T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
SAYISAL YÖNTEMLER BİLİM DALI

YÜKSEK LİSANS TEZİ

BANKA MÜŞTERİLERİNİN İNTERNET BANKACILIĞINA İLİŞKİN YAKLAŞIMLARININ VERİ MADENCİLİĞİ TEKNİKLERİ İLE İNCELENMESİ

ŞAFİYE BİLGİN

2501100823

Tez Danışmanı: Doç. Dr. Umman Tuğba ŞİMŞEK GÜRSOY

İSTANBUL, 2013

T.C. İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER
1982 ENSTİTÜSÜ



YÜKSEK LİSANS
TEZ ONAYI

ÖĞRENCİNİN

Adı ve Soyadı : **ŞAFİYE BİLGİN**
Numarası : **2501100823**
Danışman : **YRD .DOÇ U.TUĞBA ŞİMŞEK GÜRSOY**
Anabilim/Bilim Dalı : **SAYISAL YÖNTEMLER**
Tez Savunma Tarihi : **23.10.2013**
Tez Savunma Saati : **11:00**

Tez Başlığı : **BANKA MÜŞTERİLERİNİN İNTERNET BANKACILIĞINA İLŞKİN YAKLAŞIMLARININ VERİ MANDENCİLİĞİ TEKNİKLERİ İLE İNCELENMESİ**

TEZ SAVUNMA SINAVI, Lisansüstü Öğretim Yönetmeliği'nin **36. Maddesi** uyarınca yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin **KABULÜ'NE** OYBİRLİĞİ / ~~OYÇOKLUĞUYLA~~ karar verilmiştir.

JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
1-PROF.DR MEHPARE TİMOR		Kabul
2-DOÇ.DR U.TUĞBA ŞİMŞEK GÜRSOY		Kabul
3-YRD.DOÇ.DR BİRGÜL KÜÇÜK ÇIRPIN		Kabul

YEDEK JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
1- DOÇ.DR SEDA TOLON		
2- YRD.DOÇ.DR ÖZLEM AKÇAY KASAPOĞLU		

BANKA MÜŞTERİLERİNİN İNTERNET BANKACILIĞINA İLİŞKİN YAKLAŞIMLARININ VERİ MADENCİLİĞİ TEKNİKLERİ İLE İNCELENMESİ

Şafiye BİLGİN

ÖZET

Bilgi 21. Yüzyılın petrolü niteliğinde değer taşımaktadır. Son yıllarda yaşanan teknolojik gelişmeler sayesinde; dünyadaki veri miktarının artışı akıl almaz boyutlara ulaşmıştır. Günümüzde bireylerin yaptığı her alışveriş, her işlem, her telefon görüşmesi, her elektronik işlem kayıt altına alınmakta ve saklanmaktadır. Gerçekleştirilen her işlemle elde edilen veriler, beraberinde verileri toplama ve saklama alanında da zorlukları getirmiştir.

Veri Madenciliği, büyük miktarda ve karmaşık olan veriler arasından değerli bilgilerin elde edilmesini sağlayan gittikçe önemi artan bir tekniktir. Veri madenciliği bilim, sağlık, pazarlama, finans gibi pek çok sektörde yaygın bir şekilde kullanılmaktadır. Birliktelik Kuralları Analizi ve Kümeleme Analizi teknikleri, günümüzde en sık uygulanan veri madenciliği tekniklerindendir.

Bu çalışmada, banka müşterilerinin internet bankacılığına ilişkin yaklaşımlarının veri madenciliği teknikleri ile incelenmesi amaçlanmıştır. Çalışmada internet bankacılığını kullanan banka müşterilerinin profilleri incelenmiştir. Birliktelik Kuralları ve Kümeleme Analizi yardımıyla internet bankacılığını kullanan banka müşterilerinin, hangi bankaların internet bankacılığını kullandıkları, ne tür işlemler gerçekleştirdikleri ve internet bankacılığını kullanmalarında etkili olan faktörler araştırılmıştır. Çalışmada ayrıca internet bankacılığı kullanmayan banka müşteri profilleri de incelenmiş ve internet bankacılığı kullanmama sebepleri de analiz edilmiştir.

Anahtar Kelimeler: Veri Madenciliği, İnternet Bankacılığı, Kümeleme Analizi, Birliktelik Kuralları Analizi

EXAMINATION OF APPROACHES OF BANK CUSTOMERS ABOUT INTERNET BANKING BY USING DATA MINING TECHNIQUES

Şafiye BİLGİN

ABSTRACT

Information is of a great value in 21th century and can be named as oil of our century. The increase in the amount of data has reached incomprehensible levels in the world thanks to technological developments and advances in recent years. Every deal, every shopping, every process, every phone call and every electronic transaction are recorded and stored electronically in today's world. Data obtained from every transaction has also brought some difficulties in the field of data collection and storage areas.

Data mining is an increasingly important technique that allows obtaining and choosing valuable information among the other large amount of data which are very complex. Data mining is widely used in many sectors such as science, health, education, marketing, finance, etc. Association Rules Analysis and Cluster Analysis techniques are some of the most widely used data mining techniques.

It is aimed in this study to examine approaches of bank customers about internet banking by using data mining techniques. The profiles of bank customers using internet banking have been examined. The banks which are used for internet banking, transactions done by customers and factors that are effective on bank customers to use internet banking are researched with the help of Cluster Analysis and Association Rules. In addition, profiles of bank customers who do not use internet banking are examined in this study and the reasons and causes of these bank customers not to use internet banking have also been analyzed.

Keywords: Data mining, Internet banking, Cluster Analysis, Association Rules Analysis

ÖNSÖZ

Veri Madenciliği pek çok alanda yaygın olarak kullanılan, Türkiye'de ve dünyada özellikle son yıllarda oldukça önem kazanmış olan bir tekniktir. Yaygın kullanımı sebebiyle tez konusu olarak Veri Madenciliği seçilmiştir. Bu tez çalışmasında Türkiye'deki internet bankacılığı kullanıcı profillerin belirlenmesi ve internet bankacılığı kullanıcılarının Birliktelik Kuralları ve Kümeleme Analizi incelenmesi amaçlanmıştır.

Tezimin her aşamasında değerli katkılarıyla beni yönlendiren, tez konusunun belirlenmesi ve yürütülmesi aşamasında, her türlü yardımı ve desteğini esirgemeyen, fikirleriyle bana ışık tutan değerli hocam Sayın Doç. Dr. Umman Tuğba Şimşek Gürsoy'a en içten teşekkürlerimi sunmayı bir borç bilirim.

Üniversite hayatımın büyük bir bölümünde aynı evi paylaştığım, 7 yıllık arkadaşlığımız boyunca her an her konuda bana destek olan, yanımda olduğunu hissettiren, dostum, can arkadaşım Gözde Taştan'a ve yanımda olan bütün arkadaşlarıma sonsuz teşekkürlerimi iletirim.

Çalışmamın her alanında yanımda, benimle birlikte koşturan, bana moral kaynağı olan, her konuda en büyük destekçim hayat arkadaşım Ali Akın'a çok teşekkür ederim.

Bugüne gelmemde en büyük paya sahip, her zaman yanımda olan, sevgili babam Aydın Bilgin'e, sevgili annem Remziye Bilgin ve kardeşim Alper Bilgin'e, emekleri, gösterdikleri sonsuz sabır ve anlayış için, hiçbir yardımı esirgemeden, daima yanımda oldukları için tüm kalbimle teşekkür eder ve hazırlamış olduğum bu çalışmamı ithaf ederim.

ŞAFİYE BİLGİN

İÇİNDEKİLER

ÖZET i
ABSTRACT ii
ÖNSÖZ iii
İÇİNDEKİLER iv
TABLOLAR LİSTESİ vii
ŞEKİLLER LİSTESİ xi
KISALTMALAR xiii
GİRİŞ 1
1.VERİ MADENCİLİĞİ 2
1.1. VERİ MADENCİLİĞİ KAVRAMI 3
1.2. VERİ MADENCİLİĞİ SÜRECİ 6
1.2.1. Problemin Tanımlanması 9
1.2.2. Verinin Hazırlanması 11
1.2.3. Modelleme 14
1.2.4. Modelin Değerlendirilmesi 15
1.2.5. Modelin Uygulanması 17
1.3. VERİ MADENCİLİĞİ UYGULAMA ALANLARI 18
1.3.1. Pazarlama 19
1.3.2. Finans 20
1.3.3. Tıp 20
1.3.4. Diğer Uygulamalar 21
1.4. VERİ MADENCİLİĞİNDE KULLANILAN MODELLER 23
1.4.1. Birliktelik Kuralları Analizi 24
1.4.2. Ardışık Zamanlı Örüntüler 25
1.4.3. Kümeleme Analizi 25
1.4.4. Sınıflandırma Teknikleri 26
1.4.4.1. Karar Ağaçları 27
1.4.4.2. Genetik Algoritmalar 29
1.5. GELİŞMİŞ VERİ MADENCİLİĞİ TEKNİKLERİ 30

1.5.1. Web Madenciliği 31
1.5.2. Metin Madenciliği 34
1.5.3. Multimedya Madenciliği 35
1.5.4. Büyük Veri (Big Data) 39
1.6. LİTERATÜR ARAŞTIRMASI 43
2. İNTERNET BANKACILIĞI 60
2.1.İNTERNET BANKACILIĞI KAVRAMI 61
2.2. İNTERNET BANKACILIĞININ DÜNYA'DA VE TÜRKİYE'DEKİ GELİŞİMİ 62
2.3. İNTERNET BANKACILIĞI İLE SUNULAN HİZMETLER 64
2.4. İNTERNET BANKACILIĞININ AVANTAJLARI VE DEZAVANTAJLARI 65
2.5. TÜRKİYE'DEKİ İNTERNET BANKACILIĞI İSTATİSTİKLERİ 68
3.BİRLİKTELİK KURALLARI VE KÜMELEME ANALİZİ 76
3.1. BİRLİKTELİK KURALLARI 76
3.1.1. Apriori Algoritması 79
3.1.2. Birliktelik Kuralları Tekniğinin Güçlü Ve Zayıf Yönleri 84
3.2. KÜMELEME ANALİZİ 84
3.2.1. Kümelemenin Amacı Ve Uygulama Adımları 88
3.2.2. Kümeleme Analizinde Kullanılan Benzerlik Ve Uzaklık Ölçütleri 89
3.2.3. Kümeleme Analizinde Kullanılan Yöntemler 92
3.2.3.1. Hiyerarşik Yöntemler 93
3.2.3.2. Bölümlemeli Yöntemler 94
3.2.3.3. Yoğunluk Tabanlı Yöntemler 99
3.2.2.4. Izgara Tabanlı Yöntemler 99
3.2.3.4. Model Tabanlı Yöntemler 100
4.UYGULAMA 101
4.1.VERİ MADENCİLİĞİ SÜRECİ 103
4.1.1. İnternet Bankacılığı Kullanmayan Müşterileri İçin Veri Madenciliği Süreci 104
4.1.1.1.Problemin Tanımlanması 104

4.1.1.2.Verilerin Hazırlanması 104
4.1.2. İnternet Bankacılığı Kullanan Banka Müşterileri İçin Veri Madenciliği Süreci.....116
4.1.2.1.Problemin Tanımlanması.....116
4.1.2.2.Verilerin Hazırlanması 116
4.1.2.3.Modelleme 132
4.1.2.4. Modelin Kurulması 133
4.1.2.5. Modelin Kullanılması.....159
4.1.2.6. Modelin İzlenmesi.....160
SONUÇ.....161
KAYNAKÇA 165
EKLER 185

TABLOLAR LİSTESİ

Tablo 1.1.Veri Madenciliği Sürecinde Tanımlanabilecek Problemler.....10
Tablo 1.2.Veri Tabanlarında Fazla Değişken Bulunması.....13
Tablo 1.3. Web Madenciliği Kategorileri ve Sınıflandırılması 32
Tablo 2.1. İnternet Bankacılığını Kullanan Müşteri Sayıları 69
Tablo 2.2.İnternet Bankacılığında Yatırım İşlemleri 71
Tablo 2.3. İnternet Bankacılığında Finansal İşlemler 72
Tablo 2.4. İllere Göre İnternet Bankacılığı Kullanan Bireysel Müşteri Sayısı.....73
Tablo 3.1. Birliktelik Kuralı Analizine İlişkin Örnek 81
Tablo 3.2. Destek Değerinin Hesaplanması 82
Tablo 3.3. İkili İşlem Gruplarının Destek Değerleri 82
Tablo 3.4. Üçlü İşlem Gruplarının Destek Değerleri 83
Tablo 3.5. Örnek Müşteri Verileri.....87
Tablo 3.6. Veri Seti 96
Tablo 3.7. Kümelerin Ortalama Değerleri.....97
Tablo 3.8. Kümelerin Ortalama Değerleri.....98
Tablo 3.9.Küme Ortalamalarının Uzaklık Kareleri.....98
Tablo 4.1. Veri Seti 106
Tablo 4.2. IBM Modeler Programına Ait Bir Arayüz ve Değişkenler.....107
Tablo 4.3. Değişkenlere İlişkin Bazı Kodlamalar.....108
Tablo 4.4. Banka Müşterilerinin Demografik Özellikleri İçin Hazırlanan Veri Setinin Değişkenlerinin Dağılımı ve Temel İstatistikleri.....109
Tablo 4.5. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Cinsiyetlerine İlişkin Dağılımları.....109
Tablo 4.6. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Medeni Durumlarına İlişkin Dağılımları 110
Tablo 4.7. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Yaşlarına İlişkin Dağılımları.....110
Tablo 4.8. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Eğitim Düzeylerine İlişkin Dağılımları 111

Tablo 4.9. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Aylık Net Gelirlerine İlişkin Dağılımları.....112
Tablo 4.10. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Çalıştıkları Sektörlere İlişkin Dağılımları 112
Tablo 4.11. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin İnternet Kullanım Yıllarına İlişkin Dağılımları.....113
Tablo 4.12. Banka Müşterilerinin İnternet Bankacılığını Kullanmama Nedenlerine İlişkin Verdikleri Cevapların Dağılımı.....114
Tablo 4.13. İnternet Bankacılığı İnternet Bankacılığını Kullanan Banka Müşterilerine İlişkin Veri Seti 117
Tablo 4.14.İnternet Bankacılığı Kullanılan Bankalar 118
Tablo 4.15. İnternet Bankacılığında Kullanılan Hesaplar 119
Tablo 4.16. İnternet Bankacılığında Yapılan İşlemler 119
Tablo 4.17. İnternet Bankacılığı Kullanılan Bankalarda Aranılan Özellikler 120
Tablo 4.18. İnternet Bankacılığı Kullanan Banka Müşterilerine İlişkin Özet Tablo 121
Tablo 4.19. İnternet Bankacılığı Kullanan Banka Müşterilerinin Cinsiyetlerine İlişkin Dağılımları.....121
Tablo 4.20. İnternet Bankacılığı Kullanan Banka Müşterilerinin Medeni Durumlarına İlişkin Dağılımları 122
Tablo 4.21. İnternet Bankacılığı Kullanan Banka Müşterilerinin Yaşlarına İlişkin Dağılımları.....122
Tablo 4.22. İnternet Bankacılığı Kullanan Banka Müşterilerinin Eğitim Düzeylerine İlişkin Dağılımları 123
Tablo 4.23. İnternet Bankacılığı Kullanan Banka Müşterilerinin Aylık Net Gelir Düzeylerine İlişkin Dağılımları 123
Tablo 4.24. İnternet Bankacılığı Kullanan Banka Müşterilerinin Çalıştıkları Sektörlere İlişkin Dağılımları 124
Tablo 4.25. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Kullanım Yıllarına İlişkin Dağılımları.....125
Tablo 4.26. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığı Kullanım Yıllarına İlişkin Dağılımları 126

Tablo 4.27. İnternet Bankacılığı Kullanan Banka Müşterilerinin Hangi Sıklıkla İnternete Bağlandıklarına İlişkin Dağılımları....126
Tablo 4.28. İnternet Bankacılığı Kullanan Banka Müşterilerinin, İnternet Bankacılığına Nereden Bağlandıklarına İlişkin Dağılımları....128
Tablo 4.29. İnternet Bankacılığı Kullanan Banka Müşterilerinin, Kaç Adet İnternet Bankacılığı Hesabı Olduklarına İlişkin Dağılımları....130
Tablo 4.30. İnternet Bankacılığı Kullanan Banka Müşterilerinin Banka Tercihine İlişkin Dağılımları....131
Tablo 4.31. Birliktelik Tespiti ve Kural Çıkarımı için Verinin Hazırlanması....133
Tablo 4.32. Verinin Birliktelik Kuralları Analizi İçin Hazırlanması....134
Tablo 4.33. İnternet Bankacılığı Kullanan Banka Müşterilerinin, Bankacılık İşlemlerinde En Çok Hangi Hesapları Kullandıklarına İlişkin Kural Setleri....136
Tablo 4.34. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bankacılık İşlemlerinde En Çok Hangi İşlemleri Yaptıklarına İlişkin Birliktelik Kuralları Analizi İçin Veri Seti....138
Tablo 4.35. İnternet Bankacılığını Kullanan Banka Müşterilerinin Bankacılık İşlemlerinde En Çok Hangi İşlemleri Yaptıklarına İlişkin Kural Setleri....140
Tablo 4.36. İnternet Bankacılığı Kullanan Banka Müşterilerinin, İnternet Bankacılığını Kullanmalarına Etki Eden Faktörlere İlişkin Verdikleri Cevapların Dağılımı....142
Tablo 4.37. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığı İşlemlerinde Karşılaşılan Sorunlara İlişkin Verdikleri Cevapların Dağılımı....145
Tablo 4.38. İnternet Bankacılığı Kullanan Banka Müşterilerinin Gelecekte İnternet Bankacılığının Geleneksel Bankacılığın Yerini Alacağını Düşünüp Düşünmediklerine İlişkin Dağılımları....146
Tablo 4.39. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bir Bankanın İnternet Sitesinde Aradıkları Özelliklerine İlişkin Verdikleri Cevapların Dağılımları....147
Tablo 4.40. Kümeleme Analizi için Verilerin Hazır Hale Getirilmesi....150
Tablo 4 41. Önem Dereceleri....151
Tablo 4.42. K-Ortalamalar Yöntemi İle Elde Edilen Kümeler....151
Tablo 4.43. Cinsiyetin Kümeler Üzerindeki Etkisi....152
Tablo 4.44. Medeni Durumun Kümeler Üzerindeki Etkisi....153
Tablo 4.45. Yaşın Kümeler Üzerindeki Etkisi....154

Tablo 4.46. Eğitim Düzeyinin Kümeler Üzerindeki Etkisi 155
Tablo 4.47. Aylık Gelir Düzeyinin Kümeler Üzerindeki Etkisi 156
Tablo 4.48. Çalışılan Sektörün Kümeler Üzerindeki Etkisi 157
Tablo 4.49. Kümelere Ait Dağılım.....158

ŞEKİLLER LİSTESİ

Şekil 1.1.CRISP-DM Süreci 8
Şekil 1.2. CRISP-DM Metodunun Aşamalara Göre Özellikleri....18
Şekil 1.3. Web Madenciliği Sınıflandırması 32
Şekil 1.4. Parmak İzi Tarama İşlemleri 36
Şekil 1.5. Yüz Tanıma Sistemleri 37
Şekil 1.6. Yüz Tanıma Sistemleri 38
Şekil 1.7. Geliştirilmiş Yüz Tanıma Sistemleri 39
Şekil 1.8. Büyük Veri (Big Data)....41
Şekil 2. 1.İnternet Bankacılığı İle Yapılan Hizmetler....64
Şekil 2.2. İnternet Bankacılığında Aktif Bireysel Kullanıcıların İllere Göre Dağılımı ..74
Şekil 2.3. İnternet Bankacılığı Kullanan Aktif Kurumsal Kullanıcıların İllere Göre Dağılımı 74
Şekil 3.1. 2^N Kümesi Gösterilişi 80
Şekil 3.2. Küme Yapısı....85
Şekil 4.1. CRISP-DM Süreci Ara Yüzü 102
Şekil 4.2. CRISP-DM Süreci Ara Yüzü 103
Şekil 4.3. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin İnternet Kullanmama Nedenlerine Göre Dağılım Grafiği....115
Şekil 4.4. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığına Nereden Bağlandıklarına İlişkin Dağılım Grafiği....129
Şekil 4.5. İnternet Bankacılığı Kullanan Banka Müşterilerinin Hangi Bankaların İnternet Bankacılığını Tercih Ettiklerine İlişkin Dağılım Grafiği 132
Şekil 4.6. Web Grafiği Apriori Algoritması ile Kural Setlerinin Elde Edilmesi 135
Şekil 4.7. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bankacılık İşlemlerinde En Çok Hangi İşlemleri Yaptıklarına İlişkin Web Grafiği....139
Şekil 4.8. İnternet Bankacılığı Kullanan Müşterilerin, İnternet Bankacılığını Kullanmalarına Etki Eden Faktörlere İlişkin Verdikleri Cevapların Dağılım Grafiği..144
Şekil 4.9. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığı İşlemlerinde Karşılaşılan Sorunlara İlişkin Verdikleri Cevapların Dağılım Grafiği 146

Şekil 4.10. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bir Bankanın İnternet Sitesinde Aradıkları Özelliklere İlişkin Verdikleri Cevapların Dağılım Grafiği.....149

KISALTMALAR

AIS: Adaptive Importance Sampling

ADK: Alternatif Dağıtım Kanalları

ATM: Automatic Teller Machine

CHAID: Chi-squared Automatic Interaction Detector (Ki-Kare Otomatik Etkileşim Dedektörü)

CRISP-DM: Cross Industry Standard Process for Data Mining

ÇBÖ: Çok Boyutlu Ölçekleme

DBSCAN: Density-Based Spatial Clustering Of Applications With Noise

DENCLUE: Density Based Clustering

EFT: Elektronik Fon Transferi

FBI: Federal Bureau of Investigation (Federal Soruşturma Bürosu)

GYTE: Gebze Yüksek Teknoloji Enstitüsü

GPS: Global Positioning System (Küresel Konumlama Sistemi)

HTML: Hyper Text Markup Language (Zengin Metin İşaret Dili)

IBM: International Business Machines

IMKB: İstanbul Menkul Kıymetler Borsası

NCR: National Cash Register

OLAP: Online Analytical Processing (Çevrimiçi Analitik İşleme

ROI: Return on Investment(Yatırım Getirisi)

SPRINT: Scalable.Parallelizable Induction of Decision Tree (Karar Ağacının Ölçeklenebilir Paralel İndüksiyonu)

SPSS: Statistical Package For Social Sciences

WEKA: Waikato Environment for Knowledge Analysis

GİRİŞ

Büyük miktarda veri içerisinden gelecekle ilgili tahmin yapmamızı sağlayacak modellerin veya bağlantıların oluşturulmasını sağlayan veri madenciliği hayatımızın hemen hemen her alanında kullanılmaktadır. Veri madenciliğinde amaç, kolaylıkla mantıksal kurallara ya da görsel sunumlara çevrilebilecek nitel modellerin elde edilmesini sağlamaktadır. Gelişen dünya, artan veri hacmi ve teknolojik yenilikleri, ham veriden işlenmiş bilgi elde edebilmekte için veri madenciliğinden yararlanılmaktadır.

Tezin birinci bölümünde Veri Madenciliği kavramı ele alınmış; Veri Madenciliği Süreci, Veri Madenciliğinin yaygın olarak kullanıldığı alanlar, veri analizinin yapılabilmesi için kullanılan teknikler ayrıntılarıyla ele alınmıştır. Veri Madenciliği ve Veri Madenciliği teknikleriyle ilgili geniş bir literatür çalışmasına yer verilmiştir.

Tezin ikinci bölümünde veri madenciliğinin etkin bir şekilde kullanıldığı sektörlerden biri olan bankacılık sektöründe; internet bankacılığı kavramı, internet bankacılığının gelişimi, internet bankacılığının avantaj ve dezavantajları ile Türkiye'deki internet bankacılığı kullanımına ilişkin bilgilere yer verilmiştir.

Tezin üçüncü bölümünde uygulamaya temel oluşturan veri madenciliği tekniklerinden Birliktelik Kuralları Ve Kümeleme Analizi Yöntemleri''ne ayrıntılı olarak değinilmiştir.

Tezin son bölümünde, internet bankacılığı kullanan banka müşterilerinin profilleri incelenmiş, hangi internet bankacılığı ürünlerini kullandıkları ve bu ürünleri kullanmalarında etkili olan faktörler araştırılmıştır. Çalışmada ayrıca internet bankacılığı kullanmayan kişilerin profilleri de belirlenmiş ve internet bankacılığı kullanmama nedenleri saptanmıştır. Çalışmada Birliktelik Kuralları Analizi ile müşterileri profilleri belirlenmeye çalışılmış; Kümeleme Analizi ile kullanıcıların internet bankacılığı kullanımına ilişkin yaklaşımları demografik özelliklerine göre ayrılmıştır.

1.VERİ MADENCİLİĞİ

Günümüzde modern veri tabanları çok büyük miktarlarda veriyi kapsamaktadır. Sistemde var olan değişimlere ayak uydurabilmek ve hızla yanıt verebilmenin sağlanması kolay bir iş değildir. Bu durum işlenmemiş ham veriden fayda sağlamanın oldukça zor olduğunu, bilgiye ulaşımın ancak verinin iyi analiz edilmesi ile gerçekleşebileceğini göstermektedir. Özellikle bilim, işletme, pazarlama, otomotiv, tıp gibi pek çok alanda ortaya çıkan ve hızla artan veri sorunu, beraberinde verileri toplama ve saklama alanında da zorluklar çıkarmıştır. Yaşanan veri artışı problemine çözüm bulabilmek için ise veri madenciliğinden yararlanılmaktadır. Veri madenciliği çeşitli araştırma sahalarından farklı teknik ve metodları içermektedir. [1]

Dünya genelinde 1960'larda veri toplama sistemleri, 1970 ve 1980'lerde ilişkisel veri tabanları, 1990 ve 2000'lerde ise bilgisayar ve teknolojik gelişmelere paralel olarak veri madenciliği, veri ambarları, multimedya ve web veri tabanlarının oldukça popüler olduğu görülmektedir. Veri tabanlarında tutulan veri depolarının kullanılması, artan veri sayısı ve bilgi toplumu olma yolundaki gelişmeler veri madenciliği disiplinin daha fazla gündeme gelmesini sağlamıştır. Özellikle son 10 yıldır dünyada hızla yaygınlaşan veri madenciliği kavramının bir disiplinler arası disiplin olarak kullanıldığı görülmektedir.[2] Yurt dışında yaygın bir biçimde kullanılan veri madenciliği, ülkemizde ise yeni yeni tanınmaya ve kullanılmaya başlanmıştır.

[1] Harry S. Singh, **Data Warehousing-Concepts, Technologies, Implementations and Management**, U.S.A., Prentice Hall PTR, 1998, s. 1-2.
[2] John Wang, **Encyclopedia of Data Warehousing and Mining**, 2nd Edition, USA; Idea Group Inc., 2009, s.126.

1.1. VERİ MADENCİLİĞİ KAVRAMI

Bilgi teknolojilerinin hayatımızın her alanında yer almasıyla, sadece bilgiye ulaşmanın yeterli olmadığı, bilginin gerekli şartlar altında üretilmesinin de önemli olduğu görülmüştür. Sürekli büyüyen sayısal veri ortamları arasından faydalı ve gerekli olan bilgiye ulaşmada veri madenciliğinin önemi gittikçe artmaktadır.

Bilgi teknolojilerindeki gelişmelere paralel olarak donanımların ucuzlaması, verilerin uzun süre depolanabilmesini sağlamakta; böylelikle büyük kapasiteli veritabanları oluşmaktadır. Büyük veritabanlarından istenilen anlamlı bilgiye ulaşabilmek, yeni bir disiplin olan Veri Madenciliğinin doğmasına neden olmuştur. Veri madenciliği ile veriler çeşitli istatistiksel tekniklerle analiz edilir ve kurumların karar verme sürecinin etkinliği artar.[3]

İlk olarak 1989 yılında bir atölye çalışmasında tanımlanan veri madenciliği kavramı, veri işleme sürecinde bilginin son ürün olduğunu belirtmek için Veri Tabanlarında Bilgi Keşfi olarak ifade edilmiştir.

Veri madenciliğinde, büyük miktardaki veriden, ayrı sorgular yapılarak yararlı bilgi, desen ve eğilimler (genelde önceden tahmin edilemeyen) ortaya çıkarılmaktadır. Veri madenciliği bilgisayar desteğiyle çözümlenebilmekte, verinin içerisinde yararlı olacak anlamlı bilgi ve bağlantılar bilgisayar programları aracılığıyla elde edilmektedir.[4]

Veri madenciliği, araştırmacıya anlamlı ve faydalı olabilecek biçimde veri kümesinin içindeki ilişkileri ortaya çıkarmak ve yeni veriyi özetleyebilmek için veri kümelerinin incelenmesidir.[5]

---

[3] Halil Kaya, Kemal Köymen, "Veri Madenciliği Kavramı ve Uygulama Alanları", **Fırat Üniversitesi Doğu Anadolu Bölgesi Araştırmaları Dergisi**, Cilt: 2, Sayı:3, 2008, s. 159-164.

[4] Anarbek Kalıkov, "**Veri Madenciliği ve Bir E-Ticaret Uygulaması**", Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2006, s.21.

[5] Daniel.T Larose, **Discovering Knowledge in Data: An Introduction to Data Mining,** John Wiley & Sons Inc., 2005, s.42-70.

Veri tabanlarında yer alan çok büyük ölçekli verilere sahip olan birçok organizasyonun, elde bulunan bilgileri yönetebilmesi ve veriler içerisinden anlamlı bilgilerin çıkarılması, veri madenciliği olarak tanımlanmaktadır.[6]

Aşağıda veri madenciliği ile ilgili çeşitli tanımlara yer verilmektedir:

Jacobs'a göre veri madenciliği, ham verinin bütünüyle orta çıkarılamadığı durumda bilginin tam olarak sunulmasını sağlayan veri analizi sürecidir.

David'e göre veri madenciliği, çok büyük ölçekli veritabanlarındaki örüntüleri araştıran ve matematiksel algoritmaları kullanan bir bilim olmanın yanı sıra aynı zamanda bir sanattır.

Akpınar'a göre veri madenciliği, büyük miktarlardaki verinin içinden geleceğin öngörülmesine yardımcı olacak anlamlı ve faydalı bağlantı ve kuralların, bilgisayar programlarıyla desteklenerek aranması ve analizidir.[7]

Han ve Kamber'e göre veri madenciliği, veritabanlarında bilgi keşfinin bir aşamasını simgelemektedir. Han ve Kamber ayrıca veri madenciliğini, veri tabanları, veri ambarları veya diğer veri havuzlarında depolanmış büyük veri yığınlarından ilginç bilgilerin keşif süreci olarak ifade etmektedir.[8]

DuMouchel ise veri madenciliği sayesinde geniş veritabanlarındaki birlikteliklerin araştırılabildiğini belirtmiştir.[9]

---

[6] Umman Tuğba Şimşek Gürsoy, **Veri Madenciliği ve Bilgi Keşfi**, 1.Baskı, Ankara, Pegem Akademi, 2009, s.27.

[7] Akpınar, **a.g.e.**,, s. 1-22.

[8] Jiawei Han, Micheline Kamber, **Data Mining Concepts and Techniques**, U.S.A., Morgan Kaufman Publishers, 2001, s.14.

[9] William DuMouchel, "Bayesian Data Mining In Large Frequency Tables, With An Application To The FDA Spontaneous", **American Statistician**, 53 (3), 1999, s.177. .

Vahaplar ve İnceoğlu veri madenciliğini, mevcut verilerden çok net olmayan, önceden tahmin edilemeyen ancak potansiyel olarak yararlı bilginin çıkarılabilmesi olarak görmektedir.[10]

Yukarıda ifade edilen veri madenciliği tanımlarından da görüldüğü üzere veri madenciliği; istatistik, veritabanı yönetimi, görselleştirme gibi alanlardan yararlanan disiplinler arası bir yaklaşımdır.[11]

Gartner Group'a göre veri madenciliği, istatistik ve matematiksel tekniklerle birlikte örüntü tanıma (Pattern Recognition) teknolojilerini kullanarak, depolama ortamlarında saklanmış bulunan veri yığınlarının çıkarılması ile anlamlı yeni korelasyon, örüntü ve eğilimlerin keşfedilmesi süreci olarak ifade edilmektedir.[12]

Diğer bir tanıma göre veri madenciliği, kullanıcıların iş yapma biçimleri için önemli olabilecek bilgileri veri tabanlarından kolaylıkla elde etmelerine yardımcı olan bir teknik olarak ifade edilmiştir.[13]

Baykal ise veri madenciliğini, büyük veri yığınlarının oluşturduğu veri tabanı sistemleri içinden gizli kalmış bilgilerin çekilmesini sağlayan bir işlem olarak ifade etmektedir. Baykal aynı zamanda bu işlemlerin; istatistik ve matematik disiplinleri, modelleme teknikleri, veri tabanı teknolojisi ve çeşitli bilgisayar programları kullanılarak yapılabildiğini belirtmiştir.[14]

SPSS firması veri madenciliğini, tahmin edici yöntemleri kullanarak verilerdeki örüntüleri keşfetme işlemi olarak tanımlamaktadır. [15]

---

[10] Alper Vahaplar, Mustafa Murat İnceoğlu, "**Veri Madenciliği ve Elektronik Ticaret**", (Çevrimiçi), www.manisabilisim.org/dokuman/inceoglu.doc, 20 Ocak 2012.
[11] Ad Feelders, Hennie Daniels, and Marcel Holsheimer, "Methodological and Practical Aspects of Data Mining", **Information & Management**, Vol:37, 2000, s.271-281.
[12] **The Gartner Grup**, (Çevrimiçi), www.gartner.com, 14 Ocak 2012.
[13] Alex Berson, Stephen J. Smith, **Data Warehousing, Data Mining & OLAP**, New York, McGraw-Hill Companies, 1998, s. 333.
[14] Abdullah Baykal, "Veri Madenciliği Uygulama Alanları", **D.E.Ü. Ziya Gökalp Eğitim Fakültesi Dergisi,** Sayı:7, 2006, s.95-107.
[15] **"Data Mining",** (Çevrimiçi) http://www.spss.com/data_mining/, 20 Ocak 2012.

Veri madenciliğinde asıl amaç, kolay bir biçimde mantıksal kurallara veya görsel sunumlara dönüştürülebilecek nitel modellerin ortaya çıkarılmasıdır.[16]

Nobel Ödüllü Ünlü Fizikçi Arno Allan Penzias, 1999 yılında Computer World'e verdiği röportajında işletmeler açısından artan bir öneme sahip olan veri madenciliği hakkında şunu söylemiştir:

"*Veri madenciliği şimdikinden daha önemli bir hale gelecektir ve işletmeler müşterilerine ait hiçbir şeyi dışarı atamayacaktır. Çünkü bunlar işletmeler için çok büyük şey ifade edecektir. Eğer bunu yapmıyorsanız, iş hayatı dışında kalmışsınız demektir.* [17]

1.2. VERİ MADENCİLİĞİ SÜRECİ

Veri madenciliği keşfedilmemiş bilgilerin ortaya çıkarılma sürecidir. Veri madenciliğinin bir süreç olarak ifade edilebilmesi için, sürecin her bir aşamasının dikkatle izlenip gerçekleştirilmesi gerekmektedir. Çünkü veri madenciliğinde her aşama bir önceki aşamanın sonuçlarıyla ilişkilidir.

Veri madenciliği için belirlenen standart bir süreç söz konusudur. Bu süreç ile ilgili olarak en yaygın kullanılan model CRISP-DM (Cross Industry Standard Process for Data Mining) modelidir. [18]

Temel bir takım ortak uygulamaları içermekle birlikte, veri madenciliği süreci başlangıçta farklı uygulamacılar tarafından farklı şekilde yürütülmekteydi. Bu nedenle, uygulamada meydana gelen farklılıkları en aza indirecek bir standart süreç geliştirilmesi ihtiyacı olduğu görülmüştü. Bu ihtiyaç doğrultusunda düzenlenen çeşitli çalıştaylar ile veri madenciliği uygulamalarının gelişmesi ve tüm

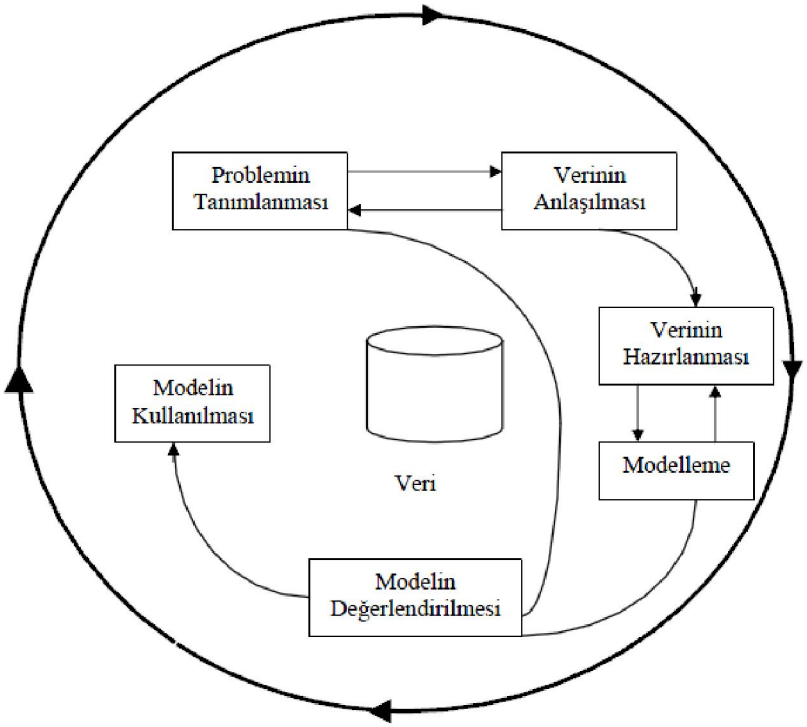
[16] Seçil Dönmez, "**Bayi Performans Değerlendirmesinde Bir Veri Madenciliği Uygulaması**", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, 2008, s.17.
[17] Teh Ying Wah et. al., "Data Mining in Computer Auditing, Insite "Where Parallels Intersect", **Informing Science**, 2002, s.1565.
[18] Gürsoy, **a.g.e.**, s.30.

uygulamacılara yardımcı olacak standart bir süreç modeli oluşturulması amacıyla CRISP–DM konsorsiyumu oluşturulmuştur.[19]

The Cross-Industry-Standard Process for Data Mining (CRISP-DM) konsorsiyumu, 1996 yılının sonlarına doğru üç firma tarafından kurulmuştur. Bu üç firmanın ilki olan Daimler Chrysler veri madenciliği tekniklerini uygulama konusunda öncü olarak gösterilmektedir. SPSS (Statistical Package for the Social Sciences) firması 1990 yılından bu yana veri madenciliği üzerine çeşitli hizmetler sağlamaktadır. SPSS firması ilk ticari veri madenciliği çalışma platformu olan Clementine'i 1994 yılında faaliyete sunmuştur. NCR(National Cash Register) firması ise müşterilerini daha da artırmak, müşteri memnuniyetini sağlamak ve alıcılarının ihtiyaçlarını daha iyi karşılayabilmek için birçok veri madenciliği danışmanlığı ve teknoloji uzmanlığı takımları oluşturmuştur.

Bu gelişmelerden bir yıl sonra, sözcüklerin baş harfi "Cross –Industry Standard Process for Data Mining" açılımında olan CRISP-DM konsorsiyumu oluşturulmuş, Avrupa Komisyonundan fon elde edilmiş ve başlangıç fikirleri oluşturulmaya çalışılmıştır. Altı adımdan oluşan CRISP-DM süreci etkileşimli ve yinelemeli bir süreçtir.[20]

[19] Colin Shearer, "The CRISP-DM Model: The New Blueprint for Data Mining," **Journal of Data Warehousing,** No:5, 2000, s.13-22.
[20] Gürsoy, **a.g.e.**, s.30.



Şekil 1.1.CRISP-DM Süreci[21]

CRISP-DM, veri madenciliği ile ilgilenen kişi veya organizasyonlar tarafından en çok kullanılan yöntembilimdir. Bir veri madenciliği sürecinde atılması gereken adımların tümünün detayını açıklamaktadır.

Birçok kurum sahip oldukları kaynaklarına, kurum ve sisteme ilişkin problemlerine, sahip oldukları veriler ve diğer kaynaklara göre kendi veri madenciliği sürecini oluşturmaktadır. Fakat veri madenciliği sürecinin oluşturulması esnasında yapılan yanlışlıklar, sürecin etkinliğine zarar verici boyutta olabilmektedir. [22]

[21] Umman Tuğba Şimşek, "**Veri Madenciliği Ve Müşteri İlişkileri Yönetiminde (CRM) Bir Uygulama",** Doktora Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2006, s.8.
[22] Larose, **a.g.e.**, s.215.

1996 yılında veri madenciliği için sektörler arası standart süreç olarak tanımlanan daha sonraları CRISP-DM adıyla anılan veri madenciliği süreci aşağıdaki aşamalardan oluşmaktadır:

- Problemin tanımlanması
- Verinin Anlaşılması
- Verilerin hazırlanması
- Modelleme
- Modelin değerlendirilmesi
- Modelin uygulanması.[23]

1.2.1. Problemin Tanımlanması

Veri madenciliği disiplinler arası bir yapıya sahip olduğu için; veri tabanları, makine öğrenimi, istatistik ve veri görselleştirme gibi pek çok farklı disiplinden yararlanmaktadır. Bu durum göz önüne alındığında; veri tabanlarında güçlü bir veri yönetiminin olmaması ve makine öğreniminin sağladığı pratik veri analiz yöntemlerinin yer almaması; veri madenciliğinin pozitif sonuçlar üretememesine neden olmaktadır.[24]

Veri madenciliği çalışmalarında problemin tanımlanma süreci, ilk ve en önemli aşamadır. Bu süreçte başarılı olmanın en temel şartı, projenin hangi işletme amacı için yapılacağının açıkça belirlenmesi ve elde edilen sonuçların başarı düzeylerinin ne şekilde ölçüleceğinin açıkça tanımlanmasıdır. Bunun yanı sıra yanlış tahminler sonucunda katlanılacak olan maliyetlerin belirlenmesi, doğru tahminlerde

[23] Hüseyin Özçınar, "**KPSS Sonuçlarının Veri Madenciliği Yöntemleriyle Tahmin Edilmesi**", Yüksek Lisans Tezi, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Denizli, 2006, s.7.
[24] Zhi-Hua Zhou. "Three Perspectives Of Data Mining", **Artificial Intelligence,** Vol:143**,** 2003, s.139–146.

kazanılacak maksimum faydalara ilişkin tahminler bu aşama sürecinde dikkate alınmalıdır.[25]

Tablo 1.1'de[26] veri madenciliğinin cevaplanmasına yardımcı olduğu bazı sorulara yer verilmiştir.

Tablo 1.1.Veri Madenciliği Sürecinde Tanımlanabilecek Problemler[27]

Müşteri Davranışı	▪ *Firmadan sıklıkla ürün alan müşterilerin ortak özellikleri nelerdir?* ▪ *Karlı müşterilerin sıklıkla aldıkları ürünler nelerdir?* ▪ *Karlı müşterilerin ücret ödeme şekli nasıldır?*
Müşterilerin Demografik Özellikleri	▪ *Karlı müşterilerin cinsiyetleri nelerdir?* ▪ *Karlı müşterilerin ikamet ettikleri yer neresidir?* ▪ *Karlı müşterilerin sahip oldukları çocuk sayısı nedir?*
Zamana Bağlılık	▪ *Karlı müşteriler hangi sıklıkla alışveriş yapmaktadır?*

Problemin tanımlanması aşamasında, uygulamaların hangi sektörde ne amaçla kullanılacağı, hedeflerin neler olduğu, işletmenin amaçlarına ulaşabilmesi için nelere ihtiyaç duyduğu ve işletmenin çalışma sonucunda elde edilecek bilgiyi nasıl değerlendireceğinin net bir şekilde belirlenmesi gerekmektedir. Çünkü ortaya koyulan amaç, ihtiyaçlar ve sınırlamalar sonucunda, veri madenciliği problem tanımı formuna çevrilir ve belirlenen hedeflere ulaşmak için stratejiler oluşturulur.[28]

[25] Sinem Akbulut, "**Veri Madenciliği Teknikleri İle Bir Kozmetik Markanın Ayrılan Müşteri Analizi Ve Müşteri Segmentasyonu**", Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Anabilim Dalı, Ankara, 2006, s.15.
[26] Şimşek **a.g.e**, s.8.
[27] Gürsoy, **a.g.e.**, s.32.
[28] Özçınar, **a.g.e.**, s.8.

1.2.2. Verinin Hazırlanması

Verinin hazırlanması aşaması da veri madenciliği için en önemli aşamalardan biridir, çünkü ham veriden kullanılacak en son veri kümesini oluşturmak için gerçekleştirilen bütün işlemleri kapsayan bir süreçtir. Bu kısım, analistin toplam zaman ve enerjisinin % 70'inden fazlasını harcamasına neden olmaktadır. Bu aşamada daha önceden tespit edilen sorunlarla ilişkin kayıt ve değişken seçimi yapılarak tablolar oluşturulur. Değişkenlere uygun olan verilerin temizliği yapılır.

Veri ambarında toplanan verilerde hatalar, gereksiz değerler yer alabilir. Bunun yanı sıra eksik kalan değişkenler ve bu değişkenlere ilişkin aşırı değerler söz konusu olabilir.

Verinin hazırlanması aşaması veri toplanması, verilerin birleştirilmesi ve temizliği, verinin seçilmesi ve gerekli durumlar için yeni değişkenler oluşturma ve modelleme araçları için verileri dönüştürme işlemlerinden oluşmaktadır.

Bu süreç; öncelikle veriyi toplamakla başlar ve veri kümesinde hangi değişkenlerin yer alacağının saptanması, bu değişkenlerin ve değerlerinin neleri ifade ettiklerinin belirlenmesi ile devam eder.[29]

Verilerin Toplanması: Veri madenciliği modeli oluşturma sürecinde ilk olarak, tanımlanan problem için gerekli olduğu düşünülen veriler toplanır. Bu aşamada toplanan bu verilerin kaynaklarının belirlenmesi gerekmektedir. Verilerin nereden, nasıl ve hangi şartlar altında toplandığı, işletme ve organizasyon için büyük önem taşımaktadır. Çünkü gereğinden az veri kaynağının olması çalışmayı eksik bırakabilirken, gereğinden çok veri olması sürecin uzamasına neden olabilmektedir. Bu durum ise veri kirliliğine yol açabilmektedir.

[29] Mustafa Aykut Göral, "**Kredi Kartı Başvuru Aşamasında Sahtecilik Tespiti İçin Bir Veri Madenciliği Modeli**", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2007, s. 14.

Verilerin toplanması aşamasında işletmenin sahip olduğu veri kaynaklarının dışında çeşitli veri tabanları ve veri pazarlayan kurumların veri tabanlarından yararlanılabilir. Ancak bu kurumlardan yararlanılırken, toplanan verilerin ne ölçüde doğru ve uyumlu oldukları incelenmeli, güvenilir olmayan verilerin kullanılmamasına dikkat edilmelidir.[30] Farklı veri kaynaklarından elde edilen veriler, uyumsuzluklar ve farklılıklara neden olabilmekte, verilerin güvenilir olmayan veri kaynaklarından kullanımı ise tüm veri madenciliği sürecinin güvenirliliği etkileyebilir.

Verilerin Birleştirilmesi ve Temizlenmesi: Veri madenciliğinin bu aşamasında farklı veri kaynaklarından toplanan veride belirlenen problemler ve uyumsuzluklar çözümlenerek, verilerin tek bir veri tabanında toplanması gerçekleştirilir. Ancak basit yöntemlerle ve gelişi güzel yapılacak problem giderme işlemlerinin, gelecek aşamalarda daha büyük problemlerin kaynağı olacağı göz ardı edilmemelidir.[31]

Veri madenciliğinde farklı veri kaynaklarından toplanan verilerin kullanımından kaynaklanan veri uyumsuzluklarının büyük çoğunluğu kodlama farklılıklarından, verilerin farklı zamanlara ait olmalarından, farklı ölçü birimlerine sahip olmalarından oluşur. Süreç içerisinde farklı kaynaklardan elde edilen verilerin birleştirilmesi ve temizlenmesi gerekir.

Örneğin bir veri tabanında kayıtlı kişilerin hem yaşları hem de doğum tarihleri yer alıyorsa, bu değişkenlerin biri fazla olabilir. Bu durumda elimizde bulunan verilerin birleştirilmesi ve tek bir değişken gibi işleme alınması gerekmektedir.[32] Tablo 1.2 'de bu duruma ilişkin bir örnek yer almaktadır.

---

[30] Akbulut, **a.g.e**. s. 17.
[31] Nur Karadeniz, "**Müşteri İlişkileri Yönetimi Açısından Veri Madenciliği Yöntemi Ve Hizmet Sektörü Üzerine Bir Uygulama**", Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı, İstanbul, 2008, s.39.
[32] Gürsoy, **a.g.e.**, s. 33.

Tablo 1.2.Veri Tabanlarında Fazla Değişken Bulunması[33]

Doğum Tarihi	Yaşı
10.03.1978	35
05.04.1986	27
04.01.1987	26

Tablo 1.2.'de görüldüğü üzere, veri tabanındaki müşterilere ilişkin hem doğum tarihi hem yaş değişkenleri yer alıyorsa, hangi değişken analiz için anlamlıysa o değişken seçilir ve diğer değişken analizden çıkartılır.

Eğer elimizdeki verilerde eksiklikler söz konusu ise veriler zaman içinde tamamlanmaya çalışılır. Örneğin kullanılan veri tabanı küçük ise ve kayıp verilere ulaşma imkanı ve zamanı varsa, en önemlisi bu eksik verilere kesinlikle ihtiyaç duyuluyorsa, veriler elle teker teker doldurulabilir.[34] Eksik verilerin tamamlanması için diğer bir alternatif ise bilgilerin tahmin yöntemiyle tamamlanmasıdır.

Bunların yanı sıra veride bulunan eksik veya kayıp bilgiler de gözden geçirilmelidir. Örneğin bir veri tabanında yer alan kişilerin medeni hali tanımlıyken, bazı kayıtlarda bu bilgi eksik belirtilmiş olabilir. Bu durumda eksiklik "**kayıp veriler**" olarak ifade edilmektedir. Bunun dışında, bazı kayıtlarda aşırı uç değerler (outlier) veya yanlış girilmiş değerler söz konusu olabilir. Bu tür bilgilere de **gürültü (noise)** adı verilmektedir.[35]

Süreç içerisinde diğer bir durum ise veri girme aşamasındaki hatalardır. Değerlerin yanlış girilmesi veya göz ardı edilmesi sonucunda, analiz sonuçları çok farklı bir biçimde değişebilir. Bu sebeple aykırı değerler gözden geçirilerek atılır, analizde yapılabilecek değişiklikler hesaplanarak, uygun ise veri tabanından silinir. Böylece bütün veriler tek bir veri tabanında düzenli bir şekilde tutulabilir.

---

[33] **A.e.**, s 34.
[34] Gökhan Silahtaroğlu, **Kavram ve Algoritmalarıyla Temel Veri Madenciliği,** 1. Basım, İstanbul, Papatya Yayıncılık, Mayıs, 2008, s.21.
[35] Han Kamber, **a.g.e**, s.107.

Verinin Seçilmesi: Verinin seçilmesi adımında, model kurulumu gerçekleştirilmeden önce veri seçimi yapılır. Burada tahmin edici bir model için örnek vermek gerekirse, bu adım bağımsız ve bağımlı değişkenlerin belirlenmesi ve modelde ele alınacak veri kümesinin seçilmesi demektir.

Sıra numarası, kimlik numarası gibi anlamlı olmayan ve diğer değişkenlerin modeldeki etkinliğinin azalmasına da sebep olabileceğinden dolayı, bu değişkenlerin modele girmemesi gerekmektedir. Bazı veri madenciliği algoritmaları konu ile ilgisi olmayan bu tür değişkenleri otomatik olarak elemekte ancak, pratikte bu işlemin kullanılan yazılıma bırakılmaması daha mantıklı olmaktadır.[36]

Bağımsız değişkenlerin seçilmesinde, verilerin görselleştirilmesine imkan sağlayan grafik araçlar ve bunların sunduğu ilişkiler önemli faydalar sağlayabilir. Yaygın olarak yanlış veri girişinden ya da bir defaya özgü bir olayın gerçekleşmesinden kaynaklanan verilerin, önemli bir bilgiyi içerip içermediği kontrol edildikten sonra veri kümesinden atılması tercih edilir.[37]

1.2.3. Modelleme

Modelin kurulma aşaması, veri madenciliğinde bilgi kaynaklarından maksimum verimin alınabilmesinin sağlanması açısından çok önemlidir. İyi kurulmuş bir model, analiz sonucunda elde edilecek sonuçların kalitesini de değiştirecek, etkileyecek bir faktördür. Analiz sonucunda hangi örüntülerin bulunabileceğini tahmin edebilmesi için iyi bir veri madenciliği uygulayıcısına ihtiyaç duyulmaktadır. Çünkü model kurma süreci, analiz için kullanıcıya hangi verilerin elde hazır bulunduğunu sunar. Bu süreçte model doğru kurulmadığı takdirde, veri seti içerisinde bulunabilecek kritik ilişkilerin doğru bir şekilde

[36] Şengül Doğan, "**Veri Madenciliği Kullanarak Biyokimya Verilerinden Hastalık Teşhisi**",Yüksek Lisans Tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Biyomühendislik Anabilim Dalı, Elazığ, 2007, s. 19-20.
[37] Karadeniz, **a.g.e**., s.40.

sunulması ve önemli örüntüler tespit edilmesi pek mümkün değildir. Bu sebepten dolayı modelden başarılı sonuç elde etme olasılığı da azalır.[38]

Modelleme aşamasında, farklı modelleme yöntemleri seçilerek uygulanması ve optimum değerlere ulaşabilmek için parametrelerin ayarlanması gerçekleştirilmektedir. Bu aşamada, seçilen yöntemler veri setleri üzerinde uygulanarak ortaya çıkan matematiksel denklemler yorumlanır. Bu süreçte yinelemeler oldukça performans iyileşmekte, sonuçlar daha güvenilir bir hal almaktadır. Süreçte seçilen veri madenciliği aracının özellikli olarak hangi algoritmaları, hangi yöntemleri kullanacağına ve hangilerinin modele en uygun olacağına karar verilmeye çalışılır.[39]

Modelde kullanılan veri tabanı çok büyük olarak karşımıza çıkabilir. Bu durumda tesadüfiliği bozmayacak şekilde örnekleme yapılması daha uygundur. Günümüzde hesaplama olanakları, işletim sistemleri ne kadar gelişmiş olursa olsun, çok büyük veri tabanları üzerinde çok sayıda modelin denenmesi zaman faktörü göz önüne alındığında, çok mümkün değildir. Bu sebeple veri tabanının tümü kullanılarak bir kaç model denemek yerine, tesadüfi olarak örneklenmiş bir veri tabanı ele alınarak, bu örneğin üzerinde birçok modelin denenmesi ve bunlar arasından en güvenilir ve güçlü modelin seçilmesi daha uygun görülmektedir.[40]

1.2.4. Modelin Değerlendirilmesi

Bir veri madenciliği problemi için çeşitli pek çok yöntem kullanılmaktadır. Problem için uygun yöntem veya yöntemler belirlenerek, bu yöntemler arasından en uygun olanlar seçilir. Model oluşturulduktan sonra uygulanacak yöntemin

---

[38] Michael J.A Berry, Gordon Linoff, **Data Mining Solutions**, U.S.A., Wiley Computer Publishing, 1998, s.25.

[39] Barış Kiremitçi, "**Veri Ambarlarında Veri Madenciliği Ve Ulaştırma-Lojistik Sektöründe Bir Uygulama**", Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2005, s.52.

[40] Gülşah Aynekin, "**İnternet İçerik Madenciliğinde Yapay Sinir Ağları Ve Bir Uygulama**", Yüksek Lisans Tezi, Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Bursa, 2006, s.26.

gereksinimlerine uygun olarak veri hazırlanma aşamasına tekrar dönülüp incelenerek gerekli görülen değişiklikler uygulanabilir.[41]

Klasik yöntemler kullanıldığında tanımlanan problem için en uygun modelin belirlenebilmesi, mümkün olduğunca çok sayıda modelin kurularak denenmesi ile gerçekleşebilmektedir. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele ulaşılıncaya kadar tekrarlanan bir süreçtir. Günümüzde ise son teknolojileri bünyesinde barındıran bazı yazılımlar sayesinde artık karşılaştırmalara gerek kalmadan en iyi ve uygun olduğu düşünülen tek bir model kullanıcılara sunulabilmektedir.[42]

Kurulan modelin doğruluk derecesi ne denli yüksek olursa olsun, modelin bütünüyle gerçek hayatı modellediğini kesin olarak söylemek mümkün değildir. Yapılan testler sonucunda model kurulumu aşamasında kabul edilen varsayımlar ve modelde kullanılan verilerin doğru olmaması modelin geçerli olmamasına sebep olabilir. Örneğin modelin kurulumu esnasında, bireylerin satın alma davranışları araştırılırken, varsayılan enflasyon oranının zaman içerisinde değişmesi, bireyin satın alma davranışlarını açık bir şekilde etkileyecektir.

Modelin değerlendirilmesi aşamasında, modelin kurulup, geçerliliğine karar verilmesinden sonra, modelin iş amacına uygun olup olmadığı değerlendirilir. Sonuçlar elde edildikten sonra geçerlilik, yenilik, yararlılık ve basitlik açılarından oluşan sonuçlar yorumlanır. Bu aşamanın sonunda ise, ulaşılan veri madenciliği sonuçlarının kullanılıp kullanılmayacağı kararı verilmektedir.

Veri madenciliği analisti modelleri belirli değerleme kriterlerine göre sıralamaktadır. Mümkün olduğunca iş hedeflerini ve iş başarı kriterlerini de göz önünde bulunduran veri analisti, pek çok veri madenciliği projelerinde farklı alternatiflere ait teknikleri birden fazla biçimde modele uygulamaktadır. Çıkan

[41] Derya Ay,"**Veri Madenciliği ve Apriori Algoritması İle Bir Süpermarket Uygulaması**", Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Sakarya, 2009, s.13.
[42] Göral, **a.g.e.**, s.24.

sonuçları karşılaştırmaktadır. Bu karşılaştırmalar en iyi modelin bulunduğu düşünülene kadar devam ettirilmektedir.[43]

1.2.5. Modelin Uygulanması

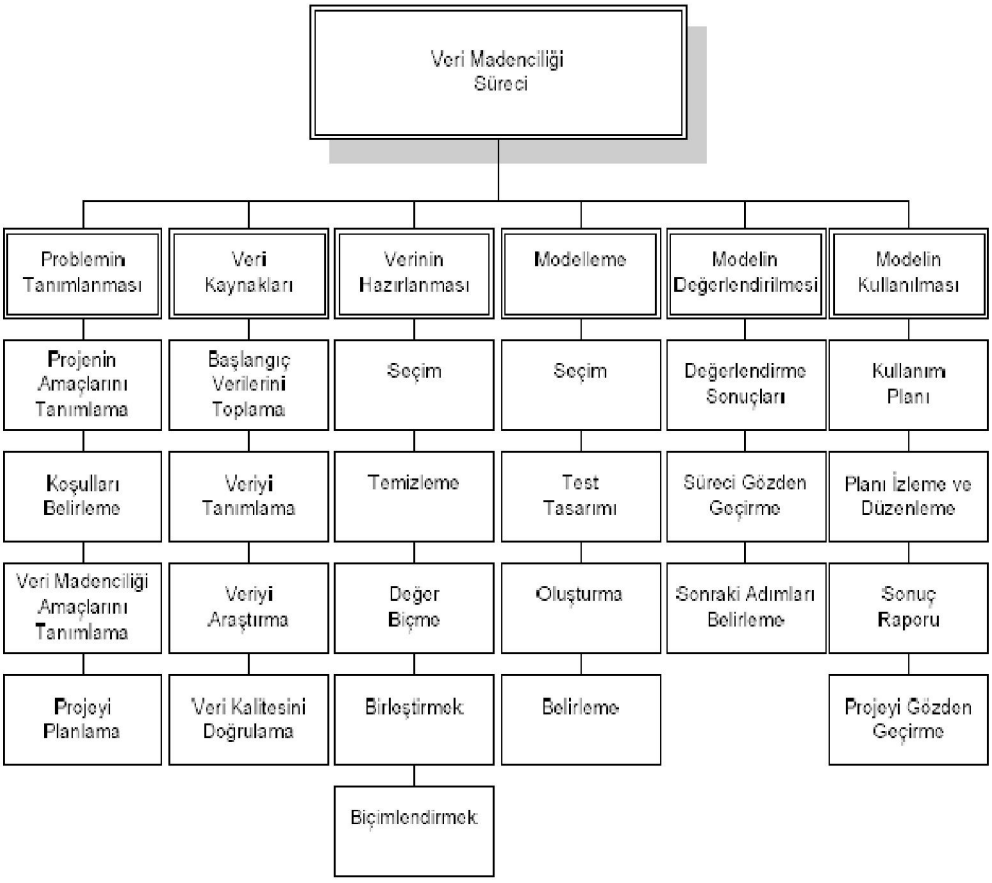
Kurulan ve geçerliliği kabul edilen modelin uygulanması veri madenciliğin son aşamasıdır. Burada model doğrudan bir uygulama olarak ya da başka bir modelin alt parçası olarak da kullanılabilir. Yaşayan bir süreç olması açısından uygulanan modelin zaman içerisinde izlenip, ortaya çıkan değişikliklerin modele yansıtılması gerekmektedir. Bu, süreç için vazgeçilmez bir koşuldur.[44]

Ayrıca geçerliliği kabul edilen bu modeller risk analizi, kredi değerlendirme, dolandırıcılık tespiti gibi işletme uygulamalarında kullanılabilmektedir.[45]

Tüm bu süreç içinde gerçekleşen adımlara ilişkin olarak Şekil 1.2 incelenirse; veri madenciliği sürecinin ayrıntıları açık bir şekilde görülebilir.

CRISP-DM metodunun aşamalara göre özellikleri Şekil 1.2'de gösterilmiştir.

[43] Shearer, **a.g.e.**, s.16-18.
[44] Akpınar, **a.g.e.**, s.11.
[45] Aynekin, **a.g.e.**, s.26.



Şekil 1.2. CRISP-DM Metodunun Aşamalara Göre Özellikleri[46]

1.3. VERİ MADENCİLİĞİ UYGULAMA ALANLARI

Veri madenciliği pek çok konuyu ele alan disiplinler arası bir yaklaşım olarak bilinmektedir. Bu sebeple yeni bir disiplin olmasına karşın uygulama alanı oldukça geniştir. Günümüzde tüm işletmeler sahip oldukları müşterilerin davranışlarını tahmin etmek istemekte ve bunun için uğraşmaktadır. Bu noktada veri madenciliği en çok yararlanılabilecek tekniklerden biridir.

Her geçen gün büyüyen ve gelişmeye çalışan işletmeler veri madenciliğini etkin olarak kullanmaktadır. Özellikle veri madenciliği günümüzde pazarlama,

[46] Turgut Özseven, (Çevrimiçi) http://www.crispdm. org/CRISPWP-0800.pdf,, 20 Nisan 2012.

perakendecilik, bankacılık ve finans, ulaşım, tıp ve daha pek çok farklı alanın içerisindedir. Aşağıda veri madenciliğinin yaygın olarak kullanıldığı alanlar ele alınmıştır. [47]

1.3.1. Pazarlama

Günümüz serbest rekabet ortamında, satış ve pazarlama faaliyetleri açısından zaman ve piyasa verilerinin karar destek amaçlı bilgi haline dönüştürülmesi oldukça önemli bir yere sahiptir. Bu kararların hayata geçirilmesi ve doğru bir şekilde kullanılması, işletmelerin varlıklarını devam ettirmesine yardımcı olmaktadır.

Veri madenciliği pazarlama alanında;

- Müşteri gruplaması,
- Müşterilerin demografik özellikleri arasındaki bağlantıların kurulması
- Müşterilerin satın alma alışkanlıklarının belirlenmesi,
- Çeşitli pazarlama kampanyalarının oluşturulması,
- Mevcut müşterilerin elde tutulması için geliştirilecek pazarlama stratejilerinin belirlenmesi,
- Yeni müşterilere ulaşabilmek için etkili faktörlerin belirlenmesi,
- Hedef pazarın belirlenmesi,
- Pazar sepeti analizi,
- Çapraz satış analizleri,
- Müşteri değerleme,
- Müşteri ilişkileri yönetimi,
- Müşteri analizleri ve satış tahminlerinde kullanılmaktadır.[48]

---

[47] Chris Rygielski, Jyun-Chen Wang, David C. Yen, "Data Mining Techniques for Customer Relationship Management", **Technology in Society**, Volume 24, Issue 4, 2002, s.488.

[48] Özge Sevda Zeyrek, "**Müşteri İlişkileri Yönetiminde Veri Madenciliği Ve Bir Uygulama**", Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2006, s.84.

1.3.2. Finans

Veri madenciliği finans sektöründe pek çok alanda etkin bir şekilde kullanılmaktadır. Finans sektöründe veri madenciliği:

- İşletme riskinin azaltılması,
- Risk derecelendirme tahminlerinin yapılması,
- Müşteri profillerinin analiz edilerek çapraz satış tahminlerinin yapılması,
- Finansal ürünlerin veya poliçe türünün tespit işlemleri,
- Karlılık analizleri,
- Firma maliyetlerin düşürülmesi,
- ATM'lere gün içinde dağıtılacak para miktarının tespiti,
- Finansal işlemler için doğru ve etkin kredi karar verebilme,
- Müşteri ve çalışan memnuniyetinin arttırılması için kullanılmaktadır[49]

Bankacılıkla ilgili olarak ise;

- Farklı finansal göstergeler arasındaki gizli korelasyonların bulunması,
- Müşteri dağılımların belirlenmesi,
- Kredi taleplerinin değerlendirilmesi,
- Kredi kartı dolandırıcılıklarının tespiti,
- Usulsüzlük tespiti,
- Risk analizleri ve
- Risk yönetiminde veri madenciliği uygulamaları sıklıkla kullanılmaktadır.[50]

1.3.3. Tıp

- Test sonuçlarının tahmini,
- Ürün geliştirme,

[49] Tuğba Özel, **Finans Sektöründe İstatistik ve Veri Madenciliği Uygulamaları**, SPSS sunumu, 27 Ocak 2012.
[50] Kalıkov, **a.g.e**., s.10.

- Ameliyat prosedürlerinin belirlenmesi,
- Hastalıkların tespiti,
- Tedavi sürecinin belirlenmesi,
- Hekimlerin teşhis ve tedaviye ilişkin kararlarını etkileyen verilere erişimin sağlanması,
- Magnetik rezonans verileri ile sinir sistemi bölge ilişkilerinin ortaya çıkarılması ve
- Laboratuar sonuçlarının, röntgen, MR gibi görüntü dosyalarının tek kayıt içerisinde düzenlenmesinde kullanılmaktadır.[51]

1.3.4. Diğer Uygulamalar

Sigortacılık ile ilgili olarak veri madenciliği;

- Dolandırıcılıklarının tespiti,
- Yeni poliçe talep edecek müşterilerin tahmin edilmesi ve
- Riskli müşteri tipinin belirlenmesinde kullanılmaktadır.

Veri madenciliğinin telekomünikasyon alanındaki kullanım amaçları aşağıdadır:

- Kalite ve iyileştirme analizleri,
- Hisse tespitleri,
- Hatların yoğunluk tahminleri,
- İletişim desenlerinin belirlenmesi,
- Kaynakların daha iyi kullanılması,
- Servis kalitesinin arttırılması,[52]
- Ses iletimi, faks, cep telefonu hizmetleri, fotoğraf aktarımı, e-posta, internet erişimi, kullanıcı arama süresi, zamanı vb. çeşitli bilgilerin elde edilmesi,
- Müşteri hesaplarına sahte girişlerin tespit edilmesi,

[51] Zeyrek, **a.g.e**., s.86.
[52] Akbulut, **a.g.e**., s.10.

- Belirli bölgelerdeki kullanımların incelenmesi, verimliliğin artırılması, müşteri sayısının artırılması, servisin kalitesini arttırmak gibi hizmetlerde veri madenciliği kullanılmaktadır [53]

Veri madenciliği perakendecilik alanında;

- Satış noktası veri analizleri,
- Özel uygulama etkinliklerin gerçekleştirilmesi,
- Alışveriş sepeti analizleri,
- Tedarik ve mağaza yerleşim optimizasyonunda kullanılmaktadır.

Endüstri alanında veri madenciliği;

- Kalite kontrol analizlerinde kullanılmaktadır.

Lojistik ile ilgili olarak veri madenciliği;

- Üretim süreçlerinin optimizasyonunda kullanılmaktadır.

Spor alanında veri madenciliği kullanılarak;

- Bir basketbol takımına alınacak oyuncunun gelecekteki performansı ve takıma sağlayacağı yararlar tahmin edilebilir.

Kütüphanecilik alanında:

- Bir müşterinin aldığı kitabı ne zaman getireceği ve bir sonraki gelişinde hangi kitabı seçeceği konusunda tahminler yapılabilir.

[53] Gaurav Gupta, Geetika Hans, Tamanna Sehgal, Applications & Trends In Data Mining, (Çevrimiçi) www.rimtengg.com/coit2007/.../94.pdf . 20 Şubat 2012.

Turizm alanında:

- Bir bölgeye turistlerin niçin geldiğini tespit ederek reklam kampanyası düzenlemek ve bir sonraki sezonda turist sayısını arttırmak için veri madenciliği kullanılabilir.[54]

Web alanında;

- Kullanıcıların profilini belirleyip onlara uygun ürünlerin reklam kampanyalarını sayfalara koyma,
- En sık beraber ziyaret edilen çift sayfaların belirlenmesi için veri madenciliği kullanılabilir.[55]

Bilim ve Mühendislik ile ilgili olarak veri madenciliği;

- Ampirik veriler üzerinde modeller kurarak bilimsel ve teknik problemlerin çözümlenmesinde, NASA atom ve uzay araştırmalarında kullanılabilir.

1.4. VERİ MADENCİLİĞİNDE KULLANILAN MODELLER

Veri madenciliği kullanıldıkları alanlara göre değişik modellere ayrılmaktadır. Bu modeller literatürde operasyon veya yöntem isimleriyle de anılmaktadır. Bu operasyonlar ya da yöntemler uygulamada değişik amaçlar için kullanılırken birçok teknik ve algoritmalardan yararlanılmaktadır[56].

En çok kullanılan teknikler;

- Birliktelik Kuralları Analizi
- Ardışık Zamanlı Örüntüler
- Kümeleme Analizi ve
- Sınıflandırma Teknikleridir.

---

[54] Larose, **a.g.e.**, s.42-70.
[55] Kalıkov, **a.g.e.**, s.14.
[56] Peter Cabena et. a.l, **Discovering Data Mining: From Concept to Implementation**, U.S.A., International Business Machines Corporation, 1998, s.12.

1.4.1. Birliktelik Kuralları Analizi

Birliktelik kuralları belirli türdeki veri ilişkilerini tanımlayan bir yaklaşımdır. Birliktelik kuralında geçmiş veriler analiz edilerek, bu veriler arasında birliktelik ilişkileri ortaya koyularak, geleceğe yönelik çalışmalar için yardımcı olunur.

Herhangi bir ürün satın alındığında, bu ürünün yanında farklı başka bir ürünün de satın alınması bir birliktelik kuralını vermektedir. Birliktelik kuralı daha çok perakendecilik sektöründe faaliyet gösteren işletmelerde uygulanmaktadır. Bir süpermarkette yapılan alışverişler incelenerek hangi ürünlerin yanında hangi ürünlerin satın alındığı, müşterilerin daha çok nelere yöneldiğinin belirlenmesinde çoğunlukla birliktelik kuralı analizinden yararlanılır.[57]

İşletmelerde önceleri satış ve müşteri verileri ender olarak yer almakta, yaygın olarak mali açıdan önemli gelir verilerinin depolanması söz konusuydu. Ancak 90'lı yıllardan bu yana günümüzde, artık teknolojinin ilerlemesi ve veri toplamadaki yeni gelişmeler doğrultusunda işletmeler bu sürece ayak uydurmaktadır. Örneğin süpermarketlerde otomatik ürün veya müşteri tanıma sistemleri (bar kod ve manyetik kart okuyucular) yaygınlaşmaya başlayarak, satış esnasında veriler elektronik ortama aktarılmakta ve bu veriler değerlendirilerek müşteri analizleri, satış tahminleri yapılabilmektedir.[58]

Birliktelik kuralları en çok Pazar Sepet Analizi altında kullanılmaktadır. Pazar Sepeti Analizi ile müşterilerin alışveriş alışkanlıklarının veri tabanındaki bilgiler aracılığıyla ortaya çıkarılması sağlanmaktadır. Bu alışveriş alışkanlıklarının ortaya çıkartılması, alışveriş merkezindeki ürünlerin raflara yerleştirilmesi, düzenlenmesi, ürün katalog tasarımları oluşturulması, market alanının tasarımı ve markette birlikte sergilenecek, satılacak ürünlerin belirlenmesinde etkili bir faktördür.

---

[57] Silahtaroğlu, **a.g.e.**, s 33.

[58] Fatih Şen, "**Veri Madenciliği İle Birliktelik Kurallarının Bulunması**", Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Ve Bilişim Mühendisliği Anabilim Dalı, Sakarya, 2009, s.23.

Birliktelik kuralına ilişkin olarak geliştirilen bazı algoritmalardan en yaygın olarak bilineni Apriori algoritmasıdır. Birliktelik kurallarını bulmak için genellikle problem iki adımda çözümlenir. Önce sık tekrarlanan öğeler bulunur. Bu öğelerin her biri en az, önceden belirlenen minimum destek sayısı kadar sık tekrarlanırlar. Sonra sık tekrarlanan öğelerden güçlü birliktelik kuralları elde edilir. Bu kurallar minimum destek ve minimum güven değerlerini karşılamalıdır.[59]

1.4.2. Ardışık Zamanlı Örüntüler

Örüntü sözcüğü, herhangi bir çizim, ses, resim, parmak izi gibi bir şekil olarak ifade edilebilmekte ya da bir kişinin yaptığı işler olarak da tanımlanabilmektedir.[60] Birbirleri ile ilişkisi olan ancak birbirini takip eden dönemlerde gerçekleşen ilişkilerin tanımlanmasında Ardışık Zamanlı Örüntüler kullanılmaktadır. Örneğin; X ameliyatı yapıldığında 15 gün içinde %45 olasılıkla Y enfeksiyonu oluşacaktır. Çamaşır makinesi alan bir müşteri 3 ay ile 9 ay içerisinde, %70 olasılıkla kurutma makinesi de alacaktır. Mağaza %70 olasılığı arttırmak için çamaşır makinesinin alımından 4 ay sonra, kurutma makinelerinde indirimler uygulayabilir, böylece müşterinin kurutma makinesi alma olasılığı artırılmaktadır.[61]

1.4.3. Kümeleme Analizi

En temel amacı, "gözlem birimlerini benzer özelliklerine göre gruplamak" olan Kümeleme Analizi, çok değişkenli istatistik tekniklerinden biridir. Bu istatistik tekniğinde; elde edilen bir küme içindeki gözlem birimleri, daha önceden belirlenmiş bir özellik bakımından birbirine benzemektedir. Bu sebepten dolayı elde edilen kümedeki gözlem birimleri homojendir.[62] Genel olarak bakıldığında; kümeleme analizinin amacı küme üyelerinin birbirine çok benzediği, fakat özellikleri

[59] Silahtaroglu, **a.g.e.**, s.33-34.
[60] **A.e**. s.35.
[61] Akpınar, **a.g.e.**, s. 6.
[62] Joseph F. Hair et. al., **Multivariate Data Analysis**, Prentice Hall, U.S.A., 1998, s. 473.

birbirinden çok farklı olan kümelerin elde edilmesi ve veri tabanındaki kayıtların bu farklı kümelere bölünmesi olarak ifade edilebilir.[63]

Kümeleme Analizi, araştırma konusu olan birey ya da nesneleri, aralarındaki benzerlikler göz önünde bulundurarak bir araya getirir. Bu sayede birey ya da nesneler, farklı özellikleri itibariyle kümelere ayrılırken, kümeler içi homojenlik ve kümeler arası heterojenlik en üst düzeyde tutulmuş olur. Yani, bir kümeyi oluşturan bireyler ya da nesneler birbirleri ile benzeşirken, diğer kümelerin bireyleriyle/nesneleriyle benzeşmeyecektir.[64]

Kümelemenin bir diğer amacı ise benzer elemanların gruplanmasıyla veri setinin küçülmesini sağlamaktadır. Bu amaçlardan dolayı Kümeleme Analizi satış hareketleri veya çağrı merkezi kayıtları gibi çok fazla parametrenin bir arada olduğu çok büyük miktarlardaki verileri analiz etmede en uygun yöntemlerden biridir. Kümeleme analizinde ağırlıklı olarak K-means algoritması kullanılmakla birlikte, AIS ve SETM algoritmalarından da yararlanılmaktadır.

Kümeleme analizi, kümelerde grafikler sayesinde görsel sonuç benzerliklerinin kolay tespit edilmesini, aykırı durumların saptanmasını, verinin bir bölümünün seçilmesi ve uygun başlangıç noktalarının belirlenmesini sağlamaktadır. Kümeleme analizinin dezavantajı ise, takip edilmesi gereken belirli kuralların olmaması sebebiyle tahminlerin gerçekleşmesi ve farklı veri tiplerinde özellik içeren nesnelerin karşılaştırılmasının zorluğudur. [65]

1.4.4. Sınıflandırma Teknikleri

Veri madenciliğinde en çok kullanılan tekniklerden bir diğeri de sınıflandırma teknikleridir. Sınıflandırma tekniklerinin sıklıkla kullanıldığı alanlar hastalık tanıları, pazarlama konuları, resim- örüntü tanıma, dolandırıcılık tespiti, kalite

[63] Akpınar, **a.g.e.**, s. 7.
[64] Mahir Nakip, **Pazarlama Araştırmaları**, Seçkin Yayınevi, Ankara, 2003, s.418.
[65] Claude Seidman, **Data Mining with Microsoft® SQL Server™ 2000 Technical Reference**, Microsoft Press, Redmond, WA, USA, 2001.

kontrol uygulamalarıdır. Sınıflandırma havanın bir sonraki gün nasıl olacağının ya da bir kutuda kaç tane kırmızı top olduğunun tahmin edilmesi gibi işlemlerdir. Yani sınıflandırma bir tahminleyici modeldir.[66]

İnsanlar verileri sürekli sınıflandırdıkları, kategorize ettikleri ve derecelendirdikleri için sınıflandırma, hem veri madenciliğinin temeli olarak hem de veri hazırlama aracı olarak da görülmekte ve kullanılmaktadır.

En çok kullanılan sınıflandırma algoritmaları ise Karar Ağaçları, Genetik Algoritmalar, Naive Bayes Sınıflayıcıları, Vektör Makinesi ve K-En Yakın Komşu Algoritmalardır.[67]

1.4.4.1. Karar Ağaçları

Çok çeşitli seçeneklerin gerçekleşmesinin belirli ya da belirsiz olduğu bir problemle ilgili en iyi karara ulaşabilmek, bazı işlemlerin yerine getirilmesini sağlayabilmek için karar verici, çeşitli yöntemlere veya araçlara gereksinim duymaktadır. Seçenek sayısının birden fazla olduğu ve/veya ardışık aşamalarda karar almanın söz konusu olduğu problemlerin analizinin gerçekleştirilmesi, modellerin kurulması, değerlendirilmesi ve çözümlenmesi işlemlerinde karar vericiler bu araçlardan birisi olan "Karar Ağacı Analizi"ni kullanabilirler.[68]

Karar ağaçları sınıflama amaçlı uygulamalarda da çok sıklıkla kullanılan yöntemlerden birisidir. Tahmin edici ve tanımlayıcı özelliklere sahip olan karar ağaçları veri madenciliği projelerinde,

- Meydana getirilmelerinin ucuz olması,
- Yorumlanmalarının kolay ve basit olması,

---

[66] Silahtaroğlu, **a.g.e**., s.45.
[67] Yasemin Koldere Akın, "**Veri Madenciliğinde Kümeleme Algoritmaları Ve Kümeleme Analizi**", Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı, İstanbul, 2006, s.55.
[68] Hayrettin Kemal Sezen, **Yöneylem Araştırması**, Bursa, Ekin Kitapevi, 2004, s.5.

- Veri tabanı sistemleri ile kolayca entegre olabilmeleri,
- Güvenilirliklerinin daha yüksek olması gibi sebeplerden dolayı tercih edilmekte ve kullanılmaktadır. Ayrıca karar ağaçları sınıflama modelleri arasında en yaygın kullanıma sahip olan tekniktir.[69]

Çok basit bir alt yapıya sahip olmalarına karşın çok önemli bilgileri uygulayıcılara sunma imkanına sahip olan karar ağaçları, çok sık olarak kümeleme uygulamalarında kullanılmaktadır. Karar ağaçları diğer tekniklerle karşılaştırıldığında; yorumlanması, anlaşılması ve yapılandırılması çok kolay olan bir tekniktir.[70] Bu teknikte sınıflama yapılırken öncelikli olarak veri setinden bir ağaç meydana getirilir. Bu ağaç oluşturulduktan sonra veri setindeki her bir kayıt bu ağaca uygulanarak bu kayıtlar sınıflandırılır.

Karar ağacı oluşturulurken, mevcut olan veri setindeki verilerin bir bölümü öğrenme amaçlı olarak kullanılacak ve bu durumdan yola çıkarak karar ağacı oluşturulacaktır. Veri setinin bir bölümü ise oluşturulan bu karar ağacını test etmek için kullanılmaktadır.

Ağaç meydana getirilirken, kurulan sistemin doğru şekilde çalışıp çalışmadığı belirlenmektedir. Kurulan sistemin istenen seviyede olduğunun anlaşıldığı durumda ise ağacın dallanmasına son verilir. Durdurma olarak adlandırılan, ağacın dallanmasının sonlandırılması ağacın hassasiyetini gösterir. Çünkü geç durdurulan karar ağacı daha fazla dallandırıldığı takdirde, istenmeyen sonuçların ortaya çıkmasına ya da önemli ilişkilerin, sınıfların, örüntülerin gözden kaçmasına neden olabilecektir. Erken durdurulan karar ağacı ise tam öğrenmenin gerçekleşmemiş olma olasılığını taşımaktadır. Bu nedenle karar ağacının dallanmasının sonlandırılması önemlidir. [71]

---

[69] Akpınar **a.g.e**. s.12.

[70] Rakesh Agrawal, Tomasz Imielinski, Arun Swami, Mining Association Rules Between Sets of Items in Large Databases", **In Proceedings of the ACM SIGMOD International Conference on Management of Data**, USA,1993, s. 915.

[71] Margaret Dunham, **Data Mining Introductory and Advanced Topics**, Prentice Hall, USA, 2003, s.96.

Karar ağacı oluşturulduktan ya da ağaç kurulumu tamamlandıktan sonra budama olarak adlandırılan, ağaçta oluşmuş, karar verirken sonucu değiştirmeyen, sınıflama için bir faydası olmayan dalların ağaçtan alınması işlemi yapılabilmektedir.[72]

1.4.4.2. Genetik Algoritmalar

Genetik algoritmalar ya da daha geniş ifade ediliş biçimiyle evrimsel algoritmalar doğadaki evrimsel süreçleri model olarak ele alan bilgisayara dayalı problem çözme yöntemleridir. Geleneksel modelleme yöntemleriyle çözülmesi zor olan özellikle sınıflandırma ve çok boyutlu optimizasyon problemleri, genetik algoritmalar yardımıyla daha kolay ve hızlı çözülebilmektedir.[73]

Öğrenme temelli bir model olan genetik algoritmalar sınıflandırma amaçlı kullanılır. Bu model farklı bir biyolojik benzetme üzerine kurulmuştur. Genetik algoritma Darwin'in evrim teorisinden ilham alınarak oluşturulmuştur. Genetik algoritma geleneksel yöntemlerle çözümü güç veya imkansız olan problemlerin çözümünde kullanılmakta, açılanabilir sonuçlar üretmektedir. Çözümü güç olan bu problemlerin neredeyse tamamına yakını geniş bir çözüm alanının taranmasıyla geçekleşebilmekte ve bu işlem çok uzun sürebilmektedir.[74]

Genetik algoritmalar yardımıyla ise kısa sürede kabul edilebilir sonuçlara ulaşılmaktadır. Ayrıca çeşitli türdeki verileri işleme özelliğine sahip olan algoritmalar en iyileme çalışmalarında, yapay sinir ağlarının eğitilmesi, bellek tabanlı yöntemlerle fonksiyonların oluşturulması gibi işlemlerde kullanılabilirler. Tüm bu olumlu yönlerin yanı sıra ise genetik algoritmaların kullanımlarında bazı sorunlar da

---

[72] Peter Cabena, et. al., **a.g.e**, s.73.
[73] Pelin Tezcanlar," **Müşteri İlişkileri Yönetimi Veri Madenciliği ve Bir Uygulama**",Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2007, s.97.
[74] M. Özgür Dolgun, "**Büyük Alışveriş Merkezleri İçin Veri Madenciliği Uygulamaları**", Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Ankara, 2006, s. 33.

yaşanmaktadır. Bunlardan en göze çarpanı karmaşık sorunların genetik kodlanmasının zorluğudur.[75]

1.5. GELİŞMİŞ VERİ MADENCİLİĞİ TEKNİKLERİ

Ülkemizin ve dünyanın gelişen ve değişen bilişim teknolojileri, gereksinimleri ve beklentileri karşılayamamakta ve farklı seçeneklerin ele alınmasını gerektirmektedir. 1990'lı yıllardan bu yana her geçen gün önemi giderek artan veri madenciliği de gelişerek bilişim teknolojilerindeki yerini derinleştirmektedir. İletişim ve bilgisayar teknolojilerindeki gelişmeler insanlara daha fazla bilgiye erişim imkanı sağlamaktadır. Bu durum bilginin daha kısa sürede toplanması, depolanması ve sunulması alanında kolaylık sağlamaktadır.

Gelişmiş veri madenciliği teknikleri de bu süreç içerisinde istenilen bilgilerin aktarılması konusunda yeni disiplinlerin ortaya çıkmasını sağlamıştır. Özellikle bilgisayar dilbilimi (Computer Linguistics) alanında web madenciliği (Web Usage Mining), metin madenciliği (Text Mining) ve vaka temelli çıkarım (Case Based Reasoning) veri madenciliğinde önemli rol oynayan alanlardan bir kaçıdır.[76]

Resim, video/ses, grafik, somut veriler, ilişkili ve kategorisel veriler üzerinde veri madenciliği tekniklerini kullanarak verilerin içerisindeki gizli bilgilerin amaç doğrultusunda elde edilmesine çalışılmaktadır. Gelişmiş veri madenciliği teknikleri ile ses (örnek telefon kayıtları) ve video verileri, e-posta içerikleri, kelime işlemci dokümanları, web sayfalarında yer alam forum verileri, anket cevapları, müşteri, kamu kurumları vs. bildirileri, öneri ve şikayet bilgileri incelenmekte ve yapısal olmayan bu verilere gelişmiş veri madenciliği teknikleri sayesinde daha kolay bir şekilde ulaşılmaktadır.[77]

---

[75] Gökçe Kolay, "**İşletmelerde Bilgi Sistemleri Verimliliğini Arttırmada Veri Madenciliği Yöntemi: Bir Simülasyon Çalışması"**, Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Zonguldak, 2006, s. 48.
[76] Akpınar, **a.g.e**., s. 14.
[77] Anıl Karadağ, Hüseyin Takçı, "**Metin Madenciliği ile Benzer Haber Tespiti**", AB 2010, Akademik Bilişim, Muğla Üniversitesi, Muğla, 10-12 Şubat 2010, s.3.

Gelişmiş veri madenciliği tekniklerinin en yaygın olarak kullanılanları web, metin ve multimedya madenciliğidir.

1.5.1. Web Madenciliği

Veri madenciliği tekniklerinin world wide web verileri üzerinde uygulanmasına web madenciliği adı verilmektedir. Web madenciliği ilk kez 1996 yılında Etzioni tarafından ortaya atılmıştır. Günümüzde ise bilgisayarların ve teknolojinin hayatımızda büyük yere sahip olması sebebiyle artık pek çok çalışmada web madenciliğinden yararlanılmaktadır.[78]

Web madenciliği sayesinde kullanıcıların web sitelerindeki davranışları incelenebilmekte, kullanıcı istekleri belirlenebilmektedir. Bu şekilde elde edilen veri sayesinde web sitelerinin tasarımlarının düzenlenmesi veya geliştirilmesi konusundaki karar alım süreci kolaylaşmaktadır. Çünkü hangi sayfalara daha sıklıkla girildiği, hangi sayfaların birlikte ziyaret edildiği gibi bilgiler sitenin yeniden düzenlenmesinde etkili olan bilgilerdir. Bu uygulamaların amacı, istatistiksel yöntemlerle kullanıcı davranışlarının öğrenilmesi, böylece site içeriği ve tasarımının bu bilgiler ışığında yenilenmesidir.[79]

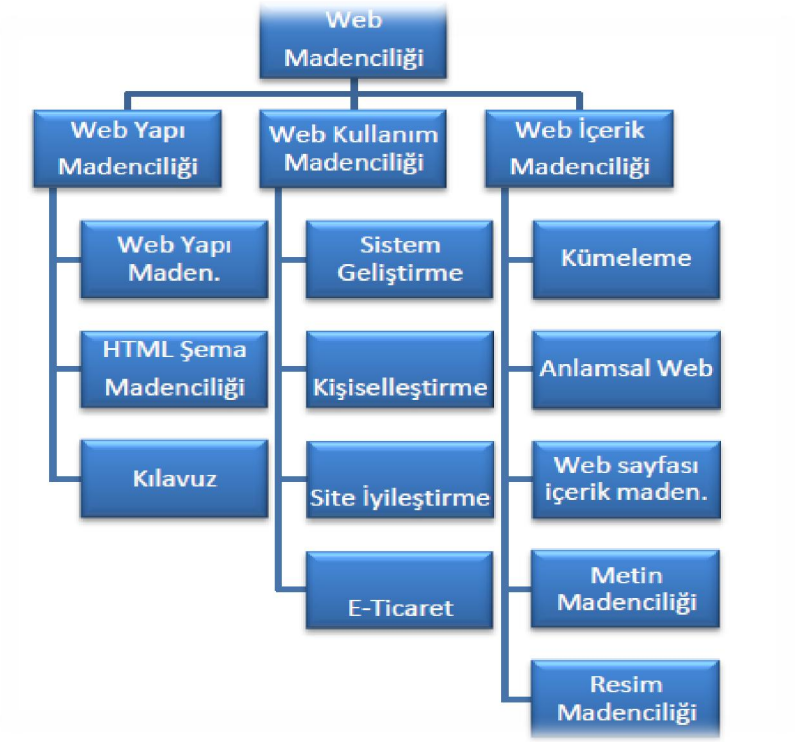
Web madenciliği; web içerik madenciliği, web yapı madenciliği ve web kullanım madenciliği olmak üzere üç ana bölümden oluşmaktadır.[80] Şekil 1.3'te web madenciliği sınıflandırılması görülmektedir.

---

[78]Hikmet Mukadder Oktay, "**Web Kullanım Madenciliğinde Birliktelik Kurallarının Uygulanması**",Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı İstanbul, 2009, s.14.

[79] İsmail Haberal, "**Veri Madenciliği Algorimaları Kullanılarak Web Günlük Erişimlerinin Analizi Uygulaması**", Yüksek Lisans Tezi, Başkent Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Ankara, 2007, s.2.

[80] Oren Etzioni, **The World Wide Web: Quagmire Or Gold Mine. Communications of the ACM**, Vol.39, 1996, s.64-67.



Şekil 1.3. Web Madenciliği Sınıflandırması [81]

Bu kategorilerin kendi aralarında temel farkları göz önüne alınarak kıyaslanması Tablo1.3'te gösterilmektedir.

Tablo 1.3. Web Madenciliği Kategorileri ve Sınıflandırılması[82]

	Web İçerik Madenciliği	**Web Yapı Madenciliği**	**Web Kullanım Madenciliği**
Veri	**Metin belgeleri ve HTML**	**HTML linkleri**	**Sunucu ve tarayıcı kayıtları (logs), çerezler, kullanıcı profilleri, kullanıcı sorguları, meta data**
Verinin Şekli	**Yapısız ve karışık**	**Link yapısı**	**Kullanıcı etkileşimi ve davranışı**
Verinin Gösterimi	**İlişkisel ve sınıflandırmalı**	**Grafiksel**	**İlişkisel tablolar ve grafiksel**

[81] Güncel Sarıman, "**Paralel Programlama İle Web Madenciliğinde Log Analizi**" Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Isparta, 2011, s.17.
[82] Oktay, **a.g.e**.,s.6.

"Web içerik madenciliği" web kaynaklarından içeriklerine göre otomatik bilgi arama tekniklerinin tanımlanmasını sağlamakta ve ana konusunu site içeriği oluşturmaktadır. Web içerik madenciliğinde temel nokta internette saklı bilgiyi bulmaktır.

"**Web yapı madenciliği'nde**" ise web sitesi ve web sayfası hakkında bağlantı verisine bakılarak bilgi üretimi amaçlanmaktadır. Teknik olarak bakıldığında, web içerik madenciliği dokümanın içeriği, yapı madenciliği ise dokümanlar arası bağlantılar üzerinde durur. Web yapı madenciliği, link kümelerine dayanarak farklı siteler arasındaki benzerlik ve ilişki gibi bilgileri oluşturur.[83]

"**Web kullanım madenciliği'nde"** ise amaç kullanıcının siteyi kullanımı sırasında, geride bıraktığı erişim verilerinden bilgi üretilmesidir. Oluşan bu veriler bütünüyle kullanıcıdan bağımsızdır. Bir yerde yazılan ya da kullanıcının isteğiyle oluşan bir veri değildir. Bu veriler istemcilerde, sunucularda ve Proxy sunucularında depolanır.[84] Web kullanım madenciliği, uyarlanabilir web sunucu tasarımı, kullanıcı sınıflandırması, kullanıcı profil değişimleri gibi alanlarda kullanılmaktadır. Web kullanım madenciliğinde büyük web veri depolarındaki kullanım logları üzerinde veri madenciliği teknikleri işlemi yapılır.[85]

Analizlerde kullanılan veri kaynakları olarak bakıldığında ise sunucu erişim kayıtları, referrer kayıtları, kullanıcı profilleri, istemci tarafında bulunan çerezler, sayfa özellikleri ve içerik özellikleri için kullanılan meta data verileri örnek verilebilir.[86]

---

[83] Fatih Gürcan, "**Web İçerik Madenciliği Ve Konu Sınıflandırması**" Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Trabzon, 2009, s.9.

[84] Ayşegül Ezerçe, **"Müşteri İlişkileri Yönetimi (Crm) Ve Veri Madenciliği (Data Mining) : Tekstil Sektöründe Bir Uygulama"**, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2008, s.66.

[85] Oktay, **a.g.e**., s.14.

[86] Ezerçe **a.g.e.**, s.67.

1.5.2. Metin Madenciliği

Günümüzde internet ve kişisel bilgisayarların hayatımızın her aşamasında kullanımının yaygınlaşmasıyla beraber, dijital depolama ortamlarının kapasitelerinde önemli bir artış yaşanmış ve bu artış beraberinde depolanan veri miktarlarının büyük boyutlara ulaşmasına neden olmuştur.[87] Depolanan bu veriler sadece yapısal verilerden oluşmamakta; araştırma yazıları, makaleler, kitaplar, sayısal kütüphaneler, telefon kayıtları gibi ses ve video verileri, e-mailler, kelime işlemci dokümanları, web sayfalarındaki forum verileri, anket cevapları, öneri ve şikayet bilgileri, çevrimiçi sohbet gibi yapısal olmayan verilerden de oluşmaktadır.[88]

Web madenciliği web içerikleri, sayfa yapıları ve web bağlantı istatistiklerinin de içinde olduğu web ile ilişkili olan verinin analizini kapsarken, metin madenciliği çok büyük belgelerin analizi ve metin tabanlı veriler içindeki bilgilerin gizli kalıplarının elde edilmesini içermektedir.[89]

Metin madenciliği veri madenciliğinin bir alt bölümü olarak gelişmektedir. Metin madenciliğinde amaç, metinler arasında elle bir ayrıma ihtiyaç duyulmadan benzerliklerin hesaplanmasıdır. Bu işlem genellikle otomatik olarak çıkarılan sözcüklerin tekrar sayıları ile gerçekleştirilir. Metin madenciliğinde e-posta, yazılı metinler, arşiv ve internet uygulamaları, haberler gibi çeşitli uygulamalar bulunmaktadır.[90]

---

[87] Karadağ, Takçı, **a.g.e.** s.4.

[88] Keziban Seçkin, "**Metin Madenciliğinde Kullanılan Yöntemlerin Karşılaştırılması: Siyasi Parti Liderlerinin Grup Genel Toplantı Konuşmaları İle Bir Uygulama**",Yüksek Lisans Tezi, Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, Sakarya, 2011, s. 4.

[89]M. Özgür Dolgun, Tülin Özdemir, Doruk Oğuz, Veri Madenciliğinde Yapısal Olmayan Verinin Analizi: Metin Ve Web Madenciliği, **İstatistikçiler Dergisi,** Sayı. 2, 2009, s.48-58.

[90] Yılmaz Akgüden, Burak Erşahin, **Veri Madenciliği, Veriden Bilgiye Masraftan Değere Online Kitap**, s.35, (Çevrimiçi) http://www.arge.com/Userfiles/File/Veri%20Madencili%C4%9Fi.pdf, 12 Mart 2012.

1.5.3. Multimedya Madenciliği

Teknolojinin gelişmesi ve internetin geniş kitlelere ulaşmasıyla birlikte yapısal olmayan verilerde artış olmakta, bu durum beraberinde yapısal olmayan verinin analizini gündeme getirmektedir. İnternet sayfaları, web siteleri, elektronik postalar gibi yapısal olmayan veriler veri madenciliğinin önemini daha da özelleştirdiğinden dolayı, metin madenciliği ve multimedya madenciliği kavramlarının önemi daha da artırmıştır.

Multimedya madenciliğinde çoklu ortam verileri bir arada kullanılmaktadır. Çoklu ortam kavramı; metin, görüntü, çizim, grafik, ses, video ve animasyonların bilgisayarda gösterilmesi, dosyalarda saklanması, bilgisayar ağından iletilmesi ve sayısal olarak işlenmesi ile ilgili bir kavramdır.[91]

Kişilerin fiziksel veya davranışsal özellikleri kullanılarak tanınması/onaylanması olarak açıklanabilen biyometri bilimi, parmak izi, yüz, iris, retina, kulak, el geometrisi, ses gibi kişiyi fizyolojik veya davranışsal bir özelliğiyle temsil edebilecek kadar bireye özgü, bireyden bireye kolay kolay aktarılamayacak kadar güvenilir, hayatın başlangıcından sonuna kadar değişmeyen kalıcı özellikleri içermektedir.[92] Bu alanda verilerin bilgisayar ortamında saklanması ve işlenmesi multimedya madenciliğinin etkisini ortaya çıkarmaktadır.

Multimedya madenciliğinin en etkin kullanılmaya başladığı alanlardan biri parmak izi işlemleridir.

Parmakların uç bölgesi insan biyometrisine uygun olarak benzersiz olarak yaratılmıştır. Her insanda farklı farklı ve birbirine benzemeyen bu sistem en güvenilir şifreleme yöntemlerinden biri olarak kullanılmaktadır. Parmak izi tanımlaması yapılırken, parmak ucu bölgesinde bulunan bir çeşit haritaya benzeyen

---

[91] Özgür Yılmazel, **Türkçe Metin Arama Sistemleri ve Veritabanlarının Karşılaştırılması**, (Çevrimiçi) http://seminer.linux.org.tr/wp-content/uploads/dilisimSunu.pdf, 25.09.2012.

[92] Şeref Sağıroğlu, "Parmak İzinden Yüz Tanıma", **Çerçeve Dergisi**, Sayı: 52, s.104, (Çevrimiçi) http://www.musiad.org.tr/contentimages/arastirmalaryayin/pdf/cerceve_dergisi_52.pdf, 10 Mart 2013.

yapı, aslında girinti ve çıkıntılardan oluşmaktadır.[93] Parmak izi okuma cihazları yardımıyla bu girinti ve çıkıntıların birbirine olan mesafesi, kıvrımların açıları ve belli noktaların birbirlerine uzaklığı algılanarak bir algoritma oluşturulur. Oluşturulan algoritma belli bir şifrelemeyle elde edilir. Bu algoritmaya dayalı olarak daha önceden cihaz içerisinde kaydedilen bir parmak izi tekrar cihaza okutulduğunda eşleştirme gerçekleşmiş olur.



Şekil 1.4. Parmak İzi Tarama İşlemleri [94]

Günümüzde pek çok sektörde giriş-çıkış kapılarındaki kimlik tespiti ve benzer güvenlik fonksiyonlarında, bilgisayar ağ güvenliği ve bilgisayar ağlarında erişim düzeninin sağlanmasında ve kaynakların paylaşımında, havaalanlarında pasaportlarda, suçluları tespit etmede bu parmak izi tarama sistemleri çok yaygın bir şekilde kullanılmaktadır. Bunların yanı sıra özellikle bankaların ATM sistemlerinde de bu sisteme geçilmiş olup yine banka güvenliğinde elektronik fon transferi veya ATM güvenliği gibi işlemlerde, çek ve kredi kartı işlemlerinde, ödeme işlemlerinde bu sistemler tercih edilmektedir. [95]

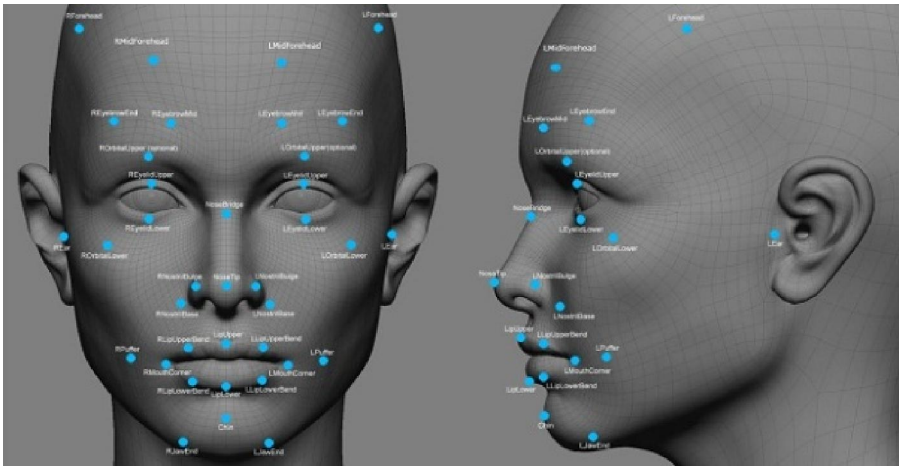
Parmak izi tarama işlemleri mücevher firmaları gibi çok yüksek güvenlik önlemlerinin alınması gereken kuruluşlarda kullanılmakta, gümrük ve göç işlemlerinde daha güvenli ve hızlı işlem için, ulusal kimliklendirme sistemlerin daha güvenilir yapılabilmesinde, adli suçların belirlenmesinde, seçmen ve sürücü

[93] Biyometrik Tanıma Sistemleri, (Çevrimiçi) http://www.ardbilisim.com.tr/projeler/biyometrik/ 15 Ocak 2013.
[94] Parmak İzi Tarama Sistemleri, (Çevrimiçi) http://www.meyer.com.tr/parmakizi.htm, 11 Ocak 2013.
[95] Necla Özkaya, Şeref Sağıroğlu, "Parmak İzinden Yüz Tanıma Sistemleri", **Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi**, Cilt: 23, Sayı. 4, 2008, s. 787-789.

kayıtlarında, e-alışverişin verimli bir şekilde uygulandığı alanlarda bu sistemler çok yaygın olarak kullanılmaktadır.[96]

Multimedya madenciliği özellikle yüz tarama sistemlerinde de kullanılmaktadır. Bu sistemlerde tıpkı parmak izinde olduğu gibi yüz yapısında da insana özgü olan benzersiz yapılar vardır. Fakat yüzdeki bu benzersiz yapının varlığı insanlarca pek bilinmez. Sistemde bilgisayar ortamında; yüzlerin simetrisi olarak bilinen yüz hatları eşleştirme yoluyla incelenir ve yüz tanıma işlemi yapılmış olur. Bilgisayar ortamında yapılan bu biyometrik eşleştirme sonucu yüz yapısı bilgisayar içinde bulunan veritabanındaki yüzlerle karşılaştırılır. Tıpkı parmak izi gibi yüz yapısı da taklit edilemez.[97]

Yüz tanıma sisteminin temelini her insanın farklı ve benzersiz bir yaratılışa sahip olması oluşturmaktadır. Özellikle büyük kuruluşlar, kamu binaları ve çok üst düzey güvenlik alanlarıyla korunan yapı ve eşyalar için bu sistemler günümüzde etkin bir şekilde kullanılır. Yüz tarama sistemleri dolandırıcılık tespiti için yaygın bir şekilde kullanılmaktadır. Yetkisiz kişilerin erişimini engellemek amacıyla gerçekleştirilen bu sistemler günümüzde her alanda yer almaya başlamıştır.



Şekil 1.5. Yüz Tanıma Sistemleri

[96] Necla Özkaya, Şeref Sağıroğlu, "Otomatik Parmakizi Tanıma Sistemlerinde Kullanılan Önişlemler İçin Yeni Yaklaşımlar", **Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi**, Cilt: 21, Sayı. 1, 2006, s. 13-19.

[97] **Yüz Tanıma Sistemleri** (Çevrimiçi), http://www.yuztanimasistemi.net/, 13 Ocak 2013.

New Scientist tarafından yayınlanan bir haberde; Amerika Federal Soruşturma Bölümü FBI'ın "Gelecek Nesil Tanımlama" olarak adlandırılan sistem ile birlikte FBI'ın takip sürecinde DNA analizi, göz taraması, ses tanıma gibi insandan insana değişkenlik gösteren biyolojik veriler etkin bir şekilde kullanılarak bir sistem geliştirmiş ve bu sistemleri veri toplamak için sokaklarda yer alan kameralardan faydalanılmıştır. bu sayede veri tabanı genişletilerek güvenliğin daha üst boyutlara taşınacağını belirtilmiştir. [98]

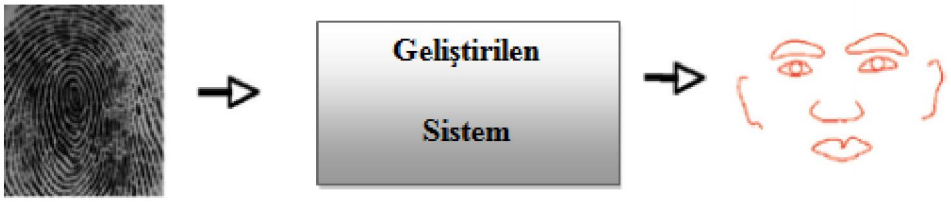
Günümüzde parmak izinden yüz tanıma sistemleri de geliştirilmektedir. Geliştirilen bu sistemlerde; kişilerin parmak izi ve yüz resimlerinden oluşan biyometrik özellikler, bir veri tabanı kullanılarak oluşturulur ve bu özellikler mevcut çözüm teknikleri ve geliştirilen teknikler kullanımıyla işlenir.



Şekil 1.6. Yüz Tanıma Sistemleri [99]

İşlenmiş parmak izleri ve yüz setlerine ait setler birbirilerine entegre edilerek, kişilerin parmak izlerinden yüz resimlerine ihtiyaç duyulmadan yüz setlerinin elde edilmesini sağlamaktadır.

[98] (Çevrimiçi) http://www.donanimhaber.com/kurumsal-diger/haberleri/FBiin-1-milyar-dolarlik-yuz-tanima-sistemi-2014te-devreye-giriyor.htm; 15 Şubat 2013.
[99] Sağıroğlu, **a.g.e.**, s.105.



Şekil 1.7. Geliştirilmiş Yüz Tanıma Sistemleri

Bu süreç özellikle suçların tespitinde son derece etkili olabilecek bir uygulama alanıdır.

Multimedya madenciliğinin diğer kullanım alanlarından biri de ses tarama sistemleridir. Telefon dinlemelerinde sesin kime ait olup olmadığını tespit etmekte kullanılır. Bunların yanı sıra trafikte her yerde karşımıza çıkan Mobese kameralarında da bu sistemler kullanılmakta olup özellikle plaka tarama işlemleri için kullanılmaktadır.

Multimedya madenciliğine eğitim alanında da geniş yer verilmektedir. Teknolojinin gelişmesi paralelinde yeni yenilikleri getirmektedir. Günümüzde artık elektronik ansiklopediler ve görsel kitaplar, ders araçları sayesinde eğitim alanında pek çok farklılıklar bulunmakta, multimedya madenciliği sayesinde sözlü anlatımlar, uygulamalar, animasyonlar ile birleştirilerek öğrencilerin öğrenim yetkinliklerini, akademik başarılarını artırmaktadır.[100]

1.5.4. Büyük Veri (Big Data)

Günümüzde hızla artan veri kalabalıklığı sonucunda; bilginin anlamlı hale gelmesi için çeşitli çalışmalar yapılmaktadır. Bu çalışmaların en başında veri madenciliği yer almaktadır. Büyük veri kavramının geçmişi çok eskilere

[100] M.Emre Sezgin, **"İkili Kodlama Kuralına Dayalı Olarak Hazırlanan Multimedya Ders Yazılımının Fen Bilgisi Öğretimindeki, Akademik Başarıya, Öğrenme Düzeylerine ve Kalıcılığına Etkisi",** Yüksek Lisans Tezi, Çukurova Üniversitesi, Sosyal Bilimler Enstitüsü, Eğitim Ana Bilim Dalı, 2002, s. 14.

dayanmamaktadır. 2004 yılında Google'ın kullanmaya başladığı bir teknoloji olan büyük veri (big data), Youtube, Facebook, Twitter gibi büyük ölçekte verisi olan firmalar tarafından da telaffuz edilmeye başlamıştır.

Büyük verinin bu kadar dile getirilmesinin tek sebebi bu firmaların bu kavramı kullanması değildir. İnternet, medya ve büyük kuruluşların gözüne çarpan, çarpacak kadar büyüyen veri miktarı ve çeşitliliğidir. Hızla artan veri hacmi ve çeşitliliği karşısında geleneksel veri tabanları yetersiz hale gelmiştir. Bunun sebebi artan veri çeşitliliğinin farklı alanlarda kullanılmak istenmesidir.[101]

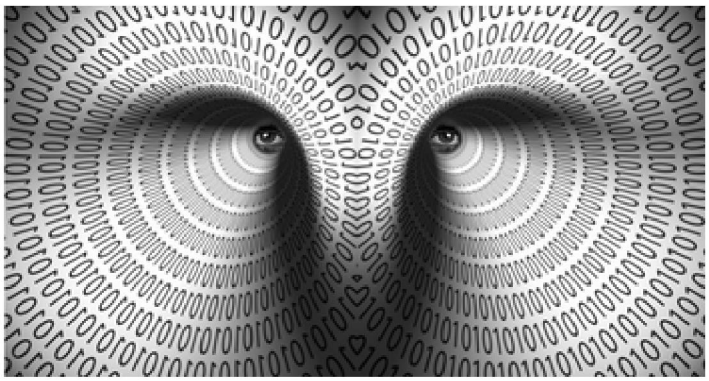
Dünya nüfusu 7 milyarı geçerken ve bu 7 milyar kişinin neredeyse yarısı internet bağlantısına sahip iken bu sonsuz veri setlerinin oluşmasına neden olmakta, bu durum zaman zaman veri çöplüğü de yaratmaktadır.[102] Örneğin sosyal ağlarda Twitter'da günlük 100 milyon aktif kullanıcı yer almaktadır ve günlük 12 milyon terabyte log data üretilmektedir. Yine dünyada 6 milyar mobil abone bulunmaktadır.[103]

Büyük veri; zamanla toplanan büyük miktardaki verilerin ekonomik ve ölçülebilir, ölçeklenebilir bir şekilde ve ilişkisel veri tabanlarının yetersiz kaldığı noktalarda kullanılan ve anlamlı işlenebilir veri elde edilmesi kavramıdır.

Büyük veri aynı zamanda büyüklüğü terabyte'lara ulaşan verilerin özellikle toplumsal medya paylaşımları, ağ günlükleri, bloglar, fotoğraf, video, log dosyaları gibi değişik kaynaklardan toparlanan tüm verinin, anlamlı ve işlenebilir biçime dönüştürülmüş biçimi olarak adlandırılmaktadır.[104]

---

[101] http://www.oganozdogan.com/2013/07/buyuk-veri-big-data-nedir.html. (Çevrimiçi) 01 Ağustos 2013.
[102] Necmi Gürsakal, **Büyük Veri**, Bursa, Dora Yayın Evi, 2003, s.55.
[103] http://eticaretmag.com/big-data-buyuk-veri-nedir/ (Çevrimiçi) 30 Mayıs 2013.
[104] http://www.economist.com/node/15557443. (Çevrimiçi) 28 Mayıs 2013.



Şekil 1.8. Büyük Veri (Big Data)[105]

İnternetin çok defa fazla kişi tarafından kullanılır olması, günlük yaşamda alınan hizmetlerin, internet üzerine taşınan uygulama yazılımları sayesinde kolayca erişilebilir olması nedeniyle oluşan yaygın kullanım, özellikle hizmet sektöründe satış sonrası müşteri memnuniyeti sağlama ihtiyacına yönelik olarak, son kullanıcıların her türlü bilgisinin firmalarca alınıp saklanması gerektiğini göstermektedir. Saklanacak ayrıntılı bilgilerin, üzerinde tutulacağı sayısal ortamlarda baş gösteren "yer darlığı" yeni arayışlar doğurmuş ve son yıllarda büyük veri dediğimiz olgu oluşmuştur.

Büyük verinin uygulandığı alanlara ilişkin örnekler aşağıda sıralanmıştır:

- Sağlık sektöründe; hastaneler; hastalarına yönelik etkili, bireysel, kişiye özel tıbbi hizmetler sunabilmek için, bireysel bazdaki verileri kendi sayısal ortamlarında depolamaktadır. Yine sağlık sektöründe büyük genomik veritabanları, araştırmacıların sürekli erişimine açık olmak durumundadır. Özellikle ilaç sektöründe araştırmalar için bu veri tabanlarına sıklıkla başvurulmaktadır.
- Hükümetler; yurttaşlarına yönelik bilgi ve hizmetleri işleyip, saklama konusunda oluşan çok büyük ölçekli veri ile çalışmak ve bunları saklamak zorundadırlar.

---

[105] Learning To Use Big Data The Right Way **http://www.colocationamerica.com/blog/learning-to-use-big-data-the-right-way.htm**, (Çevrimiçi) 06 Ağustos 2013.

- Finans sektöründe artık bankalar, müşterileriyle ilgili olarak toplayıp sakladıkları bilgiler yoluyla, kullanıcısını tanımakta, kullanıcısının neler yapabildiğini bilmekte ve hangi işlem adımlarını gerçekleştirdiğini görebilmektedir. İnternet sayfalarını müşterileri için en etkin hale getiren bankalar, müşterisine hatırlatmalar yapan, özelleştirilebilir arayüzler sunan, zengin içerikli, hızlı ve kullanışlı bir 7/24 şube hale gelmiştir.
- Enerji firmaları, akıllı şebeke ve sayaçlar kullanarak, abonelerinin bireysel kullanımlarıyla ilgili oluşan verileri, saklayıp, buna ilişkin pazar çalışmalarını yapmaktadır.

Bunların yanı sıra uydu/Harita Dizgeleri (GPS), Akıllı MOBİL Telefonlar (GSM), çok yüksek çözünürlüklü fotoğraflar çekebilen yeni nesil kameralarla üretilen, ses ve görüntü bilgileri; saklama ortamlarının sınırlarını zorlamaktadır. Bu alanda verimlilik sağlayabilmek için yine büyük veriye başvurulmaktadır.

Sosyal medya, internet, web günlükleri, resimler, fotoğraflar, ses kayıtları, kurum verileri, GSM cihazları gibi çeşitli veriler büyük verinin kaynağı olabilir.

Büyük veri için son yıllarda en çok kullanılan kaynak sosyal medyadır. Bir çok kurum sosyal medyadan, internet veya günlüklerden veri çekip, sosyal harita analizleri yapmakta, bir çok farklı analizler elde edebilmektedir.

Özellikle her türlü gezgin aygıtlar üzerinde çalışabilen, internet tabanlı yazılım ve uygulamalarca kullanılan Facebook, Twitter gibi toplumsal medya ortamlarında kullanıcılarca üretilen bilgilerin saklanması gereği, bilişim konusu ile ilgili girişimcileri de, Büyük Veri alanına itmektedir.[106]

[106] http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the world.html?_r=2&scp=1&sq=Big%20Data&st=cse&. (Çevrimiçi) 03 Haziran 2013.

1.6. LİTERATÜR ARAŞTIRMASI

Veri madenciliği günümüzde farklı sektörlerde etkin olarak yer alan ve teknolojinin beraber getirdiği yenilikler sayesinde değeri gittikçe artan bir alandır. Özellikle pazarlama, tıp, telekomünikasyon, finans, perakendecilik gibi pek çok alanda kullanılmaktadır. Aşağıda veri madenciliği tekniklerinin kullanıldığı alanlar ve bu alanlarda yapılan çalışmalara yer verilmiştir.

Veri madenciliği bilgisayar teknolojilerindeki gelişmelerle beraber birçok alanda yaygın olarak kullanılmaktadır. Gelişen teknoloji beraberinde ham verinin saklanabilmesini ve anlamlı bilgiye ulaşılmasını sağlamaktadır. Elde edilen bu bilgilerin çok çeşitli alanlarda kullanıldığı ve çok büyük kolaylık sağladığı görülmektedir.

Bu bölümde konu ile ilgili literatür taraması yapılmıştır:

Aydoğan ve arkadaşları, bu alanda yaptıkları çalışmalarında, bir kozmetik firmasının sattığı markalardan birinin müşteri gruplarını ve ayrılma eğilimi gösteren müşteri kesitini belirleyerek; bu müşterilere özel pazarlama stratejileri geliştirmeye yönelik bir uygulama gerçekleştirmişlerdir. Kozmetik firması bünyesindeki müşterileri elde tutmak için, ücretsiz cilt bakımı, promosyon ürün dağıtımı, indirimler uygulayarak, pazarlama stratejileri geliştirmişlerdir.. Tüm bu müşteri bölümlendirme çalışmalarını uygulayabilmek için kümeleme tekniklerinden ve ayrılacak müşteri kesitini belirlemek için ise sınıflama tekniklerinden yararlanılmıştır. Çalışmada WEKA paket programı ve Churn Analizi kullanılmıştır.

Churn Analizi, Türkiye'de ilk defa bir kozmetik firmasına uygulanarak, çalışma sonucunda bulunan bulguların mağazalarda uygulanmasında ve müşterilere ilişkin pazarlama geliştirme faaliyetlerinde, satışları ve müşterilerin ayrılma

oranlarını nasıl etkilediği ve buna ilişkin yatırımın geri dönüş oranını (ROI) saptamada yol gösterebileceği düşünülmüştür. [107]

Timor ve Şimşek çalışmalarında veri madenciliği tekniklerinden biri olan "Birliktelik Kuralları Sepet Analizi"ni kullanarak Türkiye'de perakende sektöründe faaliyet gösteren büyük bir market zincirinin 2004 yılı Şubat, Mart, Nisan ve Mayıs aylarına ilişkin, 999 tüzel olmayan müşterilerden 627'sine ait 27.924 alışveriş kaydı bulunan verileri incelemişlerdir. Ele aldıkları verileri SPSS Clementine programı yardımıyla analiz ederek müşterilerin satın alma davranışlarını etkileyen değişkenleri karar ağaçları ile tespit etmişlerdir. Çalışmadan elde edilen bilgiler işletmelerin; ürünlere ilişkin reklam kampanya ve promosyonlarının belirlemesinde önemli yararlar sağlamaktadır. [108]

Timor, Ezerçe ve Şimşek perakende sektöründe yapmış oldukları bir başka çalışmada; Türkiye'nin en büyük hazır giyim perakende işletmelerinden birinin 2005 yılı Ocak, Şubat, Mart, Nisan ve Mayıs aylarına ilişkin, 5.370 müşterinin 17.164 faturasına ait 42.431 alışveriş veri kayıtlarını incelemişlerdir. Ele aldıkları verilere Birliktelik Kuralları ve Kümeleme Analizi uygulayarak müşteri profillerinin ortaya çıkarılmış, ayrıca müşterilerin satın alma davranışlarını tespit etmişlerdir. Çalışmada Kümeleme analizi ile müşterilerin demografik özelliklerine göre farklıları ortaya koyulmuştur. Çalışmadan elde edilen bilgiler işletmelerin; ürünlere ilişkin pazarlama stratejilerinin belirlemesinde etkili olmaktadır.[109]

Ay ve Çil çalışmalarında veri madenciliği tekniklerinden biri olan Birliktelik Kuralını temel alarak süpermarket yerleşim yeri düzenlemeleriyle ilgili bir uygulama geliştirmişlerdir. Çalışmada Apriori algoritması ve Çok Boyutlu Ölçekleme (ÇBÖ)

---

[107] Emel Kızılkaya Aydoğan, Cevriye Gencer, Sinem Akbulut, "Churn Analysis And Customer Segmentation Of A Cosmetics Brand Using Data Mining Techniques", **Sigma Mühendislik ve Fen Bilimleri Dergisi,** Cilt 26, Sayı:1, 2008, s.43-45.

[108] Mehpare Timor, Umman Tuğba Şimşek, "Veri Madenciliğinde Sepet Analizi İle Tüketici Davranışları Modellemesi",**Yönetim Dergisi,** Yıl 19, Sayı:59, Şubat 2008, Makale 396, (Çevrimiçi), www.iie.istanbul.edu.tr/ydsy/MakaleGoster.aspx?m=396, 13 Ekim 2011.

[109] Mehpare Timor, Ayşegül Ezerçe, U. Tuğba Şimşek, "Müşteri Profili Ve Alışveriş Davranışlarını Belirlemede Kümeleme Ve Birliktelik Kuralları Analizi: Perakende Sektöründe Bir Uygulama", **Yönetim Dergisi,** Yıl 22, Sayı:68, Şubat 2011, S. 129-148.

yöntemlerinden yararlanılmıştır. Türkiye'nin önde gelen perakende satış firmalarından biri olan Migros Türk A.Ş.'deki MMM büyüklüğündeki bir mağazaya ait 01.01.2003 ve 31.05.2003 tarihleri arasındaki alışveriş kayıtları kullanılmış ve toplamda 186.883 adet müşteri kaydı elde edilmiştir.

Çalışmada Birliktelik Kuralından yararlanılarak elde edilen sonuçlar sayesinde marketlerde kullanımı uygun olan, müşterilerin alışverişlerini hiyerarşik bir seviyede yaptığı ızgara tip market düzeni önerisi geliştirilmiş, çok boyutlu karar verme metotlarından ÇBÖ analizi sonuçları doğrultusunda firma için yeni bir yerleşim düzeni önerilmiştir. Önerilen yerleşim düzeninin, müşterilerin davranışlarını etkilediği sonucuna ulaşılmıştır. Çalışmanın sonucunda perakendecilik sektöründe birlikte satılma eğilimi gösteren ürünler ve ürünler arasındaki ilişkiler hakkında önemli bilgiler ortaya koyulmaktadır.[110]

Veri madenciliği her alanda olduğu gibi müşteri ve satış verilerinin depolanmasında, veri setleri içinde önceden bilinmeyen anlamlı olabilecek ve değerli olabilecek verilerin elde edilmesinde, bu verilerin uygun hale getirilmesinde etkili olmaktadır.

Bu alanda Özçakır çalışmasında bir firmanın 7 tane pastane satış lokasyonu verilerini inceleyerek, bu satış verilerinden bir veri tabanı oluşturmuş ve bu verileri veri madenciliği tekniklerinin uygulanabilmesi için uygun hale getirmiştir. Satış verileri üzerinde veri madenciliği tekniklerinden Birliktelik Kuralı tekniği kullanılarak Apriori algoritmasından yararlanılmıştır. Bu algoritmayı kullanabilmek için, .net platformunda web tabanlı bir uygulama yazılımı hazırlanmış, bu uygulama yazılımı algoritmanın her kademesinde yapılan işlemleri adım adım göstermesini sağlayacak şekilde düzenlenmiştir.

Uygulanan Apriori algoritması sonucunda farklı zaman dilimi ve farklı satış lokasyonu girdi değerleri ile ilişkili olarak satış alınan ürünler olduğu görülmüştür.

[110] Derya Ay, İbrahim Çil ,"Migros Türk A.Ş.'de Birliktelik Kurallarının Yerleşim Düzeni Planlamada Kullanılması", **Endüstri Mühendisliği Dergisi,** Cilt: 21 Sayı: 2, 2010, s: 14-29.

Aynı ürün grubundaki ürünlerin en sık satın alınan ürünleri oluşturduğu belirlenmiştir. Bu birliktelik kuralı bilgilerinin elde edilmesi sektörde üretim ve pazarlama faaliyetlerinde etkili olmakta, ürünlerle ilgili değişikliklere gidilmesi ve satış lokasyonunda reyon farklılığın oluşturulması gibi durumları gündeme getirmektedir.[111]

Tüzüntürk ise çalışmasında; veri madenciliği ile istatistik arasındaki ilişkiyi incelemiştir. Çalışmada veri madenciliği süreci ayrıntılarıyla anlatılmış, bu süreçte istatistiğin ne kadar önemli olduğu vurgulanmıştır. Veri madenciliği ile ilgili Türkiye'nin 81 ilinin sosyoekonomik gelişmişliği üzerine OLAP analiziyle bir uygulama yapılmış ve 2000 yılında TUİK tarafından yapılan anket verileri analizde kullanılmıştır. Yapılan ankette yer alan 59 değişken incelenmiş ve OLAP Küpleri sayesinde Türkiye'nin gelişmişlik düzeyine göre çeşitli istatistiksel raporlar elde edilmiştir.[112]

Veri madenciliği birçok alanda yaygın olarak kullanıldığı gibi sağlık sektöründe de çok etkin bir şekilde kullanılmaktadır. Sağlık sektöründe, elde bulunan büyük veri setleri içinde, klinik ve demografik bilgiler gibi öncelikli verilerin düzenlenmesi zaman ve kaynak maliyeti açısından büyük sorunlara yol açmaktadır. Özellikle sağlık sektöründe; hastalıkların tespitinde, hastanın teşhis tedavi sürecinde, hekimlerin teşhis ve tedaviye ilişkin kararlarını etkileyen verilere erişimin sağlanması aşamasında veri madenciliğinden yararlanılmakta, hastaya ilişkin tüm bilgilerin, laboratuvar sonuçlarının, röntgen, MR gibi görüntü dosyalarının tek kayıt içerisinde düzenlenmesinde veri madenciliğinin büyük önemi bulunmaktadır.

Bu konuyla ilgili olarak Koyuncugil ve Özgürbaş çalışmalarında; sağlık sektöründe veri madenciliği konusuna bir altyapı oluşturmakta, veri madenciliği ve veri tabanı kavramları, veri ambarı oluşturma, iş zekası ve veri madenciliği

---

[111] Feridun Cemal Özçakır, "**Müşteri İşlemlerindeki Birlikteliklerin Belirlenmesinde Veri Madenciliği Uygulaması**", Yüksek Lisans Tezi, Marmara Üniversitesi, Fen Bilimleri Enstitüsü, 2006, s. 6.

[112] Selim Tüzüntürk,"Veri Madenciliği ve İstatistik", **Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,** Cilt XXIX, Sayı: 1, 2010, s: 65-90.

yöntemleri konularında bilgiler ve örnekler sunmaktadır Veri madenciliği özellikle sağlık sektöründeki öncelikli konu ve sorun alanlarına yeni bir bakış açısı kazandırmaktadır.[113]

Hastane bilgi sistemleri hastalara ilişkin demografik bilgiler, klinik bilgileri hasta ve tedavi durumları, yapılan tetkikler ve diğer idari bilgileri kapsamaktadır. Bilgi sistemlerinde tıp ve sağlık alanındaki birçok verinin sayısal ortamda saklanması ve bunların düzenlenmesi noktasında klasik sorgulama yöntemleri eksik kalabilmektedir. Bu doğrultuda veri madenciliği yöntemlerinden yararlanılmaktadır. Bu sayede elde edilen bilgiler hem tıbbi araştırmalar hem de yönetim planları için önem taşımaktadır.

Yıldırım ve diğerleri, çalışmalarında dünyadaki sağlık ve tıp alanında yapılan veri madenciliğiyle ilgili çalışmalar hakkında bilgi vermiş ve Hacettepe Üniversitesi Radyoloji Bilgi Sistemleri veri tabanı üzerinde ilişkisel veri madenciliği algoritmalarını kullanarak tıbbi açıdan değerli bilgilerin ortaya çıkarılmasını amaçlamışlardır.

Çalışma sonucunda hem kaynakların daha doğru planlanması, düzenlenmesi, gelecek dönemler için planların uygulanabilmesi amaçlanmakta, hem de hastane bilgi sistemleriyle ilişkili olan radyoloji bilgi sistemi incelenerek klinik çalışmalara yön gösterilmek istenmektedir.[114]

Oğuz çalışmasında metin madenciliği tekniklerini kullanarak; Akdeniz Üniversitesi Hastanesi'nin 2002-2007 yılları arasında Kulak Burun Boğaz Hastalıkları Anabilim Dalı'ndan alınan, ameliyat geçiren 600 hastaya ait hasta bilgi formunu inceleyerek yapılandırmıştır. Yapılandırılmış formata dönüştürülen formlar, hekimler için hasta ile ilgili ihtiyaç duydukları bilgilere ulaşmada ve klinik bilgilerini

[113] Ali Serhan Koyuncugil, Nermin Özgülbaş, "Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları", **Bilişim Teknolojileri Dergisi**, Mayıs 2009, Cilt:2, Sayı: 2, s.21-23.
[114] Pınar Yıldırım, Mahmut Uludağ, Abdülkadir Görür, "Hastane Bilgi Sistemlerinde Veri Madenciliği", **Akademik Bilişim 2008,** Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, 30 Ocak - 01 Şubat 2008, s. 429-431

çıkarmada kolaylık sağlaması amacıyla analiz edilmiştir. Bu analizleri gerçekleştirebilmek için bir yazılım geliştirilmiştir. Çalışmada Birliktelik Kuralı yönteminden yararlanılmıştır.

Önce Microsoft Office Word belge formatında bulunan hasta bilgi formlarındaki veri alanları ön işlemden geçirilerek veri tablosu haline dönüştürülmüştür. Veri tablosundaki veriler sayesinde, hazırlanan metin sorgu formuyla birlikte hekimlerin hasta bilgi formlarında aradıkları özellikteki hastalara erişimleri bu ölçüde kolaylaşmıştır. Ayrıca her alana özgü oluşturulan anahtar kelime listeleriyle metin içerikleri kodlanabilmiş ve bu veriler üzerinde veri madenciliği teknikleri uygulanmıştır.[115]

Dinçer ve Duru çalışmalarında veri madenciliği kümeleme tekniklerinden biri olan k-ortalamalar algoritmasını kullanarak bir yazılım geliştirmişlerdir. Bu yazılım sayesinde gırtlak kanseri ameliyat verilerinin geçmiş kayıtlarını analiz ederek, tıp doktorlarının, hasta dosyalarına ilişkin geçmiş kayıtları tek tek incelemelerine gerek duymadan tıbbi bilgilere ulaşmalarını ve elde edilen tıbbi bilgilerle geleceğe yönelik tahminde bulunabilmelerini sağlamayı amaçlamışlardır. Bunun için Kocaeli Üniversitesi Tıp Fakültesi Hastanesi, Kulak, Burun ve Boğaz Bölümü veri tabanındaki 400 gırtlak kanseri ameliyat verisini ele almışlardır. Çalışmada gerçek veri tabanı üzerinden gidilerek, kullanılan yazılımın uygulanan etkinliği ortaya çıkarılmaya çalışılmıştır.

Çalışma sonucunda gırtlak kanseri için en etkili tedavi yönteminin, benzer hasta sonuçlarından elde edilen bilgilerle yapılan tedavi olduğu belirlenmiştir. Bu sonuç veri madenciliği yöntemlerinin sağlık sektöründe kullanılmasının tıbbi sonuçlara ulaşmada ne kadar kolaylık sağladığını göstermektedir.[116]

---

[115] Başak Oğuz, "Metin Madenciliği Teknikleri, Kullanılarak Kulak Burun Boğaz Hasta Bilgi Formlarının Analizi", **Akdeniz Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Ve Tıp Bilisimi Anabilim Dalı,** Yüksek Lisans Tezi, 2009, s: 18-21.

[116] Esra Dinçer, Nevcihan Duru, "Gırtlak Kanseri Ameliyat Verilerinin K-means Yöntemiyle Analizi", **TBV Bilgisayar Bilimleri ve Mühendisliği Dergisi**, Aralık, 2006, s:53-62.

Koh ve Tan çalışmalarında; sağlık alanında veri madenciliği uygulamalarını incelemektedir. Özellikle sağlık uygulamaları içinde davranış etkililiğini değerlendirebilmek amacıyla, tedavi etkinliği, sağlık yönetimi, müşteri ilişkileri yönetimi, sahtekârlık ve kötüye algılama gibi konular hakkında bilgiler verilmektedir. HealthOrg sağlık yönetimi programı çerçevesindeki 2050 kişi üzerinde uygulanan bu çalışmada, diyabet başlangıcı ile ilgili olan cinsiyet, yaş, vücut kitle indeksi, bel/ kalça oranı, sigara içme durumu, haftalık egzersiz yapma sayısı gibi değişkenler ele alınmıştır. Bu değişkenlerin diyabet için bir risk faktörü oluşturup oluşturmadığı belirlenmeye çalışılmıştır. Analizler SPSS Clementine programı yardımıyla uygulanmıştır. Makalede son olarak veri madenciliği sınırlamaları ve diğer yöntemler hakkında bilgilere de yer verilmiştir. [117]

Breault, Fos ve Goodall, Amerika'da yaptıkları bir araştırmada diyabet hastalığının oluşmasına neden olan faktörleri karar ağaçları ile tespit etmeye çalışmışlardır. [118]

Veri madenciliğinin bir diğer uygulama alanlarından biri finans sektörüdür. Finans sektöründe özellikle bankacılıkta veri madenciliği yaygın bir şekilde kullanılmaktadır.

Albayrak ve Yılmaz çalışmalarında, İMKB 100 endeksinde sanayi ve hizmet sektörlerinde yer alan 173 işletmenin 2004–2006 yıllık finansal göstergelerini incelemişlerdir. İncelemeler sonucunda ele alınan verilere karar ağaçları tekniği CHAID algoritması uygulanmıştır. Çalışmada seçilen finansal göstergelere göre sanayi ve hizmet sektörlerinde yer alan firmaları ayıran en önemli değişkenlerin neler olduğu belirlenmeye çalışılmıştır. İMKB100 endeksinde faaliyet gösteren firmaları ayırmada etkili olan en önemli değişkenlerin karlılık, likidite, varlık kullanım

[117] Hian Chye Koh, Gerald Tan, "Data Mining Applications in Healthcare", **Journal of Healthcare Information Management,** Vol, 2011, (Çevrimiçi) https://www.himss.org/content/files/jhim/19-2/datamining.pdf 15 Kasım 2011.

[118] Joseph Breault, Colin Goodall, Peter Fos, "Data Mining a Diabetic Data Warehouse", **Artificial Intelligence in Medicine**, 2002, s. 37-54.

etkinliği, sermaye yapısı ve işletme büyüklüğü değişkenleri olup olmadığını incelemişlerdir.[119]

Albayrak çalışmasında; "Türkiye'deki yerli ve yabancı ticaret bankaların finansal etkinliğe göre veri madenciliği teknikleriyle karşılaştırmıştır. Çalışmada yerli ve yabancı olarak önceden grup üyeliği belirlenmiş bankalar çeşitli veri madenciliği sınıflandırma teknikleri (diskriminant, lojistik regresyon ve karar ağacı modelleri) ile sınıflandırılmıştır. Çalışmada bankaların likidite, gelir-gider, karlılık ve faaliyet oranları kullanılarak karşılaştırmalar yapılmıştır.

Çalışmanın sonucunda; bankaların sınıflandırmasında karar ağacı modelinin, geleneksel diskriminant ve lojistik regresyon modellerine üstünlük sağlayabileceği ve etkili bir sınıflandırma tekniği olarak kullanılabileceği ortaya koyulmuştur.[120]

Aşan çalışmasında; kredi kartı kullanan banka müşterilerinin sosyo-ekonomik özelliklerinin gruplandırılmasını incelemiştir. Çalışmada kümeleme analizi tekniği ile kredi kartı kullanan banka müşterilerini sosyo-ekonomik özelliklerine göre üç kümede gruplamıştır. Bu üç kümeye göre müşterilerin on adet sosyo-ekonomik değişkene göre faklılık gösterdiği belirlenmiştir. Çalışma sonucunda; ilgili müşterilere verilecek bireysel bankacılık hizmetlerinde ne tür müşteriyle karşılaşabileceğini bilmek açısından fayda sağlandığı, ne tür müşterilerin hangi özelliklere ve motivasyonlara sahip olduğunu bilmek açısından da yarar elde edildiği üzerinde durulmuştur.[121]

Bilen ise çalışmasında; bankacılık sektöründe çalışan satış personellerinin performanslarının değerlendirilmesi ve kümeleme yöntemlerinden k- ortalama ile

---

[119] Ali Sait Albayrak, Şebnem Koltan Yılmaz, "Veri Madenciliği: Karar Ağacı Algoritmaları Ve İMKB Verileri Üzerine Bir Uygulama", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2009, Cilt:14, Sayı:1, s:31-52.

[120] Ali Sait Albayrak,"Türkiye'de Yerli Ve Yabancı Ticaret Bankalarının Finansal Etkinliğe Göre Sınıflandırılması: Karar Ağacı, Lojistik Regresyon Ve Diskriminant Analizi Modellerinin Bir Karşılaştırılması", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2009, Cilt:14, Sayı:1, s:113-139.

[121] Zerrin Aşan, "Kredi Kartı Kullanan Müşterilerin Sosyo Ekonomik Özelliklerinin Kümeleme Analizi İle İncelenmesi", **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, Cilt: 17,2007, s: 256-267.

personellerin performans başarı düzeylerine göre sınıflandırılmasını incelemiştir. Çalışmada çalışanların yaş, medeni hal, cinsiyet gibi demografik bilgileri, öğrenim durumu, yabancı dili, SPK belgesi gibi eğitim durumlarına ilişkin bilgileri, çalıştığı şubesine ve iş yaşamındaki pozisyonuna ilişkin bilgileri dikkate alınarak veri madenciliğinde sınıflandırma algoritmaları kullanılmıştır. Uygulanan veri madenciliği teknikleri ile çalışanların performanslarına göre değerlendirilmesi yapılmış, hangi özelliklerdeki personelin hangi şubede ne oranda başarılı olduğu belirlenmiş ve bir takım kurallar oluşturulmuştur. Bu kurallar dikkate alınarak, bir personelin özelliklerine göre hangi şubelere atanabileceği ya da ataması düşünülen şubede hangi düzeyde performans gösterebileceğinin tahmin edilebileceği vurgulanmıştır. [122]

Tosun çalışmasında; veri madenciliği yöntemlerinden faydalanarak kredi kartı müşterilerinin kaybedilme sebepleriyle ilgili sonuçlara ulaşmayı amaçlamıştır. Çalışmada müşterinin neden kaybedildiği bilgisinin yanı sıra, hangi tür müşterilerin daha sık kaybedildikleri tahmin edilmeye çalışılmış, müşterilerin kaybedilme olasılıkları incelenmiştir. Çalışma sonucunda son 3 ayda hiç alışveriş yapmayan ya da hesabı hareketli olmayan müşterilerin kaybedilme olasılığının oldukça yüksek olduğu görülmüştür.[123]

Çalış bankacılık sektörüyle ilgili olan çalışmasında; veri madenciliği yöntemlerinden kümeleme ve sınıflandırma ile mevcut bireysel kredi müşterilerinin analizi ve gelecekteki potansiyel müşterilerin ödeme durumlarını incelemiştir. Çalışmada veri madenciliği yazımı olarak SPSS Clementine kullanılarak, mevcut müşterileri kümeleme analizi ile değerlendirilmiş, bireysel kredi müşterilerinin

[122] Hamdi Bilen, "**Bankacılık Sektöründe Personel Seçimi Ve Performans Değerlendirilmesine İlişkin Veri Madenciliği Uygulaması"**, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, 2009, s. 85.

[123] Tuğba Tosun, "**Veri Madenciliği Teknikleriyle Kredi Kartlarında Müşteri Kaybetme Analizi,** Yüksek Lisans Tezi," İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, 2006, s.36.

ödeme durumları sınıflandırılarak, gelecekteki potansiyel müşteriler için çıkarımlar yapılmış ve müşteri portföyü oluşturulması amaçlanmıştır.[124]

Bankacılık alanında ise veri madenciliği son derece önemli konulardan biri haline gelmiştir. Bu alanda yapılan çalışmalardan biri de Chye ve Gerry'nin banka müşterilerinin gelecek altı ayda, rakip firmaların önüne geçip geçmeyeceğini tahmin etmeye dayalı yaptıkları uygulamadır. "Churn Analizi" ile yapılan bu çalışma; müşterilerin neden başka bankalarla çalışmak istediğinin tespit edilmesi ile ilgili önemli noktalara değinmektedir. Çalışmada ayrıca Churn analizi modellemeleri sayesinde bankacılık sektöründe müşteri, ciro tahmini yapılmasında veri madenciliğinin önemi anlatılmakta ve veri madenciliğinin diğer uygulamaları hakkında bilgiler verilmektedir. [125]

Veri madenciliğinin bir diğer uygulama alanlarından biri havacılık sektörüdür. Özellikle Türkiye'de gelişmekte olan hava yolları işletmeleri büyük miktarlarda veri ve bilgi birikimine sahiptir.

Gürbüz ve arkadaşları çalışmalarında, Türkiye'de bir hava yolu işletmesinin parça söküm raporları üzerinde veri madenciliği çalışması yapmışlardır. Çalışmada, uçaklarda kullanılan parçaların, herhangi bir arıza meydana gelmeden önce düzeltici ve önleyici işlemlerin yapılması için uyarı durumlarının belirlenmesiyle ilişkili olarak kural geliştirilmeye çalışılmıştır. Sınıflandırma analizi ve farklı veri madenciliği tekniklerinden yararlanılmıştır. Çalışma sonucunda elde edilen bulguların havacılık sektöründe önemli katkıları olacağı sonucuna ulaşılmıştır.[126]

Bilgi iletişim teknolojilerinin yoğun olarak kullanıldığı günümüzde; web madenciliğinden de yararlanılmaktadır. Özellikle internetin yaşamımızın her

---

[124] Aslı Çalış, "**Veri Madenciliği Yaklaşımı İle Bireysel Müşterilerin Kredi Ödeme Performanslarının Değerlendirilmesi,** Yüksek Lisans Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, 2013, s.62.

[125] Koh Hian Chye, Chan Kin Leong Gerry, "Data Mining and Customer Relationship Management in the Banking Industry", **Singapore Management Review**, Volume 24, Issue 2, 2002, s.1-27.

[126] Feyza Gürbüz, Lale Özbakır, Hüseyin Yapıcı, "Türkiye'de Bir Havayolu İşletmesine Ait Parça Söküm Raporlarına İlişkin Veri Madenciliği Uygulaması", **Gazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi**, Cilt 24, No 1, 2009, s:73-78.

alanında yer almasıyla bilgiye ulaşım, internet üzerinden ağırlıklı olarak sağlanmaktadır. Bu durum sonucunda Web madenciliği tekniklerine başvurulması kaçınılmaz hale gelmiştir.

Web madenciliği yöntemleri kullanılarak, web sitelerindeki yeni web teknolojileri gözden geçirilebilmekte ve elde edilen bilgiler ışığında web sitesinde görülen eksiklikler ve hatalar belirlenebilmektedir. Ayrıca web sitesinin gelişimi ve etkinliği ortaya çıkarılabilmektedir. Bu durum web sitelerinin performansını artırmakta, bunun yanı sıra kullanıcılara kullanım rahatlığı sağlanmaktadır.

Özseven ve Düğenci çalışmalarında web madenciliği yöntemini kullanarak internet kullanıcılarının web sitesi ziyareti sürecince elde ettikleri erişim kayıtlarını analiz ederek, atıl durumdaki verileri bilgiye dönüştürme işlemi gerçekleştirmişlerdir. Bu işlem için Gaziosmanpaşa Üniversitesi kurumsal web sitesine ait 15 günlük sunucu erişim kayıtlarını incelemiş, bu kayıtların analizi için "Log Yazılım" adlı bir yazılım geliştirmişlerdir. Bu yazılım web sitesine ait çeşitli istatistiksel bilgileri ortaya çıkarmaktadır. Analiz için Birliktelik Kuralı ve Apriori Algoritması uygulanmıştır.[127]

Karadağ ve Takçı çalışmalarında farklı haber sitelerinin farklı kategorilerdeki birbiriyle ilişkili haberlerini tarayarak, okuyucuya faydalı bilgiler sunan bir sistem geliştirmeye çalışmışlardır. Bu çalışmada 2007 yılında haberara.net isimli birçok kaynaktan, sekiz farklı kategoriden 50.000 haber içeren bir veri kümesi ele alınmıştır. Her haberdeki kelime etiketleri belirlenerek, etiketler diğer haberlerin kelime etiketleriyle karşılaştırılarak haberler arasındaki benzerlikler bulunmaya çalışılmıştır. Bulunan etiketler kelime köklerinden elde edilmektedir. İlişkili haberleri belirlemek için bu etiket bilgilerinden yararlanılır ve haberler arasındaki ilişkiler ortaya konur. Bu doğrultuda metin madenciliği yöntemlerinden doküman kümeleme yönteminden yararlanılmıştır.[128]

---

[127] Turgut Özseven, Muharrem Düğenci, "LOG Analiz: Erişim Kayıt Dosyaları Analiz Yazılımı Ve GOP Üniversitesi Uygulaması", **Bilişim Teknolojileri Dergisi,** Cilt: 4, Sayı: 2, Mayıs 2011,s.55.
[128] Karadağ,Takçı, **a.g.e.** s.5.

Veri Madenciliği çalışmalarını gerçekleştirmek için ticari ve açık kaynak olmak üzere birçok program kullanılmaktadır. Dener ve diğerleri, veri madenciliğinde yaygın olarak kullanılan WEKA programıyla ilgili çalışmalarında; açık kaynak kodlu Veri Madenciliği programlarından olan RapidMiner(YALE), WEKA ve R programlarına ilişkin bilgiler vermişlerdir. Çalışmada bu programların birbiriyle ilgili karşılaştırmalarına yer verilmekte, algoritma ve kullanım açısından üstünlüklerinden bahsedilmektedir.

Çalışmada ayrıca Gazi Üniversitesi Fen Bilimleri Enstitüsü verilerinden yararlanılarak, enstitüye kayıtlı olan öğrencilerin lisans eğitimlerinde mezun oldukları bölüm ile lisansüstü eğitimindeki bölümler arasındaki farklılıkların ya da aynı bölüme devam etmesinin başarıları üzerindeki etkisi incelenmiştir. Bu duruma ilişkin öğrencilerin cinsiyetleri, ikamet durumları ve derslerdeki başarı oranları ele alınmıştır. Uygulama WEKA üzerinde yapılmıştır. Uygulama sonucunda, öğrencilerin çeşitli ölçütlere göre elde edilen başarı ve başarısızlık sonuçlarının lisansüstü eğitimi veren tüm Enstitülere yararlı olacağı düşünülmüştür. [129]

Savaş ve Arıcı çalışmalarında web tabanlı uzaktan eğitimde iki öğretim modelinin öğrencilerin başarısı üzerindeki etkisini araştırmışlardır. Bu doğrultuda öğrencilerden video destekli öğretim araçları ile animasyon destekli öğretim araçlarını 7 günlük bir süreçte farklı zaman dilimlerinde evlerinde, kişisel bilgisayarlarında internet üzerinden yayınlanan modülleri kullanarak çalışmaları istenmiştir. Web tabanlı uzaktan eğitimde öğrenci davranışları veri tabanında oluşturulmuştur. Oluşturulan veri tabanında bulunan tablolar ve toplanan veriler SQL komutları yardımıyla analiz edilmiştir. Elde edilen sonuçlar video destekli öğretim araçlarının, animasyon destekli öğretim araçlarına göre öğrenci başarısında daha

[129] Murat Dener vd, "Açık Kaynak Kodlu Veri Madenciliği Programları: Weka'da Örnek Uygulama", **Akademik Bilişim'09 - XI. Akademik Bilişim Konferansı Bildirileri**, 11-13 Şubat 2009, Harran Üniversitesi, Şanlıurfa, (Çevrimiçi) ab.org.tr/ab09/bildiri/42.pdf, 18 Kasım 2011.

etkili olduğunu göstermiştir. Bu durum veri analizinde veri madenciliği tekniklerinin ne kadar etkili olduğunu açık bir şekilde yansıtmaktadır.[130]

Telekomünikasyon sektöründe web madenciliği çalışmaları söz konusudur. Dolgun ve arkadaşları, çalışmalarında bir telekomünikasyon kurumunun 2070 müşterisine ait 17 değişkenden oluşan şirketi terk etme (churn) yapısal verisini incelemiş, veriler SPSS Clementine 12. 0 programında yer alan karar ağaçları algoritmalarından C5 kullanılarak analiz edilmiştir. Amaç şirketi terk eden müşterilere ait bir profil modeli oluşturmaktır. Bunun yanı sıra çalışmada müşterilere ait metin dosyaları kullanılarak, var olan yapısal veriye eklemeler yapılarak ikinci model, müşterilerin internet üzerinden ulaşılan web log dosyaları kullanarak üçüncü model oluşturulmaktadır. Yapısal veri ile yapısal olmayan verinin metin ve web madenciliği kullanılarak yapısal hale dönüştürülmesi incelenmiş ve bu model karşılaştırmaları detaylarıyla incelenmiştir.[131]

Işık ve Çamurcu çalışmalarında, Milliyet ve Yahoo News web sayfalarındaki bazı kelimeleri çeşitli oranlarda rastgele seçmiş, bu sayede elde edilen veri setinde, benzerlik ve uzaklık ölçütleri tekniklerini incelemişlerdir. Çalışmada web sayfalarını kümelemede web sayfalarında yer alan belgelerin benzerliklerini bulan tekniklerden Öklid, Kosinüs, Pearson ve Genişletilmiş Jaccard iki ayrı veri setinde incelenmiş ve hangi benzerlik tekniklerinin kullanılmasının uygun olduğu belirlenmiştir.[132]

Baykal çalışmasında; Dicle Üniversitesi sunucuları üzerindeki Access log dosyalarının 1-8 Ocak 2006 kayıtları üzerinde incelemelerde bulunarak, web/mail sitesi günlükleri üzerindeki erişim desenlerini yorumlamış ve web erişim günlükleri kullanılarak yapılan kötü amaçlı saldırıları tespit etmeye çalışmıştır.[133]

---

[130] Serdar Savaş, Nursal Arıcı," Web Tabanlı Uzaktan Eğitimde İki Farklı Öğretim Modelinin Öğrenci Başarısı Üzerindeki Etkilerinin İncelenmesi", **5. Uluslararası İleri Teknolojiler Sempozyumu (IATS'09)**, 13-15 Mayıs 2009, Karabük, .

[131] Dolgun. Özdemir, Oğuz **a.g.e.** ,s:48-58.

[132] Meltem Işık, Ali Yılmaz Çamurcu," Web Belgeleri Kümelemede Benzerlik Ve Uzaklık Ölçütleri Başarılarının Karşılaştırılması", **Marmara Üniversitesi Fen Bilimleri Enstitüsü Dergisi**, 20 (2008), s: 35-49.

[133] Abdullah Baykal, "Web Madenciliği ile Log Analizi", **D.Ü.Ziya Gökalp Eğitim Fakültesi Dergisi 8**, 2007, s:129-145.

Carus ve arkadaşları çalışmalarında, belirledikleri iki haber sitesinin HTML yapısının analizini ön planda çeşitli kategori ve başlıklar altında incelemişlerdir. Bu doğrultuda Cep gazetesi adını verdikleri, NET ortamında cep bilgisayar benzeticisi kullanarak C# programı ile geliştirmişlerdir. Yapılan uygulama sonucunda çeşitli kategorilere ait haberlerin yer aldığı başlık ve özet verileri elde edilmiş ve bu veriler cep bilgisayarları üzerinden kullanıcıya sunulmuştur.

Web içerik madenciliğinin kullanımı sonucunda; bu uygulamanın siteyle ilgili kategorilere ait sayfalara bağlanmakta ve sayfa haberlerine ulaşmakta sabit bağlantı hızına göre çok daha hızlı olduğu, veri boyutu açısından da ciddi bir tasarruf sağladığı sonucuna ulaşılmıştır.[134]

Kalıkov çalışmasında e-ticaret amaçlı kurulan Pegem A yayınevinin web sitesinin veri tabanında yer alan verilere veri madenciliği teknikleri uygulayarak analiz etmiştir. Bu çalışmaya göre Pegem A web sitesinden alınan 6 aylık alışveriş sepet tablosunun ve ona bağlı ürün tablolarının yapıları incelenmiştir. Bu veriler klasik istatistiksel tekniklerin yanı sıra Birliktelik Kuralları ve Karar Ağaçları teknikleri kullanılarak incelenmiştir. Tekniklere ilişkin uygulama sonucunda, veri tabanında bulunan sanal ürünlerin(kitapların) kategorilerine göre doğru yerleştirilmesinde faydalı olabilecek bilgilere ulaşılmıştır.

Bu çalışma sonucunda kullanıcıların site üzerinden yapacakları alışverişlerde, hangi kitapları tercih ettikleri, ilgi alanlarına göre kitap tavsiyesinde bulunacak bir SQL sorgulaması geliştirilmeye çalışılmış, bu amaçla birliktelik kuralları tekniğinden yararlanılmıştır. Bunun yanı sıra kişilerin aylık alışveriş bilgilerine göre istatistiksel bilgiler elde edilmiştir.[135]

---

[134] Aydın Carus v.d., "Cep Gazetecisi: Mobil Cihazlar İle Çalışan Bir Web Madenciliği Uygulaması", **Elektrik-Elektronik-Bilgisayar Mühendisliği 12. Ulusal Kongresi,** 2007, Eskişehir.
[135] Anarbek Kalıkov, **a.g.e.**, s.15.

Web kullanım madenciliği ile ilişkili olarak Carus ve Aydın inceledikleri çalışmada .NET ortamında geliştirdikleri ve "Web Sunucusu Analizcisi - WSA" olarak isimlendirdikleri uygulamalarında, web sitesi tasarımcılarına, tasarladıkları sitenin kullanım ve etkinliğinin artışında, değerlendirilmesinde bilgiler üreten bir sistem geliştirmişlerdir. Bu sayede WSA veri olarak web sunucusu erişim günlük dosyalarını kullanarak, bu verilerden elde ettiği bilgileri ortaya çıkarmaktadır. Bunun yanı sıra WSA, günlük dosyalarının incelenmesi sırasında günlükteki kayıt hatalarını da bulabilmekte, W3C, IIS 4, IIS 5, IIS 6, NCSA gibi farklı günlük dosyası biçimlerini otomatik olarak tanıyabilmekte, ortaya çıkan sonuçlar grafiksel raporlar olarak görüntülenebilmekte ve çıktı olarak aktarabilme olanağını sağlamaktadır. Bunun yanı sıra günlük dosyalardan elde edilen anlamlı bilgiler, veri tabanına saklanarak daha önceden değerlendirilmiş verilerle, yeni değerlendirilen veriler arasında kıyaslama yapma kolaylığı da göstermektedir.[136]

Jansen ve Spink çalışmalarında arama motoru kullanıcılarının davranışlarını incelemiş ve sürekliliklerini araştırmışlardır. Bu amaçla Jansen ve Spink yaptıkları çalışmalar sonucunda kullanıcı davranışlarının çoğunun Amerika merkezli olduğu ve diğer ülkelerdeki kullanıcılar için bu türdeki çalışmaların çok az olduğunu görmüşlerdir. Arama motorları incelediğinde; Allthe Web arama motorunun Avrupa'nın en yüksek oranlı kullanılan arama motoru olduğunu görmüşlerdir. Çalışmada yapılan karşılaştırmaları değerlendirme çalışmaları sonucunda Avrupa'da Web'de aramaların, belirli yönlerde olduğu görülmektedir. [137]

Web içerik madenciliği ile ilişkili olarak Chuang ve Chien çalışmalarında üç arama motorunun (Dreamer, GAIS ve Openfind) kayıtlarını toplayarak bir sorgu-sınıflandırma yaklaşımı geliştirmişlerdir. Bu sınıflandırma kayıtlar üzerinde web aramaları konularını gösteren, 14 ana ve alt kategoriden oluşarak iki aşamalı olarak yapılmıştır. Çalışmanın deneysel sonuçlarında, kullanıcılar tarafından verilen farklı şekillerdeki terimlerin otomatik olarak sınıflandırılabildiği görülmektedir. Bu

[136] Aydın Carus, Altan Mesut, "Web Kullanım Madenciliği Uygulaması", **II. Mühendislik Bilimleri Genç Araştırmacılar Kongresi**, İstanbul 17–19 Kasım 2005.
[137] Bernard J. Jansen, Amanda Spink, "An Analysis Of Web Searching By European Alltheweb.Com Users.", **Information Processing and Management,** 41 (2005) s: 361–381.

yaklaşımda sorgu terimleri kullanılarak web sınıflandırmalarını yapmanın veya var olanları geliştirmenin/ iyileştirmenin iyi bir başlangıç olarak görüldüğü belirlenmiştir. [138]

Arama motorlarının değerlendirilmesi konusunda yapılan çalışmalar incelendiğinde Vaughan; arama motorlarının performanslarının değerlendirilmesi için bir dizi yeni ölçüler önermiş, buna ilave olarak bazı yeni ölçülerin de geliştirebileceğinden bahsetmiştir. Çalışmada arama motorları için kullanılan değerlendirme ölçülerinin duyarlık/hassasiyet ve hatırlatma kavramları olduğunu belirtmiştir. Web bilgi sistemleri için bu geleneksel değerlendirme ölçülerine ilave olarak bu yeni ölçülerin test edilebilmesi amacıyla Alta Vista, Google ve Teoma arama motorlarının kayıtlarından yararlanmıştır. Geleneksel ölçülerin tamamlayıcısı olarak önerilen bu yeni ölçüler, arama sonucu sıralama kalitesi ve en üstte sıralanan sayfaları getirme şeklinde ifade edilebilir.

Çalışmada bu arama motorlarının performansları karşılaştırıldığında, çalışma sonucundaki yeni ölçülerin, incelenen üç arama motorunda belirgin performans farklılıkları oluşturduğu belirlenmiştir. Bu ölçülere bakıldığında en iyi performansı Google arama motorunun gösterdiği tespit edilmiştir. Benzer bir yaklaşım ile farklı ölçülerin kullanıldığı başka bir çalışmada Google'ın yine daha iyi bir performans gösterdiği ve bunun yanı sıra Google'ın göreceli üstünlüğü olduğu, Web üzerindeki en popüler arama motoru arasında Google'ın üst sıralarda olduğu anlaşılmıştır.[139]

Günümüzde bilgisayar sistemlerine yönelik saldırılar, bilgi güvenliğinin önemini ortaya çıkarmakta ve her geçen gün tüm sektörlerde bu konuda çeşitli koruma yöntemleri üzerinde çalışılmaktadır. Şirketler açısından bilgi güvenliğinin sağlanması önem taşımaktadır. Bu doğrultuda şirketler çeşitli koruma yöntemlerine başvurmaktadır. Bunun için de güvenlik duvarları, şifreleme sistemleri, açıklık

---

[138] Shui-Lung Chuang., Lee-Feng Chien, "Enriching Web Taxonomies Through Subject Categorization Of Query Terms From Search Engine Logs.", **Decision Support Systems, Volume: 35 (2003)**, Issue: 1 s: 113-127.

[139] Liwen Vaughan, "New Measurements For Search Engine Evaluation Proposed And Tested," **Information Processing and Management ,** Volume: 40, 2004, s:677–691.

tarayıcı araçları, nüfuz tespit sistemleri, anti-virüs yazılımları gibi birçok koruma yönteminden yararlanılmaktadır.

Sistemlerde oluşan yazılım kaynaklı açıklıklar, eksik bilgilendirme ve yetkilendirme, kullanılan şifrelerdeki zayıflıklar çoğunlukla saldırıların oluşmasına zemin hazırlamaktadır. Oluşan saldırı girişimlerinden başarılı olanlarına ise nüfuz adı verilmektedir. Bu konuda yapılan saldırı önleme çalışmalarında yararlanılan nüfuz tespit sistemleri güvenlik duvarının arkasında ağ içerisinden aktiftirler.

Veri madenciliği ile yakından ilişkili olan nüfuz tespit sistemleri kullanıcı davranışları analizine dayanmaktadır. Karabağ ve diğerleri, bu konuyla ilgili yaptıkları çalışmada; kullanıcı davranışlarında veri madenciliği tabanlı nüfuz tespit sisteminde kümeleme ve sınıflandırma yöntemlerinin kullanımıyla ilgili bilgiler vermektedir. Çalışmada kullanıcılara ilişkin bir küme oluşturup ağ trafikleri incelenmiş ve saldırgan olan kullanıcılar belirlenmeye çalışılmıştır.

Kümeleme tabanlı nüfuz tespiti anormallik tabanlı yaklaşımı ile benzerlik göstermekte, kümeleme algoritması için K- en yakın komşu yöntemi kullanılmaktadır. Bu yöntem kullanılarak şüpheli saldırı tespitlerinin önceden belirlenmesi, olası bir nüfuz durumunda bunun önlenmesinde yardımcı olmaktadır.

Bunun yanı sıra çalışmada kullanıcı davranışlarına ilişkin veri akışları, kullanılan portlar, paketler, kullanım zamanları, sistem çağrıları gibi verilerin değerlendirilebildiğini, oluşturulan modeller ile kötü niyetli kullanıcıların dinlenmesi, istemci ve sunucu arasındaki iletişimlerin şifrelenebileceğine değinilmiştir. [140]

[140] Rahim Karabağ v.d, "Kullanıcı Davranış Analizi ile Nüfuz Tespit Modeli (KDA-NTM)", ABG2005, **Ağ ve Bilgi Güvenliği Sempozyumu**, 9-12 Haziran 2005, (Çevrimiçi) http://www.emo.org.tr/ekler/00871c1fc726f0b_ek.pdf. 20 Kasım 2011.

2. İNTERNET BANKACILIĞI

Dünya üzerindeki birçok bilgisayar ağının birbirine bağlanmasıyla oluşmuş bir bilgisayar ağı olan internet, çağımızın vazgeçilmez parçalarından olmuştur. İnternet yüz binlerce bilgisayar ağı ile on milyonlarca bilgisayar başındaki yüz milyonlarca kullanıcının bilgisayar kaynaklarını ve bilgiyi paylaşmasını sağlamaktadır.

1990 yılı öncesinde; yüksek hızlı veri hatları üzerinden çalışan küresel bir bilgi ağı olan internet, başlarda iletişim ağı olarak kullanılmasına rağmen, daha sonraları sadece iletişim olanakları sağlamasının yanı sıra ticari, sosyal, kültürel, ekonomik ve politik oluşumlarda da etkin bir rol üstlenmeye başlamıştır. İnternet hayatın her alanında, aranılan her türlü veriye ulaşma imkanı sağlamakta ve hayatımızı kolaylaştırmaktadır. İnternetin en sıklıkla kullanıldığı alanlardan biri de bankacılık sektörüdür.

Teknolojinin artan bir yoğunlukla insan hayatına girmesiyle oluşan rekabet ortamında, bankalar müşterilere daha çok seçenek sunmaya çalışmaktadır. Sanal bankaların olması nedeniyle, tüm işlemler evden televizyon açma-kapama kadar basit bir hale dönüşmüştür.[141] Günümüzde artık fiziki olarak paranın mümkün olduğunca az el değiştirmesi bir hayal olmaktan çıkmıştır. Teknolojiyi en yakından takip etmesi gereken sektörlerin başında bankacılık sektörü gelmektedir ve bankalar teknolojiye en fazla yatırım yapan işletmeler arasındadır.

Bankacılık yoğun bilgiye dayanan bir endüstridir ve bankacılık işlemleri herhangi bir fiziksel değiş tokuş olmaksızın gerçekleştirilebilmektedir. Bu durum bankacılığı internet aracı için mükemmel bir yolcu yapmaktadır.[142]

---

[141] Özcan Kovan, "**Orta Yaş Üzeri Banka Müşterilerine İnternet Bankacılığı Ürününü Kullandırmaya Yönelik Uygun Arayüz Tasarlanması**", Yüksek Lisans Tezi, Beykent Üniversitesi Sosyal Bilimler Enstitüsü, İletişim Ve Tasarım Ana Sanat Dalı İletişim Sanatları Ve Tasarım Bilim Dalı, İstanbul, 2012, s.26.

[142] Vignesen Perumal, Bala Shanmugam, Internet Banking: Boon or Bane?, (Çevrimiçi), http://www.arraydev.com/commerce/JIBC/2004-12/Perumal.htm, 13 Şubat 2013.

Teknolojik yeniliklerle entegre olabilen ve kullanıcıya pek çok hizmeti aynı anda sunabilen internet bankacılığı, bankacılık sektöründe bu anlamda bankaların müşteri kazanımı ve müşteri sadakatini sağlamasını sağlayan en etkili alternatif dağıtım kanallarından biridir.

Karlılığın ve verimliliğin öneminin her geçen gün daha arttığı günümüzde bankacılık sektörü, operasyonel maliyetlerin azaltılması amacıyla müşterilerini; internet bankacılığı, mobil bankacılık, postaneler, telefon bankacılığı ve çağrı merkezleri gibi alternatif dağıtım kanalları (ADK) kullanmaya yönlendirmektedir.

İnternet bankacılığı sayesinde özellikle bankalar çok geniş bir müşteri portföyüne hizmet sunabilmekte, müşteriler ise zaman tasarrufu sağlayarak çok çeşitli işlemleri kısa sürede sonuçlandırabilmektedir.[143]

2.1.İNTERNET BANKACILIĞI KAVRAMI

İnternet bankacılığı, elektronik bankacılığının gelişiminin bir uzantısı olarak açık ağ sistemlerinin kullanıldığı, bankaların sanal ortamda kurmuş oldukları fiziki para alışverişi ve kredi kullanımı dışında tüm bankacılık işlemlerinin gerçekleştirilebildiği, alternatif bir dağıtım kanalı olarak ifade edilmektedir.[144]

İnternet bankacılığı, bankacılık hizmetlerinin gerçekleştirilmesini sağlayan en yeni dağıtım kanalıdır. İnternet bankacılığı, müşterilerin evlerinden veya işyerlerinden ayrılmadan bir telekomünikasyon ağı aracılığıyla hemen hemen bütün bankacılık işlemleri gerçekleştirebilecekleri hizmetleri kapsamaktadır.[145]

Genel bir ifadeyle internet bankacılığı, şubelerde yapılan bankacılık işlemlerinin, internete taşınması olarak tanımlanabilir. İnternet bankacılığı zaman ve

[143] Emre Pala, Burak Kartal, "Banka Müşterilerinin İnternet Bankacılığı İle İlgili Tutumlarına Yönelik Bir Pilot Araştırma", **Celal Bayar Üniversitesi İ.İ.B.F. Yönetim Ve Ekonomi Dergisi,** Yıl:2010, Cilt:17, Sayı:2 Manisa, s.45.
[144] Kovan, **a.g.e.**, s.27.
[145] Nezih Erçişli,"**İnternet Bankacılığı Uygulamaları Ve İşletme Performansı Üzerine Etkileri: Denizbank Örnek Olay Çalışması**", Yüksek Lisans Tezi, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, Konya, 2007, s.57.

yer sınırlaması olmaksızın internet erişimine sahip herhangi bir bilgisayar aracılığıyla dünyanın her yerinde kullanılabilen, bankacılık hizmetlerinin sunulması için hazırlanan alternatif dağıtım kanalıdır.

İnternet sistemi bankaların çok kısa zamanda dikkatini çekmiştir. İnternet bankacılığı bankalar arası rekabette önemli bir bankacılık unsuru oluşturmuş, müşterilerine en iyi ve en kolay hizmet anlayışını sunmaya çalışan bankalar; çağın yeniliklerini yakalayan banka imajını da yaratmıştır. Günümüzde artık şube bankacılığın yanı sıra internet bankacılığı kullanımı daha da yaygınlaşmıştır.[146]

2.2. İNTERNET BANKACILIĞININ DÜNYA'DA VE TÜRKİYE'DEKİ GELİŞİMİ

Bankacılık hizmetleri, özellikle teknolojinin gelişmesiyle yaşantımızda büyük bir alana sahip olmaya başlamıştır. "İnsanoğlunun elindeki malı takas etmesiyle başlayan bankacılık hizmetleri", günümüzde dünyanın her yerinde gerçekleştirilmekte, banka şubesinin olmadığı yerlerde ise telefon veya internet bağlantısıyla para çekme ve yatırma haricindeki her türlü bankacılık işlemi yapılabilmektedir.[147]

İnternet Bankacılığı fikri, 1970'li yılların sonları ile 1980'li yılların başlarında, World Wide Web'in gelişimi ile birlikte önem kazanmaya başlamıştır. İlk olarak banka veritabanlarının programlanması ve elektronik ortamda işlenebilmesi amacıyla bankaların istihdam ettikleri bilgi işlem uzmanları tarafından ortaya atılmıştır.

---

[146] Recep Murat Dağlı, "**Banka Müşterilerinin İnternet Bankacılığına İlişkin Kanaatlerinin İncelenmesi Ve Konu İle İlgili Pilot Bir Araştırma**", Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Pazarlama Anabilim Dalı, İstanbul, 2007, s.62.
[147] **A.e.**, s.9.

ABD'deki "NetBank" ilk internet bankacılığı uygulaması olup, 1996'da "Atlanta İnternet Bank" adı altında kurulmuştur. Citibank ve Wells Fargo gibi köklü bankalar 2001'de müşterilerine bu hizmeti sunmaya başlamışlardır.[148]

l990'ların başında teknolojik ilerlemeye paralel olarak, Türk bankaları da otomasyon standartlarını dünya çapında sürekli olarak yükseltmiştir. 1995 yılından sonra internetin tüm dünyada yaygın olarak kullanılmaya başlanması ile birlikte gerçek zamanlı işlem olanağı sağlayan internet bankacılığının Türkiye'de başladığı görülmüştür.

İnternet Bankacılığı Türkiye'de ilk defa 1997 yılında Türkiye'nin ilk özel bankası olan Türkiye İş Bankası tarafından uygulanmaya başlanmış, aynı yıl Garanti Bankası bunu izlemiştir. Daha sonra sırasıyla Osmanlı Bankası, Pamukbank, Esbank, Akbank, ve Yapı Kredi Bankası internet bankacılığı hizmetini müşterilerine sunmuşlardır. Akbank ilk defa 1999 yılında bireysel müşteriler için internet bankacılığına başlamıştır. O yıllarda müşterilerin internet bankacılığına alışkın olmaması ve güvenlik endişeleri gibi sebeplerden dolayı Türkiye'de internet bankacılığının yaygınlaşması sınırlı kalmıştır.[149]

Bugün ise internet bankacılığı aktif olarak kullanılmakta; İş Bankası, Garanti Bankası, Akbank, Vakıfbank, Denizbank, Koçbank, HSBC ve diğer bankalar Türk finansal pazarlarında geniş çapta internet bankacılığı hizmetleri sunmaktadırlar.

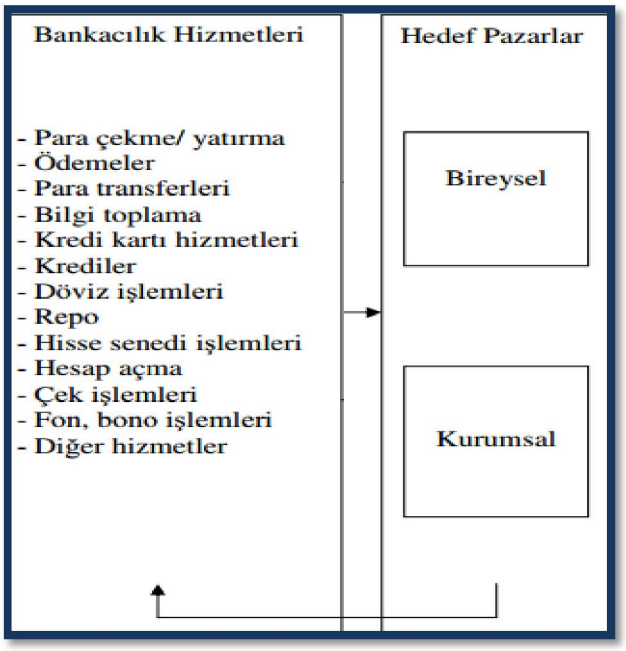
Artan bilgisayar okur-yazarlığı, finansal sektörlerdeki düzenlemeler, banka müşterilerinin internet üzerinden hizmet almak istemeleri ve sağlanan kolaylıklara bağlı olarak; internet bankacılığı 1997'den bu yana Türkiye Ticaret Bankaları tarafından alternatif dağıtım kanalı olarak algılanmaktadır. Özellikle son yıllarda bankaların ileri teknoloji kullanımlarını artırmaları, internet hizmetlerinin ve internet üzerinden yapılan ticari faaliyetlerin yoğunlaşmasıyla birlikte; internet bankacılığı

[148] David Gefen, Detmar, W. Straub, "The Relative Importance of Perceived Ease-ofUse in IS Adoption: A Study of E-Commerce Adoption," **Journal of the Association for Information Systems**, 1(8), 2000, s.1-30.
[149] Pala, Kartal, **a.g.e.**, s.45.

uygulamaları hız ve etkinlik kazanmış, gelişmiş ülkelerin bankacılık uygulamaları düzeyine ulaşılmıştır.[150]

2.3. İNTERNET BANKACILIĞI İLE SUNULAN HİZMETLER

İnternet bankacılığı, bir bankanın bireysel ve kurumsal müşterilerinin zaman ve yer sıkıntısı olmadan, mümkün olan en kısa süre ve en kolay şekilde işlem yapabilmesine olanak sağlamaktadır.



Şekil 2. 1.İnternet Bankacılığı İle Yapılan Hizmetler[151]

Bankaların müşterilerine internet bankacılığıyla sundukları bankacılık hizmetlerine bakıldığında; vadeli ve vadesiz hesap bilgilerinin dökümü, hesap hareketlerinin izlenmesi, hesap açma/kapama, kredi hesapları bakiyesi, kredi kart ödeme işlemleri, sanal kart işlemleri, kredi kartı ödemeleri, hesaplar arasında

[150] Tuğba Özdal, "**Bankacılık Sektöründe Müşteri İlişkileri Yönetimi Ve İnternet Bankacılığında Müşterilerin Seçimlerine Etki Eden Faktörler**", Yüksek Lisans Tezi, Ufuk Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Ankara, 2011, s.38.
[151] Recep Özdemir, "**KOBİ'lerin İnternet Bankacılığına Adaptasyonu: Orta Anadolu Bölgesi'nde Anket Uygulaması"**, Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı, Zonguldak, 2010, s.42.

elektronik fon transferi, havale işlemleri, fatura ödemeleri, SSK prim ödemeleri, Vergi ödemeleri, trafik cezası, üniversite harç ödemesi, KGS (Kartlı Geçiş Sistemi) ve OGS (Otomatik Geçiş Sistemi) ödemesi, cep telefonuna kontör yükleme, yatırım fonu, hisse senedi, bono-tahvil, eurobond, altın alım satımı gibi menkul kıymet işlemleri, SMS ve E-posta talimatı verme gibi daha pek çok işlem yer almaktadır.

Bir bankanın herhangi bir şubesinde yapılmakta olan fiziki para çekme ve yatırma dışında tüm bankacılık faaliyetleri gerçekleştirilebilmektedir. Bunların yanı sıra internet bankacılığında teknolojik olarak uygulama alanı bulan yöntemlerden şubesiz bankacılık, elektronik para, akıllı kart, elektronik çek gibi yöntemler en dikkat çekenleridir.[152]

2.4. İNTERNET BANKACILIĞININ AVANTAJLARI VE DEZAVANTAJLARI

Günümüzde internet bankacılığı uygulamasının tercih edilmesinin hem bankalar hem de müşteriler açısından birçok nedeni vardır.

Bankacılık sektöründe internet bankacılığı, bütün bankaların kullanmak zorunda olduğu bir kanal haline gelmiştir. Bankalar için müşteriyi elde etmek, elde tutmak ve müşteri tabanını genişletmek son derece önemlidir. Bankalar müşteriyi elde ederken, müşterilerini iyi analiz etmek ve müşteri memnuniyetini sağlamak durumdadır.

Bankaların internet bankacılığı hizmetleri sunmaları, kendileri için önemli bir ölçüde maliyet avantajı sağlamaktadır. İnternet bankacılığı kanalının bankacılık ürünlerini tanıtmak için en ucuz dağıtım kanalı olması da, bankaların internet bankacılığı hizmetini tercih etmesinde önemli bir faktördür. Bunun yanı sıra banka müşterilerinin, şubelerin çok fazla zaman ve çaba gerektirdiğini düşünerek self-servis

[152] Buket Özoğlu, "**Erciyes Üniversitesi, Tüketici Odaklı Yaklaşım İle İnternet Bankacılığı Faaliyetlerini Değerlendirmeye Yönelik Bir Uygulama**", Yüksek Lisans Tezi, Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, Kayseri, 2010, s.42.

kanallarını sağlayan bankaları tercih etmesi, bankaların müşteri portföyünü artırması için bir fırsat olmaktadır. Ayrıca bankalar bu sayede internet bankacılığı hizmeti ile şube ağlarını azaltmakta ve hizmet çalışanlarının sayısını düşürebilmektedir.[153]

Bankalar internet bankacılığının sağladığı maliyet azaltma imkânlarıyla, mevduata daha yüksek faiz verme ve işlemlerden daha az kesinti yapma imkanı sunmaktadır. Bunların yanı sıra yeni ürün ve hizmetlerin geliştirilmesi, yeni müşterilerin kazanımı, aktif müşteri sayısının arttırılması, müşteri sadakatinin arttırılması, marka imajının güçlendirilmesi, daha etkin müşteri hizmetleri sunulması, piyasadaki değişikliklere daha hızlı bir biçimde ayak uydurabilmesini de sağlamaktadır. [154]

İnternet bankacılığı sayesinde müşteri banka personeli ile yüz yüze gelme ihtiyacı duymadan ve evlerinden çıkmadan, pratik ve hızlı bir biçimde kendi işini kendi başına yapabilir. Ayrıca müşteriler ihtiyaçlarını detaylı bir şekilde takip etme olanağına sahip olmaktadır.

İnternet bankacılığı, geçen 15 yıllık periyotta, bankacılık hizmetlerinin sunulmasında müşterilere çok büyük yararlar sağlamıştır. İnternet şubeleri özellikle bütün müşterilerin aynı anda görebildiği bir ortam olduğundan fırsat eşitliği sağlamaktadır. Bunun yanı sıra bireysel ve kurumsal müşteriler, hızlı ve kolay bir şekilde ödeme ve tahsilatlarını gerçekleştirebilmekte, tüm işlemlerini takip edebilmekte, havale, EFT, yurtdışı para transferi yapabilmekte, ileri tarihli EFT talimatı verebilmektedirler. Bu işlemlere ilave olarak, hisse senedi, bono, tahvil, fon, repo, döviz alış-satış işlemlerini, fatura, vergi ve trafik cezası ödemeleri gibi işlemlerini yapabilme, firmaların bünyelerinde oluşturdukları yatırım portföylerinin değer ve değişimlerini anlık veya süreli olarak izleyebilme ve işlemlerini hızlı bir şekilde gerçekleştirebilme olanağı da bulmaktadır.

---

[153] Zeynep Özge Özcan, "**Türkiye'de Elektronik Bankacılık: İnternet Bankacılığı Üzerine Bir Çalışma"**, Yüksek Lisans Tezi, Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Muhasebe Ve Finansman Bilim Dalı, Sakarya, 2007, s.76.

[154] Robert DeYoung,, **Chicago Fed Letter**, Federal Reserve Bank of Chicago, No.163, ISSN 0895-0164, (Çevrimiçi),http://www.chicagofed.org/digital_assets/publications/chicago_fed_letter/2001/cflmar2001_163.pdf, 05 Mayıs 2013.

İnternet bankacılığından sağlanan diğer bir fayda ise, işlem komisyonlarının neredeyse sıfır düzeyinde olması, şubeden yapılan işlemlere göre internet kanalından yapılan işlemlerde daha az masraf ödemesidir.

İnternet bankacılığı sayesinde müşteriler, farklı bankaların kredi faiz oranları ile ilgili bilgilerine anında ve karşılaştırmalı olarak erişebilmekte ve bütçesine en uygun şartlardaki krediyi seçebilmektedir. Bunun yanı sıra her zaman hesabın görülebilmesi, müşterinin hesabını kolay takip etmesini, dolandırıcılık faaliyetini çok fazla zarar olmadan yakalamasını kolaylaştırır.[155]

İnternet bankacılığı kullanımlarının müşteriler için bazı dezavantajları da bulunmaktadır. Birçok güvenlik önemleri oluşturulmasına rağmen günümüzde halen kullanıcılar internet bankacılığını riskli bulmaktadır. Kullanıcılar bilgilerinin birileri tarafından ele geçirileceği ve kötü amaçlı olarak kullanılacağın düşünerek internet bankacılığını güvenli bulmamaktadır. Bu konunun üzerinde daha fazla durulması ve tedbir alınması, müşterilerin internet bankacılığı konusundaki risk algılamalarını azaltmaktadır. [156]

Müşteriler banka şubesinde tüm ihtiyaçlarının karşılandığını düşünüp, internet bankacılığına ihtiyaç duymayarak, internet bankacılığından uzak durabilmektedir. Özellikle bilgisayar ile bütünleşik ve barışık yaşayan gençlerin sorun yaşamayarak kullandıkları internet orta yaş üstü insanlar için, kullanılmaktan korkulan ve itina ile çekinilen bir araç olabilmektedir. Kullanıcılar kullanma problemleri ve bilgi eksikliği nedeniyle kullanım güçlüğü çekmeleri, kendilerine olan öz güven eksikliği nedeniyle yeni teknolojileri tanımaya karşı isteksiz olabilmekte geleneksel şube bankacılığını tercih edebilmektedir. Bankaların internet bankacılığıyla ilgili olarak kullanımı pratik ve anlaşılır bir site tasarımı yapması ve uygulamaların her müşteriye uygun tasarlanması son derece önem kazanmaktadır.[157]

---

[155] Emre Pala, **Alternatif Dağıtım Kanallarından İnternet Bankacılığına Yönelik Müşteri Tercihlerinin İncelenmesi,** Yüksek Lisans Tezi, Celal Bayar Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, 2010, s.15.
[156] Dağlı, **a.g.e**., s.71.
[157] Kovan, **a.g.e.**, s.41.

2.5. TÜRKİYE'DEKİ İNTERNET BANKACILIĞI İSTATİSTİKLERİ

Bilgi teknolojilerindeki gelişmeler internet kullanımının artmasını sağlamıştır. Dünyada İnternet kullanım oranı 2000-2012 döneminde %566.4 oranında artmıştır. Dünyada 20012 yılının tahmin edilen nüfusu olan 7.017.846.922 kişiden İnternet kullanan kişi sayısı 2.405.518.376 kişidir. Dünya nüfusunun yaklaşık %34.3'ü internet kullanmaktadır. Dünya'da internet kullanımındaki gelişmeler Türkiye'de de aktif olarak izlenmektedir.[158]

Türkiye'de internet kullanımımın artması ve bilgisayar okur-yazarlığı oranın yükselmesi sonucunda internet bankacılığı kullanımının arttığı görülmektedir. İnternet her alanda olduğu gibi bankacılık alanında da yaygınlaşmaktadır.

Türkiye Bankalar Birliği'nin üyesi, internet bankacılığı hizmeti veren toplam 26 bankadan alınan bilgilerle hazırlanan Türkiye Bankalar Birliği 2013 Haziran istatistik raporları incelendiğinde; internet bankacılığını kullanan aktif müşteri sayısın son bir yılda 1 milyon 980 bin kişi artmış olduğu görülmektedir.

Bu artışın en belirgin sebebi, internet bankacılığının kolay, kullanılabilir olması ve insanların internet bankacılığı sayesinde bankacılık işlemlerinde zaman tasarrufu elde etmiş olmalarıdır. Tablo 2.1'de internet bankacılığı kullanan müşteri sayıları verilmiştir.

[158] **World Internet Usage And Populatıon Statisıcs,** (Çevrimiçi) http://www.internetworldstats.com/stats.htm, 25 Haziran 2013.

Tablo 2.1. İnternet Bankacılığını Kullanan Müşteri Sayıları

Müşteri Sayıları	Haziran 2012	Mart 2013	Haziran 2013
Bireysel müşteri sayısı (bin kişi)			
Aktif (A) (son 3 ayda 1 kez login olmuş)	8.605	10.246	10.468
Kayıtlı (B) (en az 1 kez login olmuş)	20.509	24.012	25.313
Kayıtlı (C) (son 1 yılda en az 1 kez login olmuş)	11.828	14.654	15.707
Aktif (A) / kayıtlı (B) müşteri oranı (yüzde)	42	43	41
Kurumsal müşteri sayısı (bin kişi)			
Aktif (A) (son 3 ayda 1 kez login olmuş)	863	953	980
Kayıtlı (B) (en az 1 kez login olmuş)	2.049	2.100	2.189
Kayıtlı (C) (son 1 yılda en az 1 kez login olmuş)	1.050	1.150	1.197
Aktif (A) / kayıtlı (B) müşteri oranı (yüzde)	42	45	45
Toplam müşteri sayısı (bin kişi)			
Aktif (A) (son 3 ayda 1 kez login olmuş)	9.468	11.199	11.448
Kayıtlı (B) (en az 1 kez login olmuş)	22.557	26.112	27.503
Kayıtlı (C) (son 1 yılda en az 1 kez login olmuş)	12.878	15.805	16.904
Aktif (A) / kayıtlı (B) müşteri oranı (yüzde)	42	43	42

İnternet bankacılığı yapmak üzere sistemde kayıtlı olan ve en az bir kez giriş işlemi yapmış toplam bireysel müşteri sayısı, Haziran 2013 itibariyle, 25 milyon 313 bin kişi olmuştur. Son bir yıl içerisinde en az bir kez giriş işlemi yapmış toplam bireysel müşteri sayısı ise yaklaşık 15 milyon 707 bin kişidir. Nisan-Haziran 2013 döneminde yaklaşık 10 milyon 468 bin bireysel müşteri en az bir kez internet bankacılığı giriş işlemi yapmıştır. Bu miktar, toplam kayıtlı bireysel müşteri sayısının % 41'ini oluşturmaktadır. Nisan-Haziran 2013 döneminde, aktif bireysel müşteri sayısında bir önceki yılın aynı dönemine göre 1 milyon 864 bin kişi, bir önceki üç aylık döneme göre ise 222 bin kişi artış olmuştur. Bu dağılım incelendiği takdirde, aktif müşteri sayısında Haziran 2012'den Haziran 2013'e kadar, 1 yıllık süre içerisinde %21.66 oranlık bir artış görülmektedir.

İnternet bankacılığı yapmak üzere sistemde kayıtlı olan ve en az bir kez giriş işlemi yapmış toplam kurumsal müşteri sayısı, Haziran 2013 itibariyle, 2 milyon 189 bin kişi olmuştur. Son bir yıl içerisinde en az bir kez giriş işlemi yapmış toplam kurumsal müşteri sayısı ise yaklaşık 1 milyon 197 bin kişidir.

İnternet bankacılığı yapmak üzere sistemde kayıtlı olan ve en az bir kez giriş işlemi yapmış kurumsal müşteri sayısı Haziran 2013 itibariyle, 2 milyon 189 bin kişidir. Bunların 980 bini (% 45'i) Nisan-Haziran 2013 dönemi içerisinde en az bir kez giriş işlemi yapmıştır. Son bir yıl içerisinde en az bir kez giriş işlemi yapmış kurumsal müşteri sayısı ise 1 milyon 197 bin kişidir. Haziran 2012'de aktif müşteri sayısı 863.263 kişi iken, Haziran 2013'te bu sayısı 979.619 kişiye ulaşmıştır. Dolayısıyla internet bankacılık giriş işlemlerinde kurumsal müşteri sayısında %13.48 oranla bir artış görülmektedir.

İnternet bankacılığı için kayıt yaptıran ve en az bir kez giriş işlemi yapmış toplam (bireysel ve kurumsal) müşterilerin % 42'si Nisan-Haziran 2013 döneminde en az bir kez internet bankacılığı işlemi yapmıştır. Nisan-Haziran 2013 döneminde, toplam aktif müşteri sayısında bir önceki yılın aynı dönemine göre 1 milyon 980 bin kişi, bir önceki üç aylık döneme göre ise 249 bin kişi artış olmuştur. Haziran 2012'de aktif müşteri sayısı 9 milyon 468 kişi iken, Haziran 2013'te bu sayısı 11 milyon 448 kişiye ulaşmıştır. Dolayısıyla internet bankacılık giriş işlemlerinde toplam müşteri sayısında %20.91 oranla bir artış görülmektedir.

Tablo 2.2'de incelendiğinde; Mart-Haziran 2013 döneminde internet bankacılığı ile gerçekleştirilen 11 milyon adet yatırım işleminin hacmi 101.6 milyar TL olmuştur. Ocak-Mart 2013 dönemine göre işlem hacminde 14.2 milyar TL tutarında bir artış gerçekleşmiştir.

Tablo 2.2.İnternet Bankacılığında Yatırım İşlemleri

	Mart 2013		Haziran 2013		Net Değişme		Haziran 2013
Yatırım İşlemleri	**İşlem Adedi (Bin)**	**İşlem Hacmi (Milyon TL)**	**İşlem Adedi (Bin)**	**İşlem Hacmi (Milyon TL)**	**İşlem Adedi (Bin)**	**İşlem Hacmi (Milyon TL)**	**Ortalama İşlem Hacmi (Bin TL)**
Yatırım Fonları	2.334	21.177	2.343	23.138	9	1.961	9.9
Döviz İşlemleri	2.135	15.718	2.292	20.459	158	4.742	8.9
Vadeli hesaplar	686	22.200	695	24.517	9	2.317	35.3
Gerçekleşen hisse senedi işlemleri	3.507	15.520	4.071	19.303	564	3.783	4.7
Repo İşlemleri	80	3.101	87	3.458	8	356	39.6
Tahvil ve bono işl.	54	555	58	870	4	316	15.1
Altın	618	1.681	749	2.619	131	938	3.5
VOB	664	7.441	719	7.240	55	-201	10.1
Toplam	**10.076**	**87.393**	**11.013**	**101.605**	**937**	**14.212**	**9,2**

İkinci çeyrekte gerçekleştirilen yatırım işlemleri hacminde birinci sırayı 24.5 milyar TL tutarı ile (695 bin adet işlem) vadeli hesaplar işlemleri almıştır. Bunu yatırım fonları, döviz işlemleri ve gerçekleşen hisse senedi işlemleri izlemiştir.

En yüksek ortalama işlem hacmi 39.6 bin TL ile repo işlemlerinde gerçekleşmiştir. Vadeli hesaplar ortalama 35.3 bin TL'lik işlem hacmi ile ikinci sırayı almıştır. Toplam yatırım işlemleri için ortalama işlem hacmi 9.2 bin TL olmuştur.

Tablo 2.3. incelendiğinde; Mart-Haziran 2013 dönemi itibariyle, internet bankacılığı hizmeti kullanılarak yapılan yatırım işlemleri dışındaki finansal işlemlerin toplam adedi 110 milyon, tutarı ise 431.6 milyar TL olmuştur. EFT, havale ve döviz transferi işlemlerini kapsayan para transferleri işlemleri, yatırım işlemleri dışındaki finansal işlem hacminin yüzde 85'ini oluşturmuştur.

Tablo 2.3. İnternet Bankacılığında Finansal İşlemler

	Haziran 2012		Mart 2013		Haziran 2013	
	İşlem Adedi (Bin)	İşlem Hacmi (Milyon TL)	İşlem Adedi (Bin)	İşlem Hacmi (Milyon TL)	İşlem Adedi (Bin)	İşlem Hacmi (Milyon TL)
Para transferleri	51.095	297.076	54.836	319.925	58.684	367.380
Ödemeler	33.448	14.284	38.204	19.509	37.769	22.547
Kredi kartı işlemleri	9.987	9.383	10.741	10.515	11.084	11.419
Diğer finansal işlemleri	2.695	29.602	3.134	28.415	2.721	30.209
Toplam	**97.225**	**350.344**	**106.914**	**378.364**	**110.258**	**431.555**

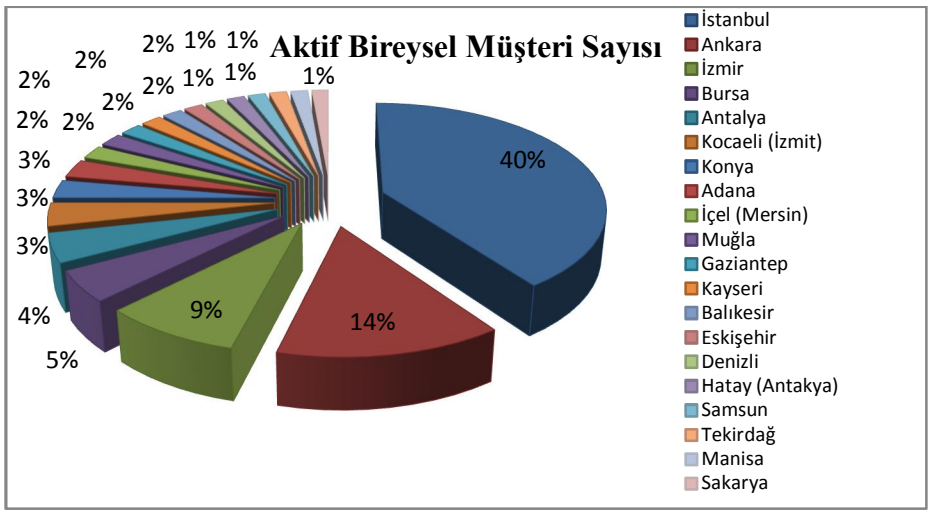
İnternet bankacılığı hizmeti kullanılarak yapılan yatırım işlemleri dışındaki finansal işlemlerin toplam işlem adedi bir önceki yılın aynı dönemine göre 13 milyon adet artarken, bir önceki üç aylık döneme göre 3.3 milyon adet artmıştır. Toplam işlem hacmi ise bir önceki yılın aynı dönemine göre 81.2 milyar TL, bir önceki üç aylık döneme göre ise 53.2 milyar TL artmıştır.

Tablo 2.4'te illere göre internet bankacılığını kullanan bireysel ve kurumsal müşteri sayıları verilmiştir. (Tablo 2.4'te ilk 20 il ele alınmıştır)

Tablo 2.4. İllere Göre İnternet Bankacılığı Kullanan Bireysel Müşteri Sayısı

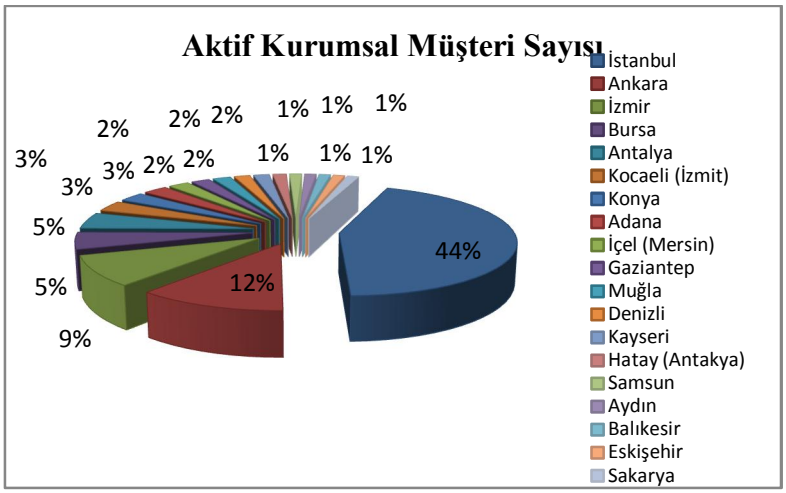
Sıra No	Bölgeler ve İller	Aktif Bireysel Müşteri Sayısı		Sıra No	Bölgeler ve İller	Aktif Kurumsal Müşteri Sayısı
1	İstanbul	3.276.918		1	İstanbul	356.355
2	Ankara	1.125.161		2	Ankara	94.538
3	İzmir	697.291		3	İzmir	71.807
4	Bursa	380.681		4	Bursa	39.956
5	Antalya	355.916		5	Antalya	38.447
6	Kocaeli (İzmit)	284.751		6	Kocaeli (İzmit)	25.341
7	Konya	223.144		7	Konya	22.144
8	Adana	223.021		8	Adana	20.986
9	İçel (Mersin)	164.886		9	İçel (Mersin)	17.838
10	Muğla	159.444		10	Gaziantep	15.924
11	Gaziantep	147.980		11	Muğla	15.482
12	Kayseri	140.077		12	Denizli	13.689
13	Balıkesir	135.449		13	Kayseri	13.660
14	Eskişehir	129.621		14	Hatay (Antakya)	11.775
15	Denizli	126.712		15	Samsun	10.564
16	Hatay (Antakya)	120.499		16	Aydın	10.047
17	Samsun	119.591		17	Balıkesir	9.983
18	Tekirdağ	119.172		18	Eskişehir	9.931
19	Manisa	117.488		19	Sakarya	9.888
20	Sakarya	112.021		20	Manisa	9.859

Tablo 2.4'teki verilere göre iller incelendiğinde; internet bankacılığını kullanan bireysel müşterilerin en fazla olduğu 20 ilin dağılımı, Şekil 2.2'de gösterilmiştir. Şekil 2.2'ye göre bireysel internet bankacılığını en fazla İstanbul, Ankara ve İzmir illerindeki müşteriler kullanmaktadır.



Şekil 2.2. İnternet Bankacılığında Aktif Bireysel Kullanıcıların İllere Göre Dağılımı

İnternet bankacılığı hizmetlerini kullanan aktif bireysel müşterilerin % 40'ı ile aktif kurumsal/tüzel müşterilerin %44'ü İstanbul'dadır. Bunu her iki müşteri grubunda da Ankara ve İzmir illeri takip etmektedir. Kurumsal müşterilerin kendi içerisinde dağılımları il bazlı incelendiğinde takdirde Şekil 2.3'teki dağılım grafiği elde edilmektedir.



Şekil 2.3. İnternet Bankacılığı Kullanan Aktif Kurumsal Kullanıcıların İllere Göre Dağılımı

Şekil 2.3'teki grafikten anlaşılacağı üzere; bireysel müşterilerde olduğu gibi kurumsal müşterilerde de internet bankacılığını kullanan en aktif iller İstanbul, Ankara ve İzmir'dir.

İstatistiklere bakıldığında Türkiye'nin üç büyük ilinin öne çıkmasının sebepleri ise, kullanımla doğru orantılı olarak bu illerdeki nüfus yoğunluğunun yanı sıra, internet kullanımının ve bilgisayar okur-yazarlığının bu illerde daha yaygın olması ile ilişkilidir. Bunun yanı sıra Türkiye'nin üç büyük ilindeki internet bankacılığı kullanıcılarının, bankacılık işlemlerini gerçekleştirmek için harcanan zaman konusunda tasarruf yapmak amacıyla internet bankacılığına yönelmiş oldukları söylenebilir. Yine söz konusu illerdeki internet kullanımının servis sağlayıcılar ve verilen telekomünikasyon hizmetleri ile ilişkisi olduğu da dikkate alınmalıdır.[159]

[159] Türkiye Bankalar Birliği, "İnternet Ve Mobil Bankacılık İstatistikleri Haziran 2013 Raporu",(Çevrimiçi),http://www.tbb.org.tr/tr/banka-ve-sektor-bilgileri/istatistiki-raporlar/Internet_ve_Mobil_Bankacilik_Istatistikleri/1274, 24 Temmuz 2013.

3.BİRLİKTELİK KURALLARI VE KÜMELEME ANALİZİ

3.1. BİRLİKTELİK KURALLARI

Bilgi her alanda gittikçe artmakta ve yaygınlaşmaktadır. Veri tabanlarındaki bilgi miktarı arttıkça birçok kurum ve kuruluş, elinde bulunan bilgiler arasındaki ilişkileri ortaya çıkarmaya çalışmaktadır. İlişki analizi veri tabanındaki bir dizi bilgi ya da kaydın diğer kayıtlar ile olan bağlantısını açıklayan işlemler dizisi olarak ifade edilmektedir. Yığınlar halindeki bilgiler arasındaki ilişkiler; kurumlar için altın değerinde sonuçlar doğurmakla kalmayıp, önemli kararların alınmasında etkileyici olmaktadır.

1990'lı yıllara kadar teknik yetersizliklerden dolayı, kurumlara ve müşterilere satış yapıldığında; o dönemde saklanan satış verilerinde ürün ve müşteri verisi çok nadir yer almaktaydı. Genelde mali açıdan önemli olan tutarsal gelir verilerinin depolaması elektronik ortamda tutularak yapılıyordu. 90'lı yılların başından itibaren ise veri toplama uygulamalarındaki gelişmeler doğrultusunda, artık firmaların satış noktalarında yeni teknoloji otomatik ürün veya müşteri tanıma sistemleri (bar kod ve manyetik kart okuyucular) kullanılmaya başlanmıştır. Böylelikle verilerin satış hareketi oluştuğu anda toplanmasına ve elektronik ortama aktarılmasına olanak sağlanmıştır. Günümüzde süper marketlerde, orta ve büyük ölçekli alışveriş mağazalarındaki satış noktalarında bu akıllı satış sistemlerinin kullanımı oldukça yaygınlaşmıştır.

"Bir mağazada gömlek satın alınan satış hareketlerinin %65'inde kravat ta satın alınmıştır." gibi birliktelik-ilişki örüntüleri ancak, örüntüde yer alan öğelerin birden fazla harekette tekrarlanmasıyla potansiyel olarak mevcut olabilirler. Burada mağaza yöneticileri söz konusu ürünleri yakın raflara yerleştirerek müşterilerin bu tür satın alma eğilimlerini değerlendirebilmekte ve satışların artırılmasını sağlayabilmektedir. Bu satışlardan elde edilen verilerde, işlem tarihi, satın alınan

ürünlere ait bilgiler (ürün kodu, miktar, fiyat, indirim vb.) yer alır ve hareket numarası tekildir.[160]

Herhangi bir ürün satın alırken, bu ürünün yanında başka bir ürün ya da ürünlerin satın alınması, bu ürünler arasındaki bağlantıyı göstermektedir. Bu bağlantıların ortaya çıkartılması ve bunun bir kural olarak ortaya konulması ise birliktelik kuralları yani ilişki analizi olarak ifade edilmektedir. Özellikle satış-pazarlamadan, ürün katalog tasarımlarına kadar birçok alanda, Finans, Tıp ve farklı olayların birbirleriyle ilişki olduğunun belirlenmesinin değerli olduğu ortamlarda yoğun bir şekilde kullanılmaktadır. Pazarlama alanında, perakende sektöründe yaygın olarak kullanılmakta ve literatürde bu tür çalışmalara Pazar sepet analizi adı verilmektedir. [161]

Birliktelik kurallarının matematiksel gösterimi 1993 yılında Agrawal, Imielinski ve Swami tarafından sunulmuştur. Pazar sepet analizi örneğini kullanarak birliktelik kuralının matematiksel modeli açıklanırsa; $I = \{i_1, i_2, \ldots . i_m\}$ kümesine "ürünler" adı verilmektedir.[162]

Birliktelik kuralları aslında olaylar arasındaki probabilistik korelasyonu tanımlamaktadır. Olaylar arasındaki korelasyon ise sık sık beraber gözlenen olaylardır. Herhangi bir veri tabanında birliktelik kurallarının tanımlanması veri tabanı bilgi keşfi sürecinin ilk adımı olarak ifade edilebilir.[163]

D, veri bütünlüğündeki tüm hareketleri, T ise ürünlerin her bir hareketini simgeler. *TID* ise, her harekete ait olan tek belirteçtir.

Birliktelik kuralı şu şekilde tanımlanabilir;

$$\mathbf{A_1, A_2, \ldots\ldots, A_m \Rightarrow B_1, B_2, \ldots\ldots, B_n} \qquad (3.1)$$

[160] Yalçın Özkan; **Veri Madenciliği Yöntemleri**, 1.Baskı, İstanbul, Papatya Yayıncılık, 2008, s. 156.
[161] Silahtaroğlu,**a.g.e.**, s.33.
[162] Agrawal, Imielinski, Swami, **a.g.e.**, s.207-216.
[163] Ezerçe, **a.g.e.**, s. 87.

Bu ifadede yer alan, Ai ve Bj, yapılan iş veya nesnelerdir. Bu kural, genellikle "A_1, A_2,....., A_m" iş veya nesneleri meydana geldiğinde, sık olarak "B_1, B_2, ….., B_n" iş veya nesnelerinin aynı olay veya hareket içinde yer aldığını belirtir.[164]

Ürün(X, "Çay") → (Ürün (Y, "Biskuvi") (3.2)

3.2 ifadesinde *X,* herhangi bir hareketteki değişkeni simgelemektedir. Çay alan bir müşterinin aynı zamanda bisküvi almaya da yöneldiği anlamı çıkarılmaktadır.

Birliktelik kuralı, kullanıcı tarafından minimum değeri belirlenmiş destek ve güvenirlik eşik değerlerini sağlayacak şekilde üretilir. Birliktelik kurallarıyla ilgili problem, birliktelik kurallarını bulmada bir eşik değerinin elde edilmesidir. İlginç birliktelik kurallarından ilginç olmayanları ayırt edebilmek için de bazı ölçütlerin belirlenmesi gereklidir. Bu ölçütler destek ve güven değerleridir. Anlamlılığı destek ve güven değerleri ile ölçülen birliktelik kuralları, "A nesnesini alan bir müşterinin muhtemelen B nesnesini de alması" tipindeki kuralların tanımlanmasını amaçlamaktadır.[165]

Pazar analizinde bunu ifade etmek gerekirse;

Destek: P(A ve B) = A ve B Mallarını Satın Almış Müşteri Sayısı / Toplam Müşteri Sayısı

Güven: P(A/B) = P(A ve B) / P(B) = A ve B Mallarını Satın Almış Müşteri Sayısı / B Malını Satın Almış Müşteri Sayısı

Bir ürün kümesindeki destek, D ile ifade edilen tüm hareketler içinde ilgili ürün kümesini içeren hareketlerin yüzdesidir. A ve B ürün kümelerinin, birliktelik kuralı "A => B" olarak gösterilirse, destek aşağıdaki gibi ifade edilir. Güven değerinin % 100 olması durumunda, kural bütün veri analizlerinde doğrudur ve bu kurallara "kesin" denir.

---

[164] Fevzi Esen **' Veri Tabanlarında Bilgi Keşfi: Veri Madenciliği Ve Bir Sağlık Uygulaması",** Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2009, s.33.

[165] Dolgun, **a.g.e.**, s.29.

Destek kriteri veride bu bağıntının ne kadar sık olduğunu, güven kriteri ise B malını almış bir kişinin hangi olasılıkla A malını alacağını belirtmektedir. İki ürünün satın alınmasındaki bağıntının önemli olması için her iki değerin de olabildiğince yüksek olması gerekmektedir.[166]

Bu durum çok bilinen bir örnek ile şöyle özetlenebilir: "Bira içeren % 30 alışverişin, % 2'si aynı zamanda çocuk bezi de içermektedir." Burada % 30 güven seviyesini, % 2 ise bu güven seviyesine olan desteği ifade etmektedir.[167]

3.1.1. Apriori Algoritması

Literatürde birliktelik kuralı elde edilebilen değişik algoritmalar yer almaktadır. Birliktelik kuralı çıkarım algoritmaları içerisinde en fazla bilinen algoritma Apriori algoritmasıdır.[168]

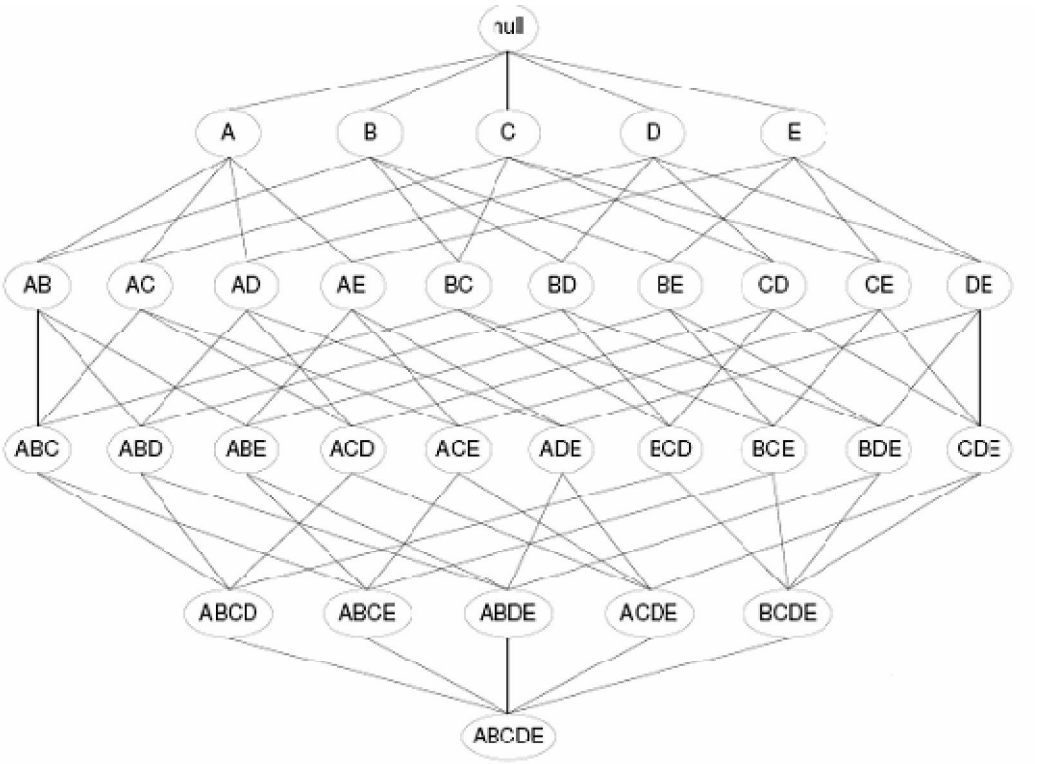
Veri kümesinde yer alan bir öğenin ele alınarak kendisinden sonra gelen ve aday olarak belirlenen öğeler ile karşılaştırılması ve aralarındaki ilişkinin ortaya çıkartılmasını sağlamaktadır. Bunun yanı sıra bu karşılaştırma sonrası bazı aday öğelerin atılarak nihai elde edilmek istenen sonuca ulaşılır.

Büyük veritabanlarında birliktelik kuralları bulunurken, şu iki işlem basamağı ele alınır;

I. adım: Sık tekrarlanan öğeler bulunur: Bu öğelerin her biri en az, önceden belirlenen minimum destek sayısı kadar sık tekrarlanırlar. Bu kümeler sık geçen öğe kümesi olarak adlandırılır. Verilen örneklemde N adet ürün (öğe) söz konusu ise, potansiyel olarak 2^N adet sık geçen öğe kümesi olabilir. Bu adımda üstel arama uzayını etkili biçimde tarayarak sık geçen öğe kümelerini bulan etkili yöntemler kullanılmalıdır.

---

[166] Ethem Apaydın, Zeki Veri Madenciliği :Ham Veriden Altın Bilgiye Ulaşma Yöntemleri", Bilişim 2000 Egitim Semineri, (Çevrimiçi) http://www.cmpe.boun.edu.tr/~ethem, 12.03.2013.
[167] Joshi Karuna Pande, "Analysis of Data Mining Algorithms"(Çevrimiçi). http://ebiquity.umbc.edu/_file_directory_/papers/462.pdf 15.04.2013.
[168] Rakesh Agrawal, Ramakrishnan Srikant, Fast Algorithms for Mining Association Rules, **20. VLDB Konferansı**, Sili, 1994, s.487-499.



Şekil 3.1. 2^N Kümesi Gösterilişi[169]

II. adım: Sık tekrarlanan öğelerden güçlü birliktelik kuralları oluşturulur: Bu kurallar minimum destek ve minimum güven değerlerini elde edilmelidir.[170] Bu adımdaki işlem şöyle gerçekleşmektedir:

Sık geçen her 1 öğe kümesi için, boş olmayan l'in tüm alt kümeleri üretilir. l'in boş olmayan alt kümeleri a ile ifade edilsin. Her a kümesi için a → (l - a) gerektirmesi, l kümesinin destek ölçütünün a kümesinin destek ölçütüne oranı minimum güvenirlik eşiği ölçütünü sağlıyorsa a → (l - a) birliktelik kuralı olarak üretilir. Minimum destek eşiğine göre üretilen çözüm uzayında, minimum güvenirlik eşiğine göre taranarak elde edilen birliktelikler kullanıcının ilgilendiği ve potansiyel olarak önemli bilgileri içeren birliktelikler olarak ifade edilmektedir.[171]

---

[169] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, **Introduction to Data Mining**, Addison Wesley,2005, Data Mining Association Analysis: Basic Concepts and Algorithms Lecture Notes, s.8.
[170] Yasin Sönmez**," Masaüstü Rehberlik Otomasyonu Gerçekleştirimi Ve Verilerinin Birliktelik Kuralı İle Değerlendirilmesi"**, Yüksek Lisans Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, , Elazığ, 2012, s.49.
[171] Agrawal,**a.g.e.** s.12.

Birliktelik sorgusu algoritmalarında performansı belirleyen adım birinci adımdır. 1. Adım gerçekleşip, sık tekrarlanan öğe kümeleri belirlendikten sonra, birliktelik kurallarının bulunması ise daha düz bir adımdır.[172]

Örneğin bir bankanın internet bankacılığında, bankacılık işlemlerini gerçekleştiren müşterilere ilişkin kayıtların tutulduğunu düşünelim. Müşterilerin bir defada yaptıkları tüm işlemler bir satır üzerinde yer almaktadır. Bu veriler kullanılarak, minimum destek oranı % 25 ve güven oranı % 60 olmak üzere Apriori algoritması uygulayalım. Tablo 3.1'de buna ilişkin bir veri seti yer almaktadır.

Tablo 3.1. Birliktelik Kuralı Analizine İlişkin Örnek

Müşteri No	Yapılan İşlemler
1	EFT-Havale İşlemleri, Bakiye İnceleme
2	Fatura Ödemeleri, Döviz İşlemleri, Kredi Kartı İşlemleri, Bireysel Emeklilik İşlemleri,
3	EFT-Havale İşlemleri, Kredi Kartı İşlemleri,Bakiye İnceleme, Repo İşlemleri, Fatura Ödemeleri,
4	EFT-Havale İşlemleri, Fatura Ödemeleri, Kredi Kartı İşlemleri
5	Bakiye İnceleme, Döviz İşlemleri, Hisse Senedi İşlemleri

Öncelikle tüm bankacılık hizmetlerinde yapılan işlemlerin destek oranı hesaplanmıştır. EFT –Havale işlemi için 3 kayıt bulunmaktadır. Toplamda ise 5 kayıt söz konusudur.

- Destek oranı 3/5 = % 60 olarak hesaplanır.

[172] Murat Karabatak, Meryem Koç, 5th International Computer & Instructional Technologies Symposium, 22-24 September 2011, Fırat University, ELAZIĞ- TURKEY, s. 3.

Tablo 3.2. Destek Değerinin Hesaplanması

İşlem	Adet	Destek
EFT-Havale işlemi	3	% 60 (3/5)
Bakiye incelemesi	3	% 60
Fatura ödemesi	3	% 60
Döviz işlemleri	2	% 40
Kredi kartı işlemleri	3	% 60
Bireysel emeklilik işlemleri	1	% 20
Repo işlemleri	1	% 20
Hisse senedir işlemleri	1	% 20
Toplam Kayıt Sayısı	5	-

İkinci aşamada ise yapılan işlemler ikili olarak ele alınır. Tablo 3.3'te destek oranı % 25'in altında olan Bireysel emeklilik işlemleri, Repo işlemleri ve hisse senedi işlemleri bu tabloda yer almaz.

Tablo 3.3. İkili İşlem Gruplarının Destek Değerleri

İşlem	Adet	Destek
EFT-Havale işlemi, Bakiye incelemesi	2	% 40
EFT-Havale işlemi, Fatura ödemesi	2	% 40
EFT-Havale işlemi, Döviz işlemleri	0	% 0
EFT-Havale işlemi, Kredi kartı işlemleri	2	% 40
Bakiye incelemesi, Fatura ödemesi	1	% 20
Bakiye incelemesi, Döviz işlemleri	1	% 20
Bakiye incelemesi, Kredi kartı işlemleri	1	% 20
Fatura ödemesi, Döviz işlemleri	1	% 20
Fatura ödemesi, Kredi kartı işlemleri	1	% 20
Döviz işlemleri, Kredi kartı işlemleri	1	%20
Toplam Kayıt Sayısı	5	

Üçüncü aşamada ise bu işlem grupları üçlü olarak ele alınır. Tablo 3.4'te destek oranı % 25'in altında olan ikili kombinasyonlar bu tabloda yer almaz.

Tablo 3.4. Üçlü İşlem Gruplarının Destek Değerleri

İşlem	Adet	Destek
EFT-Havale işlemi, Bakiye incelemesi, Fatura ödemesi	**1**	**% 20**
EFT-Havale işlemi, Bakiye incelemesi, Kredi kartı işlemleri	**1**	**% 20**
EFT-Havale işlemi, Fatura ödemesi, Kredi kartı işlemleri	**2**	**% 40**

Minimum destek oranı % 25'i geçen birliktelik "EFT-Havale işlemi, Fatura ödemesi, Kredi kartı işlemleri" üçlüsüdür. EFT-Havale işlemi, Fatura ödemesi, Kredi kartı işlemleri işlemi yapanların destek ve güven oranları aşağıda hesaplanmıştır.

- Destek Oranı : 2/5 =% 40
- Güven Oranı için;
- EFT-Havale işlemi ve fatura ödemesi işlemi yapanların; kredi kartı işlemleri yapma olasılığı= 2/2 = % 100'dür.

(Kredi kartı işlemleri-EFT-Havale işlemleri-Fatura ödemesi / EFT-Havale işlemleri-Fatura ödemesi)

- EFT-Havale işlemleri yapanların, fatura ödemesi ve kredi kartı işlemleri yapma olasılığı= 2/3 = % 67'dir.

(Kredi kartı işlemleri-EFT-Havale işlemleri-Fatura ödemesi / EFT-Havale işlemleri)

- Kredi kartı işlemleri yapanların, EFT-Havale işlemi ve fatura ödemesi yapma olasılığı = 2/3 = % 67'dir.

(Kredi kartı işlemleri-EFT-Havale işlemleri-Fatura ödemesi / Kredi Kartı işlemleri)

- Fatura ödemesi yapanların, EFT-Havale işlemler ve kredi kartı işlemleri yapma olasılığı = 2/3 = % 67'dir.

(Kredi kartı işlemleri-EFT-Havale işlemleri-Fatura ödemesi / fatura ödemesi işlemleri)

3.1.2. Birliktelik Kuralları Tekniğinin Güçlü Ve Zayıf Yönleri

Birliktelik Kuralları Analizi, özellikle web madenciliği ve perakende endüstrisinde yaygın kullanılan bir veri madenciliği tekniğidir. Oldukça pratik ve etkin sonuçlar veren bu tekniğin, güçlü yönleri yanında zayıf yönleri de bulunmaktadır.

Birliktelik kuralının güçlü yönleri incelendiğinde;

- Çeşit sayısı fazla ve büyük veri ile çalışma imkanı sağlar.
- Açık ve anlaşılır sonuçlar elde edilebilmektedir.
- Yapılan hesaplamaların anlaşılması kolaydır.

Birliktelik kuralının zayıf yönlerine bakıldığında ise;

- Çok fazla sayıda hesaplama gerektirdiği için problem boyutu artmaktadır.
- Veriler üzerindeki özelliklere sınırlı bir destek sağlar.
- Doğru özellik sayısına karar vermek zordur.
- Seyrek görülen özellikleri göz ardı eder.[173]

3.2. KÜMELEME ANALİZİ

Kümeleme yabancı kaynaklarda kümeleme (clustering) ya da bölümleme (segmentation) olarak adlandırılmaktadır. Sınıflama ile kümelemeyi birbirinden ayıran en önemli fark, kümelemede sınıflandırmanın aksine sınıflar önceden belirli değildir. Verilerin hangi gruplara ve kaç farklı gruba ayrılacağı, eldeki verilerin birbirlerine olan benzerliğine göre yapılabilmektedir.[174] Belirlenen her bir grup küme olarak ifade edilir.

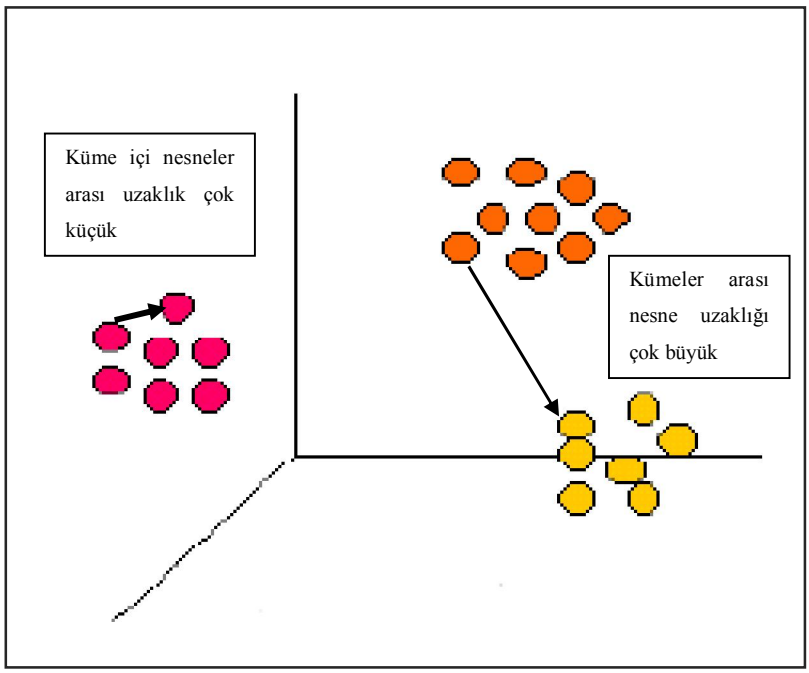
Kümeleme birbirine benzeyen veri parçalarını ayırma işlemidir. Çok değişkenli istatistik tekniklerden biri olan Kümeleme Analizi'nin (Cluster Analysis)

[173] Sujial Nagabhushana, **Data Warehousing OLAP and Data Mining**, New Delhi: New Age International Ltd. Publishers, 2006, s.293-294.
[174] Dunham, **a.g.e**,. s.8.

birincil amacı, gözlem birimlerini benzer özelliklerine göre gruplamaktır. Kümeleme analizinde elde edilen kümelerdeki gözlem birimleri; kendi içlerinde homojendir. Kümelerin kendi aralarında ise, heterojen bir yapıda olması beklenir.[175]

Tanımlamaya yönelik olan kümeleme analizi, tahminlemeye yönelik olarak kullanılmaz. Bu sebeple kümeleme analizi ile bugünün veya geçmişin verilerinin kullanılarak, geleceğin tahminlenmesi mümkün değildir.[176]

Kümeleme işleminde veri tabanındaki benzerliklere ve farklılıklara bakılmaktadır. Böylece kümeler içi homojenlik ve kümeler arası heterojenliğin maksimum düzeyde olması istenmektedir. Yani bir kümeyi oluşturan bireyler birbirleri ile benzeşirken, diğer kümelerin bireyleri ile benzer olmamaktadır.[177]



Şekil 3.2. Küme Yapısı

Şekil 3.2' de görüldüğü gibi küme içindeki nesneler arasındaki benzerlik fazla olduğu için aynı kümede yer almaktadır, ancak kümeler arası benzerlik

[175] Hair,**a.g.e**. s.473.
[176] Gürsoy, **a.g.e,** s.40
[177] Akpınar, **a.g.e**., s.6.

incelendiğinde; bir küme içindeki nesneler ile diğer küme içindeki nesneler arasındaki benzerliğin az olduğu görülmektedir.[178]

Kümeleme analizini kısaca özetlemek gerekirse;

- Her bir küme homojendir veya belirli özellikler bakımından benzer yapıdadır. Bu sebeple her bir kümedeki gözlem birimleri de birbiriyle benzerdir.
- Her bir kümenin aynı özellik açısından diğer kümelerden farklı olması gerekmektedir. Böylece bir kümedeki gözlemler, diğer kümedeki gözlemlerle farklılık oluşturmalıdır.

Kümeleme analizi pazarlama, ekonomi, biyoloji, tıp, psikoloji, eğitim, antropoloji ve telekomünikasyon gibi birçok ve birbirinden çok farklı alanlarda kullanılmaktadır.[179]

Kümeleme analizinde benzerlik veya homojenlik kavramları, yapılan çalışmaların amaçlarına bağlı olarak gelişir. Kavramlar analizden analize farklılık göstermektedir. Amaca göre farklı sayıda ve özellikte kümeler elde edilebilmesi mümkündür. [180]

Kümelemenin matematiksel tanımı ise aşağıdaki gibi yapılabilir:

Elimizde $A = \{ X_1, X_2, X_3, ..., X_n \}$, n = 1, 2, ..., m, veri tabanı olsun. Her bir X_n bir kaydı temsil etsin. $X = \{ x_1, x_2, ..., x_i \}$, i = 1, 2, ..., m her bir x_i , ad, soyad, yaş, medeni durum ve gelir gibi özellikler olsun. Kümelemedeki amaç D veri tabanını, j adet K kümesine bölmek ve $K_j \subseteq D$ koşulunun sağlanmasıdır.[181]

Örnek: Bir banka; tasarrufların değerlendirilmesiyle ilgili olarak bir reklam kampanyası düzenlemek istemektedir. Müşterilerini benzer özelliklerine göre gruplandırmak ve çıkan sonuca göre pazarlama ve reklam kampanyalarını

[178] Han, Kamber, **a.g.e.** s.1-35.
[179] Silahtaroğlu, **a.g.e.**, s.99 .
[180] Subbash Sharma, **Applied Multivariate Techniques**, John Wiley & Sons, USA, 1996, s.185.
[181] Silahtaroğlu, **a.g.e.** s.99.

yönlendirmek istemektedir. Bankada kampanya yönetimi biriminin bu gruplar hakkında önceden tanımlanmış bir fikri bulunmuyor olsun. Bankanın müşterileri hakkında elindeki veriler Tablo 4.5'teki gibidir.

Reklam kampanyasının türüne göre müşteri verilerinin hepsinin önemi yoktur. Burada tasarruflara yönelik olacağı için kampanyada çocuklu müşterilere yoğunlaşmak ve çocukları için yatırım yapan müşterilerin dikkate alınması faydalı olacaktır.

Tablo 3.5. Örnek Müşteri Verileri

Yaş	Çocuk Sayısı	Medeni Hal	Eğitim	Gelir (TL)	Küme Grupları
25	1	Evli	Lisans	1350	1.küme
25	3	Evli	Lise	1200	3. küme
22	0	Bekar	Lisans	2600	2.küme
30	0	Evli	Yüksek Lisans	6500	2.küme
35	2	Bekar	Lisans	2000	1.küme
40	0	Bekar	Lisans	2700	2.küme
45	4	Evli	Yüksek Lisans	12500	1.küme
50	2	Dul-Boşanmış	Lise	5000	3.küme
45	1	Dul-Boşanmış	Lise	1200	3.küme
35	0	Bekar	Lisans	2000	2.küme

Tablodaki kümelemede ilk grup çocukları olan eğitimi seviyesi yüksek olan gruptur. İkinci grup ilk gruba göre benzer bir grup olup çocukları yoktur. Üçüncü grup çocuğu olan lise mezunu olan bir gruptur.

Kümeleme Değişkenleri

Kümeleme analizinde nesnelerin sayısal değerler olması önemlidir. Değişken tipleri; aralık ölçekli değişkenler, ikili (0-1) değişkenler, kategorik değişkenler, oran ölçekli değişkenler, sıralı değişkenler olabilir ya da bunların kombinasyonu olabilir.

İkili (Binary) değişkenlerde değişken 1 veya 0 olmak üzere sadece iki değer alabilir.

Kategorik değişkenlerde ikiden fazla şık alabilir. Örneğin renk değişkeni: mavi, sarı, beyaz gibi şıkları alabilir.

Aralık ölçekli değişkenler lineer ölçeğin sürekli ölçümüdür. Ağırlık, uzunluk, yükseklik gibi örnekler verilebilir.

Sıralama değişkenleri büyüklüğün ne kadar olduğu belli olmadığı halde sıralama yapılabilen kategorik değişkenlerdir. Askeri rütbeler buna örnek verilebilir. (Binbaşı, Yarbay, Albay)

Oran ölçekli değişkenler doğrusal olmayan (nonlineer) ölçeğin pozitif ölçümüdür. Örneğin 2000 TL geliri olan biri, 1000 TL geliri olan birine göre iki kat fazla maaş alıyor denilebilir.[182]

Kümeleme analizinde, sınıflandırmadan farklı olarak belirlenecek kümelerin özellikleri daha öncesinde bilinmemektedir. Burada oluşabilecek küme sayısı da belirli değildir. Pek çok kümeleme algoritması bulunmakta, kullanılacak olan kümeleme algoritmasının seçimi, veri tipine, çalışmaların amacına bağlıdır.

Ortaya koyulacak sonuçların kullanılabilirliğinin artırabilmesi için, algoritmaların bir kısmı ya küme sayısını ya da her bir kümede bulunacak eleman veya bu elemanlar arasındaki minimum-maksimum benzerlik uzaklık ölçütünü kullanıcıdan öğrenmek istemektedir.[183]

3.2.1. Kümelemenin Amacı Ve Uygulama Adımları

Temel amacı gözlem birimlerini benzerliklerine göre, her biri homojen küme oluşturacak şekilde gruplamak olan kümeleme analizinin uygulanmasında şu adımlar izlenir:

---

[182]Gökhan Bilekdemir, "**Veri Madenciliği Tekniklerini Kullanarak Üretim Süresi Tahmini Ve Bir Uygulama**", Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Endüstri İşletmeleri Bilim Dalı, İzmir, 2010, s.17.
[183] Silahtaroğlu,**a.g.e**, s.100.

1. Problem formüle edilir.
2. Bir benzerlik ölçüsü belirlenir.
3. Kullanılacak olan kümeleme tekniğine karar verilir.
4. Seçilen teknik için hangi kümeleme metodunun kullanılacağı belirlenir.
5. Küme sayısına karar verilir
6. Kümeler yorumlanır ve profil tanımlaması yapılır.
7. Kümelerin geçerliliği test edilir.[184]

3.2.2. Kümeleme Analizinde Kullanılan Benzerlik Ve Uzaklık Ölçütleri

Veri tabanındaki veriler kümeler ayrılırken, uzaklık ve benzerlik kavramlarından faydalanılmaktadır. Veri tabanındaki her bir kaydın diğer bir kayıtla olan benzerliği ya da her bir kaydın veri tabanındaki diğer kayıtlardan olan uzaklığı gibi oluşturulan gerçek ve aday kümeler arasındaki mesafe ve benzerliği içermektedir. Uzaklık kavramı iki veri arasındaki mesafeyi ifade etmekte, benzerlik kavramı ise uzaklık kavramının tersine yakınlık kavramını ifade etmektedir.[185]

Aynı küme içerisindeki veriler birbirleriyle karşılaştırıldığında benzer özellikler taşımakta, diğer kümelerdeki verilerden farklı özellik taşımaktadır. Bu benzerlikler verinin nitelik değerlerine dayanarak çıkarılır ve uzaklık ölçütlerinden faydalanılır.

A olarak göstereceğimiz bir veritabanındaki A={ X_1, X_2, X_3, ..., X_n }, n = 1, 2, ..., m, X_m ile X_j arasındaki mesafe; mes(X_m , X_j), X = { x_1, x_2, ..., x_i}, i = 1, 2, ..., m iken E^n Euclid uzayında şu şekilde gösterilmektedir:

$$\text{mes}\,(X_m, X_j) = \sqrt{\sum_{i=1}^{n}(x_{mi} - x_{ji})^2} \qquad (3.3)$$

bu mesafe, Euclid mesafesi olarak ifade edilmektedir.

---

[184] Gürsoy,**a.g.e.** s.60.
[185] Songül Şekeroğlu,"**Hizmet Sektöründe Bir Veri Madenciliği Uygulaması**", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, İstanbul, 2009, s.62.

Benzerlik kavramı, iki veri arasındaki yakınlığı gösterir ve genel olarak 3.4'teki gibi ifade edilmektedir. Genel olarak;

$$ben(X_m, X_j) = \frac{1}{1 + mes(x_m, x_j)}$$ şeklinde ifade edilir. (3.4)

Benzerlik ölçüsü olarak Euclid uzaklığı dışında benzerlik ölçüsü olarak "Ki-Kare" ve " Phi Kare Uzaklık ölçüleri de kullanılmaktadır.

Ölçümler sayımla elde edilmiş değerler ise tercih edilen ölçüler "Ki-Kare" ve "Phi-Kare" Uzaklık Ölçüleri benzerlik ölçülerinden yararlanılır.

Eğer ikili gözlemlere göre veriler elde edilmiş ise birimler arasındaki benzerlikleri tespit etmek için "Euclid, Kare Euclid, Büyüklük Farkı (Size Difference), Yapı Farkı (Pattern Difference), Lance Williams Farkı (Lance and Williams Difference)" gibi benzerlik ölçüleri kullanılmaktadır.[186]

Kümelemede birçok uzaklık bağıntısı kullanılmaktadır. İki nokta arasındaki uzaklığın ölçülmesinde en çok kullanılan yöntem Öklid uzaklığıdır. Diğer yaygın kullanılan uzaklık bağıntılar, Manhattan uzaklığı ve Minkowski uzaklığıdır.

Veri matrisinde n veri sayısı, p nitelik sayısı olmak üzere aşağıdaki gibi bir veri matrisi oluşturulur.

$$\begin{bmatrix} x_{11} & \cdots & x_{1f} & \cdots & x_{1p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{i1} & \cdots & x_{if} & \cdots & x_{ip} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nf} & \cdots & x_{np} \end{bmatrix}$$

Uzaklık matrisi ise genel olarak şu şekilde elde edilir.

[186] Kazım Özdamar, **Paket Programlar İle İstatistiksel Veri Analizi I**, 5. Baskı, Kaan Kitabevi, Eskişehir, 2004, s.285.

$$\begin{bmatrix} 0 & & & & \\ d(2,1) & 0 & & & \\ d(3,1) & d(3,2) & 0 & & \\ \cdots & \cdots & \cdots & & \\ \cdots & \cdots & \cdots & & \\ d(n,1) & d(n,2) & \cdots & \cdots & 0 \end{bmatrix}$$

- **Öklid Uzaklık**

Öklid uzaklık bağıntısı uygulamada en çok kullanılan uzaklıklarından biridir. İki boyutlu uzayda Pisagor teoreminin bir uygulamasıdır. Öklid uzaklık bağıntısı 3.5'te yer almaktadır. İki veri arasındaki uzaklık (d(i,j)) ile ifade edilir.

$$d(i,j) = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \tag{3.5}$$

- **Manhattan Uzaklığı**

Manhattan uzaklığı, gözlemler arasındaki mutlak uzaklıkların toplamı alınarak hesaplanmaktadır. Manhattan uzaklık bağıntısı 3.6'da yer almaktadır.

$$d(i,j) = \sum_{k=1}^{p} \left(\left| x_{ik} \quad x_{jk} \right|^{m} \right) \quad i,j = 1,2....n; k = 1,2.....p \tag{3.6}$$

- **Minkowski Uzaklığı**

Minkowski uzaklığı p sayıda değişken göz önüne alınarak gözlem değerleri arasındaki uzaklık hesaplamasında kullanılmaktadır. Söz konusu Minkowski uzaklık bağıntısı 3.7'de yer almaktadır.

$$d(i,j)=\left[\sum_{k=1}^{p}\left(\left|x_{ik}\quad x_{jk}\right|^{m}\right)\right]^{\frac{1}{m}} i,j=1,2....n;k=1,2.....p \tag{3.7}$$

Burada m=2 yazılarak Öklid uzaklık bağlantısı da elde edilebilir.[187]

3.2.3. Kümeleme Analizinde Kullanılan Yöntemler

Literatürde birçok kümeleme algoritması bulunmaktadır. Algoritmalar birbirlerinden, kümelerin oluşturuluş sekline göre ayrılmaktadır. Kullanılacak olan kümeleme algoritmasının seçimi, veri tipine ve yapılacak çalışmanın amacına bağlıdır.

- **Hiyerarşik Yöntemler (Hierarchical Methods)**
  - Toplaşım (Agglomerative) Kümeleme Algoritmaları
  - Bölünür (Divisive) Kümeleme Algoritmaları
- **Bölümlemeli (Partitioning) Yöntemler(Hiyerarşik Olmayan Yöntemler)**
  - K Metod Yöntemi
  - K-Ortalamalar Yöntemi
  - Olasılıksal Algoritmalar
  - Yer Değiştiren Algoritmalar
  - Yoğunluğa Dayalı Algoritmalar (Density-base methods):
    - Yoğunluğa Dayalı Bağlantılı Kümeleme
    - Yoğunluk Fonksiyonlu Kümeleme
- **Izgara Tabanlı Yöntemler (Grid-Based Methods)**
- **Diğer Yöntemler**
  - Kategorik Verinin Yinelenmesine Dayanan Yöntemler

[187] Şekeroğlu, **a.g.e.**, s.63.

- Kısıtlara Dayanan Yöntemler
- Makina Öğrenmesi Alanında Kullanılan Yöntemler
- Ölçeklenebilir Kümeleme Yöntemleri[188]

3.2.3.1. Hiyerarşik Yöntemler

Hiyerarşik yöntemler, veri nesnelerini kümeler ağacı şeklinde gruplara ayırma esasına dayanır. Hiyerarşik tekniklerin ağaç diyagramları ile gösterilen sonuçlarına dendogram denir. Hiyerarşik kümeleme yöntemleri, hiyerarşik ayrışmanın aşağıdan yukarıya veya yukarıdan aşağıya doğru olmasına göre *toplaşım ve bölünür hiyerarşik* kümeleme yöntemleri olarak sınıflandırılabilir.

Hiyerarşik kümeleme algoritmaları benzerlik ve mesafe ölçütlerini kullandıkları için kullanılması son derece pratiktir. Hiyerarşik kümeleme algoritmaları hemen hemen her türlü veri türüne uygun algoritmalardır.

Toplaşım hiyerarşik kümeleme algoritmalarında; hiyerarşik ayrışma aşağıdan yukarıya doğru gerçekleşir. İlk olarak veri tabanındaki her bir nokta bir küme olarak görülür. Burada birbirine benzeyen kümeler birleşe birleşe birbirinden ayrı kümeler oluşturur. Benzerlik artıkça tüm alt gruplar tek bir kümede birleşme eğilimi göstermektedir.[189]

Toplaşım kümeleme algoritmalarında kullanılan yöntemler aşağıdaki gibidir:

1. Merkezi Kümeleme (Centroid) Yöntemi

2. En Yakın Komşu (Nearest Neighbor) veya Tek Bağlantı (Single-Linkage) Yöntemi

3. En Uzak Komsu (Farthest-neighbor) veya Tam Bağlantı (Complete-Linkage) Yöntemi

4. Ortalama Bağlantı (Average-Linkage) Yöntemi

---

[188] Haberal, **a.g.e.**, s.28.
[189] Ezerçe, **a.g.e.**, s. 98.

5. Ward Yöntemi[190]

3.2.3.2. Bölümlemeli Yöntemler

Bölünür kümeleme yöntemlerinde ise, hiyerarşik ayrışma yukarıdan aşağıya doğru olur. Toplaşım kümeleme algoritmalarında her bir nokta tek bir küme olarak görülürken, bölünür küme algoritmaları başlangıçta tüm noktaları tek bir küme olarak görmektedir. Veri tabanını taradıkça, birbirine benzemeyen noktaları kümenin içinden atarak, öncesinde karar verilmiş k kadar kümeye dağıtır. Hiyerarşik algoritmalar esnek algoritmalar olmalar rağmen, bölünür algoritmalar için k küme sayısının verilmesi dezavantaj oluşturan bir durumdur. Hiyerarşik algoritmaların diğer bir dezavantajı ise bu algoritmaların bir kümeyi oluşturduktan sonra, yapılarından dolayı oluşturulan bu kümeyi kontrol etmemeleridir.

Bölümleme yöntemleri, *n* adet nesneden oluşan veritabanını, giriş parametresi olarak belirlenen *k* adet gruba ($k \leq n$) ayırma işlemi olarak tanımlanabilir. Veritabanındaki her bir nesne bir farklılık fonksiyonuna göre *k* adet gruptan birine dahil edilir. Bu bölümlerden her biri bir küme olarak ifade edilir. Kümeler, bölümleme kriterini en iyi biçimde karşılamak üzere ayarlanmıştır. Bu bölümleme kriterine genellikle benzerlik fonksiyonu adı verilir. Bölümleme metotları k-means, k-medoids ve Clara-Clarans olarak bilinen algoritmaları kullanır.[191]

Bölümlemeli kümeleme algoritmaları, dendogram oluşturması mümkün olmayan büyük veri setleri içeren uygulamalarda etkili olmaktadır. Bölümlemeli algoritmalarla birlikte gelen bir problem ise çıktı olarak istenen kümelerin sayısını oluşturmaktadır. Algoritma birden fazla defa ve değişik giriş parametreleriyle çalıştırılarak en iyi sonuca ulaşılması mümkündür. Bölünlemeli yöntemler, hem uygulanabilirliği kolay hem de verim elde edilebilen bir yöntemdir. Bilinen en iyi ve

[190] Silahtaroğlu,, **a.g.e.**, s. 108.

[191] Cenk Gülce, " **Veri Madenciliğinde Apriori Algoritması Ve Apriori Algoritmasının Farklı Veri Kümelerinde Uygulanması"**, Yüksek Lisans Tezi, Trakya Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Edirne, 2010, s.22.

yaygın olarak kullanılan algoritmalar k-means, k-medoids ve bunların varyasyonlarıdır.[192]

K-Metod yöntemi k-ortalama yönteminden farklı olarak küme merkezi için ortalamayı değil orta noktayı kullanır. Küme merkezi belirlemedeki fark dışında k-ortalama ile aynıdır. Büyük veri setleri için doğru sonuçlar elde edilmesi mümkün olmayan K-metod yöntemi küçük kümelerde çok iyi sonuçlar üretir.[193]

K-Ortalamalar Yöntemi

Bölümlemeli kümeleme yöntemleri arasında en çok bilinen ve en yaygın olarak kullanılan K –Ortalama algoritmasıdır. K-ortalamalar yöntemi, verideki kümeleri bulmaya yarayan en açık ve etkili kümeleme algoritmasıdır. İlk olarak 1967 yılında Mac Queen tarafından ortaya atılan K-ortalama algoritması, k sayıda küme oluşmaktadır. En uygun çözüme ulaşana kadar devam eden, sürekli olarak kümelerin tekrarlandığı döngüsel bir algoritmadır. K- ortalama algoritmasında, her küme içerisinde bulunan verilerin ağırlıklı ortalaması sonucu ortaya çıkan değer ele alınır. Küme içerisinde bu değere en yakın olan nokta değeri küme merkezi (centroid) olarak kabul edilmektedir.[194]

Bu algoritmanın adımlarını aşağıdaki gibi ifade edilebilir:

1. Veri seti rassal olarak k adet başlangıç kümesine ayrılır.

2. Veri setinde yer alan örnekler; merkezi kendisine en yakın olan kümeye atanır.

3. Her atamanın sonunda küme merkezi (ortalama) yeniden hesaplanır.

4. Veri setindeki tüm örneklerin ataması yapılana kadar 2. ve 3. adımlar tekrarlanır.

Bu süreç tüm gözlemler gruplara atanıncaya kadar devam etmektedir. Bir veri noktası hangi kümeye yakın ise o küme de yer almaktadır. Bu işlem sonucunda her

[192] John Wang., "**Encyclopedia of Data Warehousing and Mining**", Idea Group Publishing, ISBN 1-59140-557-2, Hershey PA, 2006.
[193] Ezerçe, **a.g.e**, s.100.
[194] Tezcanlar,**a.g.e**. s.130.

bir veri noktası k adet kümeye yerleştirilmiş olacaktır. Oluşturulmuş yeni kümeler için merkez değeri hesaplanır. Oluşan bu merkez değeri, kümenin yeni merkezi olur. Merkez noktası değiştiği için, veri kümesinin merkez noktalarına uzaklıkları da değişecektir. K-ortalama yönteminde amaç, diğer kümeleme yöntemlerinde olduğu gibi, gerçekleştirilen kümeleme işlemi sonucunda elde edilen kümelerin, küme içi benzerliklerinin maksimum, kümeler arası benzerliklerinin ise minimum olmasıdır.[195]

Bu yöntem çok sayıda birimden elde edilmiş olan sürekli değişkenli veri setlerini küme içi kareler toplamını minimize edecek biçimde k kümeye ayırmayı amaçlamaktadır.

Aşağıda bir örnek üzerinde k-ortalamalar yöntemi incelenmiştir. A,B,C ve D için x_1 *ve* x_2 değişkenlerinin ölçülmesi istenmektedir.

Tablo 3.6. Veri Seti

	Gözlemler	
	X_1	**X_2**
A	**10**	**6**
B	**-2**	**2**
C	**2**	**-4**
D	**-6**	**-4**

Buradaki amaç, gözlemleri k = 2 kümeye ayırmak olsun. Yakın değerlere sahip olan gözlem birimleri aynı kümede yer almalıdır. İlk adımda, k = 2 – ortalama yöntemini uygulayabilmek için keyfi olarak gözlem birimleri iki kümeye ayrılır.

Örneğin (AB) ve (CD) kümeleri için bu kümelerin ortalamaları hesaplanır.

[195] Çalış, **a.g.e.** s.38.

Tablo 3.7. Kümelerin Ortalama Değerleri

	Ortalamalar	
	X_1	X_2
(AB)	**[10 + (-2)] /2 = 4**	**(6+2) /2 = 4**
(CD)	**[2+ (-6)] /2 = -2**	**[-4 + (-4)] /2 = -4**

İkinci adımda her bir gözlem biriminin küme ortalamasına olan Öklid uzaklığı hesaplanır ve gözlem birimi en yakın kümeye atanır. Eğer gözlem biriminin başlangıçta yer aldığı küme değişirse, küme ortalaması tekrar hesaplanır. Aşağıda öklid uzaklıkları hesaplanmıştır;

$$D^2 (A,(AB)) = (10 -4)^2 + (6 -4)^2 = 40$$

$$D^2 (A,(CD)) = (10 +2)^2 + (6+ 4)^2 = 244$$

A gözlem birimi, (AB) kümesine (CD) kümesinden daha yakın olduğu için tekrar atanmaz. İşleme B gözlem birimi için devam edilir.

$$D^2 (B,(AB)) = (-2 – 4)^2 + (2 - 4)^2 = 40$$

$$D^2 (B,(CD)) = (-2+2)^2 + (2 + 4)^2 = 36$$

B gözlem birimi (CD) kümesine (AB) kümesinden daha yakın olduğu için (CD) kümesine atanır. Artık elimizde olan iki küme (A) kümesi ve (BCD) kümesidir. Kümelerin ortalamaları tekrar hesaplanır.

Tablo 3.8. Kümelerin Ortalama Değerleri

	Ortalamalar	
Kümeler	X_1	X_2
(A)	**10**	**6**
(BCD)	**-2**	**-2**

Her gözlem birimi tekrar atama yapılıp yapılmaması durumu için kontrol edilir. Küme ortalamalarının uzaklık kareleri alındığında Tablo 3.9 elde edilecektir.

Tablo 3.9.Küme Ortalamalarının Uzaklık Kareleri

	Küme Ortalamalarının Uzaklık Kareleri			
Kümeler	**A**	**B**	**C**	**D**
(A)	**0**	**160**	**164**	**256**
(BCD)	**208**	**16**	**20**	**20**

Tablodan görüleceği üzere her gözlem birimi, kendisine en yakın ortalama uzaklıkta olan kümeye atanmıştır. Dolayısıyla süreç sona erer. Sonuçta elde edilen k = 2 küme (A) ve (BCD) kümeleridir.

K-ortalamalar algoritması oldukça etkin bir algoritma olmakla birlikte; sadece nümerik veri ile çalışır, ancak veri madenciliği uygulamaları sıklıkla kategorik verileri de içermektedir. K- ortalamalar algoritmasının geliştirilmesi ile elde edilen k modlar algoritması ise kategorik veriler üzerinde çalışabilen bir algoritmadır.[196]

[196] Ay, **a.g.e.** s.28.

3.2.3.3. Yoğunluk Tabanlı Yöntemler

Yoğunluk tabanlı kümeleme yöntemi Anil K. Jains tarafından 1988 yılında yayınlanan "Algorithms For Clustering Data" adlı makalesinde k-boyutlu nokta kümelerini tanımlamak için ortaya atılmıştır.[197] Yoğunluk tabanlı yöntemlerde, nesnelerin doğal dağılımı bir yoğunluk fonksiyonu aracılığı ile belirlenerek bir eşik yoğunluğunu aşan bölgeler küme olarak tanımlanır. Özellikle, düzenli olmayan ya da daha önceden belirlenemeyen şekillere sahip kümelemelerde kullanılmaktadır.

Düzgün şekilli olmayan kümeleri bulma başarısı, gürültü ve istisnalardan etkilenmeme ve tek tarama ile sonuca ulaşma avantajları ile en başarılı kümeleme metotları arasında sayılan yoğunluk tabanlı algoritmalar, DBSCAN, OPTICS, DENCLUE gibi algoritmalarla oluşturulur.[198]

3.2.2.4. Izgara Tabanlı Yöntemler

Büyük veri setlerini incelemek için kullanılan ızgara tabanlı yöntemler, kümeleme metotlarında özellik alanı ızgara şeklini alan dörtgensel hücrelere bölünmüşlerdir. Kullandıkları ızgara yapısından dolayı veritabanındaki nesne sayısından bağımsızdırlar. [199]

En önemli avantajı işlem yükü az olduğu için hızlı ve çabuk sonuca ulaşabilmeleri olan ızgara tabanlı yöntemlerin, kullandıkları kare sayısı arttıkça hesaplama zamanı artacağından performans düşüklüğü yaşanabilir.

Grid merkezli algoritmaların verileri bir kez taraması yeterli olmaktadır. Bu durum özellikle büyük veri tabanlarının işlenmesinde büyük kolaylık sağlamaktadır.

---

[197] Ömer Uçan, **"Dijital Kütüphanelerde Veri Madenciliği Uygulamaları: Akdeniz Üniversitesi Merkez Kütüphanesi Örneği"**, Yüksek Lisans Tezi, Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Antalya, 2010, s.40.
[198] Gülce, **a.g.e**, s.22.
[199] Silahtaoğlu, **a.g.e**, s.129.

Bunun yanında çok farklı geometrik şekillerle kümeleme imkanı yaratan bu yöntemde, her girdi değişkeni için modelleme oluşturulur.

Grid-merkezli kümeleme yöntemlerinde en çok kullanılan algoritmalar STING, Wavecluster ve CLIQUE algoritmalarıdır.[200]

3.2.3.4. Model Tabanlı Yöntemler

Model tabanlı yöntemler eldeki verileri bir matematiksel model ile oluşturan yöntemlerdir. Bu yöntem verilerin belirli bazı olasılık teorilerinin karışımından oluşan bir mantık ile veri uzayına yerleştiklerini farz ederler.[201]

[200] Uçan,**a.g.e.**, s.44.
[201] Gülce,**a.g.e.**, s.24.

4.UYGULAMA

İnternetin her alanda yaygınlaşmasıyla, günümüzde internet bankacılığı, hizmet sektöründe hızla gelişen faaliyetler arasında yer almaya başlamıştır.

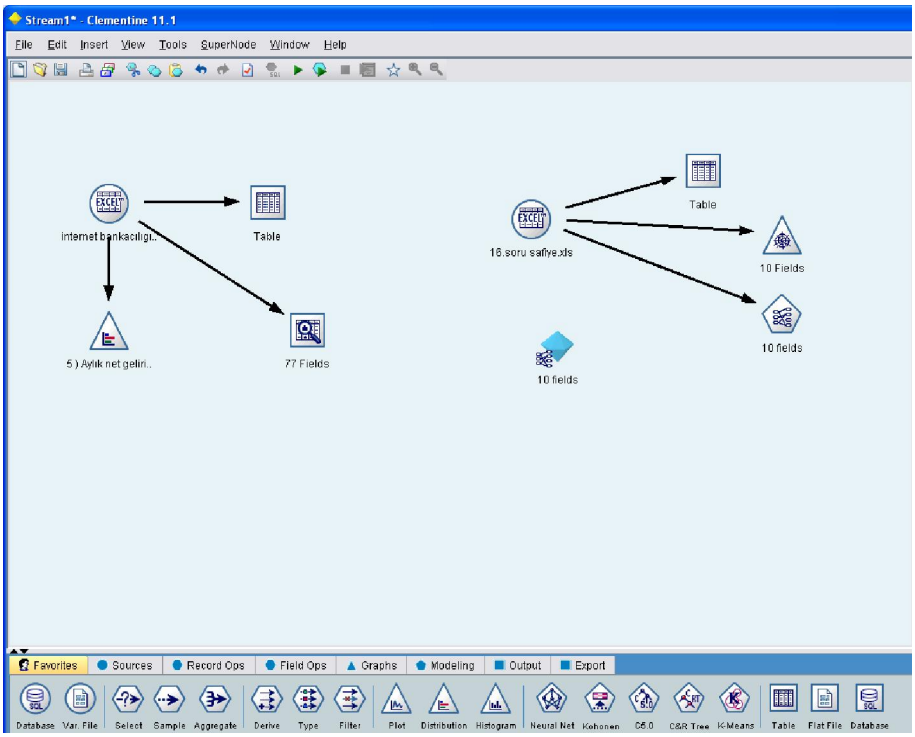
Bu çalışma banka müşterilerinin internet bankacılığına olan yaklaşımlarının veri madenciliği teknikleriyle belirlenmesi amacıyla gerçekleştirilmiştir. Çalışmaya araştırmaya katılan herhangi bir bankada en az bir hesabı bulunan toplam 1110 banka müşterisi dahil edilmiştir. Çalışmada internet bankacılığı kullanan banka müşterilerinin profilleri incelenmiş, hangi internet bankacılığı ürünlerini kullandıkları ve bu ürünleri kullanmalarında etkili olan faktörler araştırılmıştır. Çalışmada ayrıca internet bankacılığı kullanmayan kişilerin profilleri de belirlenmiş ve internet bankacılığı kullanmama nedenleri saptanmıştır.

Uygulamada IBM Modeler SPSS programı kullanılmıştır. Bu program veri madenciliğinde temel programlardan biri olması, veriye erişim aşamasında açık bir çözüm olması, veri büyüklüklerinde bir kısıtlamaya gerek duyulmaması, modelleme konusunda zengin bir içeriğe sahip olması, verinin kalitesini anlama ve verinin görsel olarak incelenmesinin kolay olması nedeniyle tercih edilmiştir.

Kullanıcı tarafından kullanımı kolay olan program, kullanıcıyla dost grafik arayüze sahip, farklı veri setlerini birleştirmeye, veri setleri üzerinde düzenlemeler yapmaya, modellemeye ve analiz etmeye, analizlerin grafiklerden iş akışı oluşturulması yöntemiyle yapılmasına olanak sağlayan bir programdır. Program sonuçların görsel teknikler kullanılarak sunulmasını sağlamaktadır. Bunun yanı sıra program, kendi veri seti hakkında detaylı bilgiye sahip fakat teknik bilgisi olmayan analizciler için de yardımcı olmaktadır.[202]

[202] Ezerçe, **a.g.e**. s.108.

Şekil 4.1'de IBM Modeler programındaki bir arayüz yer almaktadır. Çalışmanın ilerleyen aşamalarında bu sürecin, uygulamada kullanılacak olan veri seti için nasıl işlediği açıklanmıştır. [203]



Şekil 4.1. CRISP-DM Süreci Ara Yüzü

IBM Modeler programının işleyişi, kullanıcı açısından kolay ve pratiktir. Programda yapılması gereken ikonların ana panele sürüklenmesi, bunların oklarla birleştirilmesi ve analiz için modelin kurulmasıdır.

Program; kullanıcılara hızlıca tahminsel veri madenciliği modelleri geliştirmeleri yeteneğini kazandırmakta ve karar verme süreçlerini kuvvetlendirmektedir. Kullanımı kolay, farklı sistem ve veri yapılarına kolayca entegre edilebilir özelliklerinden dolayı IBM Modeler programı tezin uygulama bölümünün gerçekleştirilmesi için seçilmiştir.

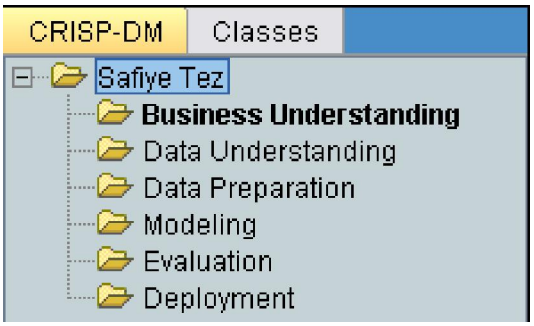
[203] Şimşek, **a.g.e.**, s.140.

Tez çalışmasında daha önce açıklanan CRISP-DM süreci dikkate alınmış ve bu sürece uygun olarak analizler gerçekleştirilmiştir.

4.1.VERİ MADENCİLİĞİ SÜRECİ

Veri madenciliği süreci çalışmanın birinci bölümde detaylı olarak ele alınmıştır. Veri madenciliğinin bir süreç olarak tanımlanabilmesi için, sürecin her bir aşamasının dikkatle izlenmesi gereklidir. Veri madenciliğinde bir aşamanın sonucu, diğer bir aşamanın girdisini oluşturmaktadır. Bu nedenle her aşama bir önceki aşamanın sonuçlarına bağımlı bir şekilde gerçekleşir.

Veri madenciliği için standart bir süreç söz konusudur. Bu standart süreç The Cross-Industry Standard Process for Data Mining (CRISP-DM) konsorsiyumu tarafından belirlenmiştir. Şekil 4.2'de bu sürecin IBM Modeler programındaki arayüzü görülmektedir.



Şekil 4.2. CRISP-DM Süreci Ara Yüzü

Uygulamada iki amaç ele alınmaktadır. Öncelikle internet bankacılığı kullanmayan müşterilerin profillerinin ve internet bankacılığını kullanmama nedenleri incelenmiştir. İkinci aşamada internet bankacılığını kullanan banka

müşterilerinin profilleri incelenmiş, Birliktelik Kuralları ve Kümeleme Analizi yardımıyla internet bankacılığını kullanan banka müşterilerinin hangi bankaların internet bankacılığını kullandıkları, ne tür işlemler gerçekleştirdikleri ve internet bankacılığını kullanmalarında etkili olan faktörler belirlenmiştir.

4.1.1. İnternet Bankacılığı Kullanmayan Müşterileri İçin Veri Madenciliği Süreci

Bu aşamada internet bankacılığı kullanmayan kişilerin profilleri araştırılmış ve internet bankacılığı kullanmama nedenleri incelenmiştir. Bu bölümde internet bankacılığı kullanmayan müşterilere ilişkin, verinin nasıl toplandığı, düzenleme aşamasında hangi işlemlere tabi tutulduğu hakkındaki detaylı bilgiler, yapılan uygulamalar ve varılan sonuçlar yer almaktadır.

4.1.1.1.Problemin Tanımlanması

Veri madenciliği çalışmalarında problemin tanımlanma süreci, ilk ve en önemli aşamadır. Uygulamada iki problem ele alınmıştır. İlk olarak internet bankacılığı kullanmayan müşterilerin profilleri ve internet bankacılığını kullanmama nedenleri belirlenmiştir.

4.1.1.2.Verilerin Hazırlanması

Verinin hazırlanması aşaması; verinin toplanması, verinin birleştirilmesi ve temizlenmesi ile verinin dönüştürülmesi aşamalarından oluşmaktadır. Birinci bölümde verinin hazırlanması aşamasından detaylı bir şekilde bahsedilmiştir. Aşağıda anket formu yardımıyla elde edilen verilerin hangi aşamalardan geçirilerek düzenlendiği açıklanmıştır.

Verinin hazırlanması aşaması veri madenciliği için en önemli aşamalardan biridir, çünkü ham veriden kullanılacak en son veri kümesini oluşturmak için

gerçekleştirilen bütün işlemleri kapsayan bir süreçtir. Verilerin hazırlanması sürecinde ilk adım verilerin toplanmasıdır.

4.1.1.2.1.Verinin Toplanması

Araştırma kapsamında 5 ilden seçilen 1500 bireye anket soruları e-mail ve yüz yüze görüşme yoluyla iletilmiş ve cevaplamaları istenmiştir. Anketlerin yaklaşık % 95'i yüz yüze görüşme yoluyla yapılmıştır.

Çalışmanın kısıtları "zaman ve maliyet" tir. Çalışmada Türkiye çapındaki bütün banka müşterilerine anketin uygulanması zaman ve maliyet kısıtları açısından mümkün olmamaktadır.

Bu nedenle araştırma Türkiye'nin 5 ilinde (İstanbul, Ankara, İzmir, Bursa, Eskişehir) uygulanmıştır. Bu seçimde Türkiye'nin 5 ilinde internet bankacılığı kullanımının yaygın olması, homojen olmayan demografik ve sosyo-kültürel yapı da etkili olmuştur.

Çalışmada veri toplama işlemi anket formuyla yapılmıştır. Anket formu 20 sorudan oluşmaktadır. Anketler kolayda örnekleme yöntemine göre yapılmıştır. Çalışmada 14.02.2013 -14.04.2013 tarihleri arasında Türkiye'nin 5 ilinde bulunan, internet bankacılığı kullanan ve kullanmayan 18 yaş ve üzeri 1500 katılımcıya ulaşılmıştır. Anketler katılımcılara e-mail ve yüz yüze görüşme yoluyla uygulanmış ve veriler elde edilmiştir. Çalışmadaki anket formu Ek 1'de yer almaktadır.

4.1.1.2.2.Verinin Birleştirilmesi ve Temizlenmesi

Çalışmada farklı neden ve sorunlar nedeniyle 247 katılımcıdan anket cevap formları alınamamış olup, çalışma 1253 katılımcıyla gerçekleştirilmiştir. 1253 katılımcıdan toplanan veriler, Microsoft Office Excel modülü ve SPSS 18 versiyonunda tek tek girilmiştir. Verilerin birleştirilmesi ve temizlenmesi aşamasında 143 örnek eksik veya hatalı olarak değerlendirilmeye alınmamış, analiz 1110 veri ile

gerçekleştirilmiştir. Araştırma verilerinin analizinde ise IBM Modeler programlarından yararlanılmıştır.

Çalışmada internet bankacılığı kullanmayan bireylere ilişkin veri setinde 14 değişken ve 100 kayıt söz konusudur. Tablo 4.1'de verilerin tamamı alınamadığı ve yer kısıtı sebebiyle 15 kayda yer verilmiştir.

Tablo 4.1. Veri Seti

Table (14 fields, 100 records)

File Edit Generate

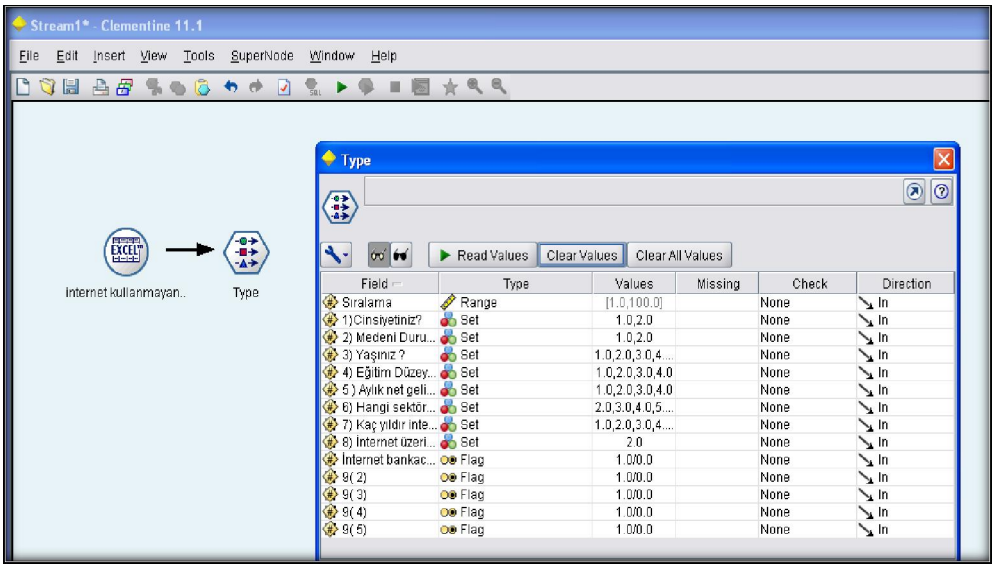
	Sıralama	1)Cinsiyetiniz?	2) Medeni Durumunuz ?	3) Yaşınız ?	4) Eğitim Düzeyiniz?	5) Aylık net geliriniz ?
1	1.000	2.000	2.000	4.000	3.000	3.000
2	2.000	2.000	2.000	4.000	2.000	3.000
3	3.000	2.000	1.000	1.000	3.000	2.000
4	4.000	1.000	2.000	4.000	2.000	3.000
5	5.000	2.000	2.000	3.000	2.000	2.000
6	6.000	1.000	2.000	3.000	2.000	2.000
7	7.000	2.000	2.000	5.000	2.000	1.000
8	8.000	2.000	2.000	5.000	3.000	3.000
9	9.000	2.000	2.000	3.000	2.000	2.000
10	10.000	2.000	1.000	1.000	3.000	2.000
11	11.000	2.000	2.000	2.000	3.000	2.000
12	12.000	2.000	2.000	4.000	2.000	4.000
13	13.000	1.000	1.000	1.000	4.000	3.000
14	14.000	2.000	2.000	3.000	4.000	4.000
15	15.000	1.000	1.000	1.000	4.000	3.000
16	16.000	2.000	1.000	2.000	4.000	3.000
17	17.000	2.000	2.000	3.000	4.000	4.000
18	18.000	1.000	2.000	3.000	2.000	2.000
19	19.000	1.000	1.000	2.000	4.000	3.000
20	20.000	1.000	1.000	1.000	3.000	2.000

Veri setinde 14 değişken bulunmaktadır. Bunlar: Sıralama, Cinsiyet, Medeni durum, Yaş, Eğitim düzeyi, Aylık net gelir, Sektör, İnternet kullanım yılı, İnternet üzerinden verilen bankacılık işlemlerinden yararlanma durumu ve İnternet bankacılığı kullanmama nedenleridir. İnternet bankacılığı kullanmama nedenleri de 5 değişken olarak ayrılmıştır. Bunlar;

- İnternet bankacılığının güvenilir olmadığının düşünülmesi,
- İnternet bankacılığı hakkında bilgi olmaması,
- İnternet bankacılığının kullanımın zor olduğunun düşünülmesi,
- Şube yoluyla bankacılık işlemleri gerçekleştirildiği için internet bankacılığına ihtiyaç duyulmaması,
- Banka müşterilerinin kendilerinin dışında, başka kişiler ilgilendiği için internet bankacılığını kullanmamasıdır.

Tabloda veri tipi alanına bakıldığında program tarafından değişkenler için "Range", "Set" ve "Flag" olmak üzere üç tanımlama yapıldığı görülmektedir. "Range" değişkenin sürekli olduğunu, "Set" değişkenin kategorik olduğunu, "Flat" ise değişkenin iki şıklı değere (Örneğin cinsiyet alanı için kadın/erkek şıklarının bulunması) sahip olduğunu ifade etmektedir. Bu tanımlamalar Tablo 4.2'de görülmektedir:

Tablo 4.2. IBM Modeler Programına Ait Bir Arayüz ve Değişkenler



Değişkenler tanımlanıp, veri temizleme ve dönüştürme işlemleri gerçekleştirilmiştir.

4.1.1.2.3.Verinin Dönüştürülmesi

Tablo 4.3'te anketteki bazı değişkenlerin kodlamaları yer almaktadır.

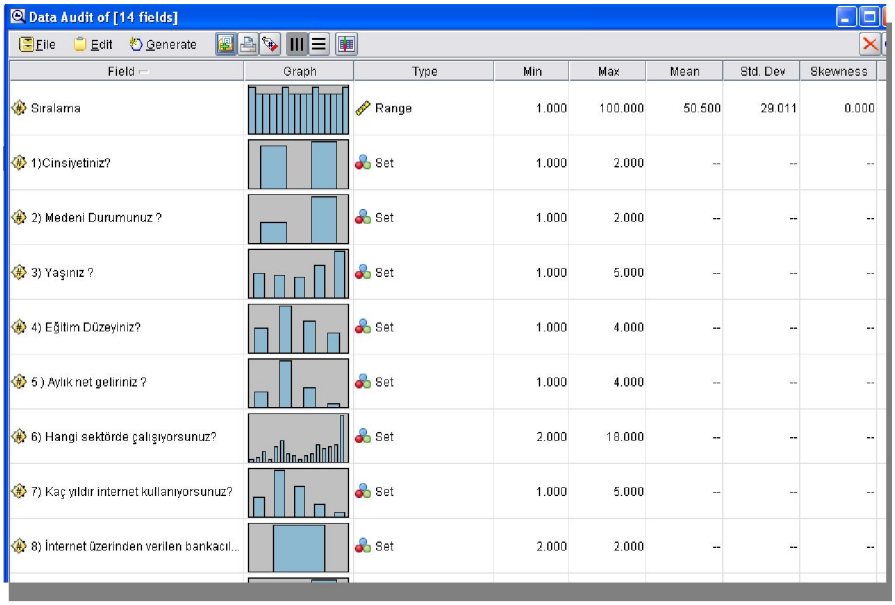
Tablo 4.3. Değişkenlere İlişkin Bazı Kodlamalar

Cinsiyet				
Kadın (1)			Erkek (2)	
Medeni Durum				
Bekar (1)			Evli (2)	
Yaş				
18-25 Yaş (1)	26-35 Yaş (2)	36-45 Yaş (3)	46-55 Yaş (4)	56 Yaş Ve Üzeri (5)
Eğitim Düzeyi				
İlkokul (1)	Ortaokul (2)	Lise (3)	Üniversite(4)	Yüksek Lisans Ve Üzeri (5)
Aylık Net Gelir				
0-750 TL'den az (1)	751-1500 TL (2)	1501-2500 TL (3)	2500 TL ve üzeri (4)	
Sektör				
Akademik (1)	Bilişim (2)	Eğitim (3)	Finans (4)	Otomotiv (5)
İnşaat (6)	Tekstil (7)	Gıda (8)	Sağlık (9)	Lojistik (10)
Perakende (11)	Turizm (12)	Kimya (13)	Öğrenciyim (14)	Çalışmıyorum(15)
Temizlik (16)	Hizmet (17)	Diğer (18)		

Değişkenler tanımlanıp, veri temizleme ve dönüştürme işlemleri gerçekleştirildikten sonra değişkenlere ait dağılımlar elde edilmiştir..

Analizin gerçekleştirilmesi için SPSS Clementine programı Data Audit nodu ile özet tablo elde edilmiştir. Tabloda değişkenlere ait grafikler, değişkenlerin tipleri, değişkenlerin minimum ve maksimum değerleri, ortalamaları, standart sapmaları ve çarpıklık değerleri görülmektedir. Özet Tablo 4.4'te yer almaktadır.

Tablo 4.4. Banka Müşterilerinin Demografik Özellikleri İçin Hazırlanan Veri Setinin Değişkenlerinin Dağılımı ve Temel İstatistikleri



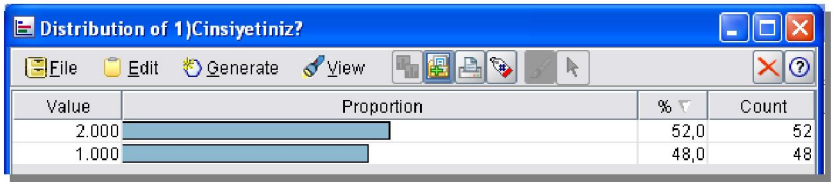
-]

(internet

bankacı ıları yer

almakta

] n

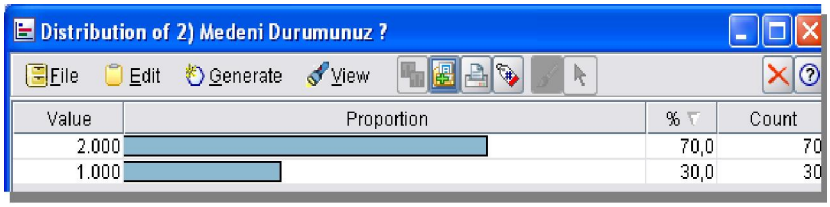


İnternet bankacılığı kullanmayan banka müşterilerinin %52'si erkek ve %48'i kadındır. Dağılımlara bakıldığında; internet bankacılığı kullanmayan kadın ve erkek

banka müşterilerinin dağılımlarının birbirine yakın olduğu, internet bankacılığı tercih edilmemesinde cinsiyete göre çok büyük farklılıklar bulunmadığı görülmektedir.

Tablo 4.6'da internet bankacılığı kullanmayan banka müşterilerinin medeni durumlarına göre dağılımları yer almaktadır.

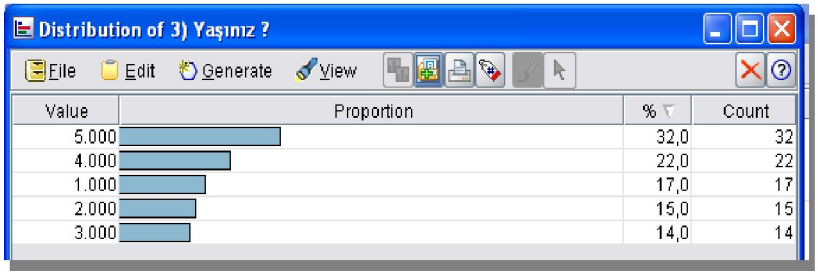
Tablo 4.6. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Medeni Durumlarına İlişkin Dağılımları



İ vli ve %30'u
bekardı müşterilerinin
dağılım lerinin bekar
katılımc ülmektedir.

inin yaşlarına
göre da

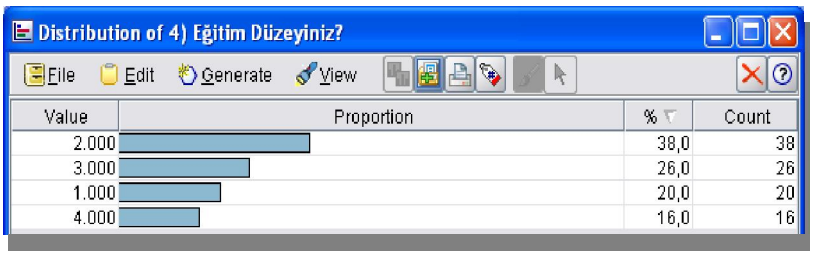
Tablo n Yaşlarına



Ankete katılan internet bankacılığı kullanmayan banka müşterilerinin %32'sinin 56 yaş ve üzeri bireylerden oluştuğu görülmektedir. Banka müşterilerinin %22'si 46-55 yaşları arasında; %17'si 18-25 yaşları arasında, %15'i 26-35 yaşları arasında, %14'ü 36-45 yaşları arasındadır. Dağılım incelendiğinde; internet bankacılığı kullanmayan banka müşterilerinin ağırlıklı olarak 56 yaş ve üzeri bireylerden oluştuğu, 46 yaş ve üzerinde banka müşterilerinin internet bankacılığını çok fazla tercih etmediği görülmektedir.

Tablo 4.8'de internet bankacılığı kullanmayan banka müşterilerinin eğitim düzeylerine göre dağılımları yer almaktadır.

Tablo 4.8. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Eğitim Düzeylerine İlişkin Dağılımları



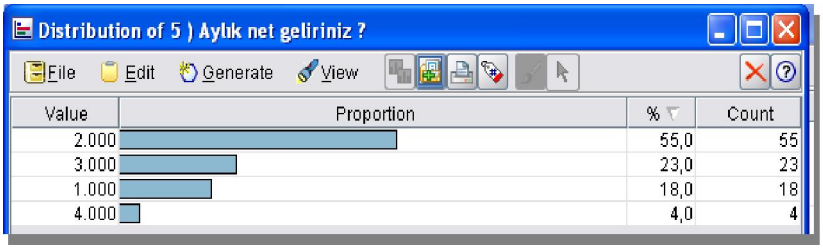
Value	Proportion	%	Count
2.000		38,0	38
3.000		26,0	26
1.000		20,0	20
4.000		16,0	16

Ankete katılan ve internet bankacılığı kullanmayan banka müşterilerinin %38'i ortaokul mezunu, %26'sı lise; %20'si ilkokul, %16'sı ise üniversite mezunudur. Banka müşterilerinin eğitim düzeylerine göre, internet bankacılığını kullanımları arasında farklılıklar bulunmaktadır. Özellikle ortaokul mezunu banka müşterilerinin internet bankacılığını çok fazla tercih etmediği görülmektedir.

.

Tablo 4.9'da internet bankacılığı kullanmayan banka müşterilerinin aylık net gelirlerine göre dağılımları yer almaktadır.

Tablo 4.9. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin Aylık Net Gelirlerine İlişkin Dağılımları

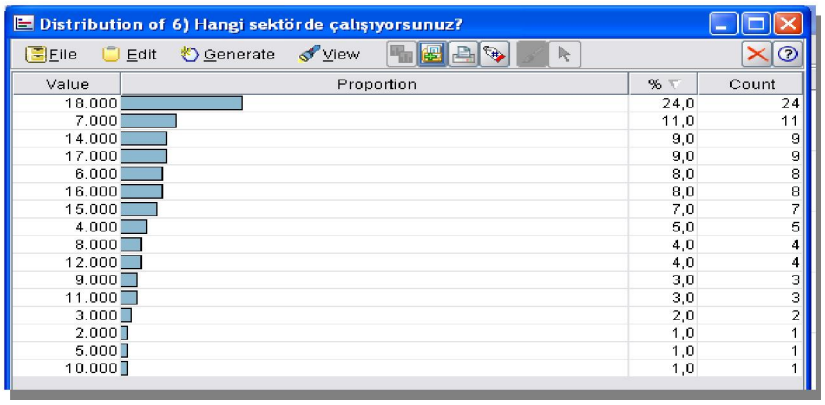


Value	Proportion	%	Count
2.000		55,0	55
3.000		23,0	23
1.000		18,0	18
4.000		4,0	4

erinin %55'i
750- 15 4'ü ise 2500
TL ve inin internet
bankacı

müşterilerinin
çalıştıkl

T erinin



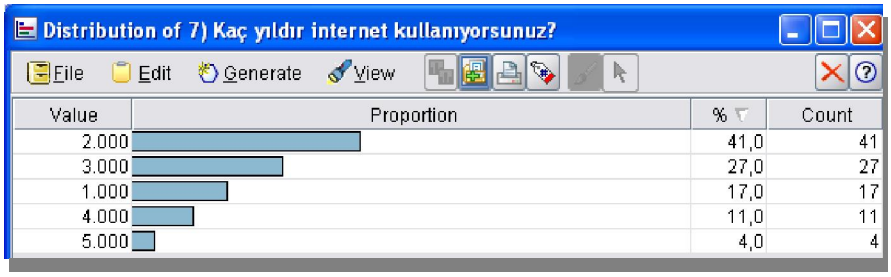
Value	Proportion	%	Count
18.000		24,0	24
7.000		11,0	11
14.000		9,0	9
17.000		9,0	9
6.000		8,0	8
16.000		8,0	8
15.000		7,0	7
4.000		5,0	5
8.000		4,0	4
12.000		4,0	4
9.000		3,0	3
11.000		3,0	3
3.000		2,0	2
2.000		1,0	1
5.000		1,0	1
10.000		1,0	1

Ankete katılan internet bankacılığı kullanmayan banka müşterilerinin çalıştıkları sektörlere göre dağılımları incelendiğinde; banka müşterilerinin % 24'ünün emekli olduğu, % 11'inin tekstil, % 9'unun öğrenci, % 8'inin hizmet, % 8'inin inşaat, % 8'inin temizlik, % 7'sinin çalışmadığı, % 5'inin finans, % 4'ünün gıda, % 4'ünün turizm, % 3'ünün sağlık, % 3'ünün perakende, % 2'sinin eğitim, % 1'inin bilişim, % 1'inin otomotiv, % 1'inin lojistik sektöründe çalıştığı belirlenmiştir.

Dağılımlar incelendiğinde; her sektörde internet bankacılığı kullanmayan bireylerin olduğu, bunun yanı sıra özellikle emekli banka müşterilerinin internet bankacılığını çok fazla tercih etmediği görülmektedir.

Tablo 4.11'de internet bankacılığı kullanmayan banka müşterilerinin kaç yıldır internet kullandıklarına ilişkin dağılımları yer almaktadır.

Tablo 4.11. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin İnternet Kullanım Yıllarına İlişkin Dağılımları



Value	Proportion	%	Count
2.000		41,0	41
3.000		27,0	27
1.000		17,0	17
4.000		11,0	11
5.000		4,0	4

Ankete katılan internet bankacılığı kullanmayan banka müşterilerinin kaç yıldır internet kullandıklarına ilişkin dağılımları incelendiğinde; banka müşterilerinin % 41'inin 1-4 yıl arasında internet kullandığı, % 27'sinin 5-9 yıl, % 17'sinin 1 yıldan az bir süredir, % 11'inin 10-14 yıl, % 4'ünün ise 15 yıl ve üstü süredir internet kullandığı belirlenmiştir. Bu durum bize; internet bankacılığı kullanmayan banka müşterilerinin, çok uzun süredir internet kullanmadığı, internet kullanmaya yeni başladığı, ortalama 1-4 yıl arasında internet kullandığını göstermektedir.

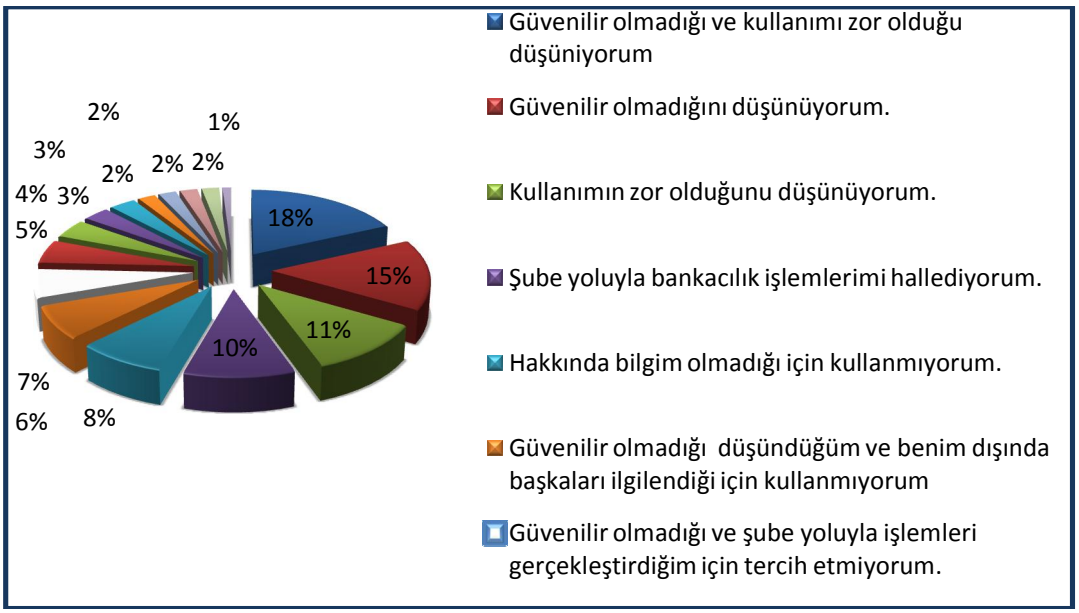
Çalışmada ankete katılan banka müşterilerine, internet üzerinden verilen bankacılık işlemlerinden yararlanıp yararlanmadıkları sorulmuştur. Araştırmaya katılan 100 banka müşterisi bu soruya hayır yanıtını vermiştir. Anketteki 9. soru katılımcıların internet bankacılığını kullanmama nedenlerine ilişkindir. Katılımcıların 8.soruya yanıtları hayır ise sadece 9. Soruyu cevaplamaları istenmiştir.

Tablo 4.12'de internet bankacılığı kullanmayan banka müşterilerinin, bu hizmeti kullanmama nedenlerine ilişkin verdikleri cevapların dağılımları yer almaktadır.

Tablo 4.12. Banka Müşterilerinin İnternet Bankacılığını Kullanmama Nedenlerine İlişkin Verdikleri Cevapların Dağılımı

İnternet Bankacılığını Kullanmama Nedenleri	Yüzde
Güvenilir olmadığı ve kullanımının zor olduğu düşünüyorum.	18
Güvenilir olmadığını düşünüyorum.	15
Kullanımın zor olduğunu düşünüyorum.	11
Şube yoluyla bankacılık işlemlerimi gerçekleştiriyorum.	10
Hakkında bilgim olmadığı için kullanmıyorum.	8
Güvenilir olmadığını düşündüğüm ve benim dışında başkaları ilgilendiği için kullanmıyorum.	7
Güvenilir olmadığı ve şube yoluyla işlemleri gerçekleştirdiğim için tercih etmiyorum.	6
Benim dışımda başkaları (eş,dost,akraba) ilgilendiği için kullanmıyorum.	5
Kullanımı zor ve şube yoluyla bankacılık işlemlerini gerçekleştirdiğim için kullanmıyorum.	4
Güvenilir olmadığı, şube yoluyla işlemleri hallettiğim ve benim yerime başkaları ilgilendiği için kullanmıyorum.	3
Kullanımı zor olduğu, benim yerine başkaları ilgilendiği için kullanmıyorum.	3
Güvenilir olmadığı hem kullanımı zor ve benim yerine başkaları ilgilendiği için tercih etmiyorum.	2
Kullanımı zor ve şube yoluyla işlemlerimi gerçekleştirdiğim için kullanmıyorum.	2
Şube yoluyla işlemlerimi gerçekleştirdiğim ve benim yerime başkaları ilgilendiği için kullanmıyorum.	2
Güvenilir olmadığı, kullanımı zor ve şube yoluyla işlemleri gerçekleştirdiğim için tercih etmiyorum.	2
Güvenilir olmadığı, bilgim olmadığı ve kullanımı zor olduğu için tercih etmiyorum.	1
Güvenilir olmadığı, bilgim olmadığı ve şube yoluyla işlerimi hallettiğim için kullanmıyorum.	1
Toplam	100

Şekil 4.3'te internet bankacılığı kullanmayan banka müşterilerinin, bu hizmeti kullanmama nedenlerine göre dağılım grafiği yer almaktadır.



Şekil 4.3. İnternet Bankacılığı Kullanmayan Banka Müşterilerinin İnternet Kullanmama Nedenlerine Göre Dağılım Grafiği

İnternet bankacılığı kullanmayan banka müşterilerinin, internet bankacılığını kullanmama nedenleri incelendiğinde; banka müşterilerinin % 18'i bu hizmetin güvenilir olmadığını ve kullanımının zor olduğunu düşündüğünü ifade etmiştir. Banka müşterilerinin %15'i sadece güvenilir olmadığını, % 11'i sadece kullanımının zor olduğunu düşündüğü için, % 10'u şube yoluyla bankacılık işlemlerini gerçekleştirdiği için, % 8'i hakkında bilgisi olmadığı için, %7'si hem güvenilir olmadığını düşündüğü ve kendisi dışında başkaları (eş, dost, akraba vb) ilgilendiği için, % 6'sı güvenilir bulmadığı ve şube yoluyla bankacılık işlemlerini gerçekleştirdiği için internet bankacılığını kullanmadığını ifade etmiştir. Dağılımlar incelendiğinde; banka müşterilerinin ağırlıklı olarak, bu hizmeti güvenilir bulmadığı ve kullanımının zor olduğunu düşündüğü için internet bankacılığı tercih etmediği görülmektedir.

Elde edilen sonuçlara bakıldığında; çoğunluğu 46 yaş ve üzerindeki ortaokul mezunu, gelir düzeyi 751-1500 TL arasında olan emekli kadın ve erkek banka müşterilerinin internet bankacılığını kullanmadıkları sonucuna ulaşılmıştır.

Çalışmada banka müşterilerinin internet bankacılığını neden tercih etmediklerine ilişkin sebepler incelenmiştir. Dağılımlara bakıldığında; müşterilerin büyük çoğunluğu internet bankacılığını güvenilir bulmamakta, internet bankacılığını kullanmanın zor olduğunu düşünmektedir.

Uygulamada ikinci aşamada internet bankacılığını kullanan banka müşterilerinin profilleri incelenmiş, Birliktelik Kuralları ve Kümeleme Analizi yardımıyla internet bankacılığını kullanan banka müşterilerinin hangi bankaların internet bankacılığını kullandıkları, ne tür işlemler gerçekleştirdikleri ve internet bankacılığını kullanmalarında etkili olan faktörler belirlenmiştir.

4.1.2. İnternet Bankacılığı Kullanan Banka Müşterileri İçin Veri Madenciliği Süreci

Bu bölümde internet bankacılığı kullanan banka müşterilerine ilişkin, verinin nasıl toplandığı, düzenleme aşamasında hangi işlemlere tabi tutulduğu hakkındaki detaylı bilgiler, yapılan uygulamalar ve varılan sonuçlar yer almaktadır.

4.1.2.1.Problemin Tanımlanması

Çalışmada internet bankacılığını kullanan banka müşterilerinin profilleri incelenmiştir. Çalışmada ayrıca banka müşterilerinin hangi bankaların internet bankacılığını kullandıkları, ne tür işlemler gerçekleştirdikleri ve internet bankacılığını kullanmalarında etkili olan faktörler belirlenmiştir.

4.1.2.2.Verilerin Hazırlanması

Verinin hazırlanması aşaması; verinin toplanması, verinin birleştirilmesi ve temizlenmesi ile verinin dönüştürülmesi aşamalarından oluşmaktadır. Bu bölümde, internet bankacılığını kullanan banka müşterilerinden anket formu yardımıyla elde edilen verilerin hangi aşamalardan geçirilerek düzenlendiği açıklanmıştır.

4.1.2.2.1.Verinin Toplanması

Veri madenciliği için en önemli aşamalardan verinin hazırlanması aşamasında; ankete katılan 1500 banka müşterisinden elde edilen veriler toplanmıştır.

4.1.2.2.2.Verinin Birleştirilmesi ve Temizlenmesi

Çalışmada farklı neden ve sorunlar nedeniyle 247 katılımcıdan anket cevap formları alınamamış olup, analiz 1253 katılımcıyla gerçekleştirilmiştir. 1253 katılımcıdan toplanan veriler, Microsoft Office Excel modülü ve SPSS 18 versiyonunda tek tek girilmiştir. Verilerin birleştirilmesi ve temizlenmesi aşamasında 143 örnek eksik veya hatalı olarak değerlendirilmeye alınmamış olup, analiz 1110 veri ile gerçekleştirilmiştir. Araştırma verilerinin analizinde ise IBM Modeler programlarından yararlanılmıştır.

Çalışmada internet bankacılığı kullanan banka müşterilerine ilişkin veri setinde 78 değişken ve 1010 kayıt söz konusudur. Tablo 4.13'te verilerin tamamı alınamadığı ve yer kısıtı sebebiyle 13 kayda yer verilmiştir.

Tablo 4.13. İnternet Bankacılığı İnternet Bankacılığını Kullanan Banka Müşterilerine İlişkin Veri Seti

Table (78 fields, 1.010 records)

File Edit Generate

	sıralama	1)Cinsiyetiniz?	2) Medeni Durumunuz ?	3) Yaşınız ?	4) Eğitim Düzeyiniz?	5) Aylık net geliriniz ?	6) Hangi sektörde çalışıyorsunuz?
1	1.000	1.000	1.000	2.000	4.000	3.000	4.000
2	2.000	2.000	2.000	2.000	4.000	3.000	7.000
3	3.000	2.000	2.000	2.000	4.000	3.000	8.000
4	4.000	2.000	2.000	2.000	4.000	3.000	7.000
5	5.000	2.000	1.000	2.000	4.000	4.000	8.000
6	6.000	2.000	1.000	1.000	4.000	3.000	9.000
7	7.000	1.000	1.000	1.000	4.000	3.000	4.000
8	8.000	1.000	1.000	1.000	4.000	3.000	4.000
9	9.000	1.000	1.000	1.000	4.000	2.000	14.000
10	10.000	1.000	1.000	1.000	4.000	2.000	14.000
11	11.000	2.000	1.000	1.000	4.000	1.000	14.000
12	12.000	2.000	1.000	1.000	4.000	3.000	8.000
13	13.000	1.000	2.000	3.000	4.000	4.000	4.000

Veri setinde 78 değişken bulunmaktadır: Bu değişkenler; Sıralama, Cinsiyet, Medeni durum, Yaş, Eğitim düzeyi, Aylık net gelir, Sektör, İnternet kullanım yılı, İnternet üzerinden verilen bankacılık işlemlerinden yararlanma durumu, İnternet bankacılığı kullanım yılı, İnternet bankacılığı için hangi sıklıkla internete bağlanıldığı, İnternet bankacılığına nereden bağlanıldığı, Kaç adet internet bankacılığı hesabının olduğu, İnternet bankacılığının geleneksel bankacılığının yerini alıp almamasıdır.

Ayrıca Birliktelik Kuralları Analizinin yapılabilmesi için verinin kategori bazında 0-1 olarak kodlanması gerekmektedir. Çalışmada internet bankacılığını kullanan banka müşterilerinin kullandıkları hesap ve işlemler için "1", kullanılmayan hesap ve işlemler için "0" kullanılmıştır. Bu uygulamanın gerçekleştirilebilmesi için veri transpoze edilerek ilgili kategoriler ikili alanlara dönüştürülmüştür.

Bunların yanı sıra; Hangi bankaların internet bankacılığı hesaplarının kullanıldığı tek bir değişken olarak ele alınmamıştır. Anket formuna göre Tablo 5.14'te yer alan bankaların her biri bir değişken olarak incelenmiştir. Bankaların ayrı ayrı değişken olarak incelenmesinde amaç; banka müşterilerinin hangi bankaların internet bankacılığını en çok kullandıklarının ortaya çıkarılmasıdır. Ele alınan bankalar değişkenleri Tablo 4.14'te görülmektedir.

Tablo 4.14.İnternet Bankacılığı Kullanılan Bankalar

• Akbank	• Yapıkredi	• Bank Asya	• İş Bankası
• Garanti	• Denizbank	• Halk Bankası	• TEB
• İngbank	• Vakıfbank	• Finansbank	• HSBC
• Ziraat Bankası	• Diğer bankalar		

İnternet bankacılığı kullanan banka müşterilerinin bankacılık işlemlerinde en çok hangi hesapları kullandıkları da ayrı ayrı değişken olarak ele alınmış olup, ele alınan değişkenler Tablo 4.15'teki gibidir.

Tablo 4.15. İnternet Bankacılığında Kullanılan Hesaplar

• Vadesiz TL Hesap	• Mevduat fon hesabı
• Vadesiz Döviz Hesabı	• Kredili mevduat hesabı
• Vadeli TL Hesap	• E- Hesap
• Vadeli Döviz Hesabı	• Maaş hesabı
• Likit fon hesabı	• Kredi Kartı Hesabı
• Altın hesabı	• Bireysel Emeklilik Hesabı
• Diğer ürün hesapları	

Çalışmada internet bankacılığı kullanan banka müşterilerinin en çok hangi işlemleri yapmak için internet bankacılığı kullandıkları ayrı ayrı değişken olarak ele alınmış olup ele alınan değişkenlere Tablo 4.16'da yer verilmiştir.

Tablo 4.16. İnternet Bankacılığında Yapılan İşlemler

• Para Transferi Yapma (EFT-Havale işlemleri)	• Döviz İşlemleri
• Hesap Bakiyesi İnceleme Ve Bilgi Alma	• Hisse Senedi İşlemleri
• Kredi Kartı İşlemleri	• Fon Alım Satımı
• Fatura Ödemeleri	• Repo İşlemleri
• Bahis Transferi	• Diğer Ödemeler

Çalışmada internet bankacılığı kullanan banka müşterilerinin, internet bankacılığını kullanmalarında etkili olan faktörler ayrı ayrı değişken olarak ele alınmış olup, ele alınan değişkenler aşağıdaki gibidir:

- 7/ 24 Hızlı ve kesintisiz bankacılık işlemleri,
- Maaşımı o bankadan almam,
- Şubeye gitmeden, sıra beklemeden kolay bankacılık işlemleri,
- Görerek ve seçerek bankacılık işlemi yapabilmek,
- Detaylı rapor ve bilgi alabilmek,
- Çok çeşitli bankacılık ürünlerini görerek bu ürünlerden faydalanabilmek,

- Bankacılık işlemlerini çok daha ucuza yapabilmek,
- Güvenilir olması,
- Pratik olması.

Çalışmada internet bankacılığı kullanan banka müşterilerinin, internet bankacılığı işlemlerinde karşılaştıkları temel sorunlar da ayrı ayrı değişken olarak ele alınmış olup, ele alınan değişkenler aşağıdaki gibidir:

- Günlük işlem limitleri yetersiz,
- Kullanışlı ve pratik değil,
- Web sayfasındaki güvenlik önlemleri yetersiz,
- Sunduğu olanaklar yetersiz,
- Hiçbir sorunla karşılaştığımı düşünmüyorum,
- Diğer

Çalışmada son olarak internet bankacılığı kullanan banka müşterilerinin, bir bankanın internet sitesinde aradıkları özellikler ayrı ayrı değişken olarak ele alınmış olup, ele alınan değişkenler Tablo 4.17'deki gibidir.

Tablo 4.17. İnternet Bankacılığı Kullanılan Bankalarda Aranılan Özellikler

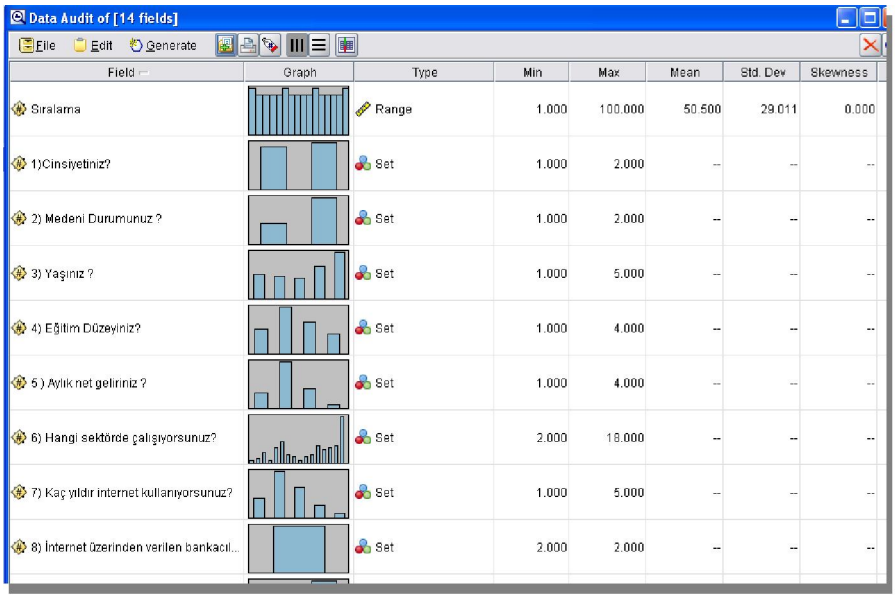
• Hizmet Kalitesi	• Hız
• Bankanın büyüklüğü	• Güvenlik
• Bankanın bilinirliği	• Kolay ve Pratik Kullanımı
• Fatura Ödemeleri	• Tavsiye
• Diğer	

Değişkenler tanımlanıp, veri temizleme ve dönüştürme işlemleri gerçekleştirildikten sonra değişkenlere ait dağılımlar elde edilmiştir.

Analizin gerçekleştirilmesi için Data Audit nodu ile özet tablo elde edilmiştir. Tabloda değişkenlere ait grafikler, değişkenlerin tipleri, değişkenlerin minimum ve

maksimum değerleri, ortalamaları, standart sapmaları, çarpıklık değerleri görülmektedir. Özet Tablo 4.18'de yer almaktadır.

Tablo 4.18. İnternet Bankacılığı Kullanan Banka Müşterilerine İlişkin Özet Tablo

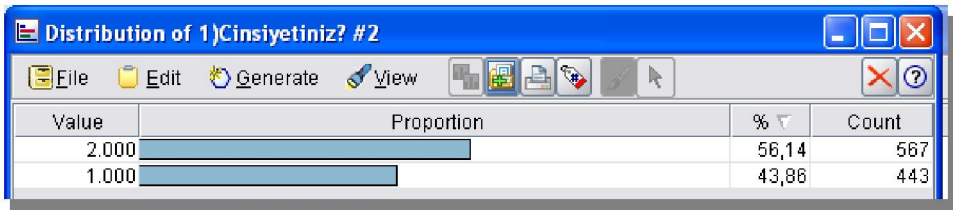


(na

ilişkin c ığı

kullanar

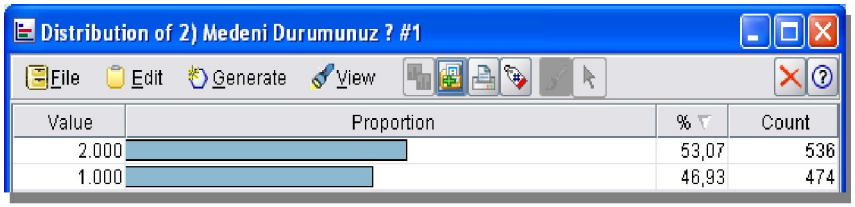
Tablo ıe



İnternet bankacılığı kullanan banka müşterilerinin %56,14'ü erkek ve %43,86'sı kadındır. Bu durum incelendiğinde internet bankacılığı kullanan kadın ve erkek banka müşterilerinin dağılımlarının birbirine yakın olduğu, internet bankacılığı kullanımında cinsiyete göre çok büyük farklılıklar bulunmadığı görülmektedir.

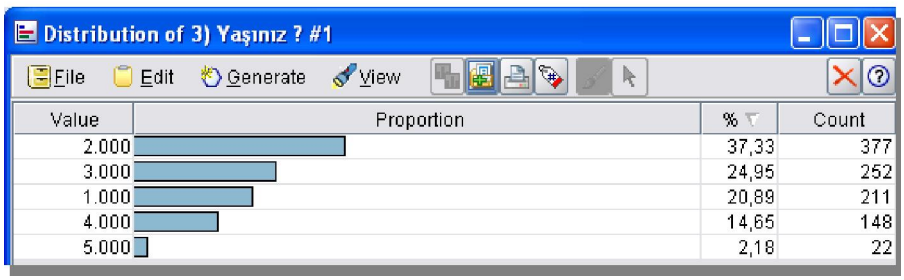
Tablo 4.20'de internet bankacılığı kullanan banka müşterilerinin medeni durumlarına göre dağılımları yer almaktadır.

Tablo 4.20. İnternet Bankacılığı Kullanan Banka Müşterilerinin Medeni Durumlarına İlişkin Dağılımları



İ Evli ve %46.93 duruma göre ço rilerinin internet banka müşteril

Tabl rına

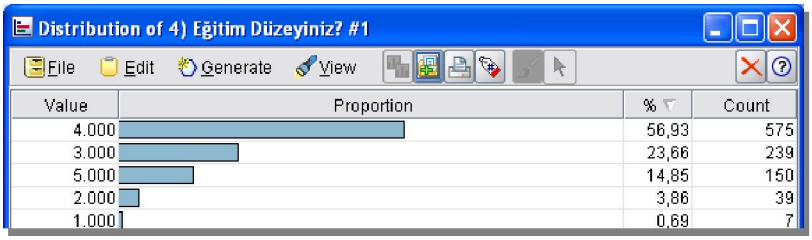


Ankete katılan internet bankacılığı kullanan banka müşterilerinin % 37.33'ü 26-35 yaşları arasında,% 24.95'i 36-45 yaşları arasında; % 20.89'u 18-25 yaşları arasında, % 14.65'i 46-55 yaşları arasında, % 2.18'i ise 56 yaş ve üzeri yaşa sahiptir. Dağılım incelendiğinde; internet bankacılığı kullanan banka müşterilerinin yaşlarına göre internet bankacılığı kullanımları farklılaşmaktadır. 56 yaş ve üzeri banka

müşterilerinin internet bankacılığı çok fazla tercih etmediği, 26-35 yaş grubu banka müşterilerinin ağırlıklı olarak internet bankacılığı kullandıkları görülmektedir.

Tablo 4.22'de internet bankacılığı kullanan banka müşterilerinin eğitim düzeylerine göre dağılımları yer almaktadır

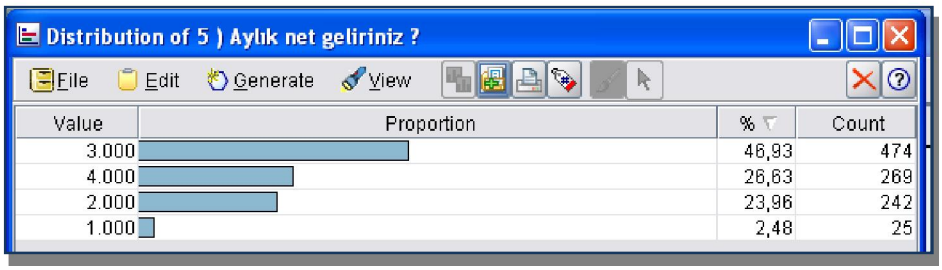
Tablo 4.22. İnternet Bankacılığı Kullanan Banka Müşterilerinin Eğitim Düzeylerine İlişkin Dağılımları



.93'ü
üniversi okul
mezunu anka
müşteri mları
farklılaş ernet
bankacı

ı net
gelir dü

Tablo 4 **Gelir**



Ankete katılan internet bankacılığı kullanan banka müşterilerinin % 46.93'ü 1501- 2500 TL gelire, % 26.63'ü 2500 TL ve üzeri gelire sahiptir. Banka müşterilerinin % 23.96'sı 751-1500 TL arasında gelire, % 2.48'i ise 750 TL'den az gelire sahiptir. Dağılım incelediğinde internet bankacılığı kullanan banka müşterilerinin gelir düzeyinin yüksek olduğu görülmektedir. Gelir düzeyi düşük olan banka müşterilerinin ise internet bankacılığı kullanımın daha az olduğu görülmektedir.

Tablo 4.24'te internet bankacılığı kullanan banka müşterilerinin, çalıştıkları sektörlere göre dağılımları yer almaktadır.

Tablo 4.24. İnternet Bankacılığı Kullanan Banka Müşterilerinin Çalıştıkları Sektörlere İlişkin Dağılımları

Distribution of 6) Hangi sektörde çalışıyorsunuz? #1

File Edit Generate View

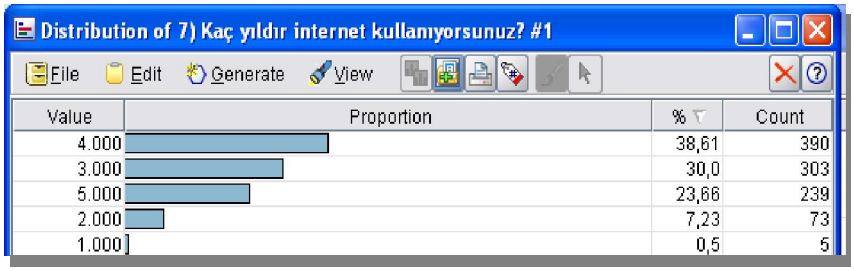
Value	Proportion	%	Count
17.000		9,9	100
4.000		9,21	93
7.000		8,12	82
14.000		8,12	82
9.000		7,92	80
3.000		7,43	75
8.000		7,13	72
11.000		6,63	67
2.000		5,35	54
5.000		5,25	53
6.000		4,75	48
12.000		4,75	48
1.000		4,55	46
13.000		4,55	46
10.000		2,48	25
16.000		2,48	25
18.000		1,09	11
15.000		0,3	3

Ankete katılan internet bankacılığı kullanan banka müşterilerinin çalıştıkları sektörlere göre dağılımları incelendiğinde; banka müşterilerinin % 9.9'unun hizmet, % 9.21'inin finans, % 8.12'sinin tekstil, % 8.12'sinin öğrenci olduğu, % 7.92'sinin sağlık, % 7.43'ünün eğitim, % 7.13'ünün gıda, % 6.63'ünün perakende, % 5.35'inin bilişim, % 5.25'inin otomotiv, % 4.75'inin inşaat, % 4.75'inin turizm, % 4.55'inin akademik, % 4.55'inin kimya, % 2.48'inin lojistik, % 2.48'inin temizlik sektöründe çalıştığı, % 1.09'unun emekli olduğu görülmüş, % 0,3'ü ise çalışmadığını ifade etmiştir. Dağılımlar incelendiğinde; sektör bakımından

hemen hemen her sektördeki banka müşterilerinin internet bankacılığı kullanmakta olduğu görülmektedir.

Tablo 4.25'te internet bankacılığı kullanan banka müşterilerinin kaç yıldır internet kullandıklarına ilişkin dağılımları yer almaktadır.

Tablo 4.25. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Kullanım Yıllarına İlişkin Dağılımları



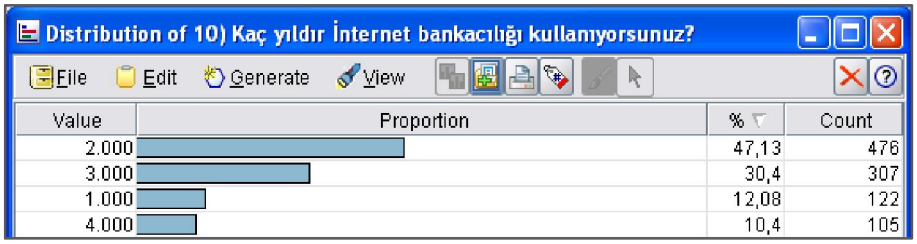
Value	Proportion	%	Count
4.000		38,61	390
3.000		30,0	303
5.000		23,66	239
2.000		7,23	73
1.000		0,5	5

Ankete katılan internet bankacılığı kullanan banka müşterilerinin kaç yıldır internet kullandıklarına ilişkin dağılımları incelendiğinde; banka müşterilerinin % 38.61'inin 10-14 yıldır internet kullandığını, % 30'unun 5-9 yıl, % 23.66'sının 15 yıl ve üzeri, % 7.23'ünün 1-4 yıl, % 0.5' inin ise 1 yıldan az süredir internet kullandığı belirlenmiştir. Bu durum; internet bankacılığı kullanan banka müşterilerinin, internet bankacılığı kullanım yıllarına göre farklılıklar olduğunu göstermektedir. Katılımcıların % 60'ından fazlası uzun süredir internet kullandığını göstermektedir.

Çalışmada ankete katılan banka müşterilerine, internet üzerinden verilen bankacılık işlemlerinden yararlanıp yararlanmadıkları sorulmuştur. Araştırmaya katılan 1010 banka müşterisi bu soruya evet yanıtını vermiştir. Anketteki 9. soru katılımcıların internet bankacılığını kullanmama nedenlerine ilişkindir. 1010 banka müşterisinin internet üzerinden verilen bankacılık hizmetlerinden yararlandıklarını belirtmeleri sebebiyle, ankette 9. Soruyu boş bırakılarak 10. Sorudan devam etmeleri istenmiştir.

Tablo 4.26'da internet bankacılığı kullanan banka müşterilerinin kaç yıldır internet bankacılığı kullandıklarına ilişkin dağılımları yer almaktadır.

Tablo 4.26. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığı Kullanım Yıllarına İlişkin Dağılımları

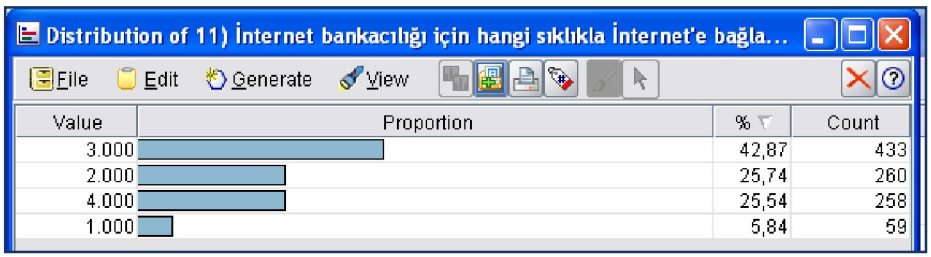


Value	Proportion	%	Count
2.000		47,13	476
3.000		30,4	307
1.000		12,08	122
4.000		10,4	105

Ankete katılan internet bankacılığı kullanan banka müşterilerinin kaç yıldır internet bankacılığı kullandıklarına ilişkin dağılımları incelendiğinde; banka müşterilerinin % 47.13'ünün 1-4 yıldır internet bankacılığını kullandığı, % 30.4' ünün 5-9 yıl, % 12.08'inin 1 yıldan az, % 10.4'ünün ise 10 yıldan uzun süredir internet bankacılığını kullandığı belirlenmiştir. Bu durum; internet bankacılığı kullanan banka müşterilerinin, internet bankacılığı kullanma yıllarına göre aralarında farklılıklar olduğunu, banka müşterilerinin 1-4 yıldır internet bankacılığını kullandığını göstermektedir.

Tablo 4.27'de internet bankacılığı kullanan banka müşterilerinin, internete hangi sıklıklarla bağlandıklarına ilişkin dağılımları yer almaktadır.

Tablo 4.27. İnternet Bankacılığı Kullanan Banka Müşterilerinin Hangi Sıklıkla İnternete Bağlandıklarına İlişkin Dağılımları



Value	Proportion	%	Count
3.000		42,87	433
2.000		25,74	260
4.000		25,54	258
1.000		5,84	59

Ankete katılan internet bankacılığı kullanan banka müşterilerinin, hangi sıklıkla internete bağlandıklarına ilişkin dağılımları incelendiğinde; banka müşterilerinin % 42.87'sinin haftada bir kez internete bağlandıkları, % 25.74'ünün haftada birkaç kez, % 25.54'ünün ayda bir kez, % 5.84' ünün ise günde birkaç kez internete bağlandığı belirlenmiştir. Bu durum; internet bankacılığı kullanan banka

müşterilerinin; ağırlıklı olarak haftada bir kez internete bağlandıklarını göstermektedir.

Çalışmada; internet bankacılığı kullanan banka müşterilerinin, internet bankacılığına erişim noktaları da belirlenmiştir. İnternet bankacılığını kullanan banka müşterilerinin birden çok seçenek işaretlemelerine olanak tanınmıştır. Buna göre;

- İnternet bankacılığı kullanan banka müşterilerinin, % 86.13'ü (878 kullanıcı) internete kişisel PC'den bağlanmaktadır.
- İnternet bankacılığı kullanan banka müşterilerinin, % 5'i (54 kullanıcı) internete şubelerdeki PC 'den bağlanmaktadır.
- İnternet bankacılığı kullanan banka müşterilerinin,% 35.54'ü (359 kullanıcı) cep telefonu ile internete bağlanmaktadır.
- İnternet bankacılığı kullanan banka müşterilerinin, % 40.4'ü (408 kullanıcı) işyerindeki bilgisayarlar ile internete bağlanmaktadır.

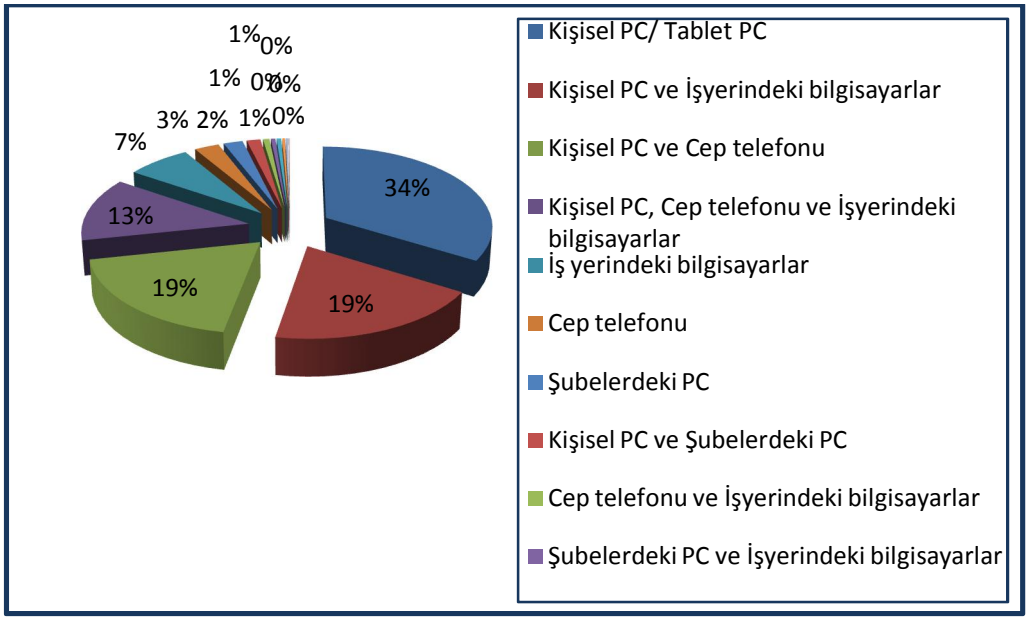
Bu sonuçlara bakıldığında; internet bankacılığı kullanan banka müşterilerinin büyük çoğunluğunun, kişisel PC, işyerindeki PC ve cep telefonlarından internet bankacılığına bağlandığı görülmektedir.

İnternet bankacılığını kullanan banka müşterilerinin internet bankacılığı erişim noktalarına ilişkin verdikleri cevaplara ayrıntılı olarak bakıldığında ise; Tablo 4.28'deki dağılımlara ulaşılmıştır.

Tablo 4.28. İnternet Bankacılığı Kullanan Banka Müşterilerinin, İnternet Bankacılığına Nereden Bağlandıklarına İlişkin Dağılımları

İnternet Bankacılığına Nereden Bağlanıyorsunuz?	Frekans	Yüzde
Kişisel PC/ Tablet PC	341	33,8
Kişisel PC ve İşyerindeki bilgisayarlar	193	19,1
Kişisel PC ve Cep telefonu	190	18,8
Kişisel PC, Cep telefonu ve İşyerindeki bilgisayarlar	129	12,8
İş yerindeki bilgisayarlar	70	6,9
Cep telefonu	27	2,7
Şubelerdeki PC	21	2,1
Kişisel PC ve Şubelerdeki PC	15	1,5
Cep telefonu ve İşyerindeki bilgisayarlar	7	0,7
Şubelerdeki PC ve İşyerindeki bilgisayarlar	5	0,5
Kişisel PC, Şubelerdeki PC, Cep telefonu ve İşyerindeki bilgisayarlar	5	0,5
Kişisel PC, Şubelerdeki PC ve Cep Telefonu	3	0,3
Kişisel PC, Şubelerdeki PC ve İşyerindeki bilgisayarlar	3	0,3
Şubelerdeki PC ve Cep telefonu	1	0,1
Toplam	1010	100,0

Şekil 4.4'te internet bankacılığını kullanan banka müşterilerinin internet bankacılığı erişim noktalarına ilişkin verdikleri cevapların dağılım grafiği yer almaktadır.

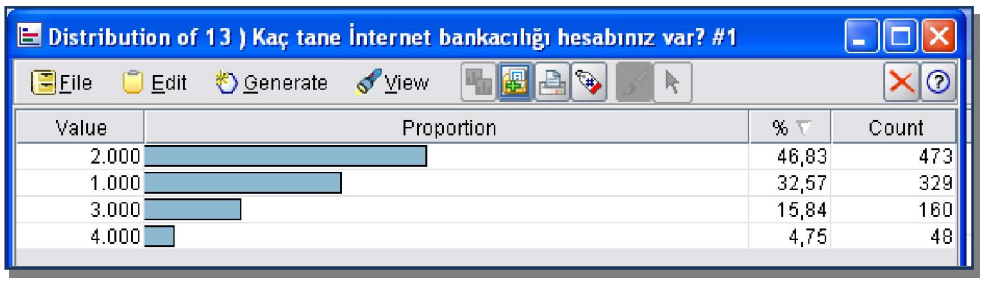


Şekil 4.4. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığına Nereden Bağlandıklarına İlişkin Dağılım Grafiği

Ankete katılan internet bankacılığı kullanan banka müşterilerinin, internet bankacılığına nerede bağlandıklarına ilişkin dağılımları incelendiğinde; banka müşterilerinin % 33,8'inin kişisel PC/ tablet PC aracılığıyla internet bankacılığına bağlandıkları, % 19.1'inin kişisel PC ve işyerindeki PC'lerden, % 18.8'inin kişisel PC ve cep telefonundan, % 12.8'inin ise kişisel PC, işyerindeki PC ve cep telefonundan internet bankacılığına bağlandığı belirlenmiştir. Bu durum; internet bankacılığı kullanan banka müşterilerinin; daha çok cep telefonları ve kişisel PC ile internet bankacılığına bağlanmayı tercih ettiğini göstermektedir.

Çalışmada internet bankacılığı kullanan banka müşterilerinin kaç adet internet bankacılığı hesabına sahip oldukları incelenmiştir. Tablo 4.29'da internet bankacılığı kullanan banka müşterilerinin, kaç adet internet bankacılığı hesabı olduklarına ilişkin dağılımları yer almaktadır.

Tablo 4.29. İnternet Bankacılığı Kullanan Banka Müşterilerinin, Kaç Adet İnternet Bankacılığı Hesabı Olduklarına İlişkin Dağılımları



Value	Proportion	%	Count
2.000		46,83	473
1.000		32,57	329
3.000		15,84	160
4.000		4,75	48

Ankete katılan internet bankacılığı kullanan banka müşterilerinin % 46.83'ü 2 adet internet bankacılığı hesabı olduğunu belirtmiştir. İnternet bankacılığını kullanan banka müşterilerinin % 32.57'si 1 adet; % 15.84'ü 3 adet, % 4.75'i ise 4 ve üzeri internet bankacılığı hesabına sahiptir. Dağılım incelendiğinde; internet bankacılığı kullanan banka müşterilerinin çoğunluğunun 2 hesaba sahip olduğu görülmektedir.

Çalışmada internet bankacılığı kullanan banka müşterilerinin; hangi bankaların internet bankacılığını kullandıkları araştırılmış ve şu sonuçlar elde edilmiştir.

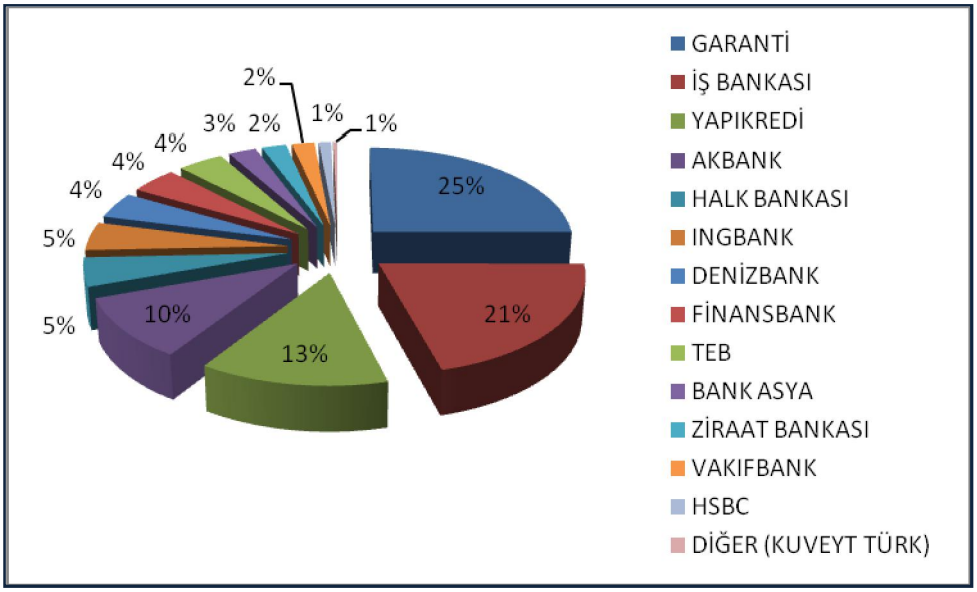
- İnternet bankacılığı kullanan banka müşterilerinin % 25'i (272 kişi) en çok Garanti Bankası'nı tercih etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 21'i (226 kişi) en çok İş Bankası'nı tercih etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 13.96'sı (145 kişi) en çok Yapı Kredi'yi tercih etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 10.89'u (112 kişi) en çok Akbank'ı tercih etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 5.15'i (52 kişi) en çok Halk Bankası'nı tercih etmiştir.

Tablo 4.30'da internet bankacılığı kullanan banka müşterilerinin hangi bankaların internet bankacılığını tercih ettiklerine ilişkin dağılımları yer almaktadır.

Tablo 4.30. İnternet Bankacılığı Kullanan Banka Müşterilerinin Banka Tercihine İlişkin Dağılımları

Bankalar/ Tercih sıralamaları	1.sırada	2.sırada	3.sırada	4.sırada	5.sırada
GARANTİ	272	131	31	5	2
İŞ BANKASI	226	105	32	3	1
YAPIKREDİ	145	87	35	6	1
AKBANK	112	65	16	10	3
HALK BANKASI	52	42	9	5	1
INGBANK	51	58	20	2	5
DENİZBANK	47	30	10	4	4
FİNANSBANK	47	40	24	7	0
TEB	45	45	9	1	0
BANK ASYA	27	14	1	2	0
ZİRAAT BANKASI	24	25	4	3	2
VAKIFBANK	22	37	11	2	1
HSBC	12	24	9	1	3
DİĞER (KUVEYT TÜRK)	4	3	1	0	1

İnternet bankacılığı kullanan banka müşterilerinin, hangi bankaların internet bankacılığını tercih ettikleri Şekil 4.5'te ayrıntılı bir şekilde gösterilmektedir.



Şekil 4.5. İnternet Bankacılığı Kullanan Banka Müşterilerinin Hangi Bankaların İnternet Bankacılığını Tercih Ettiklerine İlişkin Dağılım Grafiği

Dağılımlara bakıldığında; ankete katılan internet bankacılığı kullanan banka müşterilerinin ağırlıkla Garanti Bankası, İş Bankası, Yapıkredi, Akbank ve Halkbankası'nın internet bankacılığını tercih ettikleri görülmektedir. Kullanıcıların internet bankacılığını kullandıkları bankalara ilişkin sıralamalara bakıldığında en çok Garanti Banka'sını tercih ettikleri, en az ise Vakıfbank, HSBC ve Kuveyt Türk Bankasının internet bankacılığını tercih ettikleri görülmektedir.

4.1.2.3.Modelleme

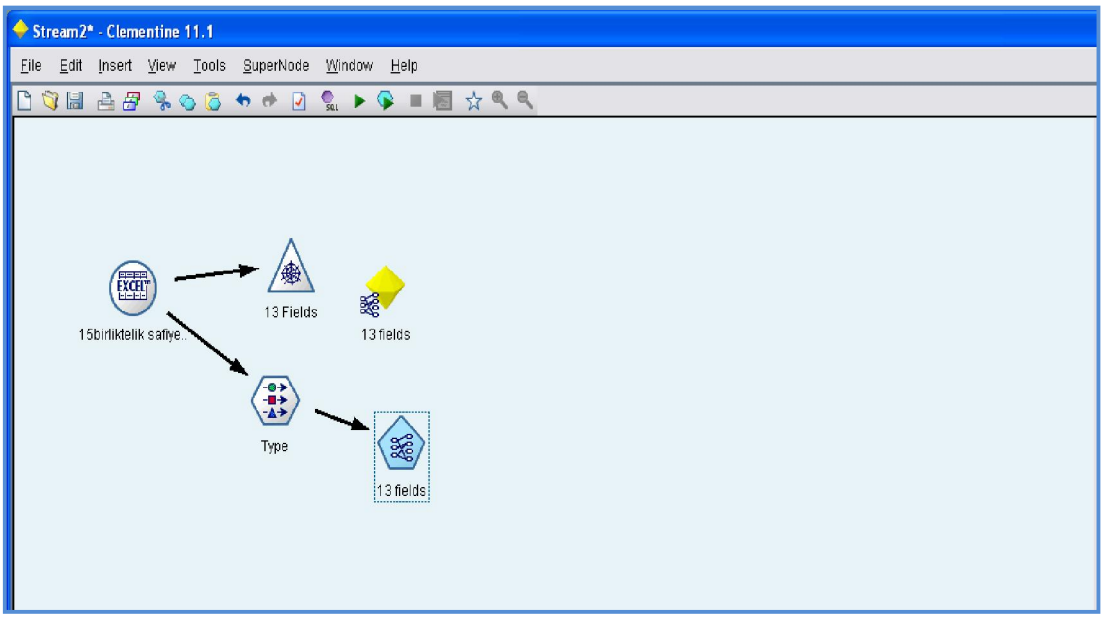
Veri madenciliğinde bilgi kaynaklarından en fazla verimin alınabilmesi iyi kurulmuş bir modele bağlıdır. İyi kurulan bir model sonucunda elde edilen sonuçların kalitesi de değişmekte, analiz sonucunda hangi örüntülerinin bulunacağı tahmin edilebilmektedir.

4.1.2.4. Modelin Kurulması

Tanımlanan problem için en uygun modelin bulunabilmesi olabildiğince çok sayıda modelin kurularak denenmesi ile gerçekleşebilmektedir. Bu sebeple veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar tekrarlanan bir süreçtir. Modelin anlaşılabilir olması buradaki önemli bir değerlendirme kriteridir. Modeller çok ender olarak yorumlanamayacak kadar karmaşık olabilir. [204]

Bu tez çalışmasında öncelikle Birliktelik Tespiti ve Kural Çıkarımı için veri hazırlanmıştır. Bu ara yüz Tablo 4.31'de görülmektedir.

Tablo 4.31. Birliktelik Tespiti ve Kural Çıkarımı için Verinin Hazırlanması

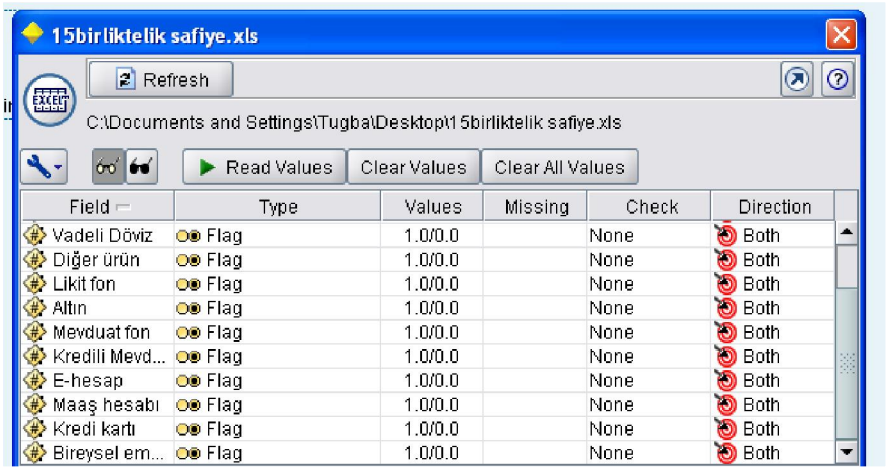


[204] Hakan Eker, "Veri Madenciliği ve Bilgi Keşfi", (Çevrimiçi) http://www.bilgiyonetimi.org, 03.03.2013.

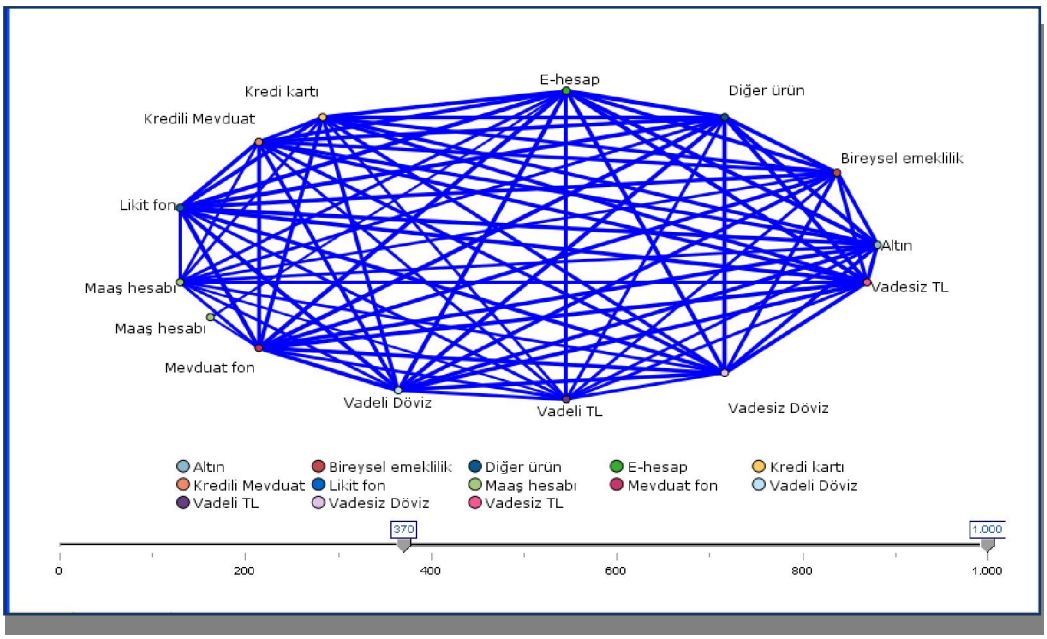
4.1.2.4.1.Birliktelik Kuralları Analizi

Birliktelik Kuralları Analizi dördüncü bölümde ayrıntılı olarak yer verilmiştir. Bu nedenle bu bölümde sadece analiz sonuçları açıklanacaktır. Birliktelik Kurallarının belirlenebilmesi için ilk olarak Web Grafiği ile kategoriler arası ilişkiler incelenmiştir. İnternet bankacılığını kullanan banka müşterilerinin bankacılık işlemlerinde en çok hangi hesapları kullandıklarına ilişkin veriler hazırlanmıştır. Buna ilişkin Tablo 4.32'de yer almaktadır.

Tablo 4.32. Verinin Birliktelik Kuralları Analizi İçin Hazırlanması



Birliktelik tespiti için dördüncü bölümde anlatılan Birliktelik Kuralları Analizi veriye uygulanmış, alanlar arasındaki ilişkilerin belirlenebilmesi için ilk olarak IBM Modeler Programının Web Grafiği aracı kullanılarak, banka müşterilerinin bankacılık işlemlerinde en çok hangi hesapları kullandıklarıyla ilgili ilişkiler incelenmiştir. Bu hesaplar arasındaki bağıntılar Şekil 4.6'da görülmektedir.



Şekil 4.6. Web Grafiği Apriori Algoritması ile Kural Setlerinin Elde Edilmesi

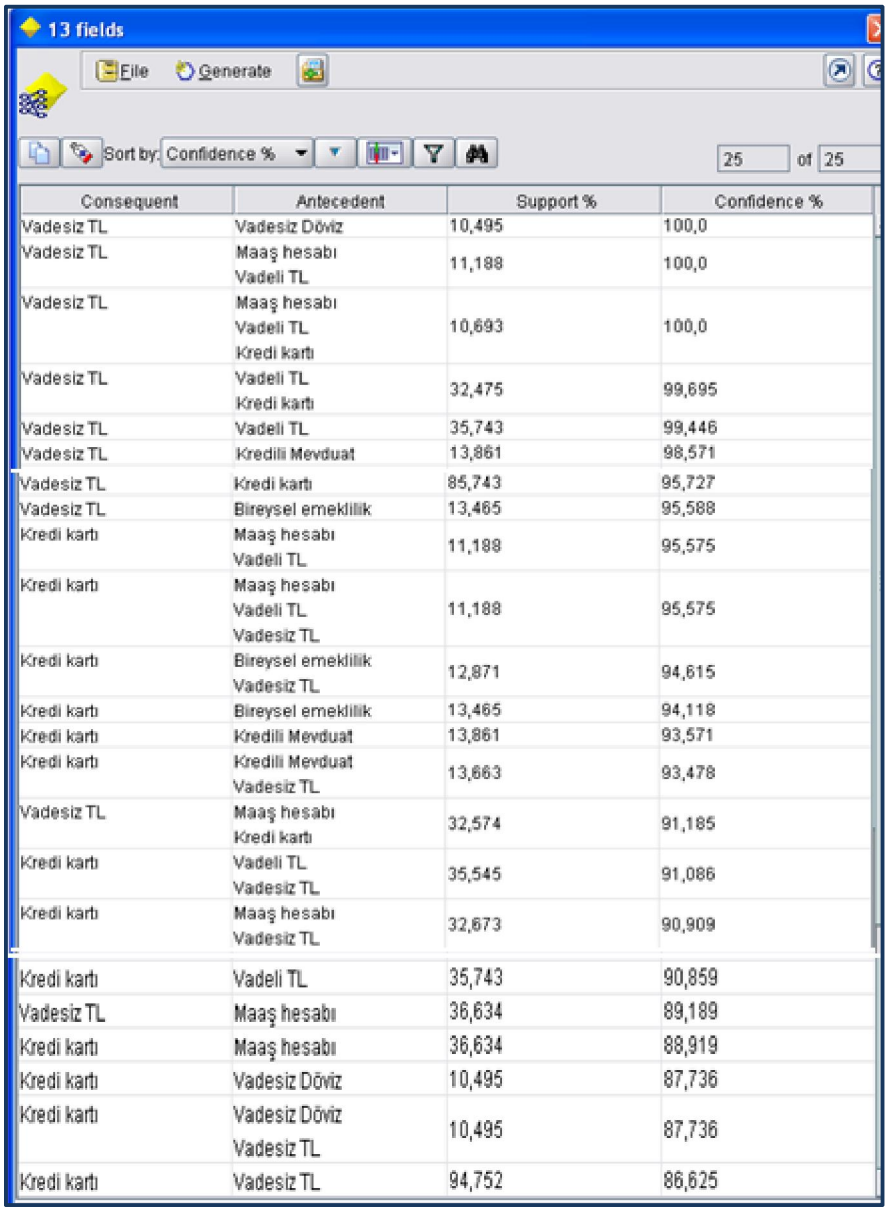
Apriori, mantıksal ilişki kuralları için geçerli bir veri madenciliği algoritmasıdır. Algoritmanın ismi, bilgileri bir önceki adımdan aldığı için "prior" anlamında "Apriori"dir. Apriori algoritması bir seviye akıllı bir şekilde dizilerin gerçekleşme sıklığını üretir. Uygulama ilk olarak kayıtlı bir dizinin tümünü sayar ve bu kayıtların içerisinde başlangıç sınırını aşan kayıtlar belirlenir. Sonra iki kayıtlı dizilerin sıklığını saptayabilmek için onların birbirleri ile ikili kombinasyonlarını alıp potansiyel olarak gerçekleşme sayısını bulur.

Hesaplama işlemi için tüm veri tabanı üzerinden geçilmeli ve iki kayıtlı dizi olmaya aday her bir hareketin kontrol edilmesi gereklidir. Üç kayıtlı dizilerin sıklığını saptayabilmek için, iki kayıtlı dizilerin kombinasyonuyla gerçekleşme sayısını bulmaya devam eder. Apriori algoritması, belirlenen dizi sıklığından daha fazla kayıt kümeleri bulunmayınca durur.[205]

Tablo 4.33'te internet bankacılığı kullanan banka müşterilerinin, bankacılık işlemlerinde en çok hangi hesapları kullandıklarına ilişkin birliktelik kural setleri yer almaktadır.

[205] Gürsoy, **a.g.e.** s.158.

Tablo 4.33. İnternet Bankacılığı Kullanan Banka Müşterilerinin, Bankacılık İşlemlerinde En Çok Hangi Hesapları Kullandıklarına İlişkin Kural Setleri



13 fields

File Generate

Sort by: Confidence % 25 of 25

Consequent	Antecedent	Support %	Confidence %
Vadesiz TL	Vadesiz Döviz	10,495	100,0
Vadesiz TL	Maaş hesabı Vadeli TL	11,188	100,0
Vadesiz TL	Maaş hesabı Vadeli TL Kredi kartı	10,693	100,0
Vadesiz TL	Vadeli TL Kredi kartı	32,475	99,695
Vadesiz TL	Vadeli TL	35,743	99,446
Vadesiz TL	Kredili Mevduat	13,861	98,571
Vadesiz TL	Kredi kartı	85,743	95,727
Vadesiz TL	Bireysel emeklilik	13,465	95,588
Kredi kartı	Maaş hesabı Vadeli TL	11,188	95,575
Kredi kartı	Maaş hesabı Vadeli TL Vadesiz TL	11,188	95,575
Kredi kartı	Bireysel emeklilik Vadesiz TL	12,871	94,615
Kredi kartı	Bireysel emeklilik	13,465	94,118
Kredi kartı	Kredili Mevduat	13,861	93,571
Kredi kartı	Kredili Mevduat Vadesiz TL	13,663	93,478
Vadesiz TL	Maaş hesabı Kredi kartı	32,574	91,185
Kredi kartı	Vadeli TL Vadesiz TL	35,545	91,086
Kredi kartı	Maaş hesabı Vadesiz TL	32,673	90,909
Kredi kartı	Vadeli TL	35,743	90,859
Vadesiz TL	Maaş hesabı	36,634	89,189
Kredi kartı	Maaş hesabı	36,634	88,919
Kredi kartı	Vadesiz Döviz	10,495	87,736
Kredi kartı	Vadesiz Döviz Vadesiz TL	10,495	87,736
Kredi kartı	Vadesiz TL	94,752	86,625

Tablo 4.33'te görülen kurallar şöyle yorumlanmaktadır.

- Vadesiz döviz hesabı olanların % 100'ü Vadesiz TL hesabına sahiptir. Bu birliktelik tüm kayıtların % 10.495'inde görülmektedir.
- Maaş ve vadeli TL hesabı olanların % 100'ü Vadesiz TL hesabına sahiptir. Bu birliktelik tüm kayıtların % 11.188'inde görülmektedir.

- Vadeli TL ve kredi kartı hesabı olanların % 99.695'i Vadesiz TL hesabına sahiptir. Bu birliktelik tüm kayıtların % 32.475'inde görülmektedir.
- Kredi mevduat hesabı olanların %98.571'i Vadesiz TL hesabına sahiptir. Bu birliktelik tüm kayıtların % 13.861'inde görülmektedir.
- Kredi kartı hesabı olanların % 95.727'si Vadesiz TL hesabına sahiptir. Bu birliktelik tüm kayıtların % 85.743'ünde görülmektedir.
- Maaş, vadeli TL ve vadesiz TL hesabı olanların % 95.575'i kredi kartı hesabına sahiptir. Bu birliktelik tüm kayıtların % 11.188'inde görülmektedir.
- Maaş ve kredi kartı hesabı olanların % 91.185'i Vadesiz TL hesabına sahiptir. Bu birliktelik tüm kayıtların % 32.574'ünde görülmektedir.
- Vadeli TL hesabı olanların % 90.859'u kredi kartı hesabına sahiptir. Bu birliktelik tüm kayıtların % 35.743'ünde görülmektedir.
- Maaş hesabı olanların % 89.919'u kredi kartı hesabına sahiptir. Bu birliktelik tüm kayıtların % 36.634'ünde görülmektedir.
- Vadesiz TL hesabı olanların %86.625'u kredi kartı hesabına sahiptir. Bu birliktelik tüm kayıtların % 94.752'sinde görülmektedir.

Elde edilen bu sonuçlar, internet bankacılığı kullanıcılarının vadesiz TL ve kredi kartı hesaplarını etkin bir şekilde kullandıklarını göstermektedir. Elde edilen veriler kullanıcıların hangi hesapları aktif olarak kullandığını göstermektedir. Ayrıca elde edilen veriler, özellikle bankacılık sektöründe, müşteri segmentasyonu ve müşterilerine yönelik kampanyalarda sıklıkla kullanılmakta, bankalar ürün aktifliklerini ve müşteri memnuniyetini elde edilen bu verileri kullanarak artırmaktadır.

İnternet bankacılığını kullanan banka müşterilerinin bankacılık işlemlerinde en çok hangi işlemleri yaptıklarına ilişkin veriler hazırlanmıştır. Buna ilişkin veri seti Tablo 4.34'te yer almaktadır.

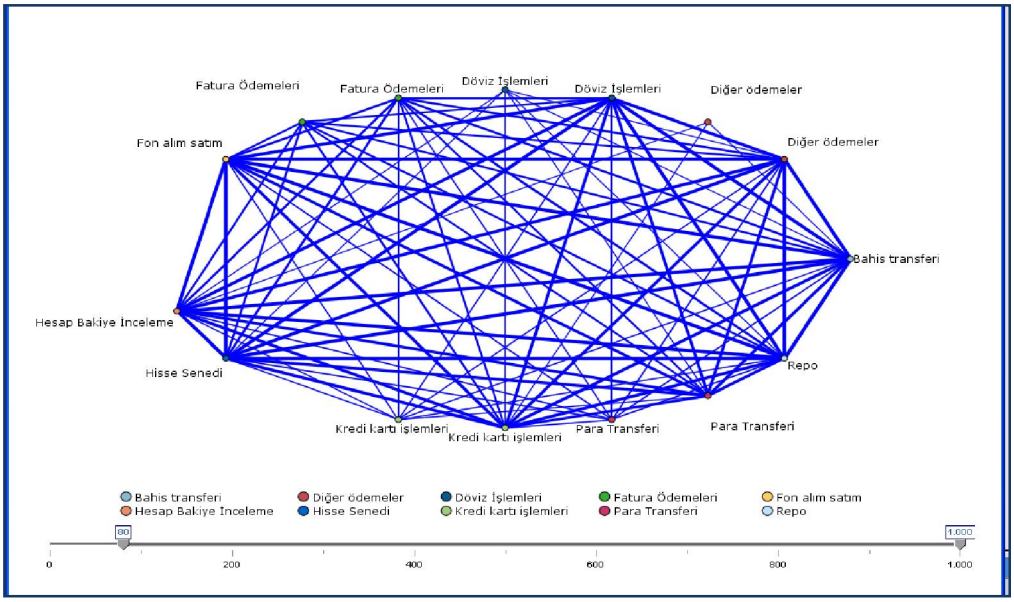
Tablo 4.34. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bankacılık İşlemlerinde En Çok Hangi İşlemleri Yaptıklarına İlişkin Birliktelik Kuralları Analizi İçin Veri Seti

Table (10 fields, 1.010 records) #1

File Edit Generate

	Para Transferi	Hesap Bakiye İnceleme	Kredi kartı işlemleri	Fatura Ödemeleri	Döviz İşlemleri	Hisse Senedi
1	1.000	1.000	1.000	1.000	0.000	0.000
2	1.000	1.000	1.000	0.000	0.000	0.000
3	1.000	1.000	1.000	0.000	0.000	0.000
4	1.000	1.000	1.000	0.000	0.000	0.000
5	1.000	1.000	1.000	0.000	0.000	0.000
6	0.000	1.000	0.000	0.000	0.000	0.000
7	0.000	1.000	1.000	0.000	0.000	0.000
8	1.000	1.000	1.000	0.000	0.000	0.000
9	0.000	1.000	1.000	1.000	0.000	0.000
10	0.000	1.000	1.000	0.000	0.000	0.000
11	1.000	1.000	1.000	0.000	0.000	0.000
12	0.000	1.000	1.000	1.000	0.000	0.000
13	1.000	0.000	1.000	1.000	0.000	0.000
14	0.000	1.000	1.000	0.000	0.000	0.000
15	1.000	1.000	1.000	1.000	0.000	0.000
16	1.000	0.000	1.000	0.000	0.000	0.000
17	1.000	1.000	1.000	1.000	0.000	0.000
18	1.000	0.000	0.000	0.000	0.000	0.000
19	0.000	0.000	0.000	0.000	1.000	0.000
20	1.000	0.000	1.000	0.000	0.000	0.000

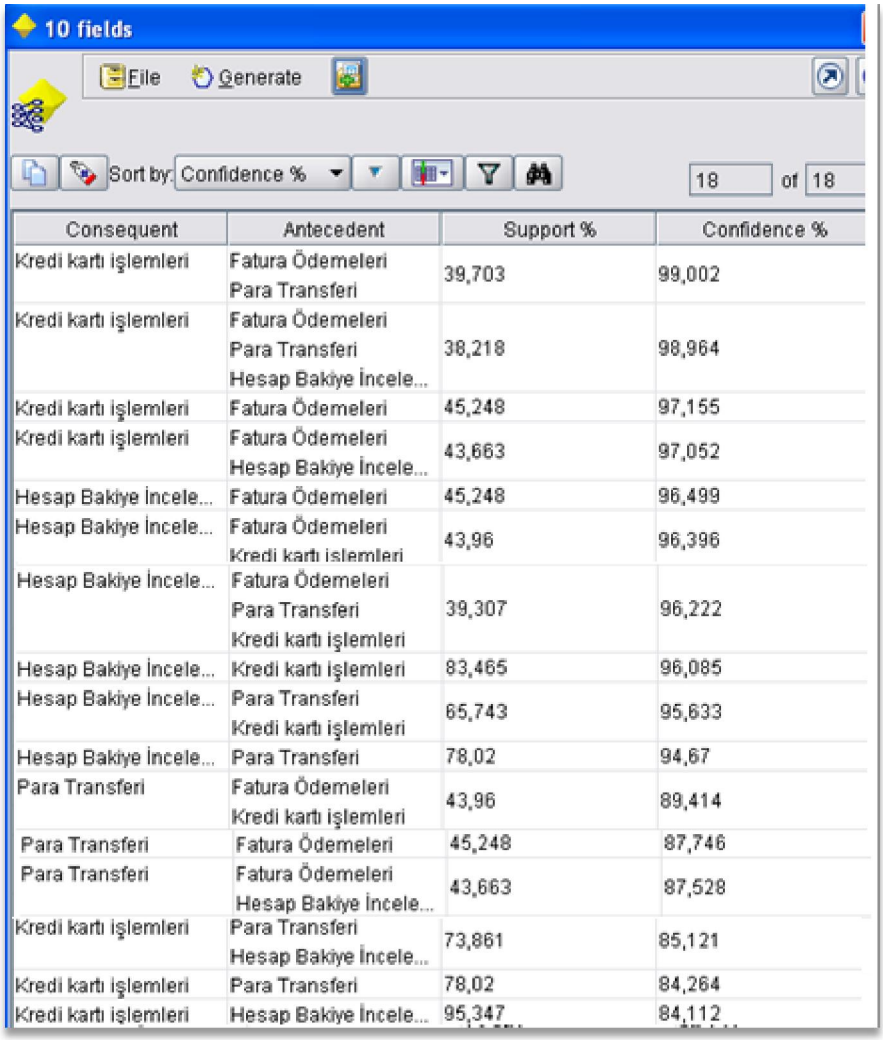
İnternet bankacılığı kullanan banka müşterilerinin bankacılık işlemlerinde en çok hangi işlemleri yaptıkları incelenmiştir. Bu işlemler arasındaki bağıntılar Şekil 4.7'de görülmektedir.



Şekil 4.7. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bankacılık İşlemlerinde En Çok Hangi İşlemleri Yaptıklarına İlişkin Web Grafiği

Tablo 4.35'te ise internet bankacılığı kullanan banka müşterilerinin bankacılık işlemlerinde en çok hangi işlemleri yaptıklarına ilişkin kural setleri yer almaktadır.

Tablo 4.35. İnternet Bankacılığını Kullanan Banka Müşterilerinin Bankacılık İşlemlerinde En Çok Hangi İşlemleri Yaptıklarına İlişkin Kural Setleri



10 fields

File Generate

Sort by: Confidence % 18 of 18

Consequent	Antecedent	Support %	Confidence %
Kredi kartı işlemleri	Fatura Ödemeleri Para Transferi	39,703	99,002
Kredi kartı işlemleri	Fatura Ödemeleri Para Transferi Hesap Bakiye İncele...	38,218	98,964
Kredi kartı işlemleri	Fatura Ödemeleri	45,248	97,155
Kredi kartı işlemleri	Fatura Ödemeleri Hesap Bakiye İncele...	43,663	97,052
Hesap Bakiye İncele...	Fatura Ödemeleri	45,248	96,499
Hesap Bakiye İncele...	Fatura Ödemeleri Kredi kartı işlemleri	43,96	96,396
Hesap Bakiye İncele...	Fatura Ödemeleri Para Transferi Kredi kartı işlemleri	39,307	96,222
Hesap Bakiye İncele...	Kredi kartı işlemleri	83,465	96,085
Hesap Bakiye İncele...	Para Transferi Kredi kartı işlemleri	65,743	95,633
Hesap Bakiye İncele...	Para Transferi	78,02	94,67
Para Transferi	Fatura Ödemeleri Kredi kartı işlemleri	43,96	89,414
Para Transferi	Fatura Ödemeleri	45,248	87,746
Para Transferi	Fatura Ödemeleri Hesap Bakiye İncele...	43,663	87,528
Kredi kartı işlemleri	Para Transferi Hesap Bakiye İncele...	73,861	85,121
Kredi kartı işlemleri	Para Transferi	78,02	84,264
Kredi kartı işlemleri	Hesap Bakiye İncele...	95,347	84,112

- Fatura ödemeleri ve para transferi yapanların % 99.002'si kredi kartı işlemlerini de gerçekleştirmektedir. Bu birliktelik tüm kayıtların % 39.703'ünde görülmektedir.

- Fatura ödemeleri yapanların % 97.155'i kredi kartı işlemlerini de gerçekleştirmektedir. Bu birliktelik tüm kayıtların % 45.248'inde görülmektedir.

- Fatura ödemeleri yapanların % 96.499'u hesap bakiyesi incelemesi gerçekleştirmektedir. Bu birliktelik tüm kayıtların % 45.248'inde görülmektedir.

- Kredi kartı işlemleri yapanların % 96.085'i hesap bakiyesi incelemesi gerçekleştirmektedir. Bu birliktelik tüm kayıtların % 83.465'inde görülmektedir.

- Para transferi yapanların ve Hesap bakiyesi inceleyenlerin % 85.121'i kredi kartı işlemlerini de gerçekleştirmektedir. Bu birliktelik tüm kayıtların % 73.861'inde görülmektedir.

- Hesap bakiyesi incelemesi yapanların % 84.112'si kredi kartı işlemlerini de gerçekleştirmektedir. Bu birliktelik tüm kayıtların %95.347'sinde görülmektedir.

Elde edilen kural setleri kullanıcıların hangi işlemleri daha çok gerçekleştirdiğini göstermektedir. Ayrıca elde edilen veriler, özellikle bankacılık sektöründe, müşteri segmentasyonu, müşterileri yönelik kampanyalarda, işlem hacimlerinin belirlenmesinde kullanılmakta, bankalar ürün aktifliklerini ve müşteri memnuniyetini elde edilen bu verileri kullanarak artırmaktadır.

Çalışmada internet bankacılığı kullanan banka müşterilerinin, internet bankacılığını kullanmalarına etki eden faktörler incelenmiştir. Ankete katılan banka müşterilerine internet bankacılığı kullanmanıza en çok etki eden faktör nedir sorusu yöneltilmiş ve faktörlere en önemlisi 1 olacak şekilde cevap vermeleri istemiştir.

Tablo 4.36'da internet bankacılığı kullanan banka müşterilerinin verdikleri cevaplara ilişkin dağılımları yer almaktadır.(Sayıların altında yüzdelik dağılımlarına yer verilmiştir.)

Tablo 4.36. İnternet Bankacılığı Kullanan Banka Müşterilerinin, İnternet Bankacılığını Kullanmalarına Etki Eden Faktörlere İlişkin Verdikleri Cevapların Dağılımı

Etki sırası / Özellikler	1	2	3	4	5	6	7	8	9	Toplam
7/24 hızlı ve kesintisiz bankacılık işlemleri	**242** (23.9)	**310** (30.7)	**155** (15.3)	**121** (12.0)	**53** (5.2)	**37** (3.7)	**45** (4.5)	**19** (1.9)	**28** (2.8)	**1010** (100)
Maaşımı o bankadan almam	**182** (18.0)	**108** (10.7)	**88** (8.7)	**83** (8.2)	**73** (7.2)	**68** (6.7)	**80** (7.9)	**73** (7.2)	**255** (25.2)	**1010** (100)
Şubeye gitmeden, kolay bankacılık işlemleri	**171** (16.9)	**138** (13.6)	**136** (13.5)	**129** (12.8)	**118** (11.7)	**92** (9.1)	**110** (10.9)	**60** (5.9)	**56** (5.5)	**1010** (100)
Pratik olması	**136** (13.5)	**204** (20.2)	**202** (20.0)	**159** (15.7)	**143** (14.1)	**69** (6.8)	**39** (3.9)	**24** (2.4)	**34** (3.4)	**1010** 100
Görerek bankacılık işlemleri yapabilmek	**93** (9.2)	**49** (4.8)	**57** (5.6)	**78** (7.7)	**79** (7.8)	**172** (17.0)	**171** (16.9)	**198** (19.6)	**113** (11.2)	**1010** (100)
Güvenilir olması	**63** (6.2)	**101** (10.0)	**217** (21.5)	**201** (19.9)	**155** (15.3)	**104** (10.3)	**61** (6.0)	**59** (5.8)	**49** (4.8)	**1010** (100)
Detaylı rapor alabilmek	**57** (5.6)	**40** (4.0)	**46** (4.5)	**64** (6.3)	**145** (14.3)	**163** (16.1)	**191** (18.9)	**174** (17.2)	**130** (12.9)	**1010** (100)
Bankacılık işlemlerinin daha ucuz olarak gerçekleştirmesi	**38** (3.8)	**28** (2.8)	**60** (5.9)	**103** (10.2)	**124** (12.3)	**167** (16.5)	**140** (13.8)	**162** (16.0)	**188** (18.6)	**1010** (100)
Çok çeşitli bankacılık işlemlerinden yararlanılması	**37** (3.7)	**31** (3.1)	**48** (4.7)	**74** (7.3)	**119** (11.8)	**135** (13.4)	**170** (16.8)	**240** (23.7)	**156** (15.4)	**1010** (100)

Çalışmada internet bankacılığı kullanan müşterilerin, internet bankacılığını kullanmalarına etki eden faktörlere ilişkin verdikleri cevapların dağılımı incelendiğinde, aşağıdaki sonuçlar elde edilmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin %23.9'u (242 kişi) 7/24 hızlı ve kesintisiz bankacılık işlemleri gerçekleştirmesinin

internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 18'i (182 kişi) maaşını o bankadan almasının internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir

- İnternet bankacılığı kullanan banka müşterilerinin % 16.9'u (171 kişi) şubeye gitmeden kolay bankacılık işlemlerinin gerçekleştirmesinin, internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 13.5'i (136 kişi), internet bankacılığının pratik olmasının internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 9,2.'si (93 kişi) görerek bankacılık işlemlerini gerçekleştirebilmesinin internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

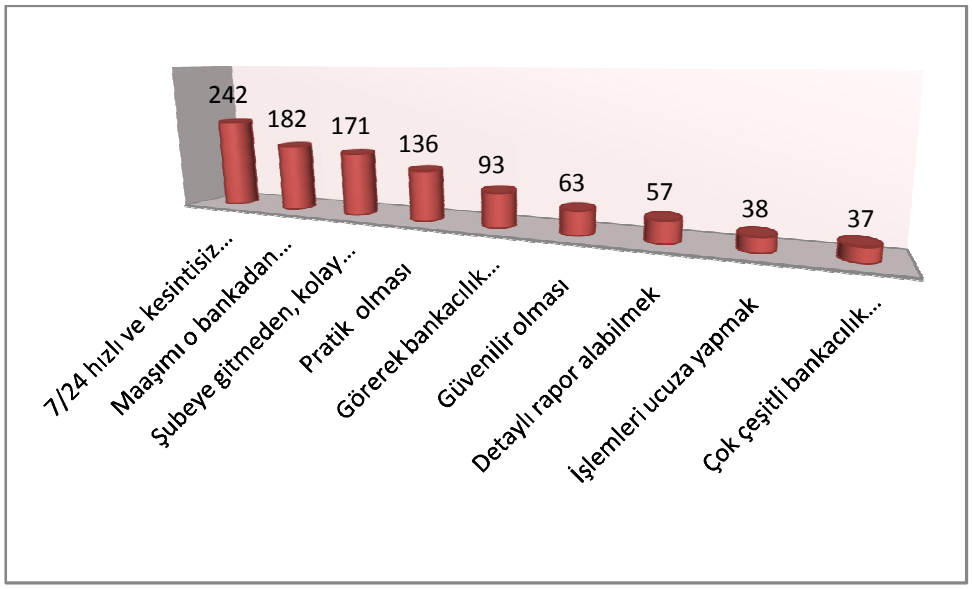
- İnternet bankacılığı kullanan banka müşterilerinin % 6.2'si (63 kişi) güvenilir olmasının internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 5.6'sı (57 kişi) detaylı rapor alınabilmesinin internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 3.8'i (38 kişi) bankacılık işlemlerini çok daha ucuza yapabilmesinin internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 3.7'si (37 kişi) çok çeşitli bankacılık ürünlerinden faydalanabilmenin internet bankacılığı kullanmasında 1. sırada etkili olduğunu ifade etmiştir.

Şekil 4.8'de internet bankacılığı kullanan müşterilerin, internet bankacılığını kullanmalarına etki eden faktörlere ilişkin verdikleri cevapların dağılım grafiği yer almaktadır.



Şekil 4.8. İnternet Bankacılığı Kullanan Müşterilerin, İnternet Bankacılığını Kullanmalarına Etki Eden Faktörlere İlişkin Verdikleri Cevapların Dağılım Grafiği

Dağılıma bakıldığında; ankete katılan internet bankacılığı müşterilerinin, internet bankacılığı kullanımında 7/24 hızlı ve kesintisiz bankacılık işlemlerini gerçekleştirebilmelerinin, maaşını o bankadan almasının ve şubeye gitmeden kolay bankacılık işlemleri gerçekleştirebilmesinin etkili olduğu görülmektedir.

Çalışmada internet bankacılığı kullanan banka müşterilerinin; internet bankacılığı işlemlerinde karşılaştıkları sorunlar araştırılmış ve Tablo 4.37'deki sonuçlar elde edilmiştir.

Tablo 4.37. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığı İşlemlerinde Karşılaşılan Sorunlara İlişkin Verdikleri Cevapların Dağılımı

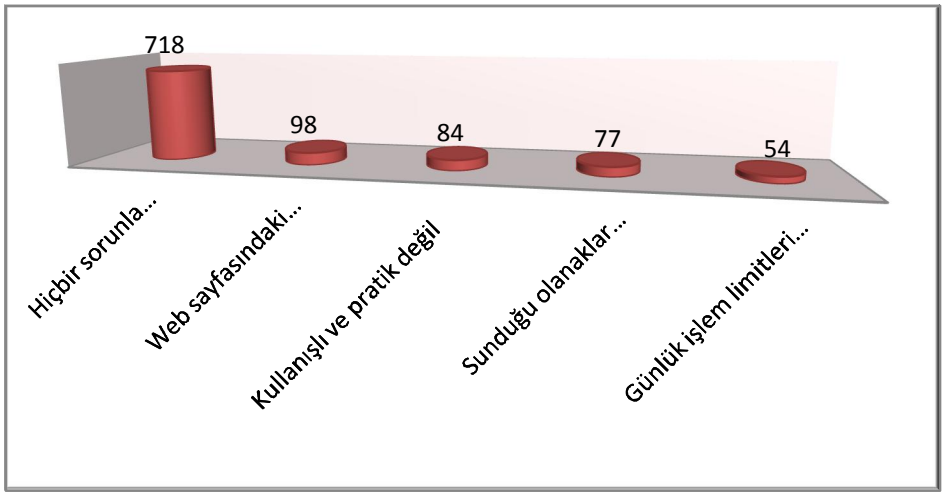
İnternet Bankacılığı İşlemlerinde Karşılaşılan Sorunlar	Frekans	Yüzde
Hiçbir sorunla karşılaştığımı düşünmüyorum	718	71,09
Web sayfasındaki güvenlik önlemleri yetersiz	98	9,70
Kullanışlı ve pratik değil	84	8,32
Sunduğu olanaklar yetersiz	77	7,6
Günlük işlem limitleri yetersiz	54	5,35

Çalışmada internet bankacılığı kullanan 1010 kişinin verdikleri cevaplar incelendiğinde, aşağıdaki sonuçlar elde edilmiştir:

- İnternet bankacılığı kullanan banka müşterilerinin % 71.09'u (718 kişi) hiçbir sorunla karşılaşmadığını belirtmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 9.70'i (98 kişi), web sayfasındaki güvenlik önlemlerinin yetersiz olduğunu düşünmektedir.
- İnternet bankacılığı kullanan banka müşterilerinin % 8.32'si (84 kişi), internet bankacılığının kullanışlı ve pratik olmadığını düşünmektedir.
- İnternet bankacılığı kullanan banka müşterilerinin % 7.6'sı (77 kişi) sunduğu olanakların yetersiz olduğunu düşünmektedir.
- İnternet bankacılığı kullanan banka müşterilerinin % 5.35'i (54 kişi) günlük işlem limitlerinin yetersiz olduğunu düşünmektedir.

Sonuçlara bakıldığında; banka müşterilerinin genel olarak bankaların internet bankacılığını kullanmaktan memnun oldukları, internet bankacılığı kullanırken bir problem yaşamadıklarını görülmektedir.

İnternet bankacılığını kullanan banka müşterilerinin internet bankacılığında yaşadıkları sorunlara ilişkin dağılım şekil 4.9'da görülmektedir.



Şekil 4.9. İnternet Bankacılığı Kullanan Banka Müşterilerinin İnternet Bankacılığı İşlemlerinde Karşılaşılan Sorunlara İlişkin Verdikleri Cevapların Dağılım Grafiği

Çalışmada internet bankacılığı kullanan banka müşterilerinin gelecekte internet bankacılığının geleneksel bankacılığın yerini alacağını düşünüp düşünmedikleri araştırılmıştır. Tablo 4.38'de cevapların dağılımı yer almaktadır.

Tablo 4.38. İnternet Bankacılığı Kullanan Banka Müşterilerinin Gelecekte İnternet Bankacılığının Geleneksel Bankacılığın Yerini Alacağını Düşünüp Düşünmediklerine İlişkin Dağılımları

Distribution of 19) İnternet bankacılığının gelecekte geleneksel bankacılı...

File Edit Generate View

Value	Proportion	%	Count
3.000		52,28	528
1.000		31,98	323
2.000		12,28	124
4.000		3,47	35

Ankete katılan internet bankacılığı kullanan banka müşterilerinin %52.28'i bu soruya kısmen yanıtını vermiştir. İnternet bankacılığı kullanan banka müşterilerinin %31.98'i evet, % 12.28'i hayır yanıtını vermiştir. % 3.47'si bu konuda fikri olmadığını belirtmiştir. Dağılımlar incelendiğinde; internet bankacılığı kullanan

banka müşterilerinin % 50'sinden fazlası kısmen de olsa, internet bankacılığının geleneksel bankacılığın yerini alacağını düşünmektedir.

Çalışmada internet bankacılığı kullanan müşterilerin, bir bankanın internet sitesinde aradıkları özelliklerin neler olduğu araştırılmıştır. İnternet bankacılığı kullanan banka müşterileri bir bankanın internet sitesinde aradıkları özellikleri en önemlisi 1 olacak şekilde cevap vermişlerdir. Tablo 4.39' da buna ilişkin dağılım yer almaktadır.

Tablo 4.39. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bir Bankanın İnternet Sitesinde Aradıkları Özelliklerine İlişkin Verdikleri Cevapların Dağılımları

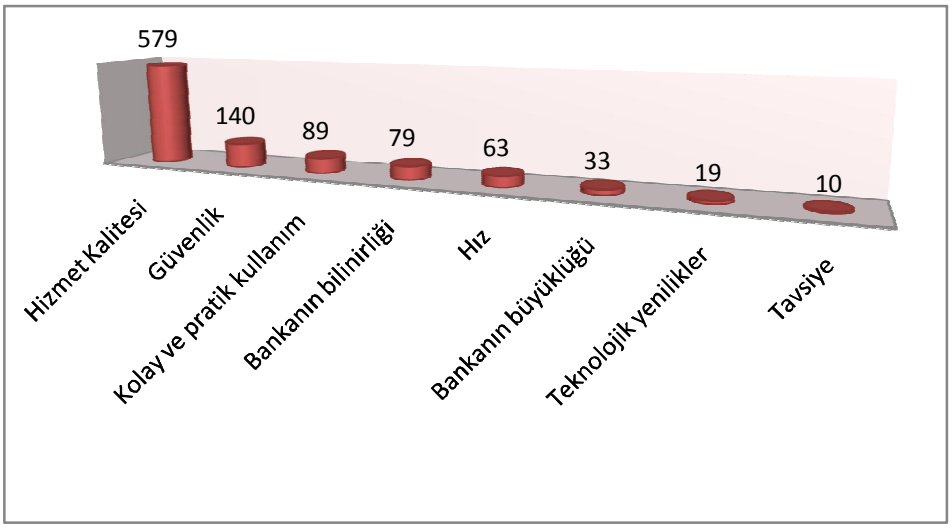
Tercih sırası / Özellikler	1.	2.	3.	4.	5.	6.	7.	8.	Toplam
Hizmet Kalitesi	579	86	126	77	50	39	28	25	1010
Güvenlik	140	316	304	93	70	38	31	18	1010
Kolay ve pratik kullanım	89	355	274	102	55	45	49	41	1010
Bankanın bilinirliği	79	37	42	98	189	154	290	121	1010
Hız	63	100	65	265	127	213	101	76	1010
Bankanın büyüklüğü	33	43	89	170	265	245	124	40	1009
Teknolojik yenilikler	19	34	86	174	193	159	175	169	1009
Tavsiye	10	39	22	31	61	116	212	518	1009

Çalışmada internet bankacılığı kullanan müşterilerin, bir bankanın internet sitesinde aradıkları özellikler incelendiğinde aşağıdaki şu sonuçlar elde edilmiştir:

- İnternet bankacılığı kullanan banka müşterilerinin % 57.33'ü (579 kişi) hizmet kalitesinin aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 13.86'sı (140 kişi) güvenliğin aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.

- İnternet bankacılığı kullanan banka müşterilerinin % 8.81'i (89 kişi) kullanışlı ve pratik olmasının aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 7.82'si (79 kişi) bankanın bilinirliğinin aradığı özellikler arasında 1.sırada olduğunu ifade etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 6.24'ü (63 kişi) hızın aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 3.27'i (33 kişi) bankanın büyüklüğünün aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 1.88'i (19 kişi) teknolojik yeniliklerin aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.
- İnternet bankacılığı kullanan banka müşterilerinin % 0.99'u (10 kişi) tavsiyenin aradığı özellikler arasında 1. sırada olduğunu ifade etmiştir.

Şekil 4.10'da internet bankacılığı kullanan banka müşterilerinin bir bankanın internet sitesinde aradıkları özelliklere ilişkin verdikleri cevapların dağılım grafiği yer almaktadır.



Şekil 4.10. İnternet Bankacılığı Kullanan Banka Müşterilerinin Bir Bankanın İnternet Sitesinde Aradıkları Özelliklere İlişkin Verdikleri Cevapların Dağılım Grafiği

Dağılım incelendiğinde; ankete katılan internet bankacılığı müşterilerinin, bir bankanın internet sitesinde aradıkları özelliklerin başında hizmet kalitesi, güvenlik, kolay ve pratik kullanılabilir olmasının geldiği görülmektedir.

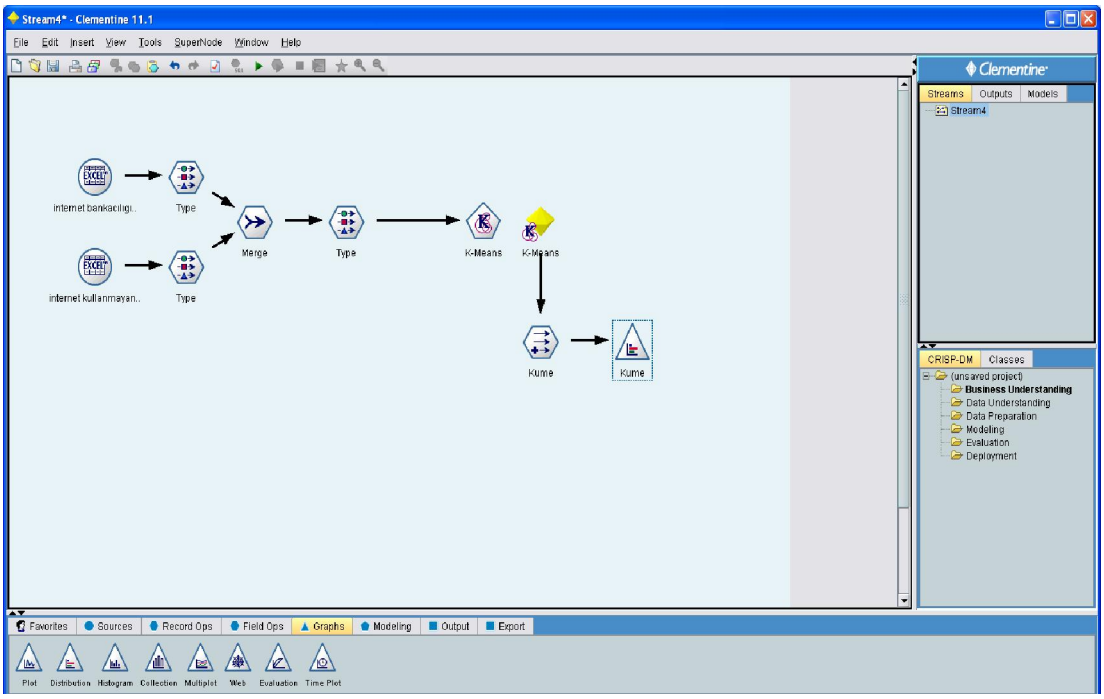
4.1.2.4.2. Kümeleme Analizi

Tezin bu bölümünde, dördüncü bölümde ayrıntılı olarak anlatılan Kümeleme Modelleri veri setine uygulanmıştır ve analiz sonuçları elde edilmiştir. Kümeleme modellerinde amaç, küme üyelerinin birbirlerine çok benzediği, ancak özellikleri birbirlerinden çok farklı olan kümelerin bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir.[206] Uygulamada, Kümeleme Analizi Tekniklerinden K-Ortalamalar (K-means) kullanılmıştır.

Kümeleme analizi uygulayabilmek için veriler hazırlanmış ve IBM Modeler programında bir model oluşturulmuştur. Bu arayüz Tablo 4.40'ta görülmektedir.

[206] Akpınar, **a.g.e.** , s.6.

Tablo 4.40. Kümeleme Analizi için Verilerin Hazır Hale Getirilmesi

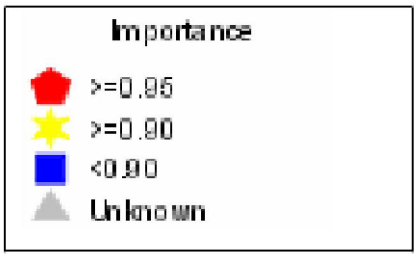


- **K-Ortalamalar Yöntemi**

K-ortalamalar yöntemi tezin dördüncü bölümünde bir örnek üzerinde ayrıntılı olarak incelenmiştir. Bu nedenle bu aşamada sadece elde edilen sonuçlara yer verilmiştir. K-ortalamalar tekniği ile yapılan, analizler sonucunda veri seti için en uygun küme sayısının dört küme olduğu tespit edilmiştir.

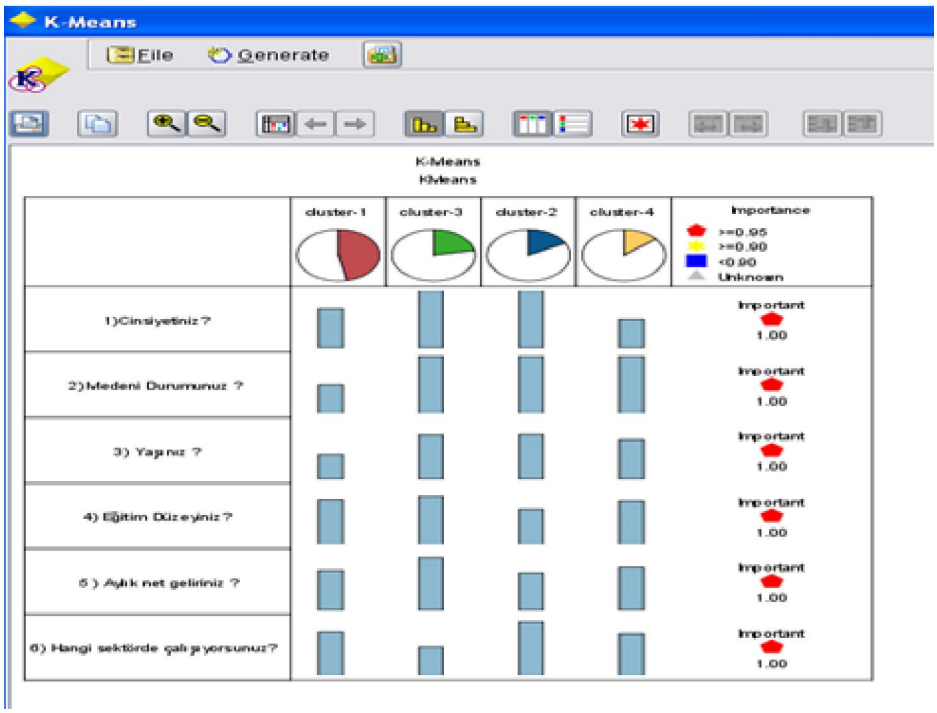
Öncelikle program tarafından önem derecesi 0.90'dan büyük olarak belirlenen değişkenler açıklanmıştır.

Tablo 4 41. Önem Dereceleri



Elde edilen dört kümede hangi değişkenlerin etkili olduğu Tablo 4.42'de görülmektedir. Değişkenlerin kümeler üzerindeki etkileri ayrı ayrı incelenmiştir. Kümeler üzerinde etkili olabilecek tüm değişkenler analize alınarak değerlendirilmiş, en etkili değişkenlerin "Cinsiyet, Medeni Durum, Yaş, Eğitim Durumu, Gelir Düzeyi, Çalışılan Sektör" olduğu belirlenmiş ve analiz uygulanmıştır.

Tablo 4.42. K-Ortalamalar Yöntemi İle Elde Edilen Kümeler



Elde edilen dört kümenin profili aşağıda açıklanmıştır:

1. Küme: İnternet bankacılığı kullanan; 26-35 yaş aralığında; Üniversite mezunu, gelir düzeyi 1501-2500 TL arasında olan, sağlık sektöründe çalışan bekar kadınlardan oluşmaktadır.

2. Küme: İnternet bankacılığı kullanan; 36-45 yaş aralığında; lise mezunu, gelir düzeyi 1501-2500 TL arasında olan, perakende sektöründe çalışan evli erkeklerden oluşmaktadır.

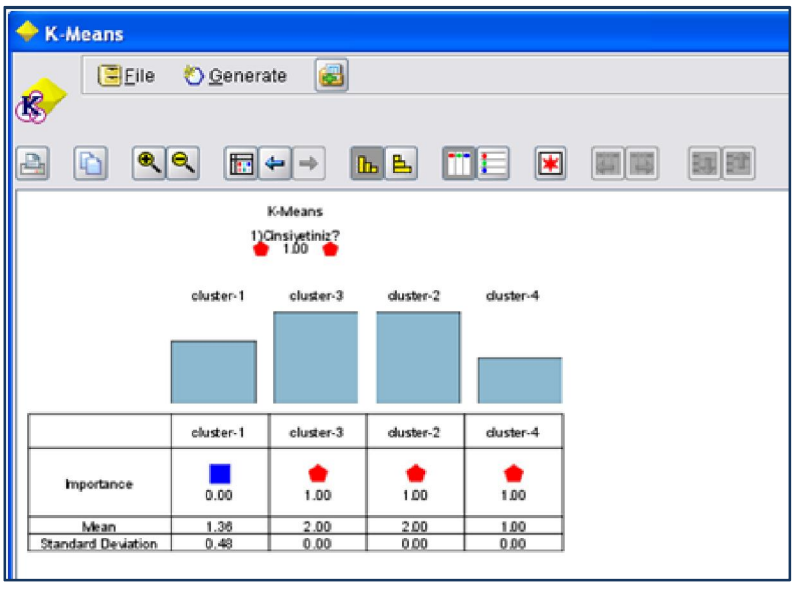
3. Küme: İnternet bankacılığı kullanan; 36-45 yaş aralığında; üniversite mezunu, gelir düzeyi 2500 TL üzerinde olan, otomotiv sektöründe çalışan evli erkeklerden oluşmaktadır.

4. Küme: İnternet bankacılığı kullanan; 36-45 yaş aralığında; üniversite mezunu, gelir düzeyi 1501-2500 TL arasında olan, gıda ve sağlık sektöründe çalışan evli kadınlardan oluşmaktadır.

Bundan sonraki bölümde, her bir değişkenin tek tek kümeler üzerindeki etkisi grafiklerle gösterilmiştir.

Tablo 4.43'te cinsiyetin kümeler üzerindeki etkisi incelenmiştir.

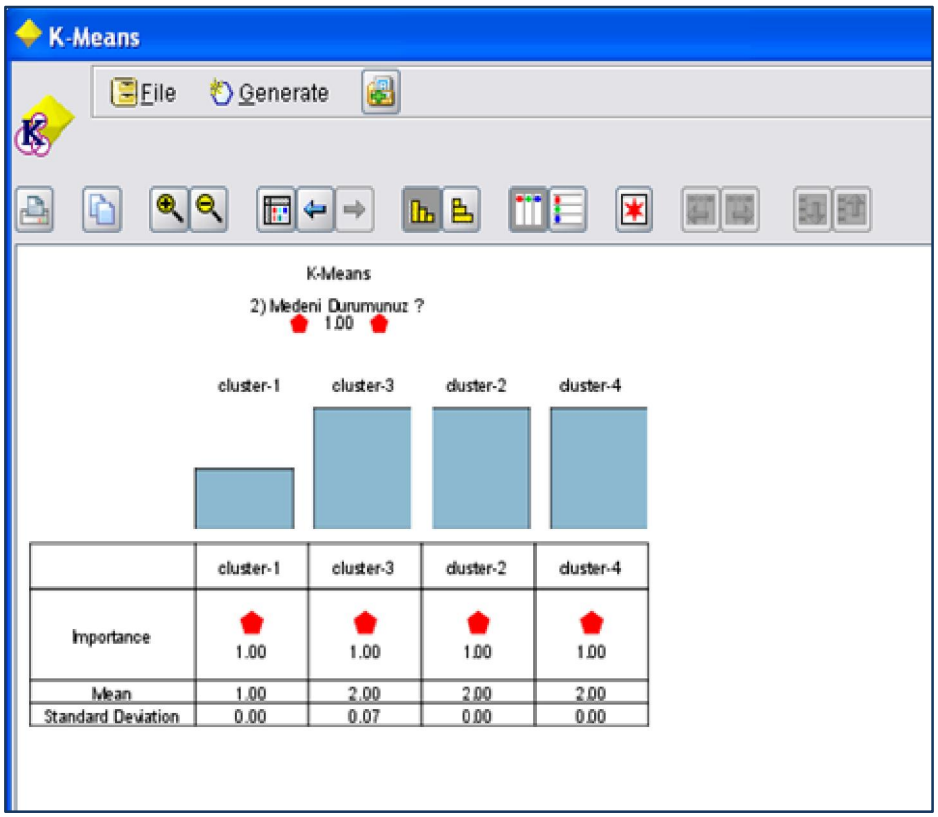
Tablo 4.43. Cinsiyetin Kümeler Üzerindeki Etkisi



Cinsiyet değişkeni her dört kümeyi de etkileyen önemli bir değişkendir. Kümeler için önem derecesi % 95'in üzerinde iken, birinci küme için önem derecesi % 90'ın altındadır ve ayırt edici bir özelliğe sahip değildir. Tablo 4.43 incelendiğinde birinci ve dördüncü kümede kadınların, ikinci ve üçüncü kümede ise erkeklerin ağırlıkta olduğu görülmektedir.

Tablo 4.44'te medeni durumun kümeler üzerindeki etkisi incelenmiştir.

Tablo 4.44. Medeni Durumun Kümeler Üzerindeki Etkisi

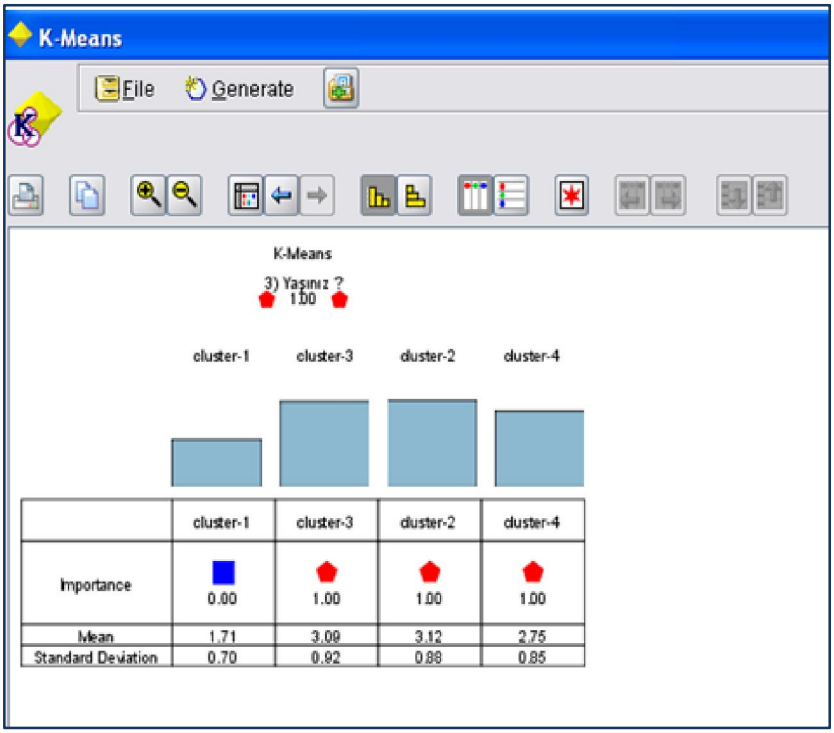


	cluster-1	cluster-3	cluster-2	cluster-4
Importance	1.00	1.00	1.00	1.00
Mean	1.00	2.00	2.00	2.00
Standard Deviation	0.00	0.07	0.00	0.00

Medeni durum değişkeni her dört kümeyi de etkileyen önemli bir değişkendir. Kümeler için önem derecesi %95'in üzerindedir. Tablo 4.44 incelendiğinde birinci kümede bekar kullanıcıların, ikinci, üçüncü ve dördüncü kümede ise evli kullanıcıların ağırlıkta olduğu görülmektedir.

Tablo 4.45'te yaş değişkenin kümeler üzerindeki etkisi incelenmiştir.

Tablo 4.45. Yaşın Kümeler Üzerindeki Etkisi

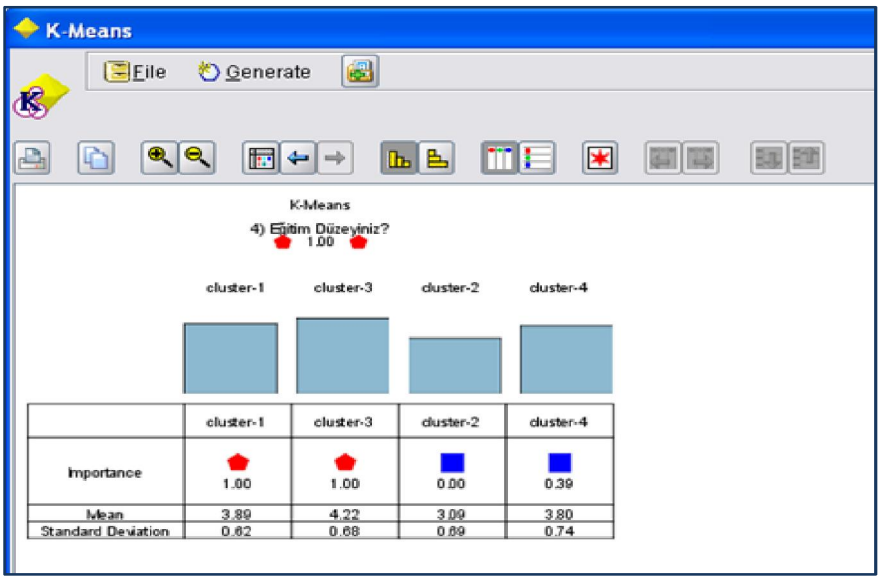


	cluster-1	cluster-3	cluster-2	cluster-4
Importance	0.00	1.00	1.00	1.00
Mean	1.71	3.09	3.12	2.75
Standard Deviation	0.70	0.92	0.88	0.85

Yaş değişkeni ikinci, üçüncü ve dördüncü kümeleri etkileyen önemli bir değişkendir. Üç küme için önem derecesi % 95'in üzerinde iken birinci küme için ayırt edici bir özelliği bulunmamaktadır. Tablo 4.45 incelendiğinde birinci kümede 26-35 yaş aralığındaki kullanıcıların, ikinci, üçüncü ve dördüncü kümede ise 36-45 yaş aralığındaki kullanıcıların ağırlıkta olduğu görülmektedir.

Tablo 4.46'da eğitim düzeyinin kümeler üzerindeki etkisi incelenmiştir.

Tablo 4.46. Eğitim Düzeyinin Kümeler Üzerindeki Etkisi

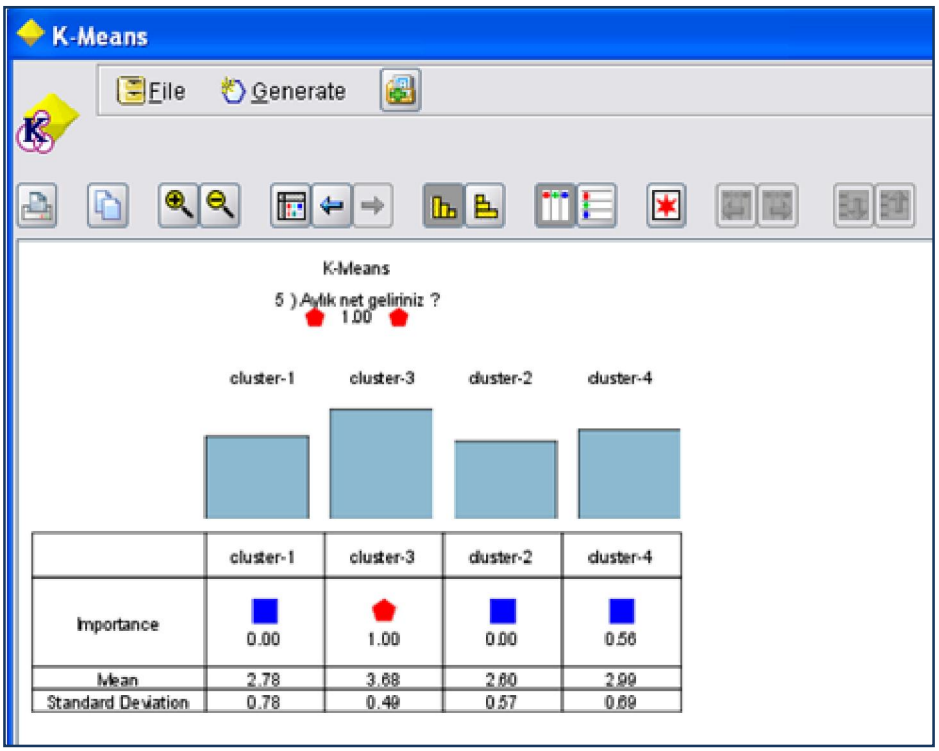


	cluster-1	cluster-3	cluster-2	cluster-4
Importance	1.00	1.00	0.00	0.39
Mean	3.89	4.22	3.09	3.80
Standard Deviation	0.62	0.88	0.89	0.74

Çalışmada eğitim düzeyinin de kümeler üzerindeki etkisi incelenmiştir. Eğitim düzeyi değişkeni birinci ve üçüncü kümeyi etkileyen önemli bir değişkendir. Kümeler için önem derecesi birinci ve üçüncü küme için %95'in üzerinde iken, ikinci ve dördüncü küme için önem derecesi % 90'ın altındadır ve ayırt edici bir özelliği bulunmamaktadır. Tablo 4.46 incelendiğinde birinci, üçüncü ve dördüncü kümede üniversite mezunu kullanıcıların, ikinci kümede ise lise mezunu kullanıcıların ağırlıkta olduğu görülmektedir.

Tablo 4.47'de aylık gelir düzeyinin kümeler üzerindeki etkisi incelenmiştir.

Tablo 4.47. Aylık Gelir Düzeyinin Kümeler Üzerindeki Etkisi

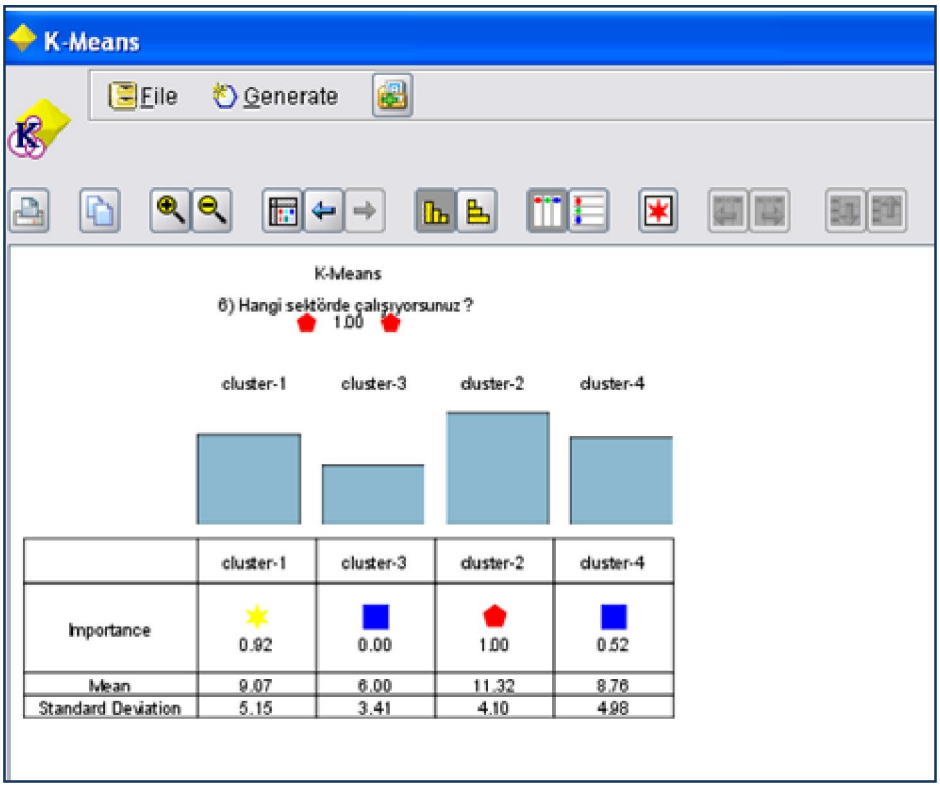


	cluster-1	cluster-3	cluster-2	cluster-4
Importance	0.00	1.00	0.00	0.56
Mean	2.78	3.68	2.60	2.99
Standard Deviation	0.78	0.49	0.57	0.69

Aylık gelir düzeyi değişkeni sadece üçüncü kümeyi etkileyen önemli bir değişkendir. Kümeler için önem derecesi sadece üçüncü küme için %95'in üzerinde iken birinci, ikinci ve dördüncü küme için önem derecesi % 90'ın altındadır ve ayırt edici bir özelliği bulunmamaktadır. Tablo 4.47 incelendiğinde birinci, ikinci ve dördüncü kümede gelir düzeyi 1501-2500 TL arasındaki kullanıcıların, üçüncü kümede ise gelir düzeyi 2500 TL üzerinde olan kullanıcıların ağırlıkta olduğu görülmektedir.

Tablo 4.48'de kullanıcıların çalıştıkları sektörün kümeler üzerindeki etkisi incelenmiştir.

Tablo 4.48. Çalışılan Sektörün Kümeler Üzerindeki Etkisi

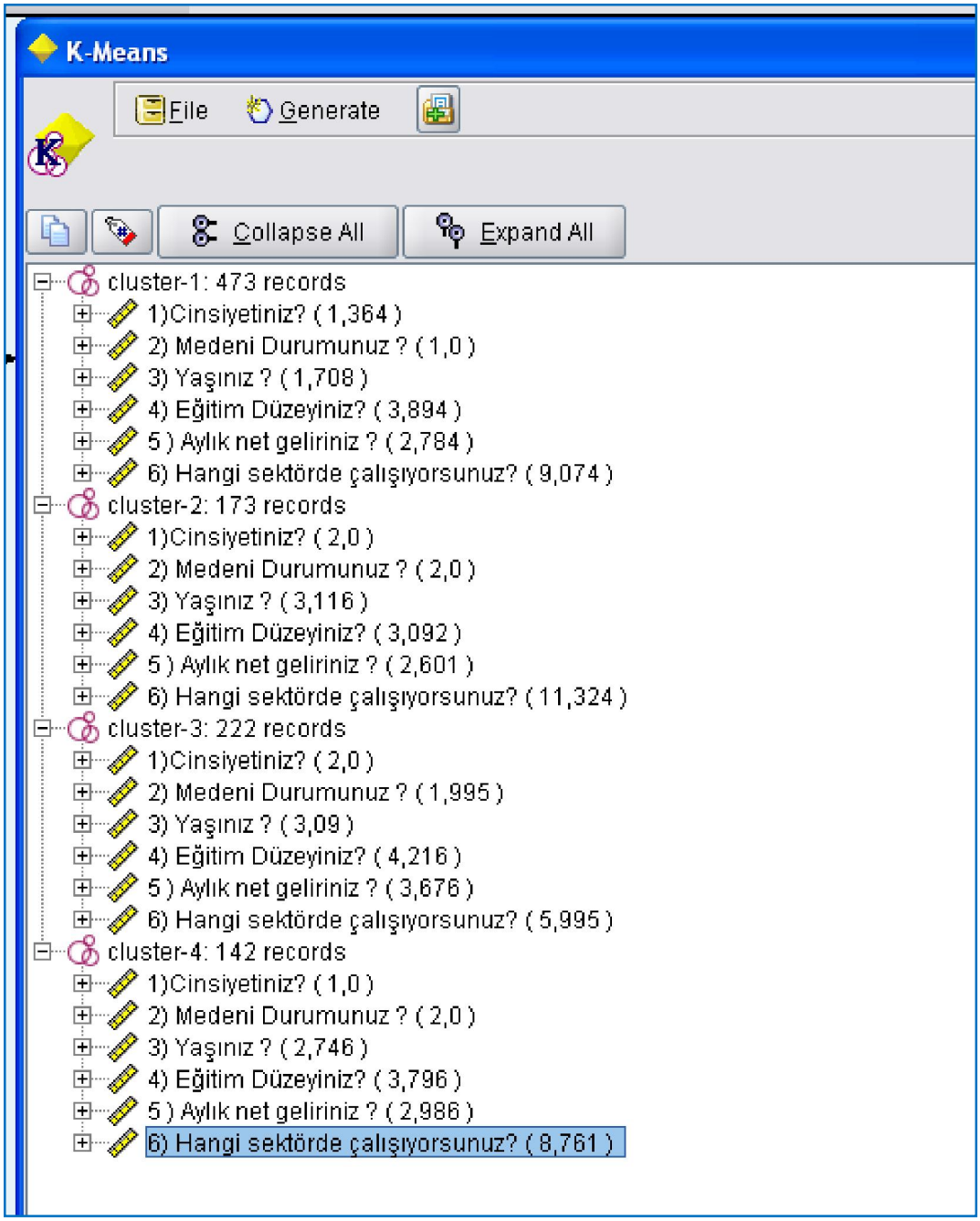


	cluster-1	cluster-3	cluster-2	cluster-4
Importance	0.92	0.00	1.00	0.52
Mean	9.07	6.00	11.32	8.76
Standard Deviation	5.15	3.41	4.10	4.98

Son olarak kullanıcıların çalıştıkları sektöre göre kümeler arasında bir farklılık olup olmadığı incelenmiştir. Analiz sonuçlarına bakıldığında; çalışılan sektör değişkeni birinci ve ikinci kümeyi etkileyen önemli bir değişkendir. Kümeler için önem derecesi sadece ikinci küme için % 95'in üzerinde iken, birinci küme için önem derecesi % 90'ın üzerindedir ve ayırt edici bir özelliğe sahiptir. Üçüncü ve dördüncü kümeler için ise önem derecesi % 90'ın altındadır ve ayırt edici bir özeliği bulunmamaktadır. Tablo 4.48 incelendiğinde birinci kümede sağlık sektöründe yer alan kullanıcıların, ikinci kümede perakende, üçüncü kümede otomotiv, dördüncü kümede ise gıda ve sağlık sektöründe çalışan kullanıcıların ağırlıkta olduğu görülmektedir.

Tablo 4.49 ‘da kümelerdeki kişi sayılara ilişkin dağılım yer almaktadır.

Tablo 4. 49. Kümelere Ait Dağılım



Küme sayılarında dağılımlara bakıldığında; 1.kümede 473 kayıt, 2. kümede 173 kayıt, 3. Kümede 222, 4. Kümede ise 142 kayıt söz konusudur. Sonuçlar incelendiğinde birinci kümede kişi sayısının daha fazla olduğu görülmektedir.

4.1.2.5. Modelin Kullanılması

Kurulan ve geçerliliği kabul edilen model, doğrudan bir uygulama olabileceği gibi bir başka uygulamanın alt parçası olarak da kullanılabilir. Kurulan modeller bankacılık sektöründe risk analizi, kredi değerlendirme, kampanya yönetimleri, CRM ile ilgili olarak bir uygulamanın içinde kullanılabilir. Bu çalışmada elde edilen sonuçlar internet bankacılığı kullanıcılarının profilini çıkartmak için kullanılmıştır. Ürünlerle ilgili kampanyalar yapıldığında hedeflenen müşterilere öncelik verilebilir.

Çalışmada elde sonuçlar bakıldığında, bankalar internet bankacılığı kullanan müşterilerinde; elde edilen bu kümeleri dikkate alarak bazı kampanyalar gerçekleştirebilir.

Örneğin; birinci kümedeki üniversite mezunu, gelir düzeyi 1501- 2500 TL arasında olan, sağlık sektöründe çalışan 26-35 yaş aralığındaki bekar kadınlara, bireysel emeklilik ürünleri ilgili bilgilendirme yapılabilir. Bunun yanı sıra alışverişlerinde tercih edebilecekleri, kredi kart promosyonlarıyla ilgili mail gönderilebilir ya da internet bankacılığı sayfasına girdiklerinde, rahatlıkla görebilecekleri çeşitli alışveriş promosyonu çalışmaları yapılabilir.

İkinci kümede, birinci kümenin aksine lise mezunu, gelir düzeyi 1501-2500 TL arasında olan, 36-45 yaş aralığındaki perakende sektöründe çalışan evli erkekler yer almaktadır. Erkeklerin alışveriş konusunda bayanlardan daha az ilgili olduğu göz önüne alınırsa, bu kümedeki banka müşterilerine, ihtiyaç kredisi, konut kredileri, kredili mevduat hesabı, otomatik ödeme kampanyaları gibi olanaklar sunulabilir.

Üçüncü kümede yine 36-45 yaş aralığındaki evli erkek müşteriler yer almaktadır. Burada ikinci kümeden farklı olarak, üniversite mezunu, otomotiv sektöründe çalışan erkek müşterilere, taşıt kredileri konusundaki kampanya ve promosyonlar sağlanabilir. Hem çalıştıkları sektör hem de gelir düzeyi olarak diğer kümelerden farklı olan bu kümedeki banka müşterileri, sektörleri dolayısıyla taşıt kampanyalarıyla ilgili olarak banka bilinirliğinin artmasında etkili olabilir. Gelir

düzeyi 2500 TL ve üzerinde olan bu kümeden müşterilere, vadeli işlemler, repo, likit hesap, altın hesabı, mevduat fon hesaplarıyla ilgili çeşitli bilgilendirmeler verilebilir ve ürün çeşitliliği sağlanabilir.

Dördüncü kümede yer alan üniversite mezunu, gelir düzeyi 1501- 2500 TL arasındaki sağlık sektöründe çalışan 36-45 yaş aralığındaki evli kadınlara, özellikle eczacı paketi, hayat sigortası gibi sektörleriyle ilgili olarak paketler hazırlanabilir, yine kadın müşterilere yönelik hazırlanacak kredi kartı promosyonları konusunda mail yollanabilir ve internet bankacılığı sayfasına girdiklerinde, rahatlıkla görebilecekleri çeşitli alışveriş promosyonu çalışmaları yapılabilir.

4.1.2.6. Modelin İzlenmesi

Sistemler her geçen gün değişmekte olduğu için, oluşan veriler üzerinde değişiklikler yapılması gerekmektedir. Zaman içinde elde edilen verilerde ortaya çıkan değişiklikler, kurulan modellerin sürekli olarak izlenmesini ve gerekiyorsa yeniden düzenlenmelidir. Tahmin edilen ve gözlenen değişkenler arasındaki farklılığı gösteren grafikler model sonuçlarının izlenmesinde kullanılan yararlı bir yöntemdir.

SONUÇ

Bu tez çalışmasında son dönemlerde yaygın olarak kullanılmakta olan Veri Madenciliği tekniği ele alınmış ve bankacılık sektöründe veri madenciliği tekniklerinin kullanımıyla ilgili bir uygulama gerçekleştirilmiştir. Veri Madenciliği ve İnternet Bankacılığı kavramları ile uygulamanın temelini oluşturan Birliktelik Kuralları Analizi ve Kümeleme analizi ayrıntılı olarak ele alınmıştır.

Veri madenciliği tekniği Türkiye'nin 5 ilinde bulunan, internet bankacılığı kullanan ve kullanmayan 18 yaş ve üzeri 1500 katılımcının verileri üzerinden uygulanmıştır. Çalışmada veri toplama işlemi anket formuyla yapılmıştır. Anketler katılımcılara e-mail ve yüz yüze görüşme yoluyla uygulanmış ve veriler elde edilmiştir. Tezin amacı veri madenciliği tekniklerinin CRISP-DM süreci ile ele alınarak; internet bankacılığı kullanıcı profilinin çıkartılması, gerekli analizlerin yapılması ve veriden anlamlı bilginin çıkarılması aşamaları için uygulanmasını gerçekleştirmektir.

Yapılan çeşitli veri düzenlemelerinden sonra verinin nasıl toplandığı, düzenleme aşamasında hangi işlemlere tabi tutulduğu hakkındaki detaylı bilgiler, yapılan uygulamalar ve varılan sonuçlar açıklanmıştır. Analizde IBM Modeler programı kullanılmıştır.

Çalışmada öncelikle müşterilerin demografik özelliklerine göre profilleri belirlenmiştir. Analizde yer alan her bir değişkenin istatistikleri incelenmiştir. Değişkenlere ait grafikler, değişkenlerin tipi, değişkenlerin minimum ve maksimum değerleri, ortalamaları, standart sapmaları ve çarpıklık değerlerine ilişkin grafikler beşinci bölümde ayrıntılı olarak sunulmuştur. İnternet bankacılığı kullanan ve kullanmayan müşteri profilleri ayrıntılı olarak incelenmiştir.

Bir sonraki aşamada müşteri verileri kullanılarak Birliktelik Kuralları Analizi ile müşterilerin internet bankacılığı kullanımna ilişkin yaklaşımları belirlenmeye çalışılmıştır. Bunu gerçekleştirmek için veriler yeniden kodlanmış ve gerekli düzenlemeler yapılmıştır.

Çalışmada Kümeleme Analizi ile müşteriler bölümlendirilmeye çalışılmış, K-ortalamalar tekniği kullanılmıştır. K-ortalamalar tekniği ile dört küme elde edilmiş ve müşterilerin demografik özelliklerinin her küme üzerindeki etkileri ayrı ayrı incelenmiştir.

Çalışma sonucunda elde edilen bilgiler ile internet bankacılığı kullanıcıları ile internet bankacılığı kullanmayan müşteriler arasında fark elde edilebilmektedir. İnternet bankacılığı kullanmayan kullanıcı profili elde edilmiş ve hangi sebeple internet bankacılığı kullanmadıkları belirlenebilmiştir.

Elde edilen sonuçlar incelendiğinde; çoğunluğu 46 yaş ve üzerindeki ortaokul mezunu, gelir düzeyi 751-1500 TL arasında olan emekli kadın ve erkek banka müşterilerinin internet bankacılığını kullanmadıkları görülmüştür. Banka müşterilerinin, internet bankacılığını tercih etmeme sebepleri incelenmiş; müşterilerin büyük çoğunluğunun internet bankacılığını güvenilir bulmadığı, internet bankacılığını kullanmanın zor olduğunu düşündüğü görülmüştür. Bu durum özellikle orta yaş üzerinde banka müşterilere yönelik yapılacak olan kampanya çalışmalarının internet üzerinde değil, şube yoluyla verilmesi gerektiğini göstermektedir.

Çalışmada internet bankacılığı kullanıcıların profilleri de incelenmiş, elde edilen sepet analizi sonucunda internet bankacılığı kullanıcılarının hangi ürünleri etkin kullandığı, hangi hesaplar üzerinden işlem yaptığı ve bu konuda nasıl farklılıklar, yenilikler yapılabileceği sonucuna ulaşılmıştır.

İnternet bankacılığı kullanan banka müşterilerine profillerine bakıldığında;

Analize katılan internet bankacılığı kullanıcılarının profilleri incelendiğinde; ağırlıklı olarak 26-35 yaş grubunda yer alan, lise ve üniversite mezunu, 1501- 2500 TL gelire sahip, kişilerden oluştuğu sonucuna ulaşılmıştır. Cinsiyetlerine göre profiller incelendiğinde; internet bankacılığı kullanan kadın ve erkek banka müşterilerinin dağılımlarının birbirine yakın olduğu, internet bankacılığı kullanımında cinsiyete göre çok büyük farklılıklar bulunmadığı görülmüştür.

Çalışmada analize katılan banka müşterileri arasında sektör bakımından çok büyük farklılıklar olmadığı, hemen hemen her sektördeki kadın ve erkek banka

müşterilerinin internet bankacılığı kullanmakta olduğu görülmüştür. Çalışma sonucunda bu kişilerin çoğunluğunun uzun yıllardır internet kullanan, ancak 1-4 yıldır internet bankacılığı kullanan, ortalama 2 internet bankacılığı hesabı olan kişiler olduğu belirlenmiştir. Bu kullanıcıların ağırlıklı olarak haftada bir kez internete bağlandıkları, büyük çoğunluğunun, daha çok cep telefonları ve kişisel PC ile internet bankacılığına bağlanmayı tercih ettiği görülmüştür.

Kullanıcıların internet bankacılığını kullandıkları bankalara ilişkin sıralamalara bakıldığında en çok Garanti Banka'sını tercih ettikleri, en az ise Vakıfbank, HSBC ve Kuveyt Türk Bankasının internet bankacılığını tercih ettikleri görülmektedir.

Bunun yanı sıra internet bankacılığı kullanıcılarının vadesiz TL ve kredi kartı hesaplarını etkin bir şekilde kullandıkları, internet bankacılığı kullanımında 7/24 hızlı ve kesintisiz bankacılık işlemlerini gerçekleştirebilmelerinin, maaşını o bankadan almasının ve şubeye gitmeden kolay bankacılık işlemleri gerçekleştirebilmesinin etkili olduğu görülmektedir.

İnternet bankacılığı müşterilerinin, bir bankanın internet sitesinde aradıkları özelliklerin başında hizmet kalitesi, güvenlik, kolay ve pratik kullanılabilir olmasının geldiği görülmektedir.

Banka müşterilerinin genel olarak bankaların internet bankacılığını kullanmaktan memnun oldukları, internet bankacılığı kullanırken bir problem yaşamadıkları görülmektedir. İnternet bankacılığı kullanan banka müşterilerinin % 50'sinden fazlası kısmen de olsa, internet bankacılığının geleneksel bankacılığın yerini alacağını düşünmektedir.

Banka müşterileri rekabet ortamında çeşitli ürünleri, daha iyi servis ve daha uygun fırsatlarla kullanmak istemektedir. Bu durum sonucunda, bankaların pazarlama tekniklerini geliştirmeleri ve müşteriye farklı alternatifler sunmaları gerekmektedir.

Elde edilen kümelerle hangi müşteri gruplarının hedeflenmesi gerektiği belirlenmiştir. Bu sonuçlar reklam kampanyalarında kullanılabilecektir.

Sonuç olarak bankaların sektörde rekabet avantajı sağlayarak uzun süre ayakta kalabilmeleri için müşterilerini doğru bir şekilde tanımaları ve doğru analizler yapmalar gerekmektedir.

Sepet analizi ve kümeleme analizi sonucunda elde edilen sonuçlar, reklam stratejileri belirlemede, CRM, müşteri profillerinin analiz edilerek çapraz satış tahminlerinin yapılması, yeni müşterilere ulaşabilmek için etkili faktörlerin belirlenmesi, hedef pazarın belirlenmesi, müşteri değerleme, müşteri segmentasyonu ve müşteri ilişkileri yönetimi için kullanılabilir.

KAYNAKÇA

Agrawal Rakesh, Srikant Ramakrishnan: Fast Algorithms for Mining Association Rules, **20. VLDB Konferansı**, Sili,1994, s.487-499.

Agrawal, Rakesh, Tomasz Imielinski, Arun N. Swami: "Mining Association Rules Between Sets of Items in Large Databases", **In Proceedings of the ACM SIGMOD International Conference on Management of Data**, USA, 1993, s.207-216.

Akbulut, Sinem: "**Veri Madenciliği Teknikleri İle Bir Kozmetik Markanın Ayrılan Müşteri Analizi Ve Müşteri Segmentasyonu**", Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Anabilim Dalı, Ankara, 2006.

Akgüden, Yılmaz, Burak Erşahin: **Veri Madenciliği, Veriden Bilgiye Masraftan Değere Online Kitap**, s.35, (Çevrimiçi) http://www.arge.com/Userfiles/File/Veri%20Madencili%C4%9Fi.pdf, 12 Mart 2012.

Akın, Yasemin Koldere: "**Veri Madenciliğinde Kümeleme Algoritmaları Ve Kümeleme Analizi**", Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı, İstanbul, 2006.

Akpınar, Haldun: "Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği". **İ.Ü. İşletme Fakültesi Dergisi**, Cilt: 29, Sayı:1, 2000, s. 1-22.

Albayrak, Ali Sait, "Veri Madenciliği: Karar Ağacı Algoritmaları Ve

Şebnem Koltan Yılmaz: İMKB Verileri Üzerine Bir Uygulama", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2009, Cilt:14, Sayı:1, s.31-52.

Albayrak, Ali Sait: "Türkiye'de Yerli Ve Yabancı Ticaret Bankalarının Finansal Etkinliğe Göre Sınıflandırılması: Karar Ağacı, Lojistik Regresyon Ve Diskriminant Analizi Modellerinin Bir Karşılaştırılması", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2009, Cilt:14, Sayı:1, s.113-139.

Apaydın, Ethem: Zeki Veri Madenciliği: Ham Veriden Altın Bilgiye Ulaşma Yöntemleri", Bilişim 2000 Egitim Semineri, (Çevrimiçi) http://www.cmpe.boun.edu.tr/~ethem, 12 Mart 2013.

Aşan, Zerrin: "Kredi Kartı Kullanan Müşterilerin Sosyo- Ekonomik Özelliklerinin Kümeleme Analizi İle İncelenmesi", **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, Cilt: 17, 2007, s. 256-267.

Ay, Derya, İbrahim Çil: "Migros Türk A.Ş.'de Birliktelik Kurallarının Yerleşim Düzeni Planlamada Kullanılması", **Endüstri Mühendisliği Dergisi,** Cilt: 21 Sayı: 2, 2010. s.14-29.

Ay, Derya: "**Veri Madenciliği ve Apriori Algoritması İle Bir Süpermarket Uygulaması**", Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Sakarya, 2009.

Aydoğan, Emel Kızılkaya, Cevriye Gencer, Sinem Akbulut: "Churn Analysis And Customer Segmentation Of A Cosmetics Brand Using Data Mining Techniques", **Sigma Mühendislik ve Fen Bilimleri Dergisi,** Cilt 26, Sayı:1, 2008, s.43-45.

Aynekin, Gülşah: "**İnternet İçerik Madenciliğinde Yapay Sinir Ağları Ve Bir Uygulama**", Yüksek Lisans Tezi, Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Bursa, 2006.

Baykal, Abdullah: "Veri Madenciliği Uygulama Alanları", **D.E.Ü.Ziya Gökalp Eğitim Fakültesi Dergisi,** Sayı:7, 2006, s.95-107.

Baykal, Abdullah: "Web Madenciliği ile Log Analizi", **D.Ü.Ziya Gökalp Eğitim Fakültesi Dergisi,** Sayı: 8, 2007, s.129-145.

Berry, Michael J.A., Gordon Linoff: **Data Mining Solutions**, U.S.A., Wiley Computer Publishing, 1998.

Berson, Alex, York, Stephen J. Smith: **Data Warehousing, Data Mining & OLAP**, New McGraw-Hill Companies, 1998.

Bilekdemir, Gökhan: "**Veri Madenciliği Tekniklerini Kullanarak Üretim Süresi Tahmini Ve Bir Uygulama**", Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Endüstri İşletmeleri Bilim Dalı, İzmir, 2010.

Bilen, Hamdi: "**Bankacılık Sektöründe Personel Seçimi Ve Performans Değerlendirilmesine İlişkin Veri Madenciliği Uygulaması"**, Yüksek Lisans Tezi, Gazi

Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, 2009. .

Biyometrik Tanıma Sistemleri: http://www.ardbilisim.com.tr/projeler/biyometri/ (Çevrimiçi) 15 Ocak 2013.

Breault, Joseph, Colin Goodall, Peter Fos: "Data Mining a Diabetic Data Warehouse", **Artificial Intelligence in Medicine**, 2002, ss. 37-54.

Cabena, Peter, et. al.: **Discovering Data Mining : From Concept to Implementation,** USA, International Business Machines Corperation, 1998.

Carus, Aydın, Altan Mesut: "Web Kullanım Madenciliği Uygulaması", **II. Mühendislik Bilimleri Genç Araştırmacılar Kongresi**, İstanbul 17–19 Kasım 2005.

Carus, Aydın, v.d.: "Cep Gazetecisi: Mobil Cihazlar İle Çalışan Bir Web Madenciliği Uygulaması", **Elektrik-Elektronik-Bilgisayar Mühendisliği 12. Ulusal Kongresi,** 2007, Eskişehir.

Chuang, Shui-Lung, Lee-Feng Chien: "Enriching Web Taxonomies Through Subject Categorization Of Query Terms From Search Engine Logs". **Decision Support Systems, Volume: 35 (2003),** Issue: 1 s.113-127.

Chye, Koh Hian, Chan Kin Leong Gerry: "Data Mining and Customer Relationship Management in the BankingIndustry", **Singapore Management Review**, Volume 24, Issue 2, 2002, s.1-27.

Çalış, Aslı: "**Veri Madenciliği Yaklaşımı İle Bireysel Müşterilerin Kredi Ödeme Performanslarının Değerlendirilmesi,** Yüksek Lisans Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, 2013.

Dağlı, Recep Murat: "**Banka Müşterilerinin İnternet Bankacılığına İlişkin Kanaatlerinin İncelenmesi Ve Konu İle İlgili Pilot Bir Araştırma**", Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Pazarlama Anabilim Dalı, İstanbul, 2007.

"Data Mining", (Çevrimiçi) http://www.spss.com/data_mining/, 20 Ocak 2012.

Dener, Murat, v.d.: "Açık Kaynak Kodlu Veri Madenciliği Programları: Weka'da Örnek Uygulama", **Akademik Bilişim'09 - XI. Akademik Bilişim Konferansı Bildirileri**, 11-13 Şubat 2009, Harran Üniversitesi, Şanlıurfa, (Çevrimiçi) ab.org.tr/ab09/bildiri/42.pdf, 18 Kasım 2011.

DeYoung, Robert: **Chicago Fed Letter**, Federal Reserve Bank of Chicago, No.163, ISSN 0895-0164, (Çevrimiçi) http://www.chicagofed.org/digital_assets/publications/chicago_fed_letter/2001/cflmar2001_163.pdf 05 Mayıs 2013.

Dinçer, Esra, Nevcihan Duru: "Gırtlak Kanseri Ameliyat Verilerinin K-means Yöntemiyle Analizi", **TBV Bilgisayar Bilimleri ve Mühendisliği Dergisi**, Aralık, 2006, s.53-62.

Doğan, Şengül: "**Veri Madenciliği Kullanarak Biyokimya Verilerinden Hastalık Teşhisi**",Yüksek Lisans Tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Biyomühendislik Anabilim Dalı, Elazığ, 2007.

Dolgun, M. Özgür, Tülin Özdemir, Doruk Oğuz: Veri Madenciliğinde Yapısal Olmayan Verinin Analizi: Metin Ve Web Madenciliği, **İstatistikçiler Dergisi,** Sayı:2, 2009, s.48-58.

Dolgun, M. Özgür: "**Büyük Alışveriş Merkezleri İçin Veri Madenciliği Uygulamaları**", Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Ankara, 2006.

Dönmez, Seçil: "**Bayi Performans Değerlendirmesinde Bir Veri Madenciliği Uygulaması**", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, 2008.

DuMouchel, William: "Bayesian Data Mining In Large Frequency Tables, With An Application To The FDA Spontaneous", **American Statistician**, 53 (3), 1999, s.177.

Dunham, Margaret: **Data Mining Introductory and Advanced Topics**, Prentice Hall, USA, 2003.

Eker, Hakan: "Veri Madenciliği veya Bilgi Keşfi", (Çevrimiçi) http://www.bilgiyonetimi.org, 03 Mart 2013.

Erçişli, Nezih: "**İnternet Bankacılığı Uygulamaları Ve İşletme Performansı Üzerine Etkileri: Denizbank Örnek Olay Çalışması**", Yüksek Lisans Tezi, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, Konya, 2007.

Esen, Fevzi: **"Veri Tabanlarında Bilgi Keşfi: Veri Madenciliği Ve Bir Sağlık Uygulaması"**, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2009.

Etzioni, Oren: **The World Wide Web: Quagmire Or Gold Mine. Communcations of the ACM**, Vol.39, 1996, s.64-67.

Ezerçe, Ayşegül: **"Müşteri İlişkileri Yönetimi (Crm) Ve Veri Madenciliği (Data Mining) : Tekstil Sektöründe Bir Uygulama"**, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2008.

Feelders, Ad, Hennie Daniels, Marcel Holsheimer: "Methodological and Practical Aspects of Data Mining", **Information & Management**, Vol:37, 2000, s.271-281.

Gefen, David, Detmar W. Straub: "The Relative Importance of Perceived Ease-ofUse in IS Adoption: A Study of E-Commerce Adoption," **Journal of the Association for Information Systems**, 1(8), 2000, s.1-30.

Göral, Mustafa Aykut: "**Kredi Kartı Başvuru Aşamasında Sahtecilik Tespiti İçin Bir Veri Madenciliği Modeli**", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2007.

Gupta, Gaurav, Geetika Hans, Tamanna Sehgal: Applications & Trends In Data Mining, (Çevrimiçi) www.rimtengg.com/coit2007/.../94.pdf . 20 Şubat 2012 .

Gülce, Cenk: **"Veri Madenciliğinde Apriori Algoritması Ve Apriori Algoritmasının Farklı Veri Kümelerinde Uygulanması"**, Yüksek Lisans Tezi, Trakya Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Edirne, 2010.

Gürbüz, Feyza, Lale Özbakır, Hüseyin Yapıcı: "Türkiye'de Bir Havayolu İşletmesine Ait Parça Söküm Raporlarına İlişkin Veri Madenciliği Uygulaması", **Gazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi**, Cilt 24, No 1, 2009, s.73-78.

Gürcan, Fatih: "**Web İçerik Madenciliği Ve Konu Sınıflandırması**" Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Trabzon, 2009.

Gürsakal, Necmi: **Büyük Veri**, Bursa, Dora Yayın Evi, 2003.

Gürsoy, U. Tuğba Şimşek: **Veri Madenciliği ve Bilgi Keşfi**, 1.Baskı, Ankara, Pegem Akademi, 2009.

Haberal, İsmail: "**Veri Madenciliği Algorimaları Kullanılarak Web Günlük Erişimlerinin Analizi Uygulaması**", Yüksek Lisans Tezi, Başkent Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Ankara, 2007.

Hair, Joseph, et. al.: **Multivariate Data Analysis**, Prentice Hall, U.S.A., 1998.

Han Jiawei, Micheline Kamber: **Data Mining Concepts and Techniques**, Morgan Kauffmann Publishers Inc., 2001.

Işık, Meltem, Ali Yılmaz Çamurcu: "Web Belgeleri Kümelemede Benzerlik Ve Uzaklık Ölçütleri Başarılarının Karşılaştırılması"**, Marmara Üniversitesi Fen Bilimleri Enstitüsü Dergisi**, Sayı:20, 2008, s.35-49.

Jansen, Bernard J., Amanda Spink: "An Analysis Of Web Searching By European Alltheweb.Com Users." **Information Processing and Management,** 41, 2005, s.361–381.

Kalıkov, Anarbek: "**Veri Madenciliği ve Bir E-Ticaret Uygulaması"**, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Elektronik ve Bilgisayar Eğitimi, Mayıs 2006, Ankara.

Karabağ, Rahim, v.d.: "Kullanıcı Davranış Analizi ile Nüfuz Tespit Modeli (KDA-NTM)", ABG2005, **Ağ ve Bilgi Güvenliği Sempozyumu**, 9-12 Haziran 2005, (Çevrimiçi) http://www.emo.org.tr/ekler/00871c1fc726f0b_ek.pdf. 20 Kasım 2011.

Karabatak, Murat, Meryem Koç: 5th International Computer & Instructional Technologies Symposium, 22-24 September 2011, Fırat University, Elazığ- Turkey, s.3.

Karadağ, Anıl, Hüseyin Takçı: "**Metin Madenciliği ile Benzer Haber Tespiti**", AB 2010, Akademik Bilişim, Muğla Üniversitesi, Muğla, 10-12 Şubat 2010.

Karadeniz, Nur: "**Müşteri İlişkileri Yönetimi Açısından Veri Madenciliği Yöntemi Ve Hizmet Sektörü Üzerine Bir Uygulama**", Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı, İstanbul, 2008.

Kaya, Halil, Kemal Köymen: "Veri Madenciliği Kavramı ve Uygulama Alanları", **Fırat Üniversitesi Doğu Anadolu Bölgesi Araştırmaları Dergisi**, Cilt:2, Sayı:3, 2008, s.159-164.

Kiremitçi, Barış: "**Veri Ambarlarında Veri Madenciliği Ve Ulaştırma-Lojistik Sektöründe Bir Uygulama**", Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2005.

Koh, Hian Chye, Gerald Tan, "Data Mining Applications in Healthcare", **Journal of Healthcare Information Management,** Vol, 2011, (Çevrimiçi)https://www.himss.org/content/files/jhim/19-2/datamining.pdf 15 Kasım 2011.

Kolay, Gökçe: "**İşletmelerde Bilgi Sistemleri Verimliliğini Arttırmada Veri Madenciliği Yöntemi: Bir Simülasyon Çalışması"**, Yüksek Lisans Tezi,

Zonguldak Karaelmas Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Zonguldak, 2006.

Kovan, Özcan: "**Orta Yaş Üzeri Banka Müşterilerine İnternet Bankacılığı Ürününü Kullandırmaya Yönelik Uygun Arayüz Tasarlanması**", Yüksek Lisans Tezi, Beykent Üniversitesi Sosyal Bilimler Enstiitüsü, İletişim Ve Tasarım Ana Sanat Dalı İletişim Sanatları Ve Tasarım Bilim Dalı, İstanbul, 2012.

Koyuncugil, Ali Serhan, Nermin Özgülbaş: "Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları", **Bilişim Teknolojileri Dergisi**, Mayıs 2009, Cilt:2, Sayı: 2, s.21-23.

Larose, Daniel T.: **Discovering Knowledge in Data: An Introduction to Data Mining,** John Wiley & Sons Inc., 2005, s.42-70.

Nagabhushana, Sujial: Data Warehousing OLAP and Data Mining, New Delhi: New Age International Ltd. Publishers, 2006, s.293-294.

Nakip, Mahir: **Pazarlama Araştırmaları**, Seçkin Yayınevi, Ankara, 2003.

Oğuz, Başak: **"Metin Madenciliği Teknikleri, Kullanılarak Kulak Burun Boğaz Hasta Bilgi Formlarının Analizi",** Yüksek Lisans Tezi, Akdeniz Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Ve Tıp Bilişimi Anabilim Dalı, 2009.

Oktay, Hikmet Mukadder: "**Web Kullanım Madenciliğinde Birliktelik Kurallarının Uygulanması**",Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı İstanbul, 2009.

Özcan, Zeynep Özge: "**Türkiye'de Elektronik Bankacılık: İnternet Bankacılığı Üzerine Bir Çalışma"**, Yüksek Lisans Tezi, Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Muhasebe Ve Finansman Bilim Dalı, Sakarya, 2007.

Özçakır, Feridun Cemal: "**Müşteri İşlemlerindeki Birlikteliklerin Belirlenmesinde Veri Madenciliği Uygulaması**", Yüksek Lisans Tezi, , Marmara Üniversitesi Fen Bilimleri Enstitüsü, 2006.

Özçınar, Hüseyin: "**KPSS Sonuçlarının Veri Madenciliği Yöntemleriyle Tahmin Edilmesi**", Yüksek Lisans Tezi, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Denizli, 2006.

Özdal, Tuğba: "**Bankacılık Sektöründe Müşteri İlişkileri Yönetimi Ve İnternet Bankacılığında Müşterilerin Seçimlerine Etki Eden Faktörler**", Yüksek Lisans Tezi, Ufuk Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Ankara, 2011.

Özdamar, Kazım: **Paket Programlar İle İstatistiksel Veri Analizi I**, 5. Baskı, Kaan Kitabevi, Eskişehir, 2004.

Özdemir, Recep: "**KOBİ'lerin İnternet Bankacılığına Adaptasyonu: Orta Anadolu Bölgesi'nde Anket Uygulaması"**,

Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı, Zonguldak, 2010.

Özel, Tuğba: **Finans Sektöründe İstatistik ve Veri Madenciliği Uygulamaları**, SPSS sunumu, 27 Ocak 2012.

Özkan, Yalçın: **Veri Madenciliği Yönetimleri**, 1.Baskı, İstanbul, Papatya Yayıncılık, 2008.

Özkaya, Necla, Şeref Sağıroğlu: "Otomatik Parmakizi Tanıma Sistemlerinde Kullanılan Önişlemler İçin Yeni Yaklaşımlar", **Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi**, Cilt: 21, Sayı. 1, 2006, s. 13-19.

Özkaya, Necla, Şeref Sağıroğlu: "Parmak İzinden Yüz Tanıma Sistemleri", **Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi**, Cilt: 23, Sayı. 4, 2008, s. 787-789.

Özoğlu, Buket: "**Erciyes Üniversitesi, Tüketici Odaklı Yaklaşım İle İnternet Bankacılığı Faaliyetlerini Değerlendirmeye Yönelik Bir Uygulama**", Yüksek Lisans Tezi, Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, Kayseri, 2010.

Özseven, Turgut, Muharrem Düğenci: "LOG Analiz: Erişim Kayıt Dosyaları Analiz Yazılımı Ve GOP Üniversitesi Uygulaması", **Bilişim Teknolojileri Dergisi,** Cilt: 4, Sayı: 2, Mayıs 2011, s.55.

Özseven, Turgut: (Çevrimiçi) http://www.crispdm. org/CRISPWP-0 800.pdf, 20 Nisan 2012.

Pala, Emre, Burak Kartal: "Banka Müşterilerinin İnternet Bankacılığı İle İlgili Tutumlarına Yönelik Bir Pilot Araştırma", **Celal Bayar Üniversitesi İ.İ.B.F. Yönetim Ve Ekonomi Dergisi,** Yıl:2010, Cilt:17, Sayı:2, Manisa, s.45.

Pala, Emre: **Alternatif Dağıtım Kanallarından İnternet Bankacılığına Yönelik Müşteri Tercihlerinin İncelenmesi,** Yüksek Lisans Tezi, Celal Bayar Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı,2010.

Pande, Joshi Karuna: "Analysis of Data Mining Algorithms", (Çevrimiçi) http://ebiquity.umbc.edu/_file_directory_/papers/462.pdf 15.04.2013.

Parmak İzi Tarama Sistemleri, (Çevrimiçi) http://www.meyer.com.tr/parmakizi.htm 11 Ocak 2013.

Perumal, Vignesen, Bala Shanmugam: Internet Banking: Boon or Bane?, (Çevrimiçi), http://www.arraydev.com/commerce/JIBC/2004-12/Perumal.htm, 13 Şubat 2013.

Rygielski, Chris, Jyun-Chen Wang, David C. Yen: "Data Mining Techniques for Customer Relationship Management", **Technology in Society**, Volume 24, Issue 4, 2002, s.488.

Sağıroğlu, Şeref: "Parmak İzinden Yüz Tanıma", **Çerçeve Dergisi**, Sayı: 52, s.104, (Çevrimiçi)

http://www.musiad.org.tr/contentimages/arastirmalaryayin/pdf/cerceve_dergisi_52.pdf, 10 Mart 2013.

Sarıman, Güncel: "**Paralel Programlama İle Web Madenciliğinde Log Analizi**" Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Isparta, 2011.

Savaş, Serdar, Nursal Arıcı: "Web Tabanlı Uzaktan Eğitimde İki Farklı Öğretim Modelinin Öğrenci Başarısı Üzerindeki Etkilerinin İncelenmesi", **5. Uluslararası İleri Teknolojiler Sempozyumu (IATS'09),** 13-15 Mayıs 2009, Karabük.

Seçkin, Keziban: "**Metin Madenciliğinde Kullanılan Yöntemlerin Karşılaştırılması: Siyasi Parti Liderlerinin Grup Genel Toplantı Konuşmaları İle Bir Uygulama**",Yüksek Lisans Tezi, Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Üretim Yönetimi Ve Pazarlama Bilim Dalı, Sakarya, 2011.

Seidman, Claude: **Data Mining with Microsoft® SQL Server™ 2000 Technical Reference**, Microsoft Press, Redmond, WA, USA, 2001.

Sezen, Hayrettin Kemal: **Yöneylem Araştırması**, Bursa, Ekin Kitapevi, 2004.

Sezgin, M. Emre: **"İkili Kodlama Kuralına Dayalı Olarak Hazırlanan Multimedya Ders Yazılımının Fen Bilgisi Öğretimindeki, Akademik Başarıya, Öğrenme**

Düzeylerine ve Kalıcılığına Etkisi", Yüksek Lisans Tezi, Çukurova Üniversitesi, Sosyal Bilimler Enstitüsü, Eğitim Ana Bilim Dalı, 2002.

Sharma, Subbash: Applied Multivariate Techniques, John Wiley & Sons, USA, 1996, s.185.

Shearer, Colin: "The CRISP-DM Model: The New Blueprint for Data Mining," **Journal of Data Warehousing,** No:5, 2000, s.13-22.

Silahtaroğlu, Gökhan: **Kavram ve Algoritmalarıyla Temel Veri Madenciliği,** 1. Basım, İstanbul, Papatya Yayıncılık, Mayıs, 2008.

Singh, Harry S.: **Data Warehousing-Concepts, Technologies, Implementations and Management**, U.S.A., Prentice Hall PTR, 1998.

Sönmez, Yasin: **"Masaüstü Rehberlik Otomasyonu Gerçekleştirimi Ve Verilerinin Birliktelik Kuralı İle Değerlendirilmesi"**, Yüksek Lisans Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elazığ, 2012.

Şekeroğlu, Songül: "**Hizmet Sektöründe Bir Veri Madenciliği Uygulaması**", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, İstanbul, 2009.

Şen, Fatih: "**Veri Madenciliği İle Birliktelik Kurallarının Bulunması**", Yüksek Lisans Tezi, Sakarya

Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Ve Bilişim Mühendisliği Anabilim Dalı, Sakarya, 2009.

Şimşek, Umman Tuğba: "**Veri Madenciliği Ve Müşteri İlişkileri Yönetiminde(CRM) Bir Uygulama",** Doktora Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2006.

Tan, Pang-Ning, Michael Steinbach, Vipin Kumar: **Introduction to Data Mining**, Addison Wesley, 2005, Data Mining Association Analysis: Basic Concepts and Algorithms Lecture Notes, s.8.

Tezcanlar, Pelin: "**Müşteri İlişkileri Yönetimi Veri Madenciliği ve Bir Uygulama**",Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2007.

The Gartner Grup: (Çevrimiçi), www.gartner.com, 14 Ocak 2012.

Timor, Mehpare, Umman Tuğba Şimşek: "Veri Madenciliğinde Sepet Analizi İle Tüketici Davranışları Modellemesi",**Yönetim Dergisi,** Yıl 19, Sayı:59,Şubat 2008,Makale 396, (Çevrimiçi), www.iie.istanbul.edu.tr/ydsy/MakaleGoster.aspx?m=396, 13 Ekim 2011.

Timor, Mehpare, Ayşegül Ezerçe, Umman Tuğba Şimşek: "Müşteri Profili Ve Alışveriş Davranışlarını Belirlemede Kümeleme Ve Birliktelik Kuralları Analizi: Perakende Sektöründe Bir Uygulama", **Yönetim Dergisi,** Yıl 22, Sayı:68,Şubat 2011,s.129-148.

Tosun, Tuğba: "**Veri Madenciliği Teknikleriyle Kredi Kartlarında Müşteri Kaybetme Analizi,** Yüksek Lisans Tezi," İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, 2006.

Türkiye Bankalar Birliği, "İnternet Ve Mobil Bankacılık İstatistikleri Haziran 2013 Raporu" (çevrimiçi) http://www.tbb.org.tr/tr/banka-ve-sektor-bilgileri/istatistikiraporlar/Internet_ve_Mobil_Bankacilik_Istatistikleri/1274, 24 Temmuz 2013.

Tüzüntürk, Selim: "Veri Madenciliği ve İstatistik", **Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,** Cilt XXIX, Sayı: 1, 2010, s.65-90.

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

Vahaplar, Alper, Mustafa Murat İnceoğlu: "Veri Madenciliği ve Elektronik Ticaret", www.manisabilisim.org/dokuman/inceoglu.doc (Çevrimiçi), 20 Ocak 2012.

Vaughan,Liwen: "New Measurements For Search Engine Evaluation Proposed And Tested," **Information Processing and Management ,** Volume: 40, 2004, s.677–691.

Wah, Teh Ying, et. al.: "Data Mining in Computer Auditing, Insite "Where Parallels Intersect", **Informing Science**, 2002, s.1565.

Wang, John: "**Encyclopedia of Data Warehousing and Mining**", Idea Group Publishing, ISBN 1-59140-557-2, Hershey PA, 2006.

Wang, John: **Encyclopedia of Data Warehousing and Mining**, 2nd Edition, USA; Idea Group Inc., 2009.

World Internet Usage And Populatıon Statıstıcs: http://www.internetworldstats.com/stats.htm (Çevrimiçi) 25 Haziran 2013.

Yıldırım, Pınar, Mahmut Uludağ, Abdulkadir Görür: "Hastane Bilgi Sistemlerinde Veri Madenciliği", **Akademik Bilişim 2008,** Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, 30 Ocak - 01 Şubat 2008, s.429-431.

Yılmazel, Özgür: **Türkçe Metin Arama Sistemleri ve Veritabanlarının Karşılaştırılması**,(Çevrimiçi)http://seminer.linux.org.tr/wp-content/uploads/dilisimSunu.pdf, 25.09.2012.

Yüz Tanıma Sistemleri (Çevrimiçi), http://www.yuztanimasistemi.net/ 13 Ocak 2013.

Zeyrek, Özge Sevda: "**Müşteri İlişkileri Yönetiminde Veri Madenciliği Ve Bir Uygulama**", Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, İstanbul, 2006.

Zhoui, Zhi-Hua: "Three Perspectives Of Data Mining", **Artificial Intelligence,** Vol:143**,** 2003, s.139–146.

http://eticaretmag.com/big-data-buyuk-veri-nedir/ (çevrimiçi) 30 Mayıs 2013.

http://www.donanimhaber.com/kurumsal-diger/haberleri/FBiin-1-milyar-dolarlik-yuz-tanima-sistemi-2014te-devreye-giriyor.htm; (Çevrimiçi) 15 Şubat 2013.

http://www.economist.com/node/15557443. (Çevrimiçi) 28 Mayıs 2013.

http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the world.html?_r=2&scp=1&sq=Big%20Data&st=cse&. (Çevrimiçi) 03 Haziran 2013.

http://www.oganozdogan.com/2013/07/buyuk-veri-big-data-nedir.html. (Çevrimiçi). 01 Ağustos 2013.

Learning to Use Big Data The Right Way: http://www.colocationamerica.com/blog/learning-to-use-big-data-the-right-way.htm (Çevrimiçi) 06 Ağustos 2013.

EKLER

EK: 1: Anket Formu

Sayın katılımcı,

Bu anket İstanbul Üniversitesi Sosyal Bilimler Enstitüsü'nde yapmakta olduğum yüksek lisans tez araştırması için hazırlanmıştır. Cevaplandıracağınız anket tamamıyla profesyonel ölçülerle değerlendirilecek, bilimsel amaçlar için kullanılacak ve hiç bir firmayla paylaşılmayacaktır. Ankette kimliğinizi belirtmeniz gerekmemektedir. Anketlerdeki bilgilerin değerlendirilmesinde gizlilik esas olacaktır. Bu nedenle kişisel görüşlerinizin objektif olarak bu ankete yansıtılması ve anketin gereği gibi doldurulması konusunda yardımcı olmanızı dilerim. Değerli katkılarınız ve vakit ayırdığınız için teşekkür ederim.
Saygılarımla,

Şafiye Bilgin

1) Cinsiyetiniz? ( ) Kadın ( ) Erkek

2) Medeni Durumunuz? ( ) Bekar ( ) Evli

3) Yaşınız? ( ) 18-25 ( ) 26-35 ( ) 36-45 ( ) 46-55 ( ) 56 ve üzeri

4) Eğitim düzeyiniz? ( ) İlkokul ( ) Ortaokul ( ) Lise
( ) Üniversite ( ) Yüksek Lisans ve üzeri

5) Aylık net geliriniz? ( ) 0-750 TL ( ) 751-1500 TL
( ) 1501-2500 TL ( ) 2500 TL üzeri

6) Hangi sektörde çalışıyorsunuz? (Lütfen tek bir şık işaretleyiniz.)

( ) Akademik ( ) Bilişim ( ) Eğitim ( ) Finans ( ) Otomotiv
( ) İnşaat ( )Tekstil ( ) Gıda ( ) Sağlık ( ) Lojistik
( ) Perakende ( )Turizm ( ) Kimya ( )Öğrenciyim ( ) Çalışmıyorum
( ) Temizlik ( )Hizmet ( ) Diğer (…….)

7) Kaç yıldır İnternet kullanıyorsunuz?

( ) 1 yıldan az ( ) 1-4 yıl ( ) 5-9 yıl ( ) 10-14 yıl ( ) 15 yıl ve üzeri

8) İnternet üzerinden verilen bankacılık hizmetlerinden yararlanıyor musunuz? (Cevabınız <u>HAYIR</u> ise lütfen <u>sadece 9. soruyu</u> cevaplayınız, cevabınız <u>EVET ise 10.sorudan</u> başlayarak cevaplayınız)

( ) Evet ( ) Hayır

9) İnternet bankacılığı <u>kullanmama</u> nedeninizle ilgili aşağıdaki ifadelerden size uygun olanı işaretleyiniz.

( ) Güvenilir olmadığını düşünüyorum.

( ) Hakkında bilgim olmadığı için kullanmıyorum.

( ) İnternet bankacılığının kullanımının zor olduğunu düşünüyorum.

( ) Şube yoluyla bankacılık işlemleri gerçekleştirdiğim için ihtiyaç duymuyorum.

( ) Benim dışımda başkaları (Eş, dost, akraba vb) ilgilendiği için kullanmıyorum.

10) Kaç yıldır İnternet bankacılığı kullanıyorsunuz?

( ) 1 yıldan az ( ) 1-4 yıl ( ) 5-9 yıl ( )10 yıl ve üzeri

11) İnternet bankacılığı için hangi sıklıkla İnternet'e bağlanıyorsunuz?

( ) Günde birkaç kez ( ) Haftada birkaç kez
( ) Haftada 1 kez ( ) Ayda bir kez

12) İnternet bankacılığına nereden bağlanıyorsunuz?

( ) Kişisel PC/Tablet PC ( ) Şubelerdeki PC/ Tablet bilgisayarlardan

( ) Cep telefonumdan ( ) İşyerindeki bilgisayarlar

13) Kaç tane İnternet bankacılığı hesabınız var?

( ) 1 **( )** 2 **( )** 3 **( )** 4 ve üzeri

14) İnternet bankacılığını kullandığınız bankalar hangileridir <u>?(En çok kullandığınız 1 olacak şekilde 1'den 5'e kadar sıralayınız)</u>

( ) AKBANK ( ) YAPI KREDİ ( ) BANK ASYA ( )İŞ BANKASI

( ) GARANTİ ( ) DENİZBANK ( ) HALK BANKASI ( ) TEB

( ) INGBANK ( )VAKIFBANK ( ) FİNANSBANK ( ) HSBC

( ) ZİRAAT BANKASI ( ) DİĞER (…….)

15) Bankacılık işlemlerinizde en çok hangi hesapları kullanıyorsunuz? (Birden fazla şık işaretleyebilirsiniz.)

( ) Vadesiz (TL) Hesap ( ) Likit fon hesabı ( ) E-hesap
( ) Vadesiz Döviz Hesabı ( ) Altın hesabı ( ) Maaş hesabı
( ) Vadeli TL hesap ( ) Mevduat fon hesabı ( ) Kredi Kartı hesabı
( ) Vadeli Döviz hesabı ( ) Kredili mevduat hesabı
( ) Bireysel emeklilik hesabı ( ) Diğer ürün hesapları

16) İnternet bankacılığını en çok hangi işlemleri yapmak için kullanıyorsunuz? (Birden fazla şık işaretleyebilirsiniz.)

( ) Para Transferi yapma (EFT-Havale işlemleri) ( ) Hisse senedi işlemleri
( ) Hesap bakiyesi inceleme ve bilgi alma ( ) Fon alım-satımı
( ) Kredi kartı işlemleri ( ) Repo işlemleri
( ) Fatura ödemeleri ( ) Bahis Transferi
( ) Döviz işlemleri ( ) Diğer Ödemeler

17) İnternet bankacılığı kullanmanıza en çok etki eden faktör nedir? (<u>En çok etki eden faktör 1 olacak şekilde 1'den 9'a kadar sıralayınız</u>: En önemli: 1,En önemsiz : 9)

( ) 7/ 24 Hızlı ve kesintisiz bankacılık işlemleri,
( ) Maaşımı o bankadan almam
( ) Şubeye gitmeden, sıra beklemeden kolay bankacılık işlemleri,
( ) Görerek ve seçerek bankacılık işlemi yapabilmek,
( ) Detaylı rapor ve bilgi alabilmek
( ) Çok çeşitli bankacılık ürünlerini görerek bu ürünlerden faydalanabilmek,
( ) Bankacılık işlemlerini çok daha ucuza yapabilmek,
( ) Güvenilir olması
( ) Pratik olması

18) İnternet bankacılığı işlemlerinde karşılaştığınız temel sorunlar nelerdir? (Birden çok seçeneği işaretleyebilirsiniz.)

( ) Günlük işlem limitleri yetersiz
( ) Kullanışlı ve pratik değil
( ) Web sayfasındaki güvenlik önlemleri yetersiz
( ) Sunduğu olanaklar yetersiz
( ) Hiçbir sorunla karşılaştığımı düşünmüyorum
( ) Diğer, (Lütfen belirtiniz)…….

19) İnternet bankacılığının gelecekte geleneksel bankacılığın (şube) yerini alacağını düşünüyor musunuz?

( ) Evet ( ) Hayır ( ) Kısmen ( ) Fikrim yok

20) Bir bankanın internet sitesinde aradığınız özellikleri, <u>en önemlisi 1 olacak şekilde 1'den 8'e kadar sıralayınız. (1: En önemli ; 8 : En az önemli)</u>

( ) Hizmet Kalitesi ( ) Hız ( ) Bankanın büyüklüğü
( ) Bankanın bilinirliği ( ) Güvenlik ( ) Teknolojik yenilikler
( ) Tavsiye ( ) Kolay ve Pratik Kullanımı ( ) Diğer ……