim Tercümanlık	**Erkek**	131	0,564
	Kadın	163	0,702
Müzik	**Erkek**	3	0,013
	Kadın	1	0,004
Müzik A.S.D.	**Erkek**	4	0,017
	Kadın	3	0,013
Odyoloji	**Erkek**	5	0,022
	Kadın	19	0,082
Ortaöğretim Sosyal Alanlar A.B.D.	**Erkek**	3	0,013
	Kadın	5	0,022

Otel Lokanta ve İkram Hizmetleri	**Erkek**	27	0,116
	Kadın	21	0,090
Özel Eğitim	**Erkek**	169	0,728
	Kadın	208	0,896
Park ve Bahçe Bitkileri	**Erkek**	1	0,004
	Kadın	1	0,004
Pazarlama ve Reklamcılık	**Erkek**	54	0,232
	Kadın	113	0,487
Peyzaj Mimarlığı	**Erkek**	19	0,082
	Kadın	59	0,254
Peyzaj Mimarlığı A.B.D.	**Kadın**	1	0,004
Piyano Anasanat Dalı	**Erkek**	5	0,022
	Kadın	10	0,043
Rekreasyon	**Erkek**	31	0,133
	Kadın	32	0,138
Rekreasyon Yönetimi A.B.D.	**Erkek**	8	0,034
	Kadın	9	0,039
Resim	**Erkek**	14	0,060
	Kadın	30	0,129
Resim A.S.D.	**Kadın**	3	0,013
Sağlık Fiziği A.B.D.	**Kadın**	3	0,013
Sağlık Yönetimi	**Erkek**	82	0,353
	Kadın	223	0,960
Sağlık Yönetimi A.B.D.	**Erkek**	21	0,090
	Kadın	30	0,129
Sanat Tarihi	**Erkek**	101	0,435
	Kadın	181	0,779
Sanat Tarihi A.B.D.	**Erkek**	10	0,043
	Kadın	11	0,047
Siyaset Bilimi ve Kamu Yönetimi	**Erkek**	327	1,408
	Kadın	387	1,666
Siyaset Bilimi ve Uluslararası İlişkiler A.B.D.	**Erkek**	100	0,431
	Kadın	71	0,306
Spor Yöneticiliği	**Erkek**	79	0,340
	Kadın	130	0,560
Tarih	**Erkek**	336	1,447
	Kadın	399	1,718
Tarih A.B.D.	**Erkek**	79	0,340
	Kadın	69	0,297

Tasarım	**Erkek**	9	0,039
	Kadın	18	0,077
Tekstil, Giyim, Ayakkabı ve Deri	**Erkek**	8	0,034
	Kadın	34	0,146
Temel Eczacılık Bilimleri A.B.D.	**Kadın**	3	0,013
Temel Eğitim	**Erkek**	175	0,753
	Kadın	868	3,737
Temel Eğitim A.B.D.	**Kadın**	9	0,039
Temel İslam Bilimleri A.B.D.	**Erkek**	4	0,017
	Kadın	1	0,004
Tıbbi Biyokimya A.B.D.	**Erkek**	3	0,013
	Kadın	6	0,026
Tıbbi Biyoloji A.B.D.	**Erkek**	1	0,004
	Kadın	3	0,013
Tıbbi Farmakoloji A.B.D.	**Erkek**	4	0,017
	Kadın	12	0,052
Tıbbi Hizmetler ve Teknikler	**Erkek**	292	1,257
	Kadın	717	3,087
Tıbbi Mikrobiyoloji A.B.D.	**Kadın**	6	0,026
Tıp	**Erkek**	549	2,364
	Kadın	754	3,246
Turizm İşletmeciliği ve Otelciliği	**Erkek**	150	0,646
	Kadın	134	0,577
Turizm İşletmeciliği ve Otelcilik A.B.D.	**Erkek**	6	0,026
	Kadın	6	0,026
Türk Dili ve Edebiyatı	**Erkek**	272	1,171
	Kadın	681	2,932
Türk Dili ve Edebiyatı A.B.D.	**Erkek**	24	0,103
	Kadın	75	0,323
Türkçe ve Sosyal Bilimler Eğitimi	**Erkek**	232	0,999
	Kadın	312	1,343
Uluslararası İlişkiler	**Erkek**	294	1,266
	Kadın	372	1,602
Uluslararası Ticaret	**Erkek**	1	0,004
	Kadın	8	0,034
Üflemeli ve Vurmalı Çalgılar Anasanat Dalı	**Erkek**	2	0,009
	Kadın	6	0,026
Yabancı Diller Eğitimi	**Erkek**	108	0,465
	Kadın	375	1,615

Yabancı Diller Eğitimi A.B.D.	**Erkek**	6	0,026
	Kadın	6	0,026
Yaylı Çalgılar Anasanat Dalı	**Erkek**	9	0,039
	Kadın	13	0,056
Yönetim ve Organizasyon	**Erkek**	108	0,465
	Kadın	182	0,784
	Toplam	23.227	100,000