

**T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI**

DOKTORA TEZİ

**VERİ MADENCİLİĞİ TEKNİKLERİNİN BİLGİ
EKONOMİSİ DEĞİŞKENLERİNE
UYGULANMASI**

Gonca YÜZBAŞI KÜNÇ

2502110339

TEZ DANIŞMANI

PROF.DR. Mehpare TİMOR

İSTANBUL – 2018



T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



DOKTORA
TEZ ONAYI

ÖĞRENCİNİN;

Adı ve Soyadı : GONCA YÜZBAŞI KÜNÇ Numarası : 2502110339
Anabilim Dalı / Anasanat Dalı / Programı : SAYISAL YÖNTEMLER Danışmanı : PROF.DR.MEHPARE TİMOR
Tez Savunma Tarihi : 20.06.2018 Saati : 14.00
Tez Başlığı : VERİ MADENCİLİĞİ TEKNİKLERİNİN BİLGİ EKONOMİSİ DEĞİŞKENLERİNE UYGULANMASI

TEZ SAVUNMA SINAVI, İÜ Lisansüstü Eğitim-Öğretim Yönetmeliği'nin 50. Maddesi uyarınca yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin KABULÜNE OYBİRLİĞİ / OYÇOKLUĞUYLA karar verilmiştir.

JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
PROF.DR.MEHPARE TİMOR		Kabul
PROF.DR.ERGÜN EROĞLU		Kabul
PROF.DR.İBRAHİM DOĞAN		Kabul
PROF.DR.TUĞBA GÜRSOY		Kabul
DOÇ.DR.ERDAL DİNÇER		Kabul

YEDEK JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
PROF.DR.AHMET METE ÇİLİNGİRTÜRK		
DOÇ.DR.SEDA TOLUN TAYALI		

ÖZ
VERİ MADENCİLİĞİ TEKNİKLERİNİN BİLGİ EKONOMİSİ DEĞİŞKENLERİNE
UYGULANMASI
Gonca YÜZBAŞI KÜNÇ

Geçmişten günümüze verilerin kaydedilmesi ve bu kayıtların büyük boyutlarda veri tabanlarını oluşturması bu veriler içerisindeki gizli bilgilerin keşfedilmesini sağlayan veri madenciliği tekniklerini yaygınlaştırmıştır. Bu tez çalışmasında ilk olarak Avrupa Birliği üye ülkeleri ve Türkiye'nin de dâhil olduğu 35 ülkenin 2013-2016 yılları için GSYİH, işsizlik ve ihracat ithalat oranı değişkenleriyle kümeleme analizi yapılmış ve Karar Ağaçları ile ekonomik seviye olarak gruplanan ülkeleri ayıran bilgi ekonomisi değişkenleri analiz edilmiştir. Daha sonra kümeleme tekniklerinden Kohonen Ağları ile ülkeler bilgi ekonomisi değişkenleri, refah seviyesi, AB üyelik yılları, ekonomik seviyeleri ve coğrafik konumları dikkate alınarak bölümlendirilmiş ve uygulamanın sonuçları bu kapsamda sunulmuştur.

Anahtar Kelimeler: Veri Madenciliği, Bilgi Ekonomisi, Kümeleme Analizi, Karar Ağaçları, Kohonen Ağları

ABSTRACT

IMPLEMENTATION OF DATA MINING TECHNIQUES ON KNOWLEDGE ECONOMY VARIABLES Gonca YÜZBAŞI KÜNÇ

The recording of data from the past to the present and the creation of large-scale databases with these records has widened data mining techniques that enable the discovery of confidential information in these data. In this thesis, clustering analysis was conducted by the 35 countries's GDP, unemployment and the import-export ratio variables between the 2013-2016 years. The 35 countries include European Union member states and Turkey. Knowledge Economy Variables which separates the countries grouped as economic level are analyzed with Decision Trees. As a complementary process, the countries are segmented based on their knowledge economy variables, welfare level, years of EU membership, economic size and geographic locations, by means of and Kohonen Maps, which are remarkable techniques of Clustering. Finally, the application results are presented within this frame .

Keywords: Data Mining, Knowledge Economy, Clustering Analysis, Decision Trees, Kohonen Networks

ÖNSÖZ

Günümüzde binlerce veri büyük boyutlarda veri tabanlarını oluşturmaktadır. Bu veri yığınlarındaki gizli bilgilerin açığa çıkarılmasını sağlayan ve geniş bir kullanım alanı olan veri madenciliği teknikleri oldukça önem kazanmıştır. Bu nedenle tez çalışmasında yöntem olarak veri madenciliği yöntemleri tercih edilmiştir.

Veri madenciliği teknikleri birçok farklı alanda faydalı bilgiler sunmaktadır. Bu tez çalışmasında veri madenciliği yöntemleri diğer alanlardan farklı olarak bilgi ekonomisi alanında kullanılmıştır. AB üye ülkelerinin ve Türkiye'nin bilgi ekonomisi verilerinin, veri madenciliği teknikleri ile analiz edilmesi amaçlanmıştır. Bu kapsamda Kümeleme Analizi, Karar Ağaçları Teknikleri ve Kohonen Ağları ile uygulama aşaması gerçekleştirilmiştir. Çalışma sonucunda tutarlı sonuçlar elde edilmiştir.

Tez çalışmasının ilerlemesinde her daim ilgilenen ve destekleyen Sevgili Danışman Hocam Prof. Dr. Mehpare TİMOR'a teşekkür ederim. Uygulama aşamasında bana yolladığı dökümanlarla önümü aydınlatan Sayın Doç. Dr. Gökhan SİLAHTAROĞLU'na teşekkürlerimi sunarım.

Tez çalışmasının tüm aşamalarında benden manevi desteğini esirgemeyen annem Feride YÜZBAŞI'ya, en umutsuz olduğum anlarda beni harekete geçiren ablam Tuğba EKMEKYAPAR'a ve eniştem Doç. Dr. Talha EKMEKYAPAR'a teşekkür ederim. Yaşamları boyunca beni destekleyen, moral veren ve insanlığıyla örnek bir şahsiyet olan anneannem Hatice YÜZBAŞI'ya, küçük yaşlarımdan itibaren bana maddi ve manevi destek olan merhum dedem Şeref YÜZBAŞI'ya teşekkür ederim. Ayrıca tezime her zaman destek olan ve her daim yanımda olan dayılarım Ethem YÜZBAŞI ve Sinan YÜZBAŞI'ya teşekkürü bir borç bilirim. Tezimin tüm aşamalarında bana karşı gösterdiği sabır ve anlayıştan dolayı eşim Soner KÜNÇ'e ve son bir yılda hayatıma renk katan ve beni güdüleyen oğlum Ahmed Taha KÜNÇ'e teşekkür ederim.

Gonca YÜZBAŞI KÜNÇ

İstanbul, 2018

İÇİNDEKİLER

ÖZ.....	iii
ABSTRACT	iv
ÖNSÖZ.....	v
İÇİNDEKİLER	vi
ŞEKİLLER LİSTESİ.....	xi
KISALTMALAR LİSTESİ.....	xii
GİRİŞ	1

BİRİNCİ BÖLÜM BİLGİ EKONOMİSİ

1.1.Bilgi ve Bilgi Toplumu	4
1.2.Bilgi Ekonomisi Nedir?	9
1.3.Bilgi Ekonomisinin Özellikleri	12
1.3.1.Bilgi Ekonomisi ve İktisadi Büyüme	15
1.3.1.1. İstihdam ve Bilgi İşçileri.....	20
1.3.1.2. Verimlilik.....	28
1.3.1.3. GSMH/GSYİH.....	29
1.3.1.4. Patent.....	31
1.3.2. Bilgi Ekonomisi, Küreselleşme ve Rekabet.....	32
1.3.3. Bilgi Ekonomisi, Ar-Ge Faaliyetleri ve İnovasyon.....	35
1.3.4. Yaşam Boyu Öğrenme (Lifelong Learning)	39
1.3.4.1. Bilgi Ekonomisinin Gelişiminde Eğitimin Yeri.....	40
1.4.Bilgi Ekonomisinin Gelişiminde Üniversitelerin Yeri.....	43
1.5.Bilgi Ekonomisi ve Bilgi İletişim Teknolojileri	44
1.5.1. Telekomünikasyon	47
1.5.2. İnternet	49
1.5.3.E-Ticaret	51
1.5.4.E-Devlet	53
1.6.Bilgi Ekonomisi Değişkenleri ile İlgili Literatür Araştırması	55
1.7.Bilgi Ekonomisinin Dünya ve Türkiye’deki Yeri.....	59

İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

2.1. Veri Madenciliği Tanımı ve Gelişimi	61
2.1.1. Veritabanlarında Bilgi Keşfi ve Veri Madenciliği	65
2.1.2. Veri Ambarı ve Veri Madenciliği	70
2.1.2.1. Veri Ambarı Temel Nitelikleri	71
2.1.2.2. Veri Ambarı Mimarisi	72
2.1.2.3. Veri Ambarının İçerdiği Veri	75
2.2. Veri Madenciliği Amaçları	77
2.3. Veri Madenciliği Uygulama Süreci	79
2.4. Veri Madenciliği Modelleri	84
2.4.1. Sınıflandırma	84
2.4.2. Kümeleme	85
2.4.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler	86
2.4.3.1. Apriori Algoritması	89
2.5. Veri Madenciliği Teknikleri	91
2.5.1. Karar Ağaçları	95
2.5.1.1. ID3 Algoritması	98
2.5.1.2. C 4.5 ve C.5 Algoritmaları	99
2.5.1.3. CART Algoritması	99
2.5.2. Yapay Sinir Ağları	100
2.5.3. Genetik Algoritmalar	103
2.5.4. Destek Vektör Makineleri	105
2.5.4.1. Doğrusal Destek Vektör Makineleri	105
2.5.4.2. Doğrusal Olmayan Destek Vektör Makineleri	107
2.5.5. Bayes Sınıflandırma Tekniği	109
2.5.6. K-En Yakın Komşu	110

ÜÇÜNCÜ BÖLÜM

KÜMELEME ANALİZİ

3.1.Kümeleme Analizi Nedir?	113
3.2.Kümeleme Analizinin Özellikleri ve Amacı.....	115
3.3.Kümeleme Analizi Adımları.....	117
3.4. Uzaklık veya Benzerlik Ölçüleri.....	118
3.5.Benzerlik Matrisinin Oluşturulması.....	120
3.6. Kümeleme Analizi Yöntemleri	122
3.6.1. Hiyerarşik Olmayan Yöntemler	122
3.6.1.1. K Ortalamalar Yöntemi.....	123
3.6.1.2. K Medoids Yöntemi.....	124
3.6.1.3. Bulanık Kümeleme	125
3.6.2.Hiyerarşik Yöntemler.....	126
3.6.2.1. Merkezi Yöntem	126
3.6.2.2.Tek Bağlantı Metodu	127
3.6.2.3. Tam Bağlantı Metodu	128
3.6.2.4. Ortalama Bağlantı Metodu.....	128
3.6.2.5.Ward Metodu	129

DÖRDÜNCÜ BÖLÜM

UYGULAMA

4.1.Ülkelerin Ekonomik Büyüklüklerine Göre Kümeleme Analizi ile Gruplanması	132
4.3. Veri Madenciliği Uygulama Süreci	136
4.3.1. Problemin Tanımlanması	137
4.3.2.Verinin Hazırlanması	137
4.3.2.1.Verilerin Toplanması	137
4.3.2.2.Verinin Önışlenmesi	138
4.3.2.3.Verilerin Dönüştürülmesi.....	143
4.3.3.Modelin Kurulması	150
4.3.3.1.Bilgi Ekonomisi ve Ekonomik Büyüklük İle İlgili Karar Ağacı Uygulaması... 150	
4.3.4.Modelin Değerlendirilmesi ve Kullanılması	164
4.3.5.Modelin İzlenmesi.....	164

SONUÇ.....	165
KAYNAKÇA	169
ÖZGEÇMİŞ.....	182

TABLÖLAR LİSTESİ

Tablo 1.1. İnternet Kullanım İstatistikleri	49
Tablo 2.1. Veri Madenciliği Gelişimi.....	64
Tablo 4.1. Kümeleme Analizi Değişkenlerinin Arasındaki Korelasyon Katsayılarının Anlamlılık Testleri.	133
Tablo 4.2. Kümeleme Analizi Değişkenlerinin Anlamlılık Testleri..	134
Tablo 4.3. Kümeleme Analizi Değişkenlerinin Küme Merkezleri.....	134
Tablo 4.4. Kümeleme Analizi Gruplama Sonuçları	136
Tablo 4.5. Bilgi Ekonomisi Değişkenlerinin Gösterimi.....	141
Tablo 4.6. Değişkenlere Ait Veri Sayıları.	142
Tablo 4.7. Kohonen Ağları ile Kümeleme Uygulaması İçin Düzenlenen Veriler. ..	144
Tablo 4.8. Verilere ait Temel İstatistikler.	146
Tablo 4.9. Logaritmik Dönüşüm Sonrası D25 Değişkenine ait Dağılım Grafiği....	147
Tablo 4.10. Logaritmik Dönüşüm Sonrası D26 Değişkenine ait Dağılım Grafiği..	147
Tablo 4.11. Logaritmik Dönüşüm Sonrası D27 Değişkenine ait Dağılım Grafiği..	148
Tablo 4.12. Logaritmik Dönüşüm Sonrası D28 Değişkenine ait Dağılım Grafiği..	148
Tablo 4.13. Logaritmik Dönüşüm Sonrası D30 Değişkenine ait Dağılım Grafiği..	149
Tablo 4.14. Logaritmik Dönüşüm Sonrası D32 Değişkenine ait Dağılım Grafiği..	150
Tablo 4.15. Karar Ağacı Analizine Ait Dendogram	152
Tablo 4.16. Karar Ağacı Analizi Sınıflandırma Performansı.....	153
Tablo 4.17. Küme Koordinatlarına Ait Grafik Gösterimi	154
Tablo 4.18. Küme Koordinatlarına ait Kayıtlar.....	155
Tablo 4.19. Kohonen Ağları ile Kümeleme Analizi Sonuçları	156
Tablo 4.20. Coğrafi Bölge Değişkeninin Kümeler Üzerindeki Etkisi.	157
Tablo 4.21. İnternet Üzerinden Satın Alım Yapan Bireylerin Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.....	158
Tablo 4.22. Çalışanlarına BİT Eğitimi Veren Girişimlerin Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.....	159
Tablo 4.23. Ar-Ge Personeli Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi. ...	160
Tablo 4.24. GSYİH’de Ar-Ge Gideri Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.	161
Tablo 4.25. Refah Seviyesi Değişkeninin Kümeler Üzerindeki Etkisi.	162
Tablo 4.26. Üyelik Yılı Değişkeninin Kümeler Üzerindeki Etkisi.	163
Tablo 4.27. Ekonomik Büyüklük Değişkeninin Kümeler Üzerindeki Etkisi.	164

ŞEKİLLER LİSTESİ

Şekil 1.1.Ekonomide Bilginin Tarihsel Önemi ve Gelişimi.	10
Şekil 1.2.Bilgi Teknolojisi Yatırımlarından Ekonomik Hasılat.	17
Şekil 1.3.Bilgi Ekonomisinin Temelleri.	46
Şekil 1.4.Bilgi Ekonomisi İndeksi ve Bilgi İndeksi.	57
Şekil 2.1.Verit Ambarı Mimarisi	72
Şekil 2.2.Verit Ambarı Mimarisi Yıldız Şema.	73
Şekil 2.3.Verit Ambarı Mimarisi Kartanesi Şema.	74
Şekil 2.4.Verit Ambarı Mimarisi Anatbalolar Şema.	75
Şekil 2.5.Verit Madenciliğindeki Önemli Adımlar.	80
Şekil 2.6.Süreç Eğiliminden Verit Madenciliğinin Analizi.	92
Şekil 2.7.Karar Ağacı Genel Görünümü.	97
Şekil 2.8.Yapay Sinir Ağları Genel Görünümü.	101
Şekil 2.9.Optimum Hiper-Düzlem ve Destek Vektörleri.	106
Şekil 2.10.Optimum Hiper-Düzlem ve Destek Vektörleri.	107
Şekil 2.11.Doğrusl Olarak Ayrılmayan Hiper-Düzlemin Belirlenmesi.	108
Şekil 2.12.Optimum Hiper- Düzlem ve Destek Vektörleri.	109
Şekil 2.13.K En Yakın Komşu Örnek Şekli	112

KISALTMALAR LİSTESİ

AB	: Avrupa Birliği
ABD	: Amerika Birleşik Devletleri
AR-GE	: Araştırma Geliştirme
BİT	: Bilgi ve İletişim Teknolojileri
DVM	: Destek Vektör Makineleri
GSMH	: Gayrı Safi Milli Hasıla
GSYİH	: Gayrı Safi Yurtiçi Hasıla
OECD	: Organisation for Economic Co-operation and Development
SQL	: Structured Query Language
VTBK	: Veri Tabanlarında Bilgi Keşfi
YSA	: Yapay Sinir Ağları

GİRİŞ

Son yıllarda bilişim teknolojilerinde meydana gelen gelişmeler oldukça hızlı bir şekilde ilerlemektedir. Bu ilerlemelerden birçok alanda yararlanılmaktadır. Yapılan her işlemin bilişim teknolojilerinin ilerlemesiyle kayıt altına alınabilmesi ve biriktirilmesi sonucu belli yöntem ve tekniklerle yararlı bilgilere ulaşılmaktadır. Bilgi ve iletişim teknolojilerinin ilerlemesiyle elde edilen çok sayıda veri içinde işletmeler, ülkeler, kurum ve kuruluşlar vb. için gizli ve faydalı bilgiler yer alabilmektedir. Özellikle bilişim alanında yer alan birçok matematiksel ve istatistiksel yöntem ve teknikler kullanılarak bu gibi birimler için birçok önemli bilgiler sunabilmektedir. Büyük miktardaki verilerin işlenmesiyle gizli bilgileri ve ilişkileri ortaya çıkaran veri madenciliği, bilişim alanındaki yöntemlerden biridir.

Günümüzde oldukça yaygın kullanım alanına sahip olan veri madenciliği teknikleri veri tabanlarının içinde bulunan önemli bilgileri kullanıcılara sunmaktadır. Ayrıca veriler arasındaki görülemeyen ilişkileri de ortaya çıkaran bir tekniktir. Veri madenciliği veri tabanlarındaki yararlı bilgiyi ortaya çıkarırken birçok istatistiksel ve matematiksel yöntemleri kullanmaktadır. Veri madenciliği birçok veriyle işlevini yerine getirirken karar ağaçları ve çok değişkenli istatistiksel yöntemlerden de faydalanmaktadır. Bu yöntemlerin başında kümeleme analizi, diskriminant analizi, regresyon analizi gelmektedir.

Veri madenciliği araçlarıyla ortaya çıkarılan bilgi birçok farklı alanda yaygın olarak kullanılmaktadır. Pazarlama, finans, mühendislik, lojistik, uzay bilimleri, tıp ve psikoloji gibi alanlar veri madenciliğinin kullanım alanlarından bazılarıdır. Bu alanların yanı sıra veri madenciliği tekniklerinin kullanımı yeni alanlarda da denenmektedir. Bu tez çalışmasında bahsedilen alanlardan farklı olarak ekonomi alanında veri madenciliği teknikleri uygulanmıştır.

Gelişen ve gelişmekte olan ülkeler arasındaki en önemli değişkenin geçmişten günümüze kadar ekonomik değişkenler olduğu açık olarak görülmektedir. Ülkelerin her anlamda ilerlemesi ve refah düzeylerinin artması için makroekonomik bazı göstergeler önemli yer tutmaktadır. Buna bağlı olarak son yıllarda bu göstergeleri

etkileyen deęiřkenlerde nem kazanmaktadır. Bilgi ekonomisi deęiřkenleri de bu gstergeleri nemli lde etkilemektedir.

Bilgi ve iletiřim teknolojilerinin hızla ilerlemesi sonucu faydalı yntemlerin kullanılmasının yansıra son yıllarda bu kanallarla bilginin yayılım hızının artması, zamanında ve doęru bilgiye ulařmanın birok birim ve birey iin hayati nem tařması topluma yeni bir boyut kazandırmıřtır. Bilginin sosyolojik boyutu ile birlikte, ekonomik alanda da nemli deęiřimleri beraberinde getirdięi grlmektedir. Ekonomik byme artık bilgiden baęımsız olarak ynetilememektedir. Bilimsel bilginin kullanımıyla lkeler ekonomik olarak rekabet etmektedirler. Bilim ve teknoloji, bilgi ve iletiřim teknolojilerinden uzay bilimlerine Ar-Ge faaliyetlerine kadar geniř bir alanı oluřturmaktadır. Bilim ve teknolojinin bu geniř alanda saęladıęı faydalar ekonomik geliřimin de hızlanmasında nemli bir rol oynamaktadır. Bilgi ekonomisi insanlara doęrudan ve dřk maliyetlerle ulařma imknı saęlamaktadır. Bu zellięinden dolayı mal ve hizmet retiminde daha etkin yollar sunmaktadır. Gnmzde bilgi ekonomisinin neminin lkeler aısından artması bu alandaki deęiřkenlerin neler olduęu, ne kadar etkin kullanıldıęı ve etkinlięi gsteren nemli deęiřkenlerin neler olduęu soruları nemini arttırmaktadır.

Bilgi ekonomisi deęiřkenleri lkelerin bilgi ve iletiřim teknolojileri ile ilgili ilerlemelerini ve bu alanda ne durumda olduklarını gstermektedir. Bilgi ekonomisi deęiřkenlerine veri madencilięi tekniklerinin uygulanmasındaki ama genel ekonomik durumu etkileyen bilgi ekonomisi deęiřkenlerini saptamak ve lkeleri bu deęiřkenlere gre gruplandırmaktır.

Tez alıřmasının amacı veri madencilięi tekniklerinin ele alınarak; genel ekonomik deęiřkenlerin ve bilgi ekonomisi deęiřkenlerinin aıklanması doęrultusunda gerekli tekniklerin uygulanması ve anlamlı bilginin ıkarılmasıdır.

Bu tez alıřmasında ilk olarak bilgi ekonomisin ele alınmaktadır. Bilgi ekonomisinin tanımı, kapsamı ve zellikleri incelenmektedir. Bu alanda yapılmıř olan istatistiksel alıřmalara deęinilerek nceki alıřmalarda kullanılmıř olan deęiřkenler belirtilmektedir.

Tez çalışmasının ikinci bölümünde veri madenciliği kavramı üzerinde durularak süreci göz önüne alınmaktadır. Ayrıca veri madenciliği alanında yapılmış çalışmalar incelenerek, yöntemleri ayrıntılı bir şekilde ele alınmaktadır.

Tezin üçüncü bölümünde ise veri madenciliği tekniklerinden ve uygulama kısmında da kullanılacak olan kümeleme analizine yer verilmektedir. Bununla birlikte kümeleme analizinin amacı, kullanımı ve yöntemleri de açıklanmaktadır.

Tez çalışmasının uygulama kısmında öncelikle bilgi ekonomisinde kullanılacak değişkenler incelenmektedir. Tez çalışmasının dördüncü bölümünde ilk olarak ülkeler genel ekonomik değişkenlere göre gruplandırılmıştır. Uygulama kısmında daha önceki bölümlerde açıklanmış olan bilgi ekonomisi değişkenleri kullanılarak, veri madenciliği tekniklerinden karar ağaçları yöntemi ile ülkelerin ekonomik gruplarını etkileyen önemli bilgi ekonomisi değişkenleri belirlenmiştir. Daha sonra ise ülkelerin ekonomik olarak gelişme ve refah düzeyini etkileyen önemli bilgi ekonomisi değişkenlerine göre AB ülkelerinin ve Türkiye'nin veri madenciliği tekniklerinden olan kümeleme analizi ile gruplandırılarak hangi ülkelerin bu konuda birbirine benzediği gösterilmektedir. Bu aşamada yapılan Kohonen ağırları ile kümeleme analizinin amacı farklı kategorik değişkenlerinde (ekonomik seviye, refah düzeyi, coğrafi konumlar, Avrupa Birliği'ne üyelik yılları) göz önüne alınarak ülkelerin nasıl gruplandığını gözlemlemektir. Ayrıca bu değişkenlerin gruplar üzerinde etkili değişkenler olup olmadığını incelemektir.

BİRİNCİ BÖLÜM

BİLGİ EKONOMİSİ

1.1.Bilgi ve Bilgi Toplumu

İnsanlık tarihini yönlendiren eksen, esas itibarıyla bilgidir; bilgideki ilerlemedir. İnsan, dolayısıyla toplumlar, davranışlarını ve stratejilerini sahip oldukları bilgiye göre ayarlamaktadır. Yaşadıkları dünya ve diğer insanlarla ilişkileri doğa, insan ve toplum hakkında ne biliyorlarsa ona bağlıdır. ¹ Bilginin tanımı; sistemli bir şekilde herhangi bir iletişim aracıyla başkalarına aktarılan, makul bir hükmü veya tecrübeye dayanan sonucu gösteren, olgu veya fikirlerle ilgili düzenli ve sistemli ifadeler bütünü olarak açıklanmaktadır. ²

Bilginin en basit formu farklı çevrelerden gelen insanlar tarafından bile yanlış anlaşılma riski en düşük olan ve bu nedenle açıkça kodlanabilen ve tamamen açık olabilen formudur. Çok uluslu firmalar bir takım bilgilerin tamamen kodlanamadığı ve açıkça dile getirilemediğini ama tecrübe ve pratikle öğrenilmesi gereken bir takım bilgiler olduğunu zamanla görmekteirler. Birçok bireysel özellik, örgütsel rutinler ve yazılı olmayan sektör normları gibi faaliyet prosedürleri tecrübeye dayalı bilgi örnekleridir. Kompleks bilgiyle ise firmaların yaşadıkları zorluklar ve gelecekte yaşayacakları potansiyel zorluklarla başa çıkmak için daha faydalı olan bilgiye gitmeyi denediklerinde karşılaşmaktadırlar. Bu bilgi endemik bilgidir ve ilk baştaki bilgilerden oldukça daha karmaşıktır. Çünkü bu bilgiyi anlamak için içeriğinin bütünleşik olduğunun anlaşılması ihtiyacı duyulmaktadır. Karmaşıklıkta en büyük seviye moda, müzik, sanat gibi kültürel alanlarda görülen varoluşsal bilgidir. Varoluşsal bilgi özel alanlarda oluşan ve üstü kapalı bilgidir. Bu bilgi türü firmayla ilgili en karmaşık bilgi türü olarak görülmektedir. Çünkü başarılı ilerlemeleri sağlayan duygu ve değerlerle çevrilidir. Bu bilgi türü bulunduğu çevrede uygulama

¹Cihan Dura, Hayriye Atik, **Bilgi Toplumu Bilgi Ekonomisi ve Türkiye**, İstanbul, 2002, s.125.

²A.e, s.134

ve tecrübe aracılığıyla onu yaşamak ve hissetmekle anlaşılmaktadır.³ Değerli bilgi sofistike ve zor olma gibi özelliklere sahiptir. Teknoloji ve pazarla ilgili değerli bilgi ise daha güç, karışık ve zor algılanmaktadır.⁴

Bilgi yönetimi literatüründe örtük ve açık bilgi arasında belli bir farklılaşma yer almaktadır. Açık bilgi dille açıkça belirtilen bilgi olarak tanımlanmaktadır. Açık bilgi diğerleriyle paylaşılabilir, hızlıca iletilebilir özellikte gösterilmektedir. Örtük bilgi ise zıt olarak bireysel tecrübeyle şekillenen kişisel bilgidir. Kişisel inançlar, yaklaşımlar, değer sistemlerinden oluşmaktadır ve tecrübeyle öğrenilmiş şeylerin toplamı olarak görülmektedir. Örtük bilgiyle iletişim kurmadan önce onun anlaşılabilen kelimelere, sayılara dönüştürülmesi gerekmektedir.⁵

Ülkeler ekonomik ve sosyal açıdan ilerlemelerini kuvvetlendirmek için bilgiye daha çok önem vermek, daha etkili şekilde ürün ve hizmet üretmek, geliştirdikleri üretim yöntemlerinin bireylere düşük seviyede maliyetlerle erişmesini sağlamak gibi faaliyetleri hayata geçirmek zorunda kalmaktadırlar. Refah seviyesinin doğal kaynaklar aracılığıyla etkin bir şekilde artırılması yerine günümüzde bireylerin bilgi düzeyleri ve üreticilik yetenekleri ile refah düzeyi arttırılmaya başlamıştır. Son yıllarda bilgi unsuru değer elde etmek için kullanılan önemli bir kaynak olarak görülmektedir.⁶

Küreselleşme, pazarların açılması ve uluslararası rekabete açık bilgi sektörlerindeki ticaretin uluslararası hale gelmesi, yetenekli çalışanlar için küresel işgücü pazarları, yeni teknolojiye erişim ve yatırım sermayesinin fonu, bilgilerin ve fikirlerin dünyadaki akışı bilgi ekonomisine geçişi hızlandırmaktadır.⁷

Bilgi çağının en temel ve ayırıcı özellikleri;

³Yves Doz, Jose Santos, Peter Williamson, **From Global to Metanational: How Companies Win in The Knowledge Economy**, Boston, Harvard Business Press, 2001, s. 77.

⁴A.e., s. 10.

⁵Philip Ternouthi, **Universities in The Knowledge Economy Higher Education Organization and Global Change**, Ed. by. Paul Temple, Newyork, 2012, s. 42.

⁶Ghirmai T. Kefela, "Knowledge-Based Economy and Society Has Become a Vita ICommodity to Countries", **International Journal of Educationa IResearch and Technology**, Vol.1, No.2, 2010, p. 69.

⁷Ian Brinkley, "The Knowledge Economy: How Knowledge is Reshaping The Economic Life of Nations", **The Work Foundation's Three Year Knowledge Economy Programme**, London, 2008, p. 15.

- 1)Toplumsal bilgiyle bütünleşik olmasıyla ortaya çıkmıştır.
- 2)Firmalar bilgi ve iletişim teknolojilerine bağlı olarak faaliyette bulunmaktadır.
- 3)Organizasyonlardaki iş akışları verimlilik ile bağdaştırılmaktadır.
- 4)Bilgi ve iletişim teknolojileri mal ve hizmetlerle bir arada yer almaktadır.⁸
- 5)Bilginin küresel sistemde artan önemi ve bilgi ile geliştirilen yeni teknolojiler toplumsal yapıyı hızla sanayi toplumundan bilgi toplumuna dönüştürmektedir.⁹

Geçmişten bu güne kadar yaşamış olan toplumlar ilk olarak tarımla uğraşarak tarım toplumunu sonraları ise sanayileşmenin artmasıyla sanayi toplumunu, en son olarak ise bilginin değerinin anlaşılmasıyla bilgi toplumunu oluşturmuşlardır. Bu dönemleri ortaya çıkaran değişkenler ise her dönemin gerektirdiği farklı teknolojiler, yönetme kültürü ve toplumsal değerler bütünü gibi değişkenlerdir.¹⁰ Dünyada toplumlar nasıl tarihsel sıçramalarla bir safhadan diğerine geçerek evrim göstermişse, bilgi dediğimiz olgu da buna çok benzer bir biçimde evrim yasalarına tabi olmuştur. Zaten toplumsal evrim safhalarını belirleyende insan bilgilerinde kaydedilen ilerlemeden başka bir şey değildir.¹¹ Temel üretim girdisinin toprak ve sermaye olmaktan çıkması ve bilginin temel üretim girdisi olması bilgi toplumunun oluşmasını sağlamıştır. Bilim ve teknolojide meydana gelen her yeni değişim bir öncekini eskitmektedir. Bu yeniliklerin kaynağında ise bilimsel bilgi ve teknolojinin olduğu meydana çıkmıştır. Enformasyon ve iletişim alanında oluşan devrimin tabanında da yine bilginin olduğu görülmektedir. Dolayısıyla enformasyon yerine bilgi toplumu kavramı daha çok kabul görmüştür.¹²

⁸ James A. Senn, **Information Technology in Business**, Prentice Hall Inc., New Jersey, 1995, p. 202.

⁹ Arzu Akolaş, "Bilişim Sistemleri ve Bilişim Teknolojisinin Küreselleşme Olgusu ve Girişimcilik Üzerine Yansımaları", 2000, s. 32-33.

¹⁰ Ünal Şentürk, "Bilgi Toplumu Parametreleriyle Türkiye'nin Sorgulanışı." **I. Ulusal Bilgi, Ekonomi ve Yönetim Kongresi**, 2002, s. 195.

¹¹ Dura ve Atik, **a.g.e.**, s. 144.

¹² Hüsnü Erkan, "Türkiye'nin Stratejik Önceliği: AB ya da Bilgi Toplumu", **Stratejik Araştırmalar Dergisi**, Sayı.13, 2009, s. 2.

Bilgiye dayalı toplum bilimsel bilgiyi yoğunlukla meydana getiren, ürettikleri bilgiyi sınıflandıran, stoklanan bilgiye hızlıca ulaşabilen, yayılmasında öncü olan, elde edilen bilgiyi doğru bir şekilde doğru alanlarda kullanılması için dönüşüme gönderen ve bilgiyi işleyebilen toplum olarak tanımlanmaktadır.¹³ Klasik üretim faktörlerinden sermayenin, hammaddenin, işgücünün bilgiye doğru dönüşüme uğradığı toplum bilgi toplumdur. Bilgi toplumunda bilgi, ekonomide yer alan hammadde ve ürün yerine geçmektedir. Ayrıca toplumda kültür olarak benimsenen bir değer haline gelmektedir. Bilgi ve iletişim teknolojileri ise hayatın her alanında aktif olarak kullanılmaktadır.¹⁴

Bilgiye dayalı toplumunda önem kazanan en önemli faaliyetlerin başında yeni teknolojilerin ilerlemesiyle bilginin ortaya çıkardığı sektörün gelişmesi, bilgi tabanlı sermayenin ve yetenekli işgücü unsurunun daha da önem kazanması, eğitimin öneminin daha da artması olarak görülmektedir. E-ticaret gibi birçok bilgi ve teknolojiye dayalı uygulamanın toplumun her kesimine yayılması sanayi toplumundan bilgi toplumuna geçişin en önemli göstergesi olarak açıklanmaktadır.¹⁵ Kısacası faydalı bilgiye ulaşan ve bu bilgiyi doğru kullanma yolunda doğru adımlar atan, gerektiğinde, zamanında ve doğru yerde istediği faydalı bilgiye ulaşan ve kullanan toplumdur.

Peter Drucker gelecek toplumun bilgi toplumu olacağını söylemektedir. Bilginin ana kaynak olacağını ve bilgi çalışanlarının iş gücünde baskın grup olacağını belirtmektedir. Bu bağlamda bilginin bazı özelliklerinden bahsetmektedir. Bu özellikler ise şunlardır: Sınırsızlık özelliği, bilginin paradan daha kolay hareket etmesidir. Artan mobilite özelliği, elde edilen eğitimin herkes için her yerde hazır olmasıdır. Başarı kadar başarısızlık potansiyeli özelliği, herkes üretim araçlarını elde edebilir ancak iş için gereken bilgiyi herkes elde edemez. Bilgi toplumunda yeni

¹³İsmail Hakkı Yücel, "Türkiye'de Bilim Teknoloji Politikaları ve İktisadi Gelişmenin Yönü", **T.C. Devlet Planlama Teşkilatı (DPT)Yayın No: 2690**, Ankara, 2006, s. 10.

¹⁴Fatih Rukancı, Hakan Anameriç, "Bilgi Toplumu ve Toplumun Bilgilenmesinde Kütüphanelerin Rolü", Kütüphaneciliğin Destanı Uluslararası Sempozyumu, Ekim 2004, s. 3.

¹⁵M. Akif Çukurçayır, Esra Çelebi, "Bilgi Toplumu ve E-Devletleşme Sürecinde Türkiye", **Uluslararası Yönetim İktisat ve İşletme Dergisi**, 2012, s. 61.

teknoloji yapısı, araştırma, inovasyon ve sürekli öğrenme süreçleri birbirine bağlanmaktadır.¹⁶

Bilgi toplumunun en ileri düzeydeki toplum olmasının nedeni bilginin ayırt edici özelliklerinden gelmektedir. Bu özelliklerin en başında bilginin stratejik olma özelliği bulunmaktadır. Klasik üretim faktörlerinden bu özelliği yönüyle ayrılan bilgi ekonomik gelişim için büyük avantajlar sağlamaktadır.¹⁷ Sosyoekonomik aktivitelere zamanla daha etkileşimli nicel iletişim bağlantılarının dahil olmasıyla veya bu bilgi ve iletişim bağlantılarının yoğunlukla kullanımının sağlanması yanında bu hedefle kullanılan her tür teknoloji ve uygulamanın elde edilmesi olarak tanımlanan bilgi toplumu, elektronik ticaretten, online eğitime, e-devlet uygulamalarından, karayollarının yönetimine kadar tüm yaşam alanlarına kadar birçok sahayı kapsamaktadır.¹⁸

Bilgiyi koruyan, dağıtan ve çoğaltan örgütlerin varlığı bilginin geliştirilmesine yardım etmektedir. Teknolojik gelişim oranı, ekonomik üretimde aktif olan insanlar tarafından etkilenmektedir ve genellikle bu oran sosyal gruplara dağılmaktadır. İnsanlar arasındaki bilgi akışı ve sosyal bilgi depolarına erişim kolaylığı geçmiş yüzyıllar üzerindeki ilerlemeyi açıklamakta en önemli nedenlerden biri olarak görülmektedir. Günümüzde geçmişten daha çok yeni teknik ve ürün üretenler, yeni yerleşik bilgi için gerekli bilgiye kolay erişimi sağlayan eğitim ve teknolojiye sahiptirler. Modern ekonomik gelişim mucizesi, modern çağın bu açıdan farklı olduğunu açıkça anlamadan anlaşılamaz.¹⁹

¹⁶Lorena Batagan,"Indicatorsfor Knowledge Economy",**Informatica Economică** 4, 2007, s. 44.

¹⁷İbrahim Güran Yumuşak, Mahmut Bilen, "Türkiye Küresel Ağa Hazır Mı? Bilgi Ekonomisi İndeksi, Beşeri Kalkınma İndeksi ve Ağa Hazırlık İndeksi Göstergeleri Üzerine Bir Değerlendirme", **Bilgi Ekonomisi ve Yönetimi Dergisi**5.2, 2010, s. 101.

¹⁸Tübitak, "Bilgi Toplumu Politikaları Üzerine Bir Değerlendirme (Dünya ve Türkiye)", Eylül 2002, s. 4.

¹⁹Christine Macleod,"The Gifts of Athena: Historical Origins of The Knowledge Economy",Ed. by. Joel Mokyr, **TheJournal of Economic History**, Princeton NJ and Oxford: Princeton University Press, 2002, s. 288.

İleri ülkelerde bilgiyi anlama ve dönüştürme yeteneğine sahip olmayan bireylerin iyi iş bulmada zorlandığı, bu bağlamda söz konusu ülkelerin bilgi toplumu olduklarını ifade etmektedir.²⁰

1.2.Bilgi Ekonomisi Nedir?

Maddi ve maddi olmayan varlıklar arasındaki farklar, ekonominin doğasında yaşanan değişim sonucu sıkça gündeme gelmektedir. Söz konusu değişim, ekonomik yapının artık maddi ürünlere/imalata değil, maddi olmayan ürünlere (bilgi, enformasyon, v.b) dayanmasından kaynaklanmaktadır.²¹ Bilgi ekonomik anlamda da bir boyuta sahiptir. Bu boyut bilgiyi üreten kişiler ya da gruplara bağlı olmasından ileri gelen mülkiyet hakkı olarak görülmektedir. Bu açıdan bilgi kişiseldir ve ait olduğu kişilere ya da gruplara sahiplik hakkı sağlamaktadır.²²

Son yıllarda ürün ve hizmetlerin üretimi geçmişten farklı olarak küreselleşmenin de etkisiyle bilgi yoğun faaliyetler gerektirmektedir. Bu nedenle ekonomi bilgiye doğru yön değiştirmiştir ve böylece “Bilgi Ekonomisi” kavramı ortaya çıkmıştır.²³ Bir toplumun ekonomik gelişmesi, onun fiziki çevre, canlı doğa, insan ve toplum hakkında sahip olduğu teknolojik bilgilere bağlıdır.²⁴

Ekonomide klasik üretim faktörleri olan hammadde, sermaye, işgücü bilgiye dayalı toplumun ilerlemesiyle yetersiz kaldığı görülmüştür. Bu durum toplumun gelişmesi için üretim faktörleri arasına en çok katkı sağlayacak bilginin de girmesini sağlamıştır.²⁵ Bilgi sektörünün sayısal analizini yapan araştırmacılar hesaplamalarına esas teşkil edecek bilgi işinin sınırlarına “bilgi üretim ve dağıtımı, eğitim-öğretim,

²⁰Thomas A. Stewart, "Intellectual Capital: The New Wealth of Nations", New York, 1997, p. 6.

²¹William E. Halal, **The Logic of Knowledge: KM Principles Support Agile System, Agile Information System: Conceptualization, Construction, and Managent**, Ed. Kevin C. Desouza, Oxford, 2006, p. 34.

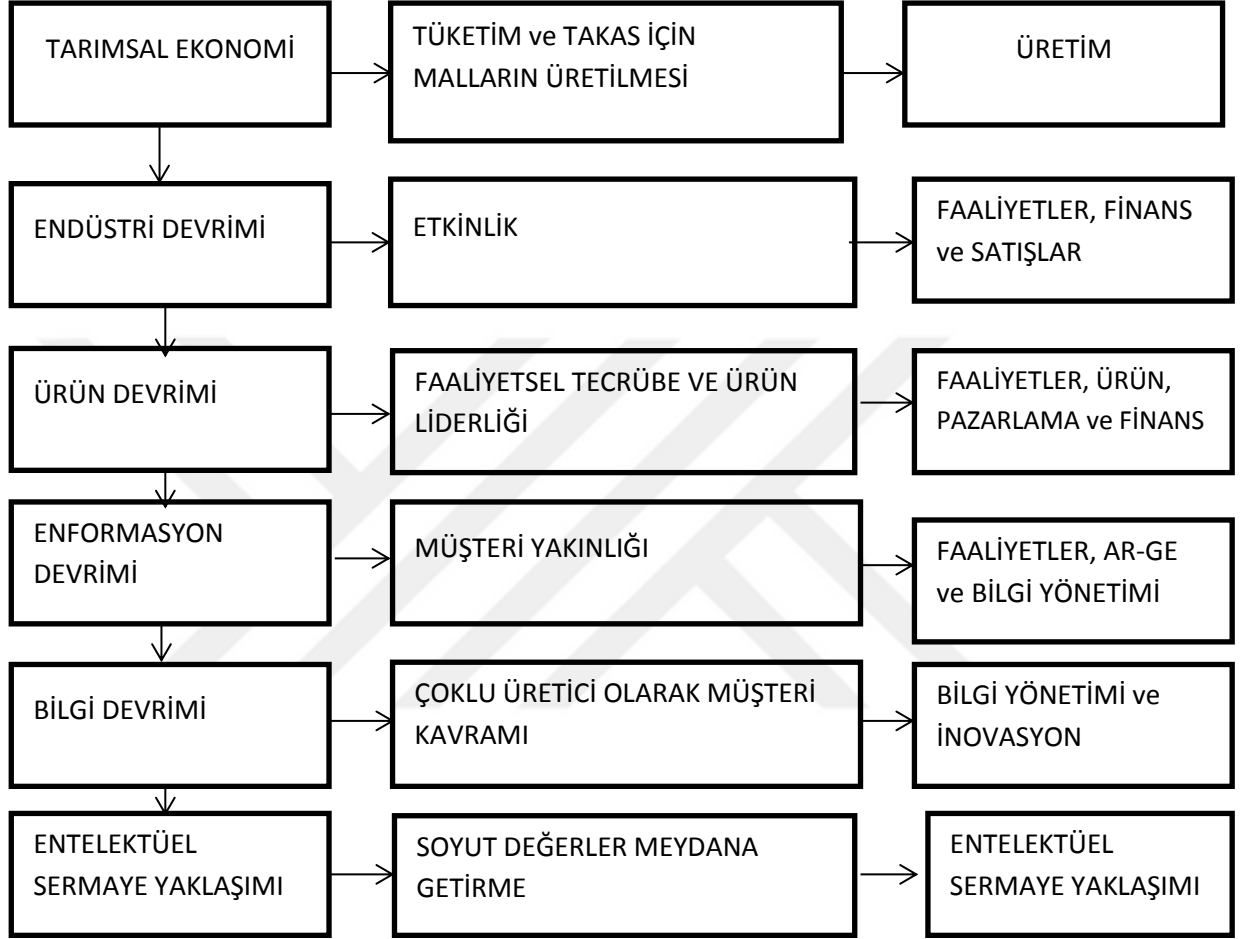
²²Daniel Bell, **The Coming of The Post Industrial Society: A Venture in Social Forecasting Newyork**, 1973, p. 7

²³Şeref Saygılı, "Bilgi Ekonomisine Geçiş Sürecinde Türkiye Ekonomisinin Dünyadaki Konumu", **Ekonomik Modeller ve Stratejik Araştırmalar Genel Müdürlüğü, Stratejik Araştırmalar Dairesi, Yayın No: 2675**, 2003, s. 5.

²⁴Nic Blakeley, Geoff Lewis, Duncan Mills, "The Economics of Knowledge: What Makes Ideas Special for Economic Growth?" **New Zealand Treasury, Polic Perspective Paper**, 2005, p. 2.

²⁵Halil Zaim, Orhan Koçak, "Bilgi Çalışanının Memnuniyeti", **Jounal of Yaşar University**, 2010, s. 1.

araştırma ve geliştirme, iletişim ve enformasyon ekipman ve hizmetlerini” dahil etmiştir.²⁶



Şekil 1.1. Ekonomide Bilginin Tarihsel Önemi ve Gelişimi.

Sanayiden bilgi ekonomisine geçişte olan bugünün küresel ekonomisi, soyut varlıklara yatırımı artan bir şekilde zenginliğin, rekabetin ve refahın kaynağı olarak tanımlanmaktadır. Aynı zamanda entelektüel sermaye, soyut sermaye, bilgi örgütleri, bilgi çalışanları, öğrenen örgütler, entelektüel varlıklar, akıllı şehirler ve bilgi bölgeleri gibi bazı terimlerin yaygın kullanımı, zihinsel modeller ve şirket modelleriyle rekabet edilirken ekonomik değer ve sosyal örgütün yeni formlarını tanımlayan yeni bir sözlüğün parçası olmaktadır. Şekil 1.1’de sanayi ekonomisinden bilgi ekonomisine geçiş anlatılmaktadır. Çok yakın zamanda entelektüel sermaye yönetimi yaklaşımı için gelişim temelli fiziksel sermayenin uygulanması ve

²⁶Cihan Dura, Hayriye Atik, "Bilgi Toplumu", TC Kültür Bakanlığı Yayınları, Ankara, 1990, s. 4.

kullanılması için ana üretim faktörü olan, bilgi fiziksel faktörlere yapılan yatırımlar tarafından çok fazla etkilenmemektedir. Odak bireysel varlıklardan; değer oluşturmada farklı varlık tiplerinin ve ortak çalıştırıldığı varlık yığınlarına doğru yön değiştirmektedir. Bilgi toplumundaki bu varlıkların payı soyuttur. Ekonomik tarih herhangi bir sektör başarısının en üretici bilgiye sahip olan firmalara geldiğini ispat etmektedir.²⁷

Bilginin üretilmesi ve yayılması ile ilgili 1962 yılında yayınlanan bir çalışmada ABD'nin o yıla ait ekonomi verileri incelenmiştir. Çalışmada o yıla ait ABD GSMH'nin yaklaşık %29'unu bilgi ekonomisinin oluşturduğu görülmüştür. Bilgi ekonomisi ABD'de ilk olarak 1960'lı yılların sonunda dikkat çekmiştir ve yapılan hesaplamalarda GSMH değerinin %46'sının ve işgücü gelirlerinin %53'ünün bilgi ekonomisi tarafından gerçekleştirildiği açıklanmıştır. Bilgi ekonomisinin ciddi bir ilerleme kaydettiği diğer zamanlar ise 1990'lı yılların başları olarak görülmektedir.²⁸

Bilgi ekonomisi, enformasyon arzının ve işlem kapasitesinin yüksek düzeylere ulaştığı, ağlarla enformasyonun yayılma hızının en yüksek düzeye çıktığı, kamu ve özel kesimin enformasyon ağları üzerinden, bilgi ve iletişim teknolojileri odaklı olarak yeniden örgütlendiği ekonomiler olarak tanımlanmaktadır²⁹

Bilgi ekonomisi entelektüel kaynakların fiziksel kaynaklardan, temel teorik araştırmaların uygulamalı araştırmalardan, öğrenimin tecrübeden daha önemli hale geldiği, değişimin ise bilimin mevcut temellerini inşa eden çok kısa bir zamanda tamamen geçersiz hale getirdiği ekonomilerdir.³⁰ Aynı zamanda işe yaramayan faaliyetlerin kaldırılması, onların yerine yeni faaliyetlerin gelmesi ve yeni bilginin oluşturulması için var olan bilginin etkin kullanımı amacıyla teşvik sağlayan ekonomik rejimdir. Yeni bilgiyi hem oluşturabilen hem de kullanabilen eğitilmiş ve

²⁷J. M.V. Marti, M. do R. Cabrita, **Entrepreneurial Excellence in the Knowledge Economy Intellectual Capital Benchmarking Systems**, Palgrave Macmillan, London, 2012, p. 16.

²⁸Benoit Godin, "The Knowledge Economy: Fritz Machlup's Construction of a Synthetic Concept" **Project on the History and Sociology of S&T Statistics Working Paper 37**, 2008, p. 4.

²⁹Eggerstedt Harald, Rebecca Brideau Hall, Swedervan Wijnbergen, **Turkey Informatics and Economic Modernization**, The World Bank, 1993, p. 45.

³⁰Numan Kurtulmuş, **Sanayi Ötesi Dönüşüm**, İstanbul, İz Yayınları, 1996, s. 100.

girişimci bir toplumdur. Bilginin etkin iletişimini, dağıtımını ve yönlendirilmesini sağlayan dinamik bilgi yapısıdır.³¹

Bilgi ekonomisi, yeni bilgi ve teknolojinin meydana getirilmesi için bilgiyi kullanan, küresel bilgi stokunun gelişimini sağlayan ve etkileşimde olan danışmanlar, araştırma merkezleri, üniversiteler ve firmaları kapsayan etkin inovasyon sistemlerini ve düşünce kuruluşlarını gerektirmektedir.³²

1.3.Bilgi Ekonomisinin Özellikleri

Verimli bir kaynak olarak bilgi temelde farklı olan özelliklere sahiptir. Bilgi ilk olarak sonsuzluk özelliğine sahiptir. İlk kez üretilen bir bilgi diğerleri tarafından tekrar kullanılabilir. Kamu alanında bilgiyi saklamak zordur. Bilgi bölünmezdir. Bununla birlikte bilgi, bir problem çözüldüğü zaman insanların onun için para ödemeye gönüllü olduğu ekonomik bir mal olmaktadır. İlk kez bir sonuç tanımlanması bilginin sonucuna ne kadar insanın ihtiyaç duyduğu o bilginin değerinin sınırlarını belirlemektedir. Bilginin bu ayırıcı özellikleri bilgi ekonomisinin organize edilebilme yolları için önemli uygulamalara sahip olmasına rağmen bilginin ekonomisi yeteri kadar anlaşılamamaktadır. Öncelikle ekonomik bir kaynak, anlayış ve içerik olarak felsefe ve bilimsel bilgiden farklıdır. Bilginin toplumsal rolü üretmek için bir kaynağa hareket etmektedir. Geleneksel olarak bilgi anlayışı bilimle yakından bağlantılıdır ve genellikle doğru olma özelliğine sahip olarak tanımlanır ve ispat edilebilir. Bugün bilgi anlayışı doğruluk geçerlilik ve güvenilirlik kavramlarına bağlıdır. Bunun yerine bilginin değeri üst performans oluşturmak yeteneğiyle bağlantılıdır. Bilgi, verimliliğe ve kaliteye bağlanacak olan küresel pazarda rekabetçi bir pozisyon sürdürme ile de bağlantılıdır. İkincisi yeni ekonominin temel özelliği olan artan getirilerin bilgi ekonomisi arasındaki anlayıştır. Üçüncüsü bilgi ölçek ve fırsat ekonomilerine bağlıdır. Onu üretmek pahalı olabilirken, sonradan çoğaltılması düşük maliyetli olmaktadır. İlk olarak üretildiğinde diğerlerine de bir yerlere kadar dağılabilir ve maliyeti sıfırdır. Bu bilgi ekonomileri son yılların zenginlik oluşumu ve artan yaşam kalitesiyle ilgili birçok şeyi açıklamaktadır; yeni inovasyonlar ve

³¹Carl Dahlman, Thomas Andersson, **Korea and The Knowledge Based Economy: Making The Transition**, World Bank Institute, 2000, s. 7.

³²Marti ve Cabrita, **a.g.e.**, s. 30.

bilgi tabanlı ürünlerin oluşmasının mümkün olduğu ekonomik gelişim için neredeyse sınırsız potansiyel olduğunu ifade etmektedirler. Ar-ge programlarında, bireylerde, teknolojilerde, sistemlerde şekillenen bilgi bir örgüt ya da bir milletin performansı için hayati derecede önemlidir.³³

Ekonomik alanda en belirgin ilerlemeyi sağlayan özellikle bilgi tabanlı toplumlarda faydalı bilgiye sahip olmaktır. Bilginin ekonomik alandaki bu etkisi onu daha da değerli hale getirmektedir. Firmalar gelişimlerini devam ettirebilmek, yeni teknolojiler geliştirmek, icatlar meydana getirmek için bilgiye ihtiyaç duymaktadırlar. İhtiyaç duyulan bu bilgi ise ekonomik ilerlemelerin temel dinamiği olarak görülmektedir.³⁴

Örgütlerin, işletmelerin ve insanların bilgiyi etkin olarak elde ettiği, meydana getirdiği, dağıttığı ve kullandığı bilgiye dayalı ekonomi kavramı, ekonomi yazınındaki yerini özellikle 1990'lı yılların başlarında almaya başlamıştır. Bilgi ekonomisindeki en önemli özellik üretim faktörü olarak bilginin ilk sırada yer alması gelmektedir. Bilgiye dayalı ekonominin en önemli unsurları ise bilgi ve iletişim teknolojileri başta olarak, inovasyon gücü, Ar-Ge faaliyetleri ve bilgi çalışanları olarak gösterilmektedir.³⁵

Son 3 yüzyıldan beri pazar ekonomilerinde zenginliğin ana kaynağı, doğal değerlerden (özellikle yer ve niteliksiz işgücü) somut oluşturulan değerler (özellikle bina, makine, ekipman, finans) aracılığıyla fiziksel değerlerde örgütlerde ya da insanlarda şekillenen soyut değerlere (özellikle bilgi ve bilişimin her türü) doğru yön değiştirmiştir.³⁶

Bilgi ekonomisinin ayırıcı özellikleri aşağıdaki gibi sıralanmaktadır:

³³Marti ve Cabrita, **a.g.e.**, s. 20.

³⁴Deniz Çalık, Özge Pelin Çınar, "Geçmişten Günümüze Bilgi Yaklaşımları Bilgi Toplumu ve İnternet", XIV. **Türkiye'de İnternet Konferansı**, İstanbul Bilgi Üniversitesi, 2009, s. 8.

³⁵Nihat Işık, Efe Can Kılıç", Bilgi Ekonomisi ve İktisadi Büyüme: OECD Ülkeleri Üzerine Bir Uygulama" **Akdeniz İİBF Dergisi** 26, 2013, s. 23.

³⁶John H. Dunning, **Regions, Globalization, and The Knowledge-Based Economy**, Oxford University Press, 2002, s. 57.

1) Bilgi, üretimin ana kaynağı olan sermaye ile yer değiştirmektedir ve soyut varlıklar firmaların katma değerinin ek bir bölümünü yansıtmaktadır.

2) Ürün ve hizmetler bilgi yoğunudur.

3) Kaynakların mülkiyeti değişmektedir. Temel ekonomik kaynak sahipliği, bilgi çalışanlarının beyinde bulunduğu için bireysel özellikte olmaktadır.

4) Üretim yapıları değişmektedir ve soyut kaynak yönetimi somut ve finansal kaynaklardan farklıdır.³⁷

Bilgi ekonomisinde verimliliğin artması önemli olduğu için yararlı bilgiye yatırım yapmak ve inovasyon gerçekleştirmek gerekmektedir. Bu ekonomide bilgi, ekonomik rekabetin ve başarının ana kaynağıdır. Yeni fikirlerin ve teknolojilerin uygulanması, verimlilikte artış yoluyla ekonomik üretime büyük çapta değer katmaktadır.³⁸

Bilgi ekonomisinin ortaya çıkması üretim faktörü olarak bilginin yenilik, öğrenme ve yetenekler üzerinde etkisinin artan rolü açısından tanımlanmaktadır. Bilginin dijitalleşmesiyle ağ kullanarak kodlanmasında büyük bir artış vardır. Buda bilginin giderek nesneleştirilmesine yol açmaktadır. Artan bilgi kodlaması bilginin depolama dengesinde yön değiştirmeye yol açmaktadır. Bilgi ekonomisi bireylerin ve firmaların sosyoekonomik durumunu belirleyen bilgi yoğun ve öğrenme yoğun ilişkileri birleştirmek için, erişim fırsatı ve yeteneği olan öğrenme oranı ve değişim oranının hızlanmasıyla işleyen network hiyerarşisidir. Firmaların öğrenen örgütler, sürekli adapte yönetimi, yeni fırsatları yakalama ve yeni teknolojilere uyum sağlama gibi özellikleri olmak zorundadır.³⁹

Bilgi tabanlı ekonomide yaparak öğrenmek en önemli faktördür. Öğrenmenin temel bir etkisi örtük bilginin kodlanmış bilgiye dönüşümüdür ve yeni örtük bilgi çeşitlerinin geliştirildiği yerler de uygulamaktır. Bilgi tabanlı ekonomi bilginin

³⁷ Daniel Andriessen, Rene Tissen, **Weightless Wealth: Find Your Real Value in a Future of Intangible Assets**, Financial Times Prentice, 2000, p. 5.

³⁸ Batagan, **a.g.e.**, s. 61.

³⁹ John Houghton, Peter Sheehan, A Primer on The Knowledge Economy, Centre of Strategic Economic Studies, Victoria University, 2002, p. 9.

oluşturulması kadar bilginin dağılımı ve kullanılmasında da önemli bir yer almaktadır. Bir bütün olarak ulus ekonomilerinin ve girişimlerinin başarı belirleyicileri bilginin toplanılması ve faydalanılması etkinliklerine bağlıdır.⁴⁰

Bilgi ekonomisinin ana bileşeni fiziksel girdiler ya da doğal kaynaklara bağımlılıktan ziyade entelektüel özelliklere bağımlılık özelliğine sahip olmasıdır. Bilgi yoğun sektörlerin büyümesi ve verimliliğin artması, 1990'larda sıra dışı makroekonomik ve finansal pazarların gelişmesi bağlamında artış meydana getirdiği için yaygın olarak literatürün çoğu bilgi ekonomisinin geçmişten farklı olarak temel bir şekilde faaliyet gösterdiğini belirtmektedir.⁴¹

1.3.1.Bilgi Ekonomisi ve İktisadi Büyüme

Tüketiciler henüz yeni olan teknolojilerin kullanıldığı serbest piyasaya güvenmemektedirler. Çünkü bilinmeyen teknolojiler firmalar için çok yeni ve belirsiz görünmektedir. Oysaki uygun ve tarafsız bir belirleyiciye ihtiyacı olan teknolojik olarak durgun bir ekonomide, yeni bilgi üzerinde yeni tekniklerin oluşturulması ve benimsenmesi ile ilişkili geri kazanım alamama gibi dezavantajlara, bilgilendirici asimetrilere ve yeni teknolojilere güvenmemek için bir neden bulunmamaktadır. Yeni teknoloji daha önce belirsizlik bulunmayan yere belirsizliği getirmektedir. Bireyler risk almama konusunda farklılaşırlar ve bu durum spesifik yeni tekniğin üreteceği her türlü amaçlanmamış sonuçları çözecek olan ek yeni bilginin elde edilip edilemeyeceği ile ilgili görülmektedir. Özellikle teknolojik değişimler süresiz bir tabiata sahip olduğu zaman, makro icatlar olarak adlandırılan faydaların ve maliyetlerin bilinmezliği yeni bilginin ekonomi politikasının merkezini oluşturmaktadır.⁴²

Sanayi toplumundan bilgi toplumuna geçişle birlikte ortaya çıkan bilgi ekonomisinde niteliksiz işgücü ya da sermayede meydana gelen artışlar yerine teknolojik değişiklikler ve bu değişikliklerin getirdiği bilginin gücü daha da önem

⁴⁰OECD, **The Knowledge Based Economy**, Paris, 1996, p. 14.

⁴¹Walter W. Powell, Kaisa Snellman. "The Knowledge Economy." **Annual Review Of Sociology**, 2004, p. 200.

⁴²Macleod, **a.g.e.**, p. 242.

kazanmıştır. Üretim faktörleri arasına en çok katkı sağlayacak bilginin de girmesini sağlamıştır.⁴³

Teknoloji daha düşük maliyetle daha fazla mal üretme imkanını sağlamaktadır. Bu nedenle toplumların hayat standartlarını yükseltmenin başlıca aracıdır. Ayrıca iktisadi gelişmenin ileri aşamalarından itibaren, sosyal eşitsizliği azaltmanın temel aracı yine o olmuştur.⁴⁴ Teknoloji, iktisadi gelişmenin ve refah artışının başlıca unsuru olarak görülmektedir. Bu nedenle bilgi toplumu, sanayi toplumuyla birlikte, aynı zamanda bir teknoloji toplumu olmaktadır. Sanayi toplumunun ve onun üzerinde yükselen bilgi toplumunun temeli ise teknolojiye dayanmaktadır. Çünkü ekonomik değişimler teknolojilerden kaynaklanmaktadır.⁴⁵

Bilgi giderek daha çok ekonominin büyümesine yön vermektedir. Ülkeler rekabet edebilmek için bilimsel bilgiyi kullanmak zorunda kalmaktadırlar. Ekonomik ve toplumsal anlamda ilerleme kaydedebilmek için bilgi ve iletişim teknolojilerinin, bilimsel bilginin, teknolojinin getirdiği yenilikleri güç olarak kullanmak gerekmektedir.⁴⁶

Yapılan çalışmalar genellikle ekonomik gelişim ve bilgi üretimi arasındaki pozitif ortaklığı vurgulamaktadır. Bilgi toplumunun altyapı parametreleriyle ekonomik gelişme arasında iki yönlü bir ilişki olduğu söylenebilmektedir. Bilgi toplumunun altyapı parametrelerinin gelişmesi ekonomik gelişmeyi hızlandıracaktır. Bununla birlikte herhangi bir toplumda bu parametrelerin yaygınlaşması da ekonomik gelişme seviyesinin yüksek olmasına bağlıdır.⁴⁷

Bilgi ekonomisinde bilginin ekonomideki önemi ve katkısı içsel büyüme modelleri ile de açıklanmaktadır. Bu modellerde iki unsur göze çarpmaktadır.

⁴³ Aykut Kibritçioglu, "İktisadi Büyümenin Belirleyicileri ve Yeni Büyüme Modellerinde Beşeri Sermayenin Yeri", **A.Ü. Siyasal Bilgiler Fakültesi Dergisi**, Cilt:53, Sayı:1-4, Ankara, 1998, s. 226.

⁴⁴ Simon Kuznets, "Croissance et Structure Economiques", Paris, 1972, s. 152.

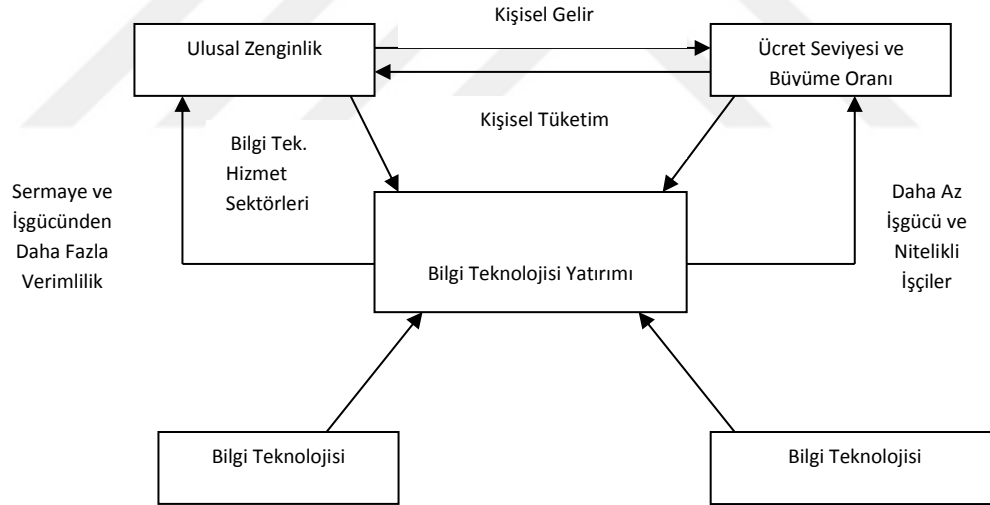
⁴⁵ John W. Kendrick, "Productivity Trends: Capital and Labour", **National Bureau of Economic Research**, 1957, p. 15.

⁴⁶ World Bank, "Lifelong Learning and The Knowledge Economy", **Summary of the Global Conference on Lifelong Learning Organized by The World Bank**, Stuttgart, Germany, 2003, p. 1.

⁴⁷ Herbert S. Dordick, Georgette Wang, **Information Society: A Retrospective View**, Sage Publications Inc., 1993, p. 34.

Bunlardan ilki, bilgi ve teknolojiye kaynaklanan pozitif dışsallık olarak açıklanırken, ikincisi ise, bilginin açığa çıkardığı getirinin ölçeğe göre artan olması olarak ifade edilmektedir. Pozitif dışsallıkla açıklanmak istenen bilginin seviyesine bakılmaksızın herhangi bir sektörde kullanılabilir olması ve değişik süreçlere dâhil olarak verimliliği artırması olarak anlatılmaktadır. Ölçeğe göre artan getiri ile anlatılmak istenen ise elde edilen bilginin faydalı alanlarda kullanma ve yayılma fırsatı olduğundan verimliliği de artırması olarak açıklanmaktadır.⁴⁸

Bilginin üretilmesi ve çoğaltılması zenginliğin elde edilmesinde baskın rol oynamaktadır. Bilginin sınırlarının genişletilmesi, bilginin ekonomik aktivitelerinin her türünde en etkin kullanımı faydalanılması ile ilgilidir.⁴⁹ Bilgi politika oluşturma sürecinde çok önemlidir. Çoğu gelişim politikası kamu yönetiminin tüm yönleri için iyi politika uygulamalarını tanımlaması ve dağıtılması üzerine yerleşmiştir.⁵⁰



Şekil 1.2. Bilgi Teknolojisi Yatırımlarından Ekonomik Hasılat.

Bilgi ve iletişim teknolojileri işletmeler arasında rekabet yoğunluklu bir ortam meydana getirmektedir. Rekabet halinde olan bu işletmeler genel olarak yaptıkları

⁴⁸ Arif Özsağır, **Bilgi Ekonomisi**, Seçkin Yayıncılık, 2013, s. 124

⁴⁹ DTI Competitiveness White Paper, "Our Competitive Future: Building The Knowledge Driven Economy", London: Department of Trade and Industry, 1998, p. 23.

⁵⁰ World Bank, "Building Knowledge Economies Advanced Strategies for Development", **WBI Development Studies**, Washington D.C., 2007, p. 9.

yatırımları üretim evresindeki çalışmalardan daha ziyade bilgi ve iletişim teknolojilerine ve telekomünikasyon faaliyetleri üzerine gerçekleştirmektedirler ki bu alandaki yatırımlar yüksek verimliliğe sahip olduğu gibi üretim sırasındaki karar alma mekanizmasını da hızlandırmaktadır.⁵¹ Üretim, istihdam ve yatırım, ileri teknoloji sektörlerinde ve eğitim enformasyon iletişim gibi bilgi yoğun hizmet sektörlerinde en hızlı gelişimi sağlanmaktadır. Aynı şekilde Ar-Ge yatırımları, iş gücü eğitimi, bilgisayar yazılımı ve teknik uzmanlık da önemlidir. Önemli OECD ekonomilerinin GSYİH'sinin %50'den fazlasının bilgi tabanlı olduğu tahmin edilmektedir. Bu ülkelerde verimlilik ve gelişim bilgi birikimi ve teknik ilerleme oranıyla belirlenmektedir.⁵²

Bilgi ve iletişim teknolojileri, ekonomik ve sosyal gelişimi sağlayan bir güç olarak tanımlanmaktadır. Bilgi ve iletişim teknolojileri her bir ülkenin özel ihtiyaçlarıyla ve kapasitesiyle uyumlu olarak yayıldığında ülke çapında ekonomik gelişimin hızlanmasını sağlayabilmektedir. Dünya bankası raporunda gelişmiş orta gelirli ülkelerin ve gelişmekte olan ülkelerin sağlıklı bir bilgi teknolojisi sektörüne sahip olarak fayda sağlayabildiklerini belirtilmiştir. Özellikle yazılım sektörleri, istihdam, ödemeler ve ekonomik performans üzerinde önemli bir etkiye sahip olabilmektedir.⁵³

Ekonomik büyüme birçok faktöre bağlı olarak sağlanmaktadır. Küreselleşen dünyada ekonomik büyümeyi sağlayan bu faktörlerden inovasyon, Ar-Ge, yenilikçi yazılımlar, farklı ürün tasarımları en önemli olanlarıdır. Bilginin ne kadar etkin kullanıldığı ülkelerin kendi güçlü taraflarını iyileştirmelerine, beşeri sermayeye ve kendi kurumlarını verimli hale getirmelerine, bilgi ve iletişim teknolojisi araçlarına, inovasyon tabanlı girişimleri arttırmalarına bağlıdır.⁵⁴ Bilgi ve iletişim teknolojileri ile ilgili yapılan yatırımların ekonomik ilerlemelerdeki etkinlikleri gelişmiş ülkelerde

⁵¹Ömer F. Durdu, "Avrupa Birliği Yolunda Bilgi Teknolojisinin Türkiye Ekonomisi İçin Önemi", **Akademik Bilişim Konferansları**, (Çevirimiçi) <http://ab.org.tr/ab03/tammetin/171.doc>, 13.05.2007

⁵²Bengt-Åke Lundvall, Dominique Foray, **The Knowledge Based Economy: From The Economics of Knowledge to The Learning Economy in Employment and Growth in The Knowledge Based Economy**, Paris,1996, p. 13.

⁵³ Paul Braund v.d., "Information and Communications Technology for Economic Development", **A Report on the Global E-Discussion, Technical Report, World Bank**, 2006, p. 9.

⁵⁴World Bank, "Measuring Knowledge in the World's Economies: Knowledge Assessment Methodology and Knowledge Economy", **Knowledge for Development Program**,2009, p. 1.

doğru orantılı olarak incelenmiştir. Gelişmekte olan ülkelerde ise sosyolojik ve kültürel açıdan, kuruluş yetersizliğinden doğru orantılı olarak ilerlemediği görülmüştür. Bu eksiklikler ortadan kaldırıldıktan sonra gelişmekte olan bu ülkelerde BİT yatırımları doğru orantılı olarak ekonomiyi etkileyecektir.⁵⁵

Ulusal ekonomi performansı üzerinde toplam bilgi ekonomisi indeksinin etkisini incelemek için, 56 ülke ve bölgesel gruplar üzerindeki Dünya Bankası bilgi ekonomisi alt bileşenleri analiz edilmiştir. Bu çalışmadaki veriler 1995-2001 yılları arasındaki verilerdir. Bilgi ekonomisi değişkenleri olarak ise ekonomik teşvikler, eğitim verileri, inovasyon yapısı incelenmiştir. Çalışma sonucunda ülkelerin ekonomik büyümeleri üzerinde bilginin çok önemli bir faktör olduğu ortaya çıkarılmıştır. Ülkelerin ekonomik büyüklük farklıları ise inovasyon, Ar-Ge, bilgi ve iletişim teknolojilerine yapılan yatırımlarla ifade edilmiştir. Çalışmanın sonuçlarında bilgi tabanını genişleten girişimlerin bir sonucu olarak ekonomik çıktı ve gelişimin arttığı belirtilmiştir.⁵⁶

OECD ülkeleri üzerine yapılmış çalışmada ise panel veri analizi uygulanarak bilgi ekonomisi ve iktisadi büyüme arasında ilişki incelenmiştir. Bilgi ekonomilerinde üreticilik, verimlilik, rekabet ve inovasyon potansiyelinin oldukça güçlü olduğunun belirtildiği çalışmada ekonomik gelişme ve bilgi ekonomisi değişkenleri arasında doğru orantılı bir ilişki olduğu açıklanmıştır. Bu incelemede bilgi ekonomisi değişkenleri olarak patent başvuru sayısı, BİT sermaye hizmetlerinin büyüme oranı, BİT mallarının ihracatı ve ileri teknoloji ihracatı artış oranı kullanılmıştır.⁵⁷

Türkiye ekonomisinde bilgi ve ekonomik büyüme arasındaki ilişkinin yönü ve büyüklüğünün araştırıldığı bir çalışmada bilgi ekonomisinin en temel göstergesi olan bilgi ve iletişim teknolojilerini temsil eden telekomünikasyon yatırımları kullanılmıştır. Büyüme değişkenleri olarak ise GSMH verileri, istihdam verilerinin

⁵⁵Rahmi Yamak, Necmettin Alpay Koçak, "Bilgi Teknolojisi Harcamalarının Ekonomik Büyüme Üzerine Etkileri: 1993-2005", **Bilgi Ekonomisi ve Yönetimi Dergisi**, 2007, s. 8.

⁵⁶Ahmed Driouchi, El Mustapha Azelmad, Gary C. Anders, An Econometric Analysis Of The Role Of Knowledge In Economic Performance., **The Journal of Technology Transfer**, **31.2**, , 2006, p.250.

⁵⁷ Işık ve Kılıç, **a.g.e.**, s. 24.

kullanıldığı çalışmada uzun dönemde ekonomik büyüme ile fiziki sermaye, işgücü, beşeri sermaye ve bilgi arasında pozitif bir ilişki saptanmıştır. Ayrıca çalışmada bilginin ekonomik büyümeye katkıda bulunduğu da değinilmiştir.⁵⁸

Ekonomik gelişim ve bilgi birbirleriyle yakından ilişkilidir. Bilgi ekonomisi indeksiyle 2008’de ölçülen bilgi birikimi ve ekonomik gelişim seviyesi arasındaki korelasyon % 87 olarak ölçülmüştür. Daha yüksek indeks değeri olan ülkeler ise ekonomik gelişimi daha yüksek ülkeler olarak görülmektedir.⁵⁹

1.3.1.1. İstihdam ve Bilgi İşçileri

Organizasyonların teknolojilerde meydana gelen yenilikleri benimsemesi, bu doğrultuda yeni ürün ve hizmetler tasarlaması ve bunları yaparken kalifiye eleman istihdamını arttırması bilgi ekonomisinde kilit rol oynamaktadır.⁶⁰

Bilgi ekonomisine geçiş süreci birçok yenilik ve değişimin meydana gelmesiyle sağlanmaktadır. Bunlardan bazıları işgücünde meydana gelen bilgi artışı, mal ve hizmetlerin farklılıklarının artması gibi değişimler olarak görülmektedir. Bu süreçteki yenilikler ve değişimler eğitim yapılarında, üretim mekanlarında, ticari kuruluşlarda gerçekleşmektedir.⁶¹Bilgi tabanlı işgücünün genişlemesinin önemli bir yönü özgün fikirlerin üretilmesi ve dağıtılmasını sağlayan bireylerin, rakip örgütlerin olduğu fiili toplumların gelişimidir.⁶²

Özellikle yeni teknolojilerin üretim sürecine başarıyla yansıtılması, standart nitelikte kitle üretimini sarsıntıya uğrattırırken, işletmelerin ihtiyacı olan işgücünün yapısını da değiştirmiştir. Teknolojik gelişme ve yayılma, rekabetin yoğunlaşması ve üretim modelindeki gelişme, vasıf düzeyi yüksek işgücüne olan talebi arttırmıştır. Küresel rekabete özellikle mikro teknolojilerdeki ve enformasyon teknolojisindeki

⁵⁸Sevda Yapraklı, Tuncay Sağlam, "Türkiye’de Bilgi İletişim Teknolojileri ve Ekonomik Büyüme: Ekonometrik Bir Analiz (1980-2008)", 2010, s. 580.

⁵⁹World Bank, **a.g.e.**, p. 7.

⁶⁰Dura ve Atik, **a.g.e.**, s. 65

⁶¹Charlie Karlsson, Börje Johansson, "Towards a Dynamic Theory for The Spatial Knowledge Economy", **Royal Institute of Technology Centre of Excellence for Studies in Science and Innovation 20**, 2004, p. 3.

⁶²Erik B. jolfsson, LorinHitt, "Information Technology as a Factor of Production: The Role of Differences Among Firms." **Economics of Innovationand New Technology**, 1995, 17.

hızlı gelişimin eşlik etmesi, sanayi ötesi ekonomilerden başlamak üzere istihdamda köklü değişimlere yol açmıştır.⁶³

Teknolojinin ilerlemesiyle birlikte istihdamda olumlu ve olumsuz, doğrudan ve dolaylı olarak etkilenmektedir. Mal ve hizmetlerin yenilenmesi, dağıtılması, üretim süreçlerinin geliştirilmesiyle istihdam doğrudan etkilenirken, teknolojinin kullanılmasıyla dolaylı olarak etkilenmektedir. Yeni ürünlerin genel olarak istihdamı geliştirdiği söylenirken, akışların yenilenmesinin ise herhangi bir etkiye sahip olup olmadığı belirsizliğini korumaktadır. Literatürde, insana yatırılan sermaye ve verimlilik arasında yer alan kuvvetli bağlantı, farklı teknolojiler ile insani sermayenin bütünleştiricilik ilişkisine sahip olduğunun da açık bir göstergesi olarak ifade edilmektedir. Çoğu ülkede ileri düzeyde bir ekonomik ilerleme performansının temellinde yer alan unsurlardan birinin, sahip oldukları ileri seviyedeki yetenekli işgücü havuzunun olduğu kabul edilmektedir.⁶⁴

Teknoloji sadece bir süreç inovasyonu olmayıp, ek işlerin meydana getirildiği tüm yeni ekonomik bölümleri ifade etmektedir. Birçok çalışma ürün inovasyonlarının ya yeni malların gelişimine ya da var olan malların farklılaştırılmasına imkân sağladığı için istihdam üzerinde pozitif etkiye sahip olduğunu göstermektedir.⁶⁵ Teknolojik ilerlemeler derecelerine bağlı olarak yetenek isteyen işgücü gerektirmektedir. Bu durum istihdamın yüksek beceri isteyen işgücü doğrultusunda artmasına yol açmaktadır ve bu becerilere sahip olmayan işgücünü ise azaltmaktadır. Teknolojinin gelişmesine bağlı olan bu beceriler sadece yetenekleri olmayan işgücü için değil, mevcut yeteneklerini teknoloji doğrultusunda ilerletmeyen işgücü içinde olumsuzluklar içermektedir.⁶⁶

Bilgi ve iletişim teknolojisi yatırımları insan kaynakları ve yeteneklerine ait yatırımları tamamlayıcıdır. Gerekli yetenekler ikame özellikte değil tamamlayıcı yeteneklerdir. Makineler sanayi çağında işgücüyle yer değiştirmiştir. Bilgi teknolojisi

⁶³TISK, "Küresel Eğilimler ve TÜİK Çalışma Kaydı", Yayın No: 164, Ankara, 1997, s. 76.

⁶⁴Mehmet Ali Kelleci, "Bilgi Ekonomisi, İşgücü Piyasasının Temel Aktörleri ve Eşitsizlik: Eğilimler, Roller, Fırsatlar ve Riskler. Devlet Planlama Teşkilatı, 2003, s. 26.

⁶⁵Mariacristina Piva, Marco Vivarelli, "Innovation and Employment: Evidence from Italian Microdata", *Journal of Economics*, 2005, p. 66.

⁶⁶Kelleci, a.g.e., s. 37.

bilgi ekonomisinde kodlanmış bilginin konumunu ifade etmektedir. Bilgi ekonomisindeki çalışmalar giderek daha çok istenilen yeni işgücü yetenekleri talep etmektedir. Bu yeni işgücü ve yetenekler kavramsal ve kişilerarası yönetim gibi iletişim yetenekleridir.⁶⁷

Birçok gözlem ekonominin bilgi tabanlı olarak karakterize edilmesine imkân sağlamaktadır. İşgücü talebinde daha fazla yetenekli çalışana doğru nispi bir yön değiştirme görülmektedir. Değişim oranındaki hızlanma bilgi ve yeteneklerin sürekli yenilendiği anlamına gelmektedir. Bu nedenle bilgi stokundaki artış bireyleri firmaları ülkeleri daha çok öğrenmeye itmektedir ve böylece öğrenen ekonomiyi oluşturmaktadır. Bireysel amaçlara ulaşmada daha başarılı olan öğrenme, bireye rekabet ve yeteneklerin edinilmesi olanağını sağlayan merkezi bir süreç olarak tanımlanmaktadır.⁶⁸

Yeni ekonomi, dünya ekonomisini hem makro hem de mikro seviyede etkilemiştir. ABD’de işgücü verimliliğinin ekonomiye katkıda bulunduğunu istatistiksel olarak açıklanmaktadır. Bilgi ve iletişim teknolojileri devrimi temel ekonomik istatistiklerin destekleyicisi görünmesi yeni ekonomiye doğru yön değişimine neden olmuştur. Bilgi ve iletişim teknolojileri sektörü Avrupa, Japonya ve ABD’de istihdamın % 3-4’ünü, katma değer % 6-9’unu, ihracatların % 10-25’ini, araştırma ve gelişim giderinin % 25-40’ını temsil etmektedir.⁶⁹

Bilgi tabanlı sektörlerde 1995-2005 yılları arasında istihdamda % 17 oranında bir artış yaşanmıştır. Yüzde olarak en büyük artış % 40 ile ileri teknoloji hizmetlerinde gerçekleşmiştir. Sağlık eğitim ve eğlence hizmetleri ise toplamda %27’lik bir artış göstermiştir. Yönetim ve iletişim hizmetleri ise % 23’lük bir artış göstermiştir.⁷⁰

⁶⁷Houghton ve Sheehan, **a.g.e.**, p. 11.

⁶⁸OECD, **Organisation for Economic Co-operation and Development, Knowledge Management in The Learning Society**, OECD Publishing and Centre for Educational Research and Innovation, 2000, p. 2.

⁶⁹HeliKoski, Petri Rouvinen, Pekka Ylä-Anttila, "ICT Clusters in Europe The Great Central Banana and The Small Nordic Potato", **Information Economics and Policy**, 2002, p. 146.

⁷⁰Brinkley, **a.g.e.**, p. 33.

İşgücü verimliliği üzerinde önemli çalışmalar yapmış olan John W. Kendrick'in hesaplarına göre, 1889-1957 döneminde ABD'nin bir saatlik işgücü başına reel üretimi, yılda ortalama % 2-2,5'luk bir artış kaydetmiştir. Bu aynı zamanda bir reel gelir artışını ifade etmekteydi, üstelik çalışma saatlerinde de %20-30 oranında bir azalma sağlanmıştır. Kendrick'e göre bu artış, yönetimin bilimsel esaslara oturtulmasının, araştırma ve geliştirmenin toplumun en iyi şekilde örgütlenmesinin eseri olarak görülmekteydi.⁷¹

Ülkeler bilgi temelli ekonomiye yaklaştıkça hizmet sektörü istihdamında da belirgin bir biçimde artış olmaktadır. Diğer taraftan hizmet sektöründe, geleneksel hizmetler giderek azalan bir seyir izlerken, bilgi temelli faaliyetlerde hızlı bir ivme ile artmaktadır. Bu anlamda hizmet sektörü içindeki bilgi işlem faaliyetlerinde artışlar gözlenmektedir. Bunun temel sebebi olarak teknoloji, bilgi ve beceri arasında tamamlayıcılık ilişkisinin varlığı olarak gösterilmektedir. Sanayi ekonomilerinde teknoloji ile işgücü arasındaki ikame ilişkisi yerini bilgi ekonomisinde teknoloji ile bilgi işçisi arasındaki tamamlayıcılık ilişkisine bırakmıştır.⁷² Bilgi ekonomisinde işlerin çoğu yüksek bilgi ve beceri gerektirdiğinden bu anlamda işgücü talebi artarken düşük becerili işgücü ikame edilmektedir. Zira bilgi ve iletişim teknolojilerine yapılan yatırımlar ile nitelikli işgücü birbirini tamamlayıcı unsurlardır.⁷³

Bilginin elde edilmesi belli bir çaba gerektirmektedir. Elde edilen bilginin uygulanması ise nitelikli işgücü ve kaynağa bağlı olarak görülmektedir. Gerekli kaynak ve işgücü ek ödemeleri de beraberinde getirmektedir. Yapılacak olan bu ek ödemeler farklılaştırmalar meydana getirmektedir. Bu durum ise ülkenin ekonomik ilerleme hızı ve seviyesiyle doğru orantılı olarak gelişmektedir. Gelişmekte olan ülkelerde genellikle mevcut iş gücü, talep edilen işgücünden oldukça fazla olarak görülmektedir. Nitelikli işgücü için iş bulmanın zor olduğu anlaşılmaktadır. Bu nedenle bu yetenek gerektiren işler için yapılan ödemeler düşük miktarlarda

⁷¹Kendrick, **a.g.e.**, p. 20.

⁷²Houghton ve Sheehan, **a.g.e.**, p. 11.

⁷³Luc Soete, "Macroeconomic and Structural Policy in The Knowledge-Based Economy; National Policy Challenges" **OECD Proceedings Industrial Competitiveness in the Knowledge-Based Economy; The New Role of Governments**, Paris, 1997, p. 136.

olmaktadır. Gelişmiş ülkelerde ise durum daha farklıdır. Nitelikli iş gücü talebi, bu ülkelerde fazla olduğu için ekonomik büyümeyle pozitif hareket etmektedir.⁷⁴ Değişen ve bilgi ekonomisi yönünde gelişen dünyada, istihdamında bu özel yeteneklere sahip bilgi işçilerine daha çok ihtiyaç duyduğu görülmektedir.

İnsan kaynağının kalifiye faaliyet göstermesi bilgi ekonomisinde oldukça önemlidir. Bu sayede insan kaynakları bilgi ekonomisinin bel kemiğini meydana getirmektedir. Bilgi çalışanları birçok özelliğe bir arada sahiptir. Kendi kendini idare eden, yüksek kalifiye ve güçlü olan, analitik yetenekleri olan ve bu yetenekleri faydalı işler üretmek için bilgi meydana getirebilmekte kullanan çalışanlardır.⁷⁵ Bilgi ekonomisinde daha bilgili çalışanlar istihdam edilmeye daha uygun çalışanlar olarak görülmektedir.

1955'te Peter Drucker tarafından dikkat çekilen bir terim olan bilgi çalışanı, güçlü bir kaynak olarak bilgiyi örgütlerin üretmesinden bizzat kendileri ortaya çıkaran bireyler olarak tanımlamaktadır. Vogt ise bilgi çalışanını fikirleri uygulamak, geliştirmek, eğitmek ve iletişim kurmak için yeteneği ve ortak çalışma kapasitesi, motivasyonu olan bir kişi olarak tanımlamaktadır. Bütün tanımların ana fikri bilginin oluşumuna ve faydalanılmasına çalışanların iştirak etmesidir.⁷⁶

Bilgi toplumu, öğrenen bir topludur. Ekonomik başarı ve sürekli inovasyon kültürü, çalışanların sürekli ve diğerlerinden daha güçlü öğrenme kapasitelerine bağlıdır. Bilgi ekonomisi makine gücü üzerinden değil, düşünme, öğrenme ve inovasyon gücünü ifade eden beyin gücü üzerinden işlemektedir.⁷⁷

Bilgi ekonomisinin bileşenleri yetenekli işgücü kapasitesi, Ar-Ge faaliyetleri, bilgi ve iletişim teknolojileri, global seviyede hareketli pazar ve iş ortamı olarak gösterilmektedir. Nitelikli insan faktörünün bilgi ekonomisinde küresel rekabet

⁷⁴Atik ve Dura, a.g.e., s. 164.

⁷⁵OECD Knowledge Management in The Learning Society, Organisation for Economic Co-operation and Development, , OECD Publishing and Centre for Educational Research and Innovation, 2000, p. 1.

⁷⁶Marti ve Cabrita, a.g.e., s. 31.

⁷⁷Andy Hargreaves, Teaching in The Knowledge Society: Education in The Age of Insecurity, Teachers College Press, 2003, p. 18-19.

açısından önemli bir faktör olduğu vurgulanmaktadır.⁷⁸ Ekonomide yaşanan gelişmelere paralel olarak ekonomide aranan, istenen emeğin niteliği de değişime uğramaktadır. Genel anlamda bilgiye dayalı ekonomiye geçmek ve özellikle bu ekonominin temelinde ağ, elektronik ve nicel aktivitelerle alakalı kısmı meydana getiren elektronik ekonomiye aktiflik kazandırılması sürecinde, temelinde bilgi teknolojileri olan yeteneklere sahip işgücü emeğine talep yükselmiştir. Bu anlatılan yeni süreçte, bilginin işlenmesi ve sorunların çözülmesi için hem bireysel hem de takım olarak iş yapabilme becerilerine, hareketli, risk alabilme cesaretine ve düşünebilme yeteneğine sahip bireylere daha çok ihtiyaç duyulmaktadır.⁷⁹

Bir bilgi işçisi, karmaşık görevleri tamamlamada kullanılan yöntemler, süreçler ve teknoloji hakkında değerli tecrübe, bilgi ve uzmanlığa sahip olan çalışandır.⁸⁰ Bilgi işçileri yararlı bilgiyi üretip, doğru zaman ve doğru mekânda kullanabilmektedirler. Bu kişiler işlerini genellikle zihinsel aktiviteler üzerinden gerçekleştirmektedirler ve bu nedenle işlerine odaklanmaları gerekmektedir. Bilgi işçilerinin özelliklerinden en önemlisi de tekrardan uzak aktivitelerle uğraşmalarıdır. İletişim yeteneklerinin kuvvetli olması, motivasyonlarının güçlü olması ve entelektüel kabiliyetleri olması, elde ettikleri bilgileri uygulama yetisine sahip olmaları bilgi işçilerinin genel özellikleridir.⁸¹

Bilgi işçisi, çalıştığı örgütte bilgiyi kendi başına bulma, anlama ve kullanma yeteneğine sahip çalışandır. Bilgi çalışanı, kendi öğrenme sorumluluğunu alır, kurumsal, ulusal ve uluslararası kaynaklardan ilgili bilimsel enformasyonu günlük çalışma hayatında müşteri sorunlarının çözümünde kullanabilir. Tipik bilgi işçileri; hukuk firmasındaki avukatlar, mühendislik firmasındaki mühendisler, imalat

⁷⁸Kamil Taşçı, Çağatay Telli, "Bilgi Ekonomisinin Kuramsal Çerçevesi" **XII'inci Türkiye'de İnternet Konferansı**, 2007, s. 19.

⁷⁹Özsağır, **a.g.e.**, s. 106.

⁸⁰Peter F. Drucker, "Knowledge-Worker Productivity: The Biggest Challenge", **California Management Review Vol.41. No.2, 1999**, p. 10.

⁸¹Taşçı, **a.g.e.**, s. 17.

firmasındaki ürün geliştiriciler, üniversitelerdeki öğretim üyeleri ve kamu dairelerindeki planlayıcılardır.⁸²

Genel olarak hizmet tabanlı sektörler içinde göz önüne alınan bilgiyi üretme faaliyeti, sağlıktan eğitime, iktisattan mühendisliğe, mimariden hukuka bir çok meslek ile alakalı çok geniş bir alanı kapsamaktadır.⁸³ Bilgi işçileri de farklı seviyelerde bilgi üretimi ile uğraşan, toplayıp, depolayıp, analiz ve koordine eden çalışanları ifade etmektedir.⁸⁴

Bir beden işçisi de çok sayıda değerli tecrübeye sahip olabilir fakat beden işçileri üretim araçlarına sahip değildir. Buna karşı, bir bilgi işçisi üretim araçlarına sahiptir. Bu bilgi, taşınabilir ve muazzam bir sermaye varlığını ifade etmektedir. Bu yüzden beden işçileri masraf olarak görülürken bilgi işçileri sermaye varlıkları olarak değerlendirilmektedir. Masraflar kontrol altında tutulmalı ve azaltılmalıdır. Buna karşı varlıklar daha da arttırılmalı, büyütülmelidir.⁸⁵ Bilgi işçileri kolayca ikame edilemezler, bu sebeple pazarlık güçleri yüksektir, dolayısıyla kendi haklarının korunması için kitle örgütlerine mavi yakalı işçiler kadar ihtiyaç duymazlar, gerektiğinde tek başına grev yapabilmektedirler, iyi eğitim almışlardır ve ücretleri yüksektir.⁸⁶

İnsani sermayeye yatırım bilgi ekonomisinde çok önemlidir. Bilim ve inovasyon sürecinin uzun dönem sürdürülebilir, idmanlı ve yeterli yeni araştırmacı ve öğreticilere bağlı olarak görülmektedir.⁸⁷ Bilgi tabanlı ekonomide, ücret primlerinden yararlanan, yetenekli, daha kalifiye çalışanlar için iş gücü pazar talebi giderek daha çok artmaktadır. Bazı ülkelerdeki çalışmalar, kalifiye eleman talebinin daha çok artmasıyla birlikte, bilgi yoğun üretim araçlarına daha hızlı giriş olduğunu göstermektedir. Farklı çalışmalar ileri seviyede teknoloji kullanan çalışanlara, daha

⁸²Petter Gottschalk, **Strategic Knowledge Management Technology**, IdeaGroup Publishing, London, 2005, s. 27.

⁸³John Garrick, Stewart Clegg, "Knowledge Work and The New Demands of Learning", **Journal of Knowledge Management** 4.4, 2000, p. 283.

⁸⁴Birsel Ersel, "Bilgi Çağında Çalışma İlişkileri ve Beyin Göçü", **2. Ulusal Bilgi, Ekonomi ve Yönetim Kongresi, Bildiriler Kitabı**, İzmit, 2003, s. 715.

⁸⁵Drucker, **a.g.e.**, s. 6.

⁸⁶Nusret Ekin, **Küresel Bilgi Çağında Eğitim-Verimlilik, İstihdam**, İTO Yayınları, İstanbul, 1999, s. 37.

⁸⁷OECD (2000), **a.g.e.**, p. 18.

yüksek ücretler ödendiğini, düşük kalifiye çalışanlara talebin azaldığını göstermektedir.⁸⁸

Bilim ve teknolojideki insan kaynağı oldukça önemlidir. Çünkü insani sermayedeki yatırım inovasyona ve gelişime göre ayarlanmaktadır. Gelişmekte olan ülkelerde bilgi ve teknoloji kullanıcıları, hayat standartlarını arttıran ve yeni meslekler meydana getiren teknoloji tabanlı ekonomik gelişimin merkezindedirler.⁸⁹

Bilgi işçileri bilgi ekonomisinin temelini oluşturmaktadır. Bilgi işçileri alanıyla ilgili teorik bilgiye, bu bilgiyi uygulama kabiliyetine sahip, iletişim yönü kuvvetli, entelektüel olarak yetenekli olan çalışanlardır. Firmaların karar almasında gerçekleşen işlemlerin maliyetleri, BİT ve inovasyon bilgisine sahip çalışanlarla azaltılmaktadır.⁹⁰ Bilgi işçileri için iş gücü piyasaları, hem bilgi ve iletişim teknolojileri hizmetleri gibi uzmanlık alanlarını hem de finans, sağlık ve eğitimden başka araştırma gibi uzmanlık alanlarını giderek daha önemli hale getirmektedir.⁹¹

Bilgi çalışanları mal ve hizmet üretmektense fikir ve bilgi üretmektedirler ve dağıtmaktadırlar. Bilgi çalışanları geçmişin hiyerarşi personelinde farklı bireylerdir ve mobildirler. Yetenekli insanları çalıştırmak zordur ve onları elde tutmak daha da zordur. Bilgi çalışanları patrona değil bilgi yöneticisine ihtiyaç duymaktadırlar. Bilgi çalışanlarının performansı sürekli gelişimsel olduğu için, değerlendirme, izleme, yönetme ve geri bildirim sunma aşmalarındaki yeteneklerini keskinleştirme ihtiyacı duymaktadırlar.⁹² BİT kullanan çalışanlar hizmetlerinde ve örgütlerinde tamamlayıcı inovasyon yaparlar. Bu çalışanlar BİT kullanımı, daha yaygın karar alma, üretim hattı çalışanları için daha geniş meslek sorumluluklarının yanı sıra daha çok kendi kendini yöneten takımları dahil eden yeni bir iş gücünün tamamlayıcısıdır. Hem BİT

⁸⁸ OECD, **The Knowledge – Based Economy, Organisation for Economic Co-operation and Development**, Paris, 1996, p. 15.

⁸⁹ Bill Lehr, Frank Lichtenberg, "Information Technology And Its Impact On Firm-Level Productivity: Evidence From Government And Private Data Sources. In: *Centre For Study Of Living Standards" Conference On Service Sector Productivity And The Productivity Paradox*, 1997, p.2.

⁹⁰ Işık ve Kılıç, **a.g.e.**, s. 23.

⁹¹ Brinkley, **a.g.e.**, s. 48.

⁹² OECD (2000), **a.g.e.**, p. 10.

hem de yeni örgüt birçok yolla ölçülebilen çalışan yeteneğiyle tamamlanabilmektedir.⁹³

Bilişim teknolojileri çalışanları ile ilgili yayınlanan bir çalışmada bu sektörlerle alakalı işçileri elde tutmanın ve onlardan verim elde edebilmenin sektördeki firmalar için stratejik önem taşıdığı ifade edilmektedir. Üretim unsurları açısından bilgi ekonomisinin en önemli faktörü olan bilgi katma değer üretmekte ve rekabet avantajı sağlamaktadır. Bilgi ve iletişim teknolojileri çalışanları, sahip olduğu yetenekler açısından katma değer sağlayan ve idare edilmesi zor olan çalışanlar olarak anlatılmıştır. Bu çalışanların işletmeler için önemli bir rekabet avantajı sağladığı çalışmada gösterilmektedir.⁹⁴

1.3.1.2. Verimlilik

Bilgi tabanlı üretim faktörleri, içsel gelişim teorileriyle birlikte özellikle 1990’larda gelişim süreci gösteren ve ekonomik anlamda geleneksel üretim faktörlerinden daha dominant hale gelmiştir. Bu değişimle birlikte bilginin rekabetçi gücünü ve önemini vurgulayan yeni ekonomik sistem yeni ekonomi diye tanımlanmaya başlanmıştır. Bununla birlikte yeni ekonomik sistem maliyet, kar gibi ekonomik kavramların yapısını da değiştirmiş ve hem makro hem de mikro seviyede geleneksel ekonomileri değişime zorlamıştır. Bilgi ve iletişim teknolojilerindeki ilerlemeler bilginin depolanmasını, analiz edilmesini ve transferini daha da kolaylaştırmıştır. Küreselleşmeyle birlikte ise bilgi ve iletişim teknolojileri, firmaların karlılığını, çıktılarının kalitesini ve verimliliğini arttırmıştır.⁹⁵

Bilgi ekonomisiyle ilgili makroekonomik araştırmaların çoğu, teknoloji ve işgücü verimliliği arasındaki bağlantıya odaklanmış durumdadır. Verimliliğin gelişimine bilgi teknolojisi yatırımlarının katkısının diğer yatırımlarınkinden fazla olduğu görülmektedir. Teknoloji uzun dönem verimliliğin artmasını sağlayan, çıktı

⁹³Timothy F Bresnahan, Erik Brynjolfsson, Lorin M. Hitt, "Information Technology, Workplace Organization and The Demand for Skilled Labor: Firm-Level Evidence", **National Bureau of Economic Research**, 1999, p. 52.

⁹⁴Zaim ve Koçak, **a.g.e.**, s. 77.

⁹⁵Arzdar Kıracı, Murad Tiryakioğlu, "Contribution of The New Economy to The Profitability of Turkish Firms: A Sectoral Analysis", Ed. By. Sedef Murat, İbrahim Güran Yumuşak, **Bilgi Ekonomisi ve Yönetimi Dergisi 5.2**, İstanbul, 2010, s. 3.

kalitesini arttıran ve maliyetleri azaltan, tamamlayıcı, örgütsel yatırımlara imkan sağlamaktadır. Ayrıca daha çok bilgisayar kullanan çalışanların sayısının artması, birlikte üretim araçlarının daha verimli olmasına imkân sağlamaktadır.⁹⁶

Hem verimlilik gelişimi hem de uzun dönem ekonomik gelişim için inovasyon ve teknoloji oldukça önemlidir. Belirleyiciler ekonomi ve istihdam üzerindeki teknolojik ilerlemenin etkilerini yakalamaya ihtiyaç duymaktadır. Özellikle verimliliği ölçmenin zor olduğu hizmet sektöründe Ar-Ge getiri oranlarının tespit edilmesinde genellikle zorluk yaşanmaktadır.⁹⁷

Yeni ekonomi ifadesi bilgi ve iletişim teknolojilerinin etkin kullanılması ve ekonominin darboğazlara girmeden aktif bir başarı sağlamasıyla açıklanmaktadır. Bu ifade ABD’de gözlenen en uzun ekonomik başarı dönemiyle birlikte 90’lı yıllarda literatüre girmiştir. BİT’lerinin kullanımıyla verimlilik açısından kayda eğer bir ilerleme kaydedilmesiyle ekonomik büyümenin de pozitif yönde gerçekleştiği belirtilmektedir. BİT sektörü bilgi ekonomisinin en önemli alanını oluşturmaktadır. Bu alandaki yeni firmaların, verimliliğini ciddi bir oranda arttırdığını ortaya çıkaran çalışmalar bulunmaktadır. Her firma hedefine ulaşmasa da sektörde hareketliliklerin olduğu gözlenmektedir.⁹⁸

İnternet teknolojileri aracılığıyla BİT’leri, organizasyonlara, kamu kuruluşlarına vb. geniş bilgi ağlarına erişim imkânı sağlamaktadır. Bu nedenle herhangi bir mekânda yeni üretilen bir bilgi oldukça hızlı bir şekilde bu birimlere aktarılabilmektedir. Bu aktiviteler dışsalılık olarak isimlendirilmektedir. Dışsalılıklar sayesinde bünyesinde bulunmayan bir bilgiden yararlanma imkânı kurum ve kuruluşlara sağlanmaktadır ve verimliliği arttırmaktadır.⁹⁹

1.3.1.3. GSMH/GSYİH

Ekonomi bir bütün olarak değerlendirildiğinde büyüklüğünün ölçülerinden olan GSMH ve GSYİH göstergeleri bilgi ekonomisi açısından da önemli değişkenler

⁹⁶Sandra E. Black, Lisa M. Lynch, "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity", **Review of Economics and Statistics**83.3, 2001, p. 440.

⁹⁷ OECD (1996), **a.g.e.**, p. 38.

⁹⁸Dura ve Atik, **a.g.e.**, s. 150.

⁹⁹Saygılı, **a.g.e.**, s. 100.

olarak görülmektedir. Bu noktada ülkelerin büyüme değişkenleri içinde yer alan bilgi ekonomisini geliştiren harcamaların değeri göz önüne alınmaktadır. Bilgi ekonomisi alanında yapılan harcamaların çoğu genel ülke ekonomisinin gelişimini ve büyümesini sağlamaktadır. Bu nedenle genel ülke ekonomisinin gelişimini gösteren GSMH ve GSYİH değerleri içindeki Ar-Ge ve inovasyon harcamaları, teknolojik yatırımlar gibi birçok gösterge son zamanlarda ekonominin gelişimine yön vermektedir. Bilgi ekonomisinin ana girdileri doğal kaynaklar ya da fiziksel girdilerdense entelektüel yetenekler üzerinde daha çok bağımlılık içermektedir. Bu entelektüel yetenekler Ar-Ge'den fabrika alanına, müşterilerden ara yüze, üretim sürecinin her aşamasındaki ilerlemeleri bütünleştirme çalışmalarıyla birleştirilmiştir. Bu değişimler soyut sermayeye atfedilebilen GSYİH'nın nispi payını yansıtmaktadır.¹⁰⁰

Bilgiye dayalı toplumun ekonomik anlamda iki ana parametresi olduğu belirtilmektedir. İlk parametre bilgi ekonomisiyle birlikte oluşan ve bilgi sektöründe gerçekleşen istihdam oranıyken, ikinci parametre bilgi sektöründe gerçekleşen getirinin GSMH'ya sağladığı katkı gösterilmektedir. Ekonomik düzen içerisinde yer alan bilgi sektörünün payı ile ilgili çalışmalar 1960'lı ve 1970'li yıllarda ilk olarak ABD'de yapılmış olup daha sonraları diğer ülkelerde devam etmiştir.¹⁰¹

Bilgi ekonomisine doğru gelişim sağlayan bir ülke, sürekli olarak verimliliğin artmasında ve ekonomik gelişimde baskın bir rol oynayan bilgiyle birlikte uzun dönemde sürdürülebilir GSYİH gelişim oranları elde edebilmektedir.¹⁰² Bilgi ekonomisinin gelişmesi daha fazla eğitime yatırım, inovasyon ve araştırma, gerektirmektedir. Bu ekonominin gelişiminin ölçülmesi GSYİH göstergesine dayanmaktadır. Yapılan hesaplamalara göre, önemli OECD ülkelerinin GSYİH'larının % 50'sinden fazlasının bilgiye dayalı olduğu tahmin edilmektedir.¹⁰³

¹⁰⁰ Jolfsson ve Hitt, **a.g.e.**, p. 17.

¹⁰¹ Dura ve Atik, **a.g.e.**, s. 102.

¹⁰² OECD (2000), **a.g.e.**,p. 3.

¹⁰³ OECD (1996), **a.g.e.**,p. 9.

1.3.1.4. Patent

Bilginin bir mal olarak niteliği fikri mülkiyet hakları uygulamalarının ekonomik mantığının anlaşılabilmesi için önem taşımaktadır. Çünkü fikri mülkiyet hakları bilginin mülkiyet konusu olmasını sağladığında aynı zamanda bir bilgi piyasası oluşmasını da sağlamaktadır. Bilgini rekabetçi olmayan bir mal olması, yeni bilgi üretilmesi için yapılan yatırımlar sonunda bilginin maliyetsiz olarak paylaşılabilmesine sebep olduğundan, bilgiyi üretenler için onu paylaşmak ve korumak arasında hep aynı çelişkiyi oluşturmaktadır. Bilginin paylaşılması karşısında bilgiyi korumak için gerekli temel mekanizmayı ise fikri mülkiyet hakları oluşturmaktadır. Fikri mülkiyet hakları, bilginin yasal mekanizmalarla dışlanılabilir bir mal olmasını sağladığı gibi bilginin bölünemez karakterinden dolayı aynı zamanda bilgi üzerinde tekeller de oluşturmaktadır. Fikri mülkiyet düzenlemelerinin karakterini açık biçimde yansıtan patent sistemi, teknoloji alanında üretilen her yeni bilgi için belirli süreler içinde tekel imtiyazı sağlamaktadır. Teknolojinin ekonominin temel değişkeni olduğu göz önüne alındığında patent sisteminin bu alanda önemli bir politika olduğu görülmektedir.¹⁰⁴

Patentler kendi fikirlerini yansıttıklarından dolayı bilgi oluşumunun doğrudan belirleyicilerine daha yakındırlar. En çokta bilgi çıktılarını doğrudan ölçmeye uygundurlar. Patent verileri, çoğu ülkenin merkezi veri tabanlarında organize edilen ulusal patent sistemlerine sahip olmaları açısından bazı avantajlara sahiptir. Bu veriler tüm teknolojik alanları kapsamaktadır ve patent belgeleri icat, teknoloji ve icat yapan kişiyle ilgili birçok bilgiyi içermektedir. Patentler, bilgideki daha genel ilerlemeler ve anlayışlardansa özel fikirlerin pratik uygulamasını temsil etmektedir.¹⁰⁵

Bilginin meydana gelmesi ve bilgi üretimi anlayışını kullanan birçok yazar süreç çıktısı olarak teknik inovasyondan ve teknolojik bilgiden bahsetmektedir. Yeni gelişim teorisinde, Ar-Ge sektörünün çıktısı ya öncekinden daha etkin olan yeni bir üretim süreci için bir tasarım, patent gibi özel mülk araçlarıyla korunabildiği

¹⁰⁴ Sinan Yıldız, "Patent Sisteminin Ekonomik Mantığı ve Tarihsel Analizi", **İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi**, 2010, s. 55.

¹⁰⁵ OECD (1996), **a.g.e.**, p. 32.

varsayılan malların, ya da rekabetçiler tarafından kolayca kopyalanamayan yeni yarı mamul malların üretimi olarak görülmektedir.¹⁰⁶

1.3.2. Bilgi Ekonomisi, Küreselleşme ve Rekabet

Gün geçtikçe firmalar iş dünyasında yeni başlangıçlar yapmaktadırlar, küresel pazarda faaliyette bulunmaktadırlar ve yeni teknolojilerle, kalifiye elemanlara yatırım yapmaktadırlar. Küreselleşme firmaların giderek daha fazla pazara girmelerini sağlamaktadır. Bunu yaparken rekabet avantajı içinde yeni potansiyeller sunmaktadır. Uzaklığın ortadan kalkması finansal pazarların küreselleşmesini sağlamaktadır. Yerel pazarlar, küresel sistemler aracılığıyla düşük maliyetlerle yirmi dört saat ticaretin yapılmasına bağlantı sağlamaktadır. Veri tabanlarında depolanmış bilgi hareket ettirildiğinde, dijitalleşmeyle araştırma maliyetlerinin de ortadan kalktığı yeni bir boyut eklenerek uzaklık faktörü kaybolmaktadır. Aynı durum giderek azalan maliyetlerle ses, video, veri ve resmin küresel aktarımını sağlayan iletişim teknolojileri içinde geçerli olmaktadır.¹⁰⁷

Bilgi ekonomisinin iki önemli tanımdan oluştuğu ifade edilmektedir. Bunlar ekonomik faaliyetlerin bilgi yoğunluğunun artması ve ekonomik faaliyetlerin küreselleşmesinin artmasıdır. Bilgi ve iletişim teknolojisindeki ilerlemeler ve teknolojik değişimlerin entegre gücü sayesinde bilginin yoğunluğunda artış meydana gelmektedir. İletişimin günden güne güçlenmesi teknolojinin büyük bir hızla giderek daha da gelişmesi küreselleşmenin temelini daha da kuvvetlendirmektedir. Aynı zamanda bilgi ekonomisi sadece bir bileşenle değil, meydana gelen bütün ekonomik yapıyla alakadar olarak görülmektedir.¹⁰⁸

Bilgi ekonomisinin başka bir özelliği ekonomik faaliyetlerin çok hızlı küreselleşmesidir. Dünya ekonomisine hızlı entegrasyon için temel yapı zorunluluğu temin eden iletişim ve bilgisayarlaşma devrimiyle birlikte, küreselleşmenin hızı ve genişliği bilgi teknolojilerindeki gelişme ve düzenlemelerin iç içe geçmiş etkilerinin

¹⁰⁶David Hargreaves, "Knowledge Management in The Learning Society", **Proceeding of Forum of OECD Education Ministers**, 2000, p. 5.

¹⁰⁷Julian Birkinshaw, Örjan Sölvell, "Preface", **International Studies of Management & Organization** 30.2, 2000, p. 84.

¹⁰⁸Houghton ve Sheehan, **a.g.e.**, p. 2.

bir sonucu olarak görülmektedir. İki ana etki küreselleşmenin evrelerini karakterize etmektedir. İlk olarak yabancı doğrudan yatırımlar ve sermaye akışları, ticaret akışlarından daha hızlı gelişmektedir. İkinci olarak bu süreç oldukça hızlı ve yaygın ilerlemektedir. Aynı zamanda tamamen anlaşılması mümkün değildir. Büyük çapta küreselleşmenin maliyetleri ve faydalarının açık bir görünüme sahip olması imkânsız olarak görülmektedir.¹⁰⁹

Bilgi ekonomisi toplumdaki zenginliğin kaynağı olarak bilginin artan öneminden bahsetmektedir. Zenginlik ilk zamanlar hammadde ve toprak sahipliği ile gerçekleştirilmekteydi. Daha sonraları sermayenin kıt kaynak olması ve sanayi devrimin gerçekleşmesiyle teknoloji ve becerinin değeri artmıştır.¹¹⁰

Teknoloji, tüketiciler tarafından gerçekleşen talep, dünyanın giderek daha da küreselleşmesi bilgi ekonomisinin itici güçleridir. Günümüzde hizmet temelli sektörler doğru talebin kaydığı görülmektedir. Mal ve hizmetlerin üretim sürecinde bilgisayar ve internet kullanımı, bilgi tabanlı ekonomileri daha da geliştirmektedir. Ayrıca bu süreçte gerçekleşen yatırım ve ulaşım maliyetleri de kayda değer ölçüde bilgisayar ve internetin kullanımıyla azalmaktadır. Bilgi yoğun sektörlerde yeni piyasaların açılması ve ticaretin uluslararası hale getirilmesi rekabetin artmasını sağlamaktadır.¹¹¹

İletişim teknolojilerindeki ilerlemeler ve serbestleşmeler sonucu dünya rekabeti genişlemektedir ve küresel rekabetin yeni bir formu meydana gelmektedir. Başarılı bir şekilde rakiplerle rekabet edebilmek için firmalar bütün pazarlarda başa baş rekabet etmek zorundadır. Bu yeni çevrede, rekabet edebilirlik büyük ölçüde uzmanlaşmış, yönetsel, ticari, teknolojik, finansal, sektörel ve kültürel özellikler arasında oluşturulan sinerji ve koordinasyona bağlıdır.¹¹² Uygulamaya yönelik, yüksek kâr marjlı bazı sektörler kendilerine has önemlerini yitirirken, uluslararası arenada ekonomi ve rekabeti değiştiren dinamikler yeni sektörlerin meydana çıkması olarak görülmektedir. Önceki sektörler son yıllarda gelişen bu ekonomik ortama

¹⁰⁹ Marti ve Cabrita, **a.g.e.**, p. 20.

¹¹⁰ Birkinshaw ve Sölvell, **a.g.e.**, p. 85.

¹¹¹ Brinkley, **a.g.e.**, p. 15.

¹¹² Houghton ve Sheehan, **a.g.e.**, p. 12.

uyum sağlayamadıklarında rekabet çevrelerini ve çalışanları yitirmektedirler. Teknolojik ilerlemeler doğrultusunda gelişen yeni sektörler ise yetenekli çalışanları barındırmaktadırlar.¹¹³

Dünyada yayılan kompleks bilgiyi kullanmak, birleştirmek ve algılamak oldukça zor görülmektedir. Bu güçlükle yüzleşmek için bilgi ekonomisinin öncüleri üç farklı alanda rakiplerini geride bırakmak zorunda kalmaktadırlar. Rekabetin ilk seviyesi dünyanın önemli yerlerinde ortaya çıkan potansiyel pazar bilgisine, pazar yeri ve ilgili teknolojilere erişme ve tanımlama yarışı olarak açıklanmaktadır. Rekabetin ikinci seviyesi firmaların küresel olarak dağılan bilgi parçalarını bir araya getirebilme ve onları yenilikçi ürünleri, hizmetleri ve süreçleri oluşturmada kullanma hızı ve etkinliği olarak görülmektedir. Rekabetin üçüncü seviyesi küresel pazarlar karşısında hızlı ve maliyet etkin bir şekilde ürün, hizmet, süreç ve inovasyonlarını dengelemek için tedarik zinciri, pazarlama dağıtım ve küresel satışların etkinliğinin optimize edilmesidir.¹¹⁴

Küresel yayılma belirgin bir rekabet avantajı oluşturmaktadır. Üretim, dağıtım ve satışların etkin bir küresel ağının bulunması gerekmektedir. Değerli bilgi giderek daha çok dağıtılmaktadır. Rakiplerinden daha iyi performans gösterme ihtiyacı hisseden firmalar, yeteneklerini dünya uzmanlaşmış farklı alanlarda artan bir şekilde yaymaktadırlar. Rakiplerinin ilerisinde yeni farklılaşma kaynağı ile ilgili bilgi geliştirebilen, değerlendirilebilen ve keşfedilebilen firmalar için yeni teknolojiler ve rakipler ortaya çıkmaktadır.¹¹⁵ Sistemlerin bilgi teknolojilerine bağımlılığı firmaların hem rekabetçilerine hem de ortaklarına odaklanmak zorunda olduğu anlamına gelmektedir.¹¹⁶

Kurumsal Ar-Ge kapasitesinin gelişimiyle araştırma merkezlerinin önemi giderek artmıştır. Bütün firmalar için, üretim fonksiyonunu optimize etmek amacıyla yeni bir potansiyel sunan küresel pazarların ortaya çıkması, bilginin değerini ve

¹¹³ Hakkı Okan Yeloğlu, "Bilgi Ekonomisi Değişkenlerine Yönelik İlk İzlenimler: Türkiye-OECD Ülkeleri Karşılaştırmaları (1995-1999)", **Bilgi Dünyası** 10.2, 2009, s. 246.

¹¹⁴ Doz, Santos ve Williamson, **a.g.e.**, p. 6.

¹¹⁵ **A.e.**, p. 10.

¹¹⁶ Carl Shapiro, Hal R. Varian, **Information Rules: A Strategic Guide to The Network Economy**, Harvard Business School Press, Cambridge, 1999, p. 10.

rekabet avantajının deęerini daha ok arttırmıř olup, firmaların kendilerine mahsus uygulamalarının da artmasını saęlamıřtır.¹¹⁷ Kreselleřme, sermaye ve iřgc piyasasındaki arz tarafının sınırlarının kaldırılmasına yardım etmektedir. Firmalar ileri teknolojiye sahip řirketlerin bulunduęu sektrlerdeki yatırım sermayesi iin dięer finansal kurumlar aracılıęıyla dnya piyasalarına eriřebilmektedir. oęu yabancı doęrudan yatırım bilgi ekonomisine odaklanmaktadır.¹¹⁸ Ekonomik bařarıyı saęlamak, rekabet avantajı iin ana kaynak olarak bilgi, yetenek ve yenilikilik potansiyeli gibi soyut varlıklardan etkin olarak faydalanmaya dayanmaktadır. Ekonomiler her zaman inovasyon ve bilgiyle ilerlemektedirler. Bilgi ekonomisinin temel dinamikleri yeni teknolojiler ve kreselleřme ile mřterilerden elde edilen pazar talebidir.¹¹⁹

Kreselleřme inovasyon adımlarını hızlandırmaktadır ve teřvik etmektedir. Birok bilgi tabanlı sektr, aęırlıklı olarak uluslararası rekabete aıktır. Btn sektrler iin inovasyondan, kresel aęlar arasındaki yeniliklerin, fikirlerin ve teknolojilerin srekli akıřından faydalanmak bilgi ekonomisindeki firmaların yeteneęine baęlıdır. Firmaların Ar-Ge ve inovasyon faaliyetlerini geliřtirmesi rekabet ve kreselleřmeyi en doęru ve hızlı řekilde takip edebilmelerini saęlamaktadır. Geleneksel iktisat kıt kaynakların etkin kullanımını optimize etmeyi arařtıran bir sistem zerine kurulmuřtur. Ancak bilginin zelliklerinden dolayı kıtlıęın anlamı giderek deęiřmektedir. Maliyet rekabeti zorunlu olmakla birliktebařarı iin yeterli deęildir. İnovasyon ve bilgi, rekabet avantajı oluřturmada ve devam ettirmede yine en nemli faktr olmaktadır.¹²⁰

1.3.3. Bilgi Ekonomisi, Ar-Ge Faaliyetleri ve İnovasyon

İnovasyon retilmiř yeni bir rnn ya da deęiřtirilerek geliřtirilmiř farklı bir rnn, srecin uygulanmasıdır. Burada deęiřtirilmiř olan řey sadece somut bir mal deęil aynı zamanda yeni bir ynetim sreci de olabilmektedir. Yeni ve farklı dřncelerin katma deęer oluřturabilecek faaliyetlere dnřtrlme sreci olarak

¹¹⁷Birkinshawve Slvell, **a.g.e.**, p. 85.

¹¹⁸Brinkley, **a.g.e.**, p. 47.

¹¹⁹**A.e.**, s. 15.

¹²⁰Houghton ve Sheehan, **a.g.e.**, p. 14.

inovasyon tanımlanabilmektedir. İnovasyon denilen bu süreç iki adımdan meydana gelmektedir. İlk adım fikir tasarlamak yani yeniliği ya da değişikliği ortaya çıkaracak olan düşünceyi tanımlamaktır. İkinci adım ise üretilen bu fikirlerin aksiyon haline gelmesini yani ticarileşmesini sağlayan aşamadır. Bu adımda gerçekleştirilecek olan faaliyetler için emek ve sermaye gerekmektedir.¹²¹

Küreselleşme bilim ve teknoloji politikalarıyla ilgili yeni zorluklar da sunmaktadır. Teknoloji giderek uluslararası bir kamu malı olarak görülmektedir, ulusal inovasyon sistemlerini giderek daha geçirgen yapmaktadır. Yeni teknik bilginin genişlik ve derinliği girişimci kültür ve inovasyon alt yapısını geliştirmek zorundadır.¹²² Ülkelerin kalkınmasının gerçekleşmesinde en büyük faydayı yeni bilgilerin ve fikirlerin üretilmesi sağlamaktadır. Bu noktada inovasyon kabiliyeti yüksek olan ülkeler avantajlı görülmektedir. Bu faaliyetlerin hızlı olarak gerçekleşmesi ve bilginin kolaylıkla elde edilmesi verimlilik ve kalitenin artmasına öncülük etmektedir.¹²³

Bilgi üretimiyle, bilgi çıktısını elde edilmesi ve inovasyonun gerçekleştirilmesi için iki neden vardır. İlki inovasyonla var olan bilgiye yeni şeylerin eklenmesinin yansıtılmasıdır. İkincisi inovasyonun aranan bilgiyi bulmaya çalışmasıdır.¹²⁴

1988-2007 yılları arasındaki döneme ait yapılan bir çalışma da inovasyon kapasiteleri ile iktisadi büyümeleri arasındaki ilişki araştırılmış ve bölgesel inovasyon kapasitesinin ekonomik etkisi belirlenmeye çalışılmıştır. Bu çalışma ABD'nin belli metropol bölgelerinde yapılmıştır. Bilgi ekonomisi fikirlerle ve üretkenlikle açıklandığı için inovasyon kapasitesinin bölgesel ekonomik gelişmeyi nasıl kolaylaştırabildiğini göstermek önem arz etmektedir. Gelişimin öncüleri olarak bilgi teknolojileri ve inovasyon kapasitesinin fiziksel altyapının yerini aldığı, internet üzerinden şirket kurulması döneminden sonra bu kapasitenin ekonomik gelişmenin

¹²¹OECD, EUROSTAT, **Guidelines for Collecting and Interpreting Innovation Data**, Oslo, 2005, s. 114.

¹²²Houghton ve Sheehan, **a.g.e.**, p. 21.

¹²³Dominique Foray, **Economics of Knowledge**, MIT Press, 2004, s. 21.

¹²⁴Brief OECD Policy, "Science, Technology and Innovation in The New Economy", **OECD Observer**, 2000, p. 2.

giderek daha güçlü olmasını sağladığı ortaya çıkmıştır.¹²⁵ Bunun yanı sıra inovasyonun faydaları birçok şekilde açıklanmaktadır. İlki geniş kullanımda mallar için maliyet azalımını ifade etmektedir. İkincisi inovasyonun daha heyecan verici sonucu olan tamamen yeni mal ve hizmetlerin geliştirilmesidir.¹²⁶ İleri ekonomiler refah ekonomisinin ana sürücüsü olarak inovasyona giderek daha çok güvenmektedirler.¹²⁷

İnovasyon bilgi ekonomisinin en önemli dinamiklerinden birisidir. Ekonomik büyümenin inovasyonla sağlanması, kalifiye işgücünün, yeni ve faydalı ürünlerin, üretim yöntemlerinin, yeni örgütlerin geliştirilmesiyle gerçekleşmektedir. İnovasyon sürecinde ülkelerin üniversiteler, teknoloji firmaları, araştırma merkezleri gibi birimleri hayati anlamda önem taşımaktadır. Bu birimler ülkelerin inovasyon faaliyetlerini geliştirirken, kapasitesini de genişletmektedirler. Teknolojiye bağlılık derecesi yani o andaki teknolojiden öteye gidememek inovasyon tabanlı ekonomiye sahip ülkelerde araştırma merkezlerinin kalitesiyle, ürünlerin katma değerleriyle ters orantılıdır. İnovasyonun ekonomik gelişmişlik açısından kilit rol oynadığının farkında olan ülkeleri Ar-Ge faaliyetlerini geliştirmekte, patent geliştirmekte ve insan kaynaklarına daha fazla ağırlık vermektedirler.¹²⁸

Ar-Ge aktiviteleri, bilgi potansiyelini geliştirmek için düzenli olarak devam eden verimli çalışma ve elde edilen bilgileri değişik uygulamalarda kullanma olarak tanımlanmaktadır. İcra edilen iş üzerine yeni düşüncelerin tasarlanması gibi resmi olmayan teknikler de bu süreçte kullanılmaktadır. İşletmelerde bilginin üretilmesini etkileyen unsurlar; Ar-Ge araştırmaları, Ar-Ge faaliyetlerinin verimliliği, beşeri sermayenin seviyesi, işletmenin örgütsel yapısı ve inovasyon için sağlanan teşvik uygulamaları şeklinde sıralanmaktadır. Bilginin elde edilmesini yakından etkileyen faktörlerin başında beşeri sermaye, Ar-Ge faaliyetleri, dışarıdaki kaynaklarla bağlantı kurma kabiliyeti ve kanıtlanan bilgiye ulaşma olarak görülmektedir. Bilgiyi

¹²⁵ Benjamin JKYeo, "DrivingThe Knowledge Economy: Explaining The Impact of Regional Innovation Capacity on Economic Performance", **Contemporary Management Research**6.1, 2010, p. 83.

¹²⁶ Jolfsson ve Hitt, **a.g.e.**, p. 17.

¹²⁷ Philip Ternouth, "Universities in The Knowledge Economy: Higher Education Organisation and Global Change", **Universities as Knowledge Producers**, London, 2012, p. 41.

¹²⁸ OECD, **The Impact of the Crisis on ICTs and Their Role in The Recovery**, 2009, p. 33.

elde ettikten sonra kullanmak ise organizasyona fayda sağlamak için yenilikçi üretim süreçlerinin uygulanması ile gerçekleştirilmektedir.¹²⁹

Ar-Ge faaliyetleri depolanan bilginin artmasını sağlayan ve depolanan bu bilginin yeni ve değişik uygulamalarda kullanımını gerçekleştiren aktiviteler olarak tanımlanmaktadır. İşletmelerin verimlilik seviyeleri gerçekleştirdikleri Ar-Ge faaliyetleri ile pozitif olarak ilerlemektedir. Ar-Ge aktiviteleri sonucu kazanılan ve depolanan bilgiyle üretilen yeni ve değişik ürünler işletmelerin karlarının da artmasını sağlamaktadır. Ar-Ge faaliyetleri sadece yeni ve değişik teknolojilerin meydana çıkmasını değil aynı zamanda gelişmiş ekonomilerde tasarım odaklı faaliyet ve süreçlerinde ortaya çıkmasını sağlamaktadır.¹³⁰

Ülkelerin teknolojik olarak sahip olduğu yetenekler ilk olarak yaptıkları Ar-Ge harcamalarıyla ortaya çıkmaktadır. Ar-Ge faaliyetlerine yapılan giderler yeni mal ve hizmet geliştirme, yeni üretim süreci geliştirme, eldeki teknolojilerin etkin olarak uygulanması, adapte edilmesi gibi aşamalarda önemli rol oynamaktadır.¹³¹

Ekonometrik çalışmalar Ar-Ge harcamasının verimlilikte bir artışla bağlantılı olduğunu tahmin etmektedir. Son çalışmalar ülkeler arasında bu noktada farklılıkların olduğunu ve yabancı Ar-Ge faaliyetlerinin ve yurt içi şirket Ar-Ge faaliyetlerinin yüksek seviyelerinin ülkelerin ekonomik gelişiminde önemli bir etkiye sahip olduğunu göstermektedir. Bu da bilgi gelişiminin boyutunun, bir ülkenin yabancı teknolojileri benimseme yeteneğine bağımlı olduğunu ortaya çıkarmaktadır.¹³²

Yapılan birçok bilimsel çalışmada bilgi ekonomisi değişkenleri olarak Ar-Ge faaliyetleri ile ilgili birçok sayısal gösterge değişken olarak kullanılmaktadır. Türkiye ve AB ülkelerinin bilgi ekonomisi göstergelerinin girdi ve çıktı değişkenleri olarak ele alındığı bir çalışmada veri zarflama analizi yapılmıştır. Girdi değişkenleri olarak BİT harcamaları, kamu Ar-Ge harcamaları, işletmelerin Ar-Ge harcamaları, fen ve mühendislik mezunları sayısı, kamu eğitim harcamalarının GSYİH içindeki payı ve

¹²⁹ Blakeley, Lewis ve Mills, **a.g.e.**, p. 8-9.

¹³⁰ Manual Frascati, **Proposed Standard Practice for Surveys on Research and Experimental Development**, 2014, p. 250.

¹³¹ Cohen ve Levinthal, Wesley M. Cohen, Daniel A. Levinthal, "Innovation and Learning: The Two Faces of R&D", **The Economic Journal**, 1989, p. 70.

¹³² Lundvall ve Foray, **a.g.e.**, p. 12.

insani gelişme indeksi alınmıştır. Çıktı değişkenleri olarak ise ileri teknoloji ürünlerinin ihracat içerisindeki payı, Amerikan patent enstitüsüne kayıtlı patent sayısı, bilgi yoğun hizmetlerde çalışanların toplam istihdama oranı, yüz kişi başına düşen cep telefonu aboneliği sayısı, internet erişimli işletmelerin yüzdesi ve internet kullananların sayısı alınmıştır. Bu girdi ve çıktı değişkenleriyle AB ülkeleri ve Türkiye'nin bilgi ekonomisi performansları ölçülmüştür.¹³³

Bilgi ekonomisinin temel bazı göstergeleri açısından Türkiye ve AB ülkelerinin karşılaştırıldığı başka bir çalışmada ise kullanılan değişkenler Ar-Ge yoğunluğu, Ar-Ge istihdamı, bilimsel yayın sayısı, ileri teknoloji ihracatı rakamları, patent başvuru sayıları, ticari marka başvuruları ve bilgi ve iletişim teknolojileri harcamalarıdır. Söz konusu değişkenlerin rakamları açısından Türkiye'nin AB ülkelerine göre çok geride yer aldığı görülmektedir. Özellikle Ar-Ge rakamları ve ileri teknoloji ihracatı rakamları açısından diğer değişkenlere göre daha da geride yer almaktadır.¹³⁴ Bilgi ekonomisinin en önemli başlıklarından olan Ar-Ge ve inovasyon faaliyetleri dolaylı yoldan ve doğrudan bilgi ekonomisinin gelişmesine ve ülke ekonomilerinin büyümelerine ciddi büyüklükte katkı sağlamaktadır.

1.3.4. Yaşam Boyu Öğrenme (Lifelong Learning)

Bilgi ekonomisine geçiş aşamasında eğitim, yetenekli personel için gerekli eğitimler ve yaşam boyu öğrenme ifadeleri daha da anlam kazanmıştır. Eğitim hane halkı, firmalar, kamu kurumları için beşeri sermayenin hayati bir bileşeni olarak görülmektedir.¹³⁵ Yaşam boyu öğrenme eğitim sürecinde meydana gelen öğrenim uygulamalarını içeren bir kavram olarak tanımlanmaktadır. Örgün eğitim sürecinin başka bir seçeneği değildir. Ancak örgün eğitimde ki eksikliklerin giderilmesini ve ortaya çıkarılmamış becerilerin açığa çıkarılması olarak açıklanabilmektedir.¹³⁶

¹³³ Yaylalı vd., "Türkiye ve Avrupa Birliğine Üye Ülkelerin Bilgi Ekonomisi Performanslarının Veri Zarflama Analizi Metoduyla Karşılaştırılması", **Marmara Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi C.21 S.1**, 2007, s. 10.

¹³⁴ Funda Rana Adaçay, "Bilgi Ekonomisine İlişkin Temel Göstergeler Açısından Avrupa Birliği ve Türkiye'nin Karşılaştırılması", **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi 19**, 2007, s. 189.

¹³⁵ Kelleci, **a.g.e.**, s. 40.

¹³⁶ Turan Selçuk Göksan, Soner Uzundurukan, S. Nilay Keskin, "Yaşam Boyu Öğrenme ve Avrupa Birliği'nin Yaşam Boyu Öğrenme Programları", **İnşaat Mühendisliği Eğitimi Sempozyumu**, 2009, 144

Sivil toplumun katılımı, bölgesel bilgi ve kapasite sunabilirken aynı zamanda gelişim için bilgi ve iletişim teknolojileri girişimlerinin sürdürülmesi ve bu teknolojilerin uygulanması önemli olmaktadır. Sivil toplum katılımı, şeffaflığın ve iyi yönetimin gelişmesine, bilgi ve iletişim teknolojileri yeniliklerinin ölçeklendirilmesi için bir güven çevresi oluşturulmasına katkıda bulunabilmektedir.¹³⁷ Yaşam boyu öğrenme hem ulusal ekonomik rekabete hem de sosyal bağlılığa göre ayarlanmaktadır. Yaşam boyu öğrenme sistemleri birçok farklı eksen boyunca sınıflandırılabilir. Bu eksenler kurumsal yapılarına göre, yönetim ve düzenleme şekillerine göre, baskın eğitim bilimlerine göre ve bilgi-öğretim programı geleneklerine göre şekillenmektedir.¹³⁸

Yaşam boyu öğrenme ile ilgili yapılan bir çalışmada ülkelerin değişkenleri ele alınarak kümeleme analizi yapılmıştır. Çalışmadaki ülkeler Avrupa Birliği üye ülkeleri ve Türkiye olarak alınmıştır. Kullanılan değişkenler ise eğitim gören 25-64 yaşa arası yetişkin nüfus değişkeni, BİT giderlerinin GSYH'deki yüzdesi, bilişim teknolojilerini aktif olarak kullanan bireylerin genel istihdam içindeki oranı, bilişim teknolojileri uzmanlık seviyesindeki çalışanların genel istihdam içindeki yüzdesidir.¹³⁹

1.3.4.1. Bilgi Ekonomisinin Gelişiminde Eğitimin Yeri

Ekonomik aktivitelerin artan bilgi yoğunluğu ile ilgili iki temel neden kabul edilmektedir. Bunlardan ilki, işgücünün artan eğitim seviyesidir. İkincisi ise yeni ürün ve hizmetlerin gelişiminin bütünleştirici bir bölümü olarak gerçekleştirilen bilimsel ve teknik bilgidir.¹⁴⁰

Bilgi ekonomisi daha komplike, daha zeki ve daha eğitilmiş müşteriler ve şirketler tarafından oluşturulan daha çok katma değerli mal ve hizmetler için talepleri kapsamaktadır.¹⁴¹ Birçok teknolojik inovasyon, eğitilmiş çalışanlar için talebin

¹³⁷ Braund vd., a.g.e., p. 15.

¹³⁸ Andy Green, "Models of Lifelong Learning and The 'Knowledge Society'", **Compare** 36.3, 2006, p. 312.

¹³⁹ Bahar Berberoğlu, "Yaşam Boyu Öğrenme İle Bilgi ve İletişim Teknolojilerinin Açısından Türkiye'nin Avrupa Birliği'ndeki Konumu", **Bilgi Ekonomisi ve Yönetimi Dergisi** 5.2., 2010, s. 15.

¹⁴⁰ Lundvall ve Foray, a.g.e., p. 12.

¹⁴¹ Brinkley, a.g.e., p. 32.

artmasına yol açan ve ilgili inovasyon teknolojisinin bilgisine ve tamamlayıcı yeteneklere sahip olan çalışanları gerektirmektedir.¹⁴²

Kalifiye yetenekler ve yetilerin bulunduğu ortamlarda bilgi ve iletişim teknolojilerinin gelişmesi ve yayılmasıyla birlikte, meydana gelen ağların dış dünya ile bağlantısının kurulması daha kolay olmaktadır. Bu yetenek ve yetilerin kazanılması noktasında bilgi ekonomisinde eğitimin ne kadar önemli olduğu görülmektedir. Dolayısıyla yükseköğretim de yeni teknolojilerin geliştirilmesi ve öğretilmesi oldukça önemlidir. Yetenek sahibi personel arzının artması ve ekonomik büyümenin sürdürülmesi de konunun başka bir boyutudur.¹⁴³

Bilgi tabanlı ekonomi kodlanmış bilgiyi ve bu bilgiyi kullanmak için yetkinlikleri sürekli öğrenme ihtiyacı ile nitelemektedir. Çalışanlar hem resmi eğitime hem de yeni, teorik ve analitik bilgiyi uygulama ve elde etme yeteneğine ihtiyaç duymaktadır.¹⁴⁴ OECD ve Dünya Bankası gibi uluslararası örgütler kendi bölgelerindeki eğitim enstitülerine daha büyük roller vermektedirler. Bu roller hem ekonomik roller hem de çevresel, toplumsal ve kültürel katılımlardır.¹⁴⁵ Bilgi ekonomisinin ilerleyişi öğrenme ekonomisini de içine almaktadır. Ekonomistler eğitim ve tecrübe yılları gibi yetiler açısından beşeri sermaye gelişimini ölçmektedirler.¹⁴⁶

Ülkelerin kendi bünyelerindeki üniversiteleri yetersiz durumdaysa bireyler kendi ülkelerindeki bu üniversitelerde eğitim almak yerine küreselleşmenin de etkisiyle dünyanın herhangi bir yerindeki daha iyi bir okulda internet aracılığıyla eğitim alabilmektedirler. Dolayısıyla küresel olarak ortaya çıkarılan bilgi artık bilgisayarlar üzerinden internet aracılığıyla sadece bir tık ötede bulunmaktadır. Globalleşen dünyada gelişmekte olan ülkeler kendi içerisindeki yetenekli bireylerini

¹⁴²Walter W. Powell, Kaisa Snellman, "The Knowledge Economy", **Annual Review of Sociology** Vol:30, 2004, s. 214.

¹⁴³Kelleci, **a.g.e.**, s. 40.

¹⁴⁴OECD (1996), **a.g.e.**,p. 14.

¹⁴⁵Fumi Kitagawa, "Universities and Regional Advantage: Higher Education and Innovation Policies in English Regions", **European Planning Studies**12.6, 2004, p. 105.

¹⁴⁶OECD (1996), **a.g.e.**,p. 41.

kaybetmemek için ve rekabet edilebilirliklerini devam ettirebilmek için bilgi üretim sürecine dâhil olmak zorundadırlar.¹⁴⁷

Temel eğitim insanların bilgiyi kullanma ve öğrenme kapasitesini arttırmak için zorunludur. Diğer yandan teknik ikinci seviye eğitim teknolojik inovasyon için zorunludur. Yeni bilginin üretimi ve onu belli bir ekonomik konuma adapte etmek genellikle daha yüksek eğitim ve araştırma gerekmektedir.¹⁴⁸ Bilgi ekonomisi gerçekte öğrenme oranı ve değişim oranının hızlanmasıyla ilerleyen ağların hiyerarşisidir.¹⁴⁹ Akademik camia araştırma ve gelişmede, yenilikçi ve başarılı bilgi ve iletişim teknolojileri uygulamalarının geliştirilmesinde, uzmanlık ve yeteneklerin inşa edilmesinde ana rol oynamaktadır.¹⁵⁰

Bilgi toplumları yeni karakterlere sahiptir ve bilgi toplumlarında eğitim ise daha karmaşık ve bundan daha fazla oranda temel bir zorunluluk haline gelmektedir. Bilgi toplumu, eğitimden beklentilerini ve eğitim ihtiyacını daha önemli kılmaktadır. Bilgi toplumunda eğitim ve yetenek oluşturma, yalnızca içerik ve ders planları ile ilgili bir konu değildir. Bu konu temel olarak yöntemlerle ilgilidir ve eğitimin gerçekleştirilme şeklinde önemli değişimler gerektirmektedir. Bu bağlamda eğitimin enformasyon işleme, bilgi işleme, elde etme ve iletme, kolektif zeka geliştirme ve ortak çalışma, yeni yetenekler inşa etme, bilgi toplumunun vatandaşlarını hazırlama gibi yeni konulara yönelmesi gerektiği görülmektedir.¹⁵¹ Ayrıca devletlerin eğitimle ilgili düzenlemelerde, herkes için eğitime eşit erişimi sağlamasının yanı sıra herkese yüksek nitelikte eğitim sunulmasını esas alması gerektiği görülmektedir.¹⁵²

¹⁴⁷Carl Dahlman, Lan Xue, Kemal Gürüz, "Lifelong Learning and The Knowledge Economy", **Summary of The Global Conference on Lifelong Learning**, Stuttgart, Germany, 2002, p. 17-18.

¹⁴⁸Derek Chen, Carl J. Dahlman, "The Knowledge Economy, The KAM Methodology and World Bank Operations", **World Bank Institute Working Paper**, 2005, p. 5.

¹⁴⁹Houghton ve Sheehan, **a.g.e.**, p. 11.

¹⁵⁰Braundvd., **a.g.e.**, p. 15.

¹⁵¹Bernard Cornu, **Collective Intelligence and Capacity Building in a Network Society, Education and The Knowledge Society: Information Technology Supporting Human Development**, Ed. by. Tom J. van Weert, New York, Springer, 2004, p. 33.

¹⁵²R. P. Singh, **Teaching Teachers for A in Knowledge Society**, Gyan Publishing House, New Delhi, 2008, p. 15.

1.4.Bilgi Ekonomisinin Gelişiminde Üniversitelerin Yeri

Eğitimin önemli bir bölümünü temsil eden üniversiteler ve yükseköğretim kurumları bilgi ekonomisinde de aynı derecede önemli olarak görülmektedir. Üniversiteler iç içe girmiş uluslararası, ulusal ve yerel rollerde birkaç farklı ölçekte faaliyet göstermektedirler. Birçok yerde üniversiteler bilgi tabanının merkezi olarak yani bilgi ekonomisinin kalbi olarak görülmektedir. İnovasyon sistemlerinin ana elemanları olarak önemli rol oynamaktadırlar. Böylece bilim ve inovasyon tabanlı bölgesel gelişimi desteklemektedirler. Üniversiteler uluslararası, ulusal ve yerel seviyedeki aktörlerle etkileşim kurarak inovasyon sistemlerinin çoklu basamak süreçlerinin bölümünü oluşturmaktadırlar.¹⁵³ Firmalar, bilim ve araştırma merkezleri, üniversiteler, düşünce kuruluşlarının yansırı küresel bilgiye erişimi sağlayan ve etkileşim içerisinde bulunan diğer kuruluşların oluşturduğu etkin bir inovasyon sistemi yerel ihtiyaçlara bilgiyi adapte edebilmektedir. İnovasyonla yeni bilgi ve teknolojiler meydana getirmektedirler.¹⁵⁴

Yükseköğretim ABD’de sanayi sonrası topluma geçiş sürecinde daha önce görülmemiş bir hızla yayılmıştır. Rakamlara göre, üniversiteye kayıtlı öğrenci sayısı 1946-1964 yılları arasında %145, dönemin son yılında %104 oranında artmıştır. Amerikan üniversitelerindeki öğrenci sayısı 1879 yılından bu yana her yirmi yılda bir iki katına çıkmıştır. Aynı dönemde, üniversitelerden mezun olanlarla daha ileri derecelerdeki öğrencilerin sayısı, üniversitelerde okumakta olan öğrencilerin sayısından daha hızlı artmıştır.¹⁵⁵

Üniversitelerde üretilen bilgi ticari kanallar aracılığıyla takas edilmektedir. Bu ticari kanallara lisanslı patentler, danışman firmalar da dâhil olmaktadır. Bilginin takas edilmesinin avantajı, hem sektörel ortaklarla bilim ve teknoloji alanlarını hem de sağlık, eğitim, sanat gibi alanlardaki farklı tip ortak ve örgütlerle birlikte insani ve sosyal bilimlerin birçok alanını sisteme dâhil etmesidir.¹⁵⁶ Üniversiteler tarihsel olarak farklı fonksiyonlara, kaynaklara, ağlara sahip olmakla birlikte, ulusal yüksek eğitim sistemlerinin çeşitlenmesinde önemli amaçlara sahiptir. Üniversiteler bölgesel

¹⁵³ Kitagawa, **a.g.e.**, p. 103.

¹⁵⁴ Dahlman ve Andersson, **a.g.e.**, p. 14.

¹⁵⁵ Dura ve Atik, **a.g.e.**, s. 83.

¹⁵⁶ Kitagawa, **a.g.e.**, p. 104.

ağlarda ve kurumsal kapasite yapılanmasında önemli bir rol oynamaktadır.¹⁵⁷ Üniversiteler ve araştırma merkezleri bilgi üretim sürecinde kilit rol oynamaktadırlar. Bu birimler bilginin üretilmesi, uyarlanması ve doğru yerde ve zamanda kullanılması işlevleri ile bilgi ekonomisinin gelişiminde önemli bir rol oynamaktadırlar.

1.5.Bilgi Ekonomisi ve Bilgi İletişim Teknolojileri

Bilgisayar ve iletişim teknolojilerinin bir araya getirilerek kullanılmasıyla ortaya bilişim teknolojileri çıkmıştır. Bilişim sektörüne dâhil olan teknolojiler sırasıyla, mikro elektronik, bilgisayarlar, mobil telefonlar, yazılımlar, televizyonlar, çevrimiçi veri tabanları vb. olarak gösterilmektedir. Karar aşamalarında kullanılmak amacıyla örgüt içerisindeki yazışmalar, departmanlardaki veriler, bilgi teknolojileri ile bir araya getirilir ve büyük veri tabanları oluşturulur. Bu veri tabanlarında ki veriler gruplandırılarak kaydedilir ve stratejik süreçlerde kullanılır. 1990'lı yılların başında yeni ekonomi kavramının iktisat yazınına (literatürüne) kazandırılmasına yol açan bilgi ve iletişim teknolojilerindeki süreç, farklı bilgi değişim modellerinin ve depolamayla ilgili yöntemlerin oluşturulmasını sağlamış ve bilgiyi kodlamayı iletmiştir.¹⁵⁸ Bilgi ve iletişim teknoloji sistemleri, bilgi tabanlı ekonomiyle üretim sistemlerinin eşleştirilmesi gibi bilginin üretim ve dağıtımı için gerekli koşulları radikal olarak değiştiren yeni ve teknolojik bir taban meydana getirmektedir.¹⁵⁹

Günümüzde bilgi ve iletişim teknolojileri istihdamın gelişmesinde, ekonomik büyümenin artmasında, inovasyonun gelişmesinde önemli rol oynamaktadır. Son yıllarda özellikle 2006 yılından itibaren özel sektördeki BİT firmalarının istihdama katkısı %5.5, katma değere katkısı ise %8 seviyesinde gerçekleşmiştir. Özel sektördeki BİT organizasyonlarının ürünleri genellikle Ar-Ge yoğunluğuna sahiptir. Ekonomik olarak altyapının sağlam olması da BİT sektörünün yayılımına ve internete bağlı olarak görülmektedir. BİT sektörü ekonomik ve sosyolojik olarak tam anlamıyla yayıldığı zaman maksimum fayda sağlamaktadır. BİT sektöründe gerçekleştirilen büyük miktardaki yatırımlar, organizasyonların rekabet

¹⁵⁷ A.e., p. 106.

¹⁵⁸ Philip Cooke, "Regional Innovation Systems, Clusters, and The Knowledge Economy", Industrial and Corporate Change 10.4, 2001, p. 949.

¹⁵⁹ Lundvall ve Foray, a.g.e., p. 14.

edebilirliğini, üretim süreçlerinin verimli hale getirilmesini, örgütsel yenilikleri, işgücü kabiliyetlerinin artırılmasını, yeni işletmelerin meydana gelmesini sağlamaktadır. BİT sektörüne yapılan yatırımlar aynı zamanda işgücündeki verimliliğin artmasını da sağlamaktadır. BİT firmalarının daha da artarak ekonominin içinde yayılmasıyla, özel sektörün, kamu kurumlarının ve sosyal örgütlerin aktivitelerinin daha verimli hale gelmesine, yeniliklerin artmasına imkân sağlamaktadır.¹⁶⁰

Bilgi ve iletişim teknolojileri yaparak öğrenmeyle gelişebilmektedir. Bilgi ve iletişim teknolojileri öğretim görenler için kullanılabilir bilgi kaynaklarını arttırmaktadır. Öğrenciler için hızlı geribildirim sunabilmektedir ve işbirlikçi öğrenmeyi sağlayabilmektedir.¹⁶¹

Bir ülke bilgi ve iletişim teknolojilerini kullanarak ekonomik anlamda ilerleme sağlamak, verimlilik elde etmek ve tutarlı bir şekilde devam ettirmek istiyorsa, bu teknolojileri bireylerin kullanımı için daha da yaygınlaştırması ve bu altyapıyı kurması gerekmektedir. Ayrıca bu ilerlemeyi sağlamak için uygun çalışma ortamları ve hukuksal altyapı sağlanmalıdır. Bu altyapının; adil ve eşitlikçi olması, insani ve ekonomik kalkınmaya uygun bir ortam yaratması, ülkenin küresel rekabet avantajını desteklemesi, ancak hukuk devleti ilke ve normlarına uygun olarak demokratik bir ortamda geliştirilmesine bağlıdır.¹⁶²

Bilginin, zenginlik oluşturma ana gücü ve sürdürülebilir rekabet avantajının kaynağı olması süreci yeni ekonominin birçok ilişkili hareketiyle ilerlemektedir. Ekonomik aktivitelerin bilgi yoğunluğundaki artışı, ekonomik işlerin küreselleşmesi ve bilgi ve iletişim teknolojileri devrimi bilgi ekonomisinin temel dışlıları olarak tanımlanmaktadır.¹⁶³

¹⁶⁰ OECD (2009), **a.g.e.**, p. 162.

¹⁶¹ Charles D. Wislaw, William L. Bramer, **Future Work Putting Knowledge to Work in The Knowledge Economy**, The Free Press, Newyork, 1994, p. 333.

¹⁶² Özgür Uçkan, **E-Devlet, E-Demokrasi ve Türkiye**, Literatür Yayıncılık, İstanbul, 2003, s. 32.

¹⁶³ Marti ve Cabrita, **a.g.e.**, s. 23.



Şekil 1.3.Bilgi Ekonomisinin Temelleri.

Firmaların fiziksel mallarını nasıl sattığı ve firmanın sanal internet çevresine nasıl geçiş yaptığı, dolayısıyla satışlar ve finansal durumunun nasıl geliştiği, sanal bilgisayarda dizayn edilen ortamlardaki, firmaların telefon sistemleri yerine internet faaliyetlerini kullanarak maliyetleri nasıl düşürdükleri ile ilgili birçok örnek bulunmaktadır. Bilgi ve iletişim teknolojileri üretim maliyetlerini azaltan ve verimliliği arttıran yenilikçi ürünler üretmektedir.¹⁶⁴

2005-2010 yılları arasında bilgi ekonomisi indeksi esas alınarak ekonomik teşvik düzeni, eğitim yapısı, inovasyon, BİT değişkenleri arasındaki regresyon ve nedensellik ilişkisi panel veri yöntemi ile analiz edilmiştir. Bu çalışmada 19 OECD ülkesi incelenmiştir. Uzun dönemde bilgi ekonomisinden ekonomik teşvik sistemi, eğitim oranı, inovasyon ve bilgi iletişim teknolojileri değişkenlerine doğru nedensellik ilişkisi olduğu belirtilmektedir.¹⁶⁵

Bilgi ve iletişim teknolojileri ekonomik gelişimin ilerlemesinde ve yoksulluğun azalmasında hayati bir rol oynamaktadır. 56 gelişmekte olan ülkede yürütülen bir çalışmada bilgi ve iletişim teknolojilerini kullanan firmaların

¹⁶⁴ A.e., s. 32.

¹⁶⁵ Selahattin Kaynak, İbrahim Arslan, "In The Context of The Knowledge Economy of Parameters: Econometrics Analysis of the Organisation for Economic Co-operation and Development (OECD) Economies", **African Journal of Business Management** 6.12, 2012, p. 4459.

kullanmayanlardan daha hızlı geliştiği, daha çok yatırım yaptığı, daha etkin ve karlı olduğu bulunmuştur.¹⁶⁶

1.5.1. Telekomünikasyon

Bilgi ve telekomünikasyon teknolojileri ekonomik zenginliğin ana kaynağı olarak görülmektedir. Bilgi ve iletişim teknolojisi serbest girişimlerin yeni oluşumlarına imkan sağlamaktadır. Serbest girişim ve bilgi iletişim teknolojilerinin kesişimi küresel ölçekte artan zenginlikle sonuçlanan bilgi ekonomisinin gelişimine yol açmaktadır.¹⁶⁷

Bilgi ekonomisi iki sektöre ayrılmaktadır. Bunlar birincil ve ikincil bilgi sektörü olarak adlandırılmaktadır. Birincil bilgi sektörü yerleşik bir pazar için bilgiyi üretme, işleme ve dağıtmada direkt olarak kullanılan ya da bilgiyi mal ve hizmetleri üreten bütün sektörler olarak tanımlanmaktadır yani birincil sektör telekomünikasyon gibi hizmetler ve bilgisayar gibi bilgi mallarını kapsamaktadır. İkincil sektör ise bilgi özelliği olmayan firmaları ve devlet tarafından içsel tüketim için üretilen bütün bilgi hizmetlerini kapsamaktadır.¹⁶⁸

Geçmişten bugüne otomasyon, ekonominin üç sektörünü tamamen modernize etmiştir. Bunlar tarım, sanayi ve hizmetlerdir. Bilgi ekonomisinde teknoloji yeni sermayeyi oluşturmak için gelişmek zorundadır ve bilgi çalışanları enformasyon sisteminin işlemesi için öncü olmak zorundadırlar. Yeni ekonomi avcı-toplayıcı ekonomisine geri dönmektedir. Buradaki av bilgi olarak görülürken, ortam ise küresel telekomünikasyon ağları ve enformasyon ortamıdır. Daha küçük örgütler inovasyon, meydana getirme yeteneği, etkinlik arasında yeni bilgi formlarını bulmak amacıyla girişimde bulunmaktadırlar. Bu bilgiyi satmak için ve ürün/hizmet üretmek için yeni yollar bulmaktadırlar.¹⁶⁹

İnternet teknolojilerinin, telekomünikasyonun, bilgi ve iletişim teknolojilerinin giderek daha çok günlük hayata dâhil olmasıyla, verimlilik, mal ve

¹⁶⁶Braund vd.,**a.g.e.**, p. 10.

¹⁶⁷Laura Lengel, "The Information Economy and The Internet", **Journal Mass Commun** 2.,2009, p. 2.

¹⁶⁸Marc Uri Porat, **The Information Economy: Definition and Measurement**, 1977, p. 23.

¹⁶⁹Wislaw ve Bramer,**a.g.e.**, p. 331.

hizmet kalitesi, net kar ifadeleri de değişmiştir. Klasik ekonomik sistemin temel dayanma noktaları da bu teknolojilerle birlikte sorgulanır duruma gelmiştir. Tüm bu ilerlemelere ilaveten küreselleşmenin de gündeme gelmesiyle ekonominin seyrinin değişimi güçlenmiştir. Özellikle bilgi ve iletişim teknolojileri sektörü büyük bir ilerleme kaydetmiştir. Bu pazarda lider olanlar da yeni ekonomi değişkenlerini belirlemişlerdir.¹⁷⁰ Telekomünikasyonun altyapı olarak güçlendirilmesi, hızlı ve akıllı bağlantıların kurulması, ekonomik kurumların bilgiye kolay bir şekilde erişebilmesi internet ve e-ticaretten sağlanan en önemli faydalardır.¹⁷¹

Gelişmekte olan ülkeler ekonomilerini iyileştirmek için bilgi ekonomisi üzerine yönelmektedirler. Bu ülkelere 1990-2003 yılları arasında telekomünikasyon başta olmak üzere bilgi sektörlerine 200 milyar dolar yatırım yapılmıştır. Geçtiğimiz 25 yıl boyunca gelişmekte olan ülkeler özellikle telefon hizmetleri için bilgi iletişim teknolojileri erişimini önemli oranda arttırmıştır. 1980-2005 yılları arasında abonelerin sayısı 30 kat artmıştır. 1980’de gelişmekte olan ülkeler dünyadaki telefon hattının sadece %20’sini temsil etmekteydi. 2005’te telefon hatlarının %60’ı küresel olarak gelişmekte olan ülkelerde bulunmaktaydı. Bu gelişme artan özel sektör rekabeti gibi mobil telefon teknolojik devrimiyle gerçekleşmiştir. Özel sektör rekabetine açılmak, yabancı yatırımların büyük oranda içeriye akışını sağlamıştır. 1990-2003 yılları arasında 154 gelişmekte olan ülkenin 122’si sadece telekomünikasyon için yabancı yatırımlardan 200 milyon dolar kazanmıştır.¹⁷²

İletişim açısından globalleşme; kişilerin bilgisayar ağları, telefon ve kitle iletişim araçları gibi teknolojik araçlar yardımı ile bulundukları yer ve devlet sınırlarından bağımsız bir şekilde birbirleri ile kurdukları bağlantıların yoğunlaşmasını ifade ederken, organizasyonlar bakımından globalleşme; sınır ötesi ağlar şeklinde faaliyet gösteren şirket, dernek, vakıf ve kurumların sayıca artması ve çeşitlilik göstermeleri anlamına gelmektedir.¹⁷³ Bu durumda telekomünikasyon

¹⁷⁰Yeloğlu, a.g.e., s. 246.

¹⁷¹Özgür Aslan, Selcen Öner, "İnternet Ekonomisi", **İstanbul Üniversitesi İletişim Fakültesi Hakemli Dergisi 26.**, 2006, s. 17.

¹⁷²Braund vd.,a.g.e., p. 10.

¹⁷³ Coşkun Can Aktan, İstiklal Y. Vural, **Yeni Ekonomi ve Yeni Rekabet**, Türkiye İşveren Sendikaları Konfederasyonu, 2004, s. 146.

faaliyetlerinin her geçen gün önemi artmaktadır. Telekomünikasyon hizmetlerinin maliyeti ve yaygınlığı internet kullanımının kapsamını belirlemektedir. Gelişmiş ülkelerde bu hizmetler daha düşük maliyetlerle sunulmaktadır.¹⁷⁴

1.5.2. İnternet

Elektronikte meydana gelen yenilik ve değişim süreci internetin de gelişimini sağlamıştır. İnternet son yüzyıldaki en önemli icatlardan biri olarak nitelendirilmektedir. Bu bağlamda internet bilgiyle yoğunlaşan toplumun en önemli teknik ve sosyolojik kazançlarından biri olarak görülmektedir. Dünyanın tamamını çevreleyen bu ağ, fiziksel kısıtlamaları ortadan kaldırarak teknolojik ilerlemeler için uygun meydanı sunmuştur.¹⁷⁵

BÖLGELER	NÜFUS (2017)	İNTERNET KULLANICILARI (2017)	POPÜLASYON YÜZDESİ	GELİŞİM 2000-2017	TABLONUN KULLANICI YÜZDESİ
AFRİKA	1.287.914.329	453.329.534	35.2 %	9,94%	10.9 %
ASYA	4.207.588.157	2.023.630.194	48.1 %	1,67%	48.7 %
AVRUPA	827.650.849	704.833.752	85.2 %	570%	17.0 %
ORTA DOĞU	652.047.996	437.001.277	67.0 %	2,32%	10.5 %
KUZEY AMERİKA	254.438.981	164.037.259	64.5 %	4.89%	3.9 %
LATİN AMERİKA	363.844.662	345.660.847	95.0 %	219%	8.3 %
AVUSTRALYA	41.273.454	28.439.277	68.9 %	273%	0.7 %
DÜNYA TOPLAMI	7.634.758.428	4.156.932.140	54.4 %	1,05%	100.0 %

Tablo 1. 1.İnternet Kullanım İstatistikleri.¹⁷⁶

¹⁷⁴International Labour Office, **World Employment Report 2001: Life at Work in the Information Economy**, Geneva, 2001., p.1.

¹⁷⁵Braund vd.,**a.g.e.**, p. 10.

¹⁷⁶ <https://www.internetworldstats.com/stats.htm>

Gelişen dünya, bilgi ve iletişim teknolojileri yapısının yayılmasında büyük bir ilerleme gösteriyorken görüntü ileri bilgi teknolojileri kullanımı için daha da karışık hale gelmiştir. İnternete ait güvenli sunucu sayısındaki farklılıklara ve e-ticaretin kullanılması için gerekli yetkilerin olmasına rağmen 2000-2005 yılları arasında kullanımı üç kat artmıştır.¹⁷⁷ Bilgisayarların bilgilerin depolanarak geliştirildiği yıllardan sonra bu depolanan ve yeni üretilen bilgi ticari hale getirilmeye başlanmıştır. Sonraki aşamalarda ise ortaya çıkan bu bilgisayarların birbirleriyle bağlantı sağlayarak bütün dünyada kabul edilmesi ve yayılması interneti meydana getirmiştir. Geliştirilen bu yeni kavram yani internet bir süre sonrada ekonomik sistemin temelini oluşturmaya başlamıştır.¹⁷⁸

İnternet ve bilgisayarlar küresel bilgi ağlarını oluşturmaktadır. Küresel bilgi ağları bilgi ekonomisine birçok imkân sağlamaktadır ve birçok açıdan üretim sürecine yol göstermektedirler. Böylece yatırım maliyetlerini azaltmaktadırlar. Yeni teknolojiler, onlardan faydalanılması için niş pazarların oluşmasına imkân sağlamaktadır.¹⁷⁹ İnternetteki mevcut bilgi miktarı insanların zamanında tüm kaynak çeşitlerine erişmesine imkân sağlamaktadır.¹⁸⁰

İnternet geleneksel sistemlerin yavaş yavaş yerini almakta ve hayatımızı her geçen gün daha kolay hale getirmektedir. Bilgi zengin internet ekonomisinde, hız ve ulaşılabilirlik kritik faktörlerdir. Ekonomik başarı sadece en hızlı ve doğru bilgiye ulaşanların yeteneğe kalmış gibi görünmektedir. Bu anlamda zaman ve bilgi, gittikçe artarak zaman ve piyasaya doğru yön değiştirmektedir. İnternet başka bir deyişle içinde her türlü bilginin bulunduğu gelişmiş bir pazar yeri olarak da ifade edilebilir.¹⁸¹

¹⁷⁷ A.e., p. 10.

¹⁷⁸ Şule Özmen, **Ağ Ekonomisinde Yeni Ticaret Yolu: E-Ticaret**, İstanbul Bilgi Üniversitesi Yayınları, 2003, s. 6.

¹⁷⁹ Brinkley, a.g.e., p. 15.

¹⁸⁰ Jolfsson ve Hitt, a.g.e., p. 17.

¹⁸¹ Nihal Kargı, **Bilgi Ekonomisi**, Ekin Kitabevi, Bursa, 2006, s. 194.

İnternet, iletişimin her yönünü değiştirmektedir. Ekonomilerde üretimin rolü azalmaktadır. Bunun yanı sıra da hizmetlerin rolü artmaktadır. Küresel süreci devam ettiren telekomünikasyondaki olağanüstü ilerlemeler dünya çapında rekabete imkan sağlamaktadır. Yeni ulaşım teknolojileriyle birlikte harcanan zaman ve uzaklık da azalmaktadır.¹⁸²

Bilgi toplumuna dönüşen ülkelerde, başta bilgisayarlar olmak üzere bilgi işleyen tüm araçlar günlük hayatın önemli bir parçası haline gelmekte ve bu tür araçlar için yapılan harcamalar büyük bir artış göstermektedir.¹⁸³ İnternet bilgi ekonomisi kavramının da temelini oluşturmaktadır. Dijital ekonomi; internet, kurum içi ağ (intranet), kurum dışı ağ (extranet) gibi dijital iletişim araçları kanallarının dahil olduğu dijital teknoloji temeline dayalı ekonomi anlamına gelmektedir. Dijital ekonomi bazen internet ekonomisi, web ekonomisi veya yeni ekonomi olarak da isimlendirilebilmektedir. Bu yeni ekonomi ortamında dijital ağlar ve iletişim altyapıları, insanların interaktif iletişim kurabileceği, işbirliği yapabileceği ve bilgiyi arayabileceği küresel bir platform sağlamaktadır.¹⁸⁴

1.5.3.E-Ticaret

Üniversite ya da firmalardan üçüncü taraflara, oluşturulan bilginin dağılımı, faaliyetlerin yenilenmesini zorunlu kılmaktadır. Bu gibi bilgi dağılımları mekânsal olarak kısıtlanma eğiliminde olabilmektedir. Böylece küreselleşme ironisi çoğu mal ve hizmet ile ilgili coğrafik pazarda giderek küresel olmaktadır. Öncü gelişmiş ülkelerde yenilikçi faaliyetin artan önemi, nispi avantajın ana kaynağı olarak yerel bölgelerdeki canlanmayı tetiklemektedir. Küreselleşme hızlı bir şekilde ulaşım maliyetlerini azaltmaktadır. Bu maliyetlere hem hammadde maliyetleri hem de coğrafik alan üzerindeki maliyetler de dahil olarak gösterilmektedir.¹⁸⁵

Sınırları kaldıran küreselleşmenin sonuçlarından biride e-ticarettir. Verilerin internet platformunda iletilmesiyle birlikte bilgisayarlarla ve ileri yazılımlarla hızlı

¹⁸²World Bank, **Building Knowledge Economies: Advanced Strategies for Development**, World Bank Institute Development Studies, Washington, D.C., 2007, p. 9.

¹⁸³Atık ve Dura, **a.g.e.**, s. 209.

¹⁸⁴Soon-Yong Choi, Dale O. Stahl, Andrew B. Whinston, "Intermediation, Contracts and Micropayments in Electronic Commerce", **Electronic Markets** 8.1., 1998, p. 201

¹⁸⁵Birkinshawve Sölvell, **a.g.e.**, p. 77.

bir şekilde işlenerek bilgiye dönüştürülmesidir. Veri bilgiye dönüştükten sonra ağlar aracılığı ile yayılmaktadır. Bu şekilde internet, bilginin ticarileşme aşamalarını kolaylaştırmaktadır.¹⁸⁶ Elektronik ticaret, internetin içinde olduğu bilgisayar ağları vasıtasıyla ürün, hizmet ve bilginin alınıp satılma, transfer olma, yer değiştirme süreçlerini kapsamaktadır.¹⁸⁷

Yeni ekonomi globalleşmeyle birlikte uluslararası sınırları ortadan kaldırmıştır ve sınırlı ülke kavramını aşındırmıştır. Gelişmiş olan ülkelerin global pazarı kontrol altına alma yönünde tavırlarını kuvvetlendirmiştir. Bu ilerlemeler, ticarete meydana gelen alış-verişlerin boyutunu değiştirmiştir ve e-ticaretin kullanımını daha da genişletmiştir. Mobil telefonların, televizyonların ve daha birçok bilgi teknolojisi aletlerinin kullanımının ilerlemesiyle e-ticarete daha hızlı bir geçiş sağlanmıştır. İnternet teknolojilerinin yaygınlaşarak ticari ilişkilerde kullanılması e-ticaretin alanını daha da genişletmiştir.¹⁸⁸

Bilgi ekonomisinin her bir bileşeninin birbirlerini yüksek seviyede etkilediği görülmektedir. İnternetin aktif kullanımının artmasıyla ve ekonomik aktivitelere de sıçramasıyla bilgi ekonomisi şekillenmiştir. Üretilen bilginin tüm kullanıcılara açık olması, şeffaf olması ve teknik yönü e-ticareti sadece basit bir satış süreci olmaktan çıkarmaktadır.¹⁸⁹

Bilgi ve iletişim teknolojilerindeki gelişmeler en çok 1990'lı yılların başında gerçekleşmiştir. İnternet teknolojilerinin ve kablosuz bağlantı araçlarının yayılmasıyla klasik ekonomik anlayıştan yeni ekonomiye yani dijital ekonomiye geçişinde başladığı görülmektedir. Firmalar ekonomik aktivitelerini bu ilerlemelerle ortaya çıkan teknolojilerle bütünleştirmeye başlamışlardır. Bütünleşme sonucu üretici ve tüketiciler arasında daha yakın bir ilişki kuran e-ticaret meydana çıkmıştır.. Elektronik ticaret, bilgi ve iletişim teknolojileri araçlarının meydana getirdiği

¹⁸⁶ Salih Barışık, Oya Yirmibeşçik, "Türkiye'de Yeni Ekonominin Oluşum Sürecini Hızlandırmaya Yönelik Uyum Çabaları", **Uluslararası Yönetim İktisat ve İşletme Dergisi** 2.4., 2012, s. 41.

¹⁸⁷ Efraim Turban, Dave King, Jae Kyu Lee, Dennis Viehland, "Electronic Commerce: A Managerial Perspective", 2002, p. 3.

¹⁸⁸ Buket Öztuna Cox, **Avrupa Birliği Hukukunda Elektronik Ticaret ve Türkiye'deki Gelişmeler**, Pusula Yayıncılık, İstanbul, 2002, s. 2.

¹⁸⁹ **A.e.**, s. 13.

değişimin en faydalı neticelerinden biri olarak görülmektedir. İnternet aracılığıyla gerçekleştirilen e-ticaret, tüketiciler ve üreticiler arasındaki bağlantıyı güçlendirmektedir.¹⁹⁰

Firma yönetimin de çıkış açan e-ticaret, internet aracılığıyla birçok fayda sağlamaktadır. Son yıllarda daha da yoğunlaşmasıyla birlikte e-ticaret ekonomik işlemleri ve toplumsal çevreyi değişime uğratmıştır. Aynı zamanda firma aktiviteleri içinde internetin yayılması ve e-ticaret uygulamalarının ilerlemesi, verimliliğin artmasına neden olmaktadır. Sağladığı en önemli kolaylıklardan biri olarak az maliyetle daha çok ve daha hızlı bilgi iletişimi gösterilmektedir. E-ticaretin diğer önemli bir faydası da ekonomik ortam içerisindeki tüm birimleri daha rekabetçi hale getirmesidir.¹⁹¹

Pazarlama, finans, iletişim gibi alanlar internetin meydana getirdiği yeniliklerden doğru orantılı olarak faydalanmıştır. Kamu işlemleri, eğitim, sağlık gibi alanlarda bu değişimden oldukça fazla etkilenmiştir.¹⁹²

1.5.4.E-Devlet

Ekonomik sistemin artan bilgi yoğunluğu ve bu bilgiyi dağıtma yeteneğinin arttırılması, ekonomik sistemdeki tüm katılımcılar için bilginin değerini de arttırmaktadır. Bunun etkileri hem firma stratejileri ve hükümet politikaları için hem de ekonomik davranışları düzenleyen sistemler ve kuruluşlar içi önemlidir.¹⁹³

Hükümet politikaları için bilgi teknolojilerinin kullanımı günümüzde e-devlet uygulamalarının geliştirmesine yol açmıştır. Hükümetler, ekonomik ve sosyal gelişimin merkezindeki bilgi ve iletişim teknolojilerinde yer alan e-stratejilerin, kapsamlı, şeffaf, düzenleyici yapıların uygulanmasında ve gelişmesinde önemli rol

¹⁹⁰Aslan ve Öner, **a.g.e.**, s. 6.

¹⁹¹**A.e.**, s. 17.

¹⁹²Andreas Schleicher, **Measuring Student Knowledge and Skills: A New Framework for Assessment**, Organisation for Economic Co-Operation and Development (OECD), Paris, 1999, p. 9.

¹⁹³Houghton ve Sheehan,**a.g.e.**, p. 1.

oynamaktadır. Etkin bir biçimde e-uygulamaları kullanmak için tarafların örgütsel ve insani kapasitelerini sağlamlaştırılması etkili olabilmektedir.¹⁹⁴

Devlet hizmet etmekle görevli olduğu vatandaş ve firmalarla olan ilişkisini elektronik ağlarla kurmaktadır. Bu yeni yapıya kısaca e-devlet denmektedir. Kamuda sağlanan hizmetlerin çok daha kaliteli ve düşük maliyetle olması için elektronik devlet uygulamaları kullanılmaktadır. Elektronik devlet sürecinde bilgi ve iletişim teknolojilerinin sağladığı kolaylıklar ve elektronik sektörünün faydaları çok önemli olarak görülmektedir. Son yıllarda hayatın her aşamasında varlığını hissettiren bilgi ve iletişim teknolojileri kamuda var olan gereksiz işlemlerden ve süreçlerden devleti kurtarmaktadır.¹⁹⁵

E-devleti oluşturan temel unsurlar e-vatandaş, e-şirket ve e-kurum'dur. Burada asıl odaklanılan mesele vatandaşa, yönetime ve iş yaşamına internetin sunduğu fırsatlarla hayatı kolaylaştırmak ve daha üretken hale getirmektir. Süreçleri hızlandırmak suretiyle ve vatandaşın mümkün olduğu kadar daha az yorularak daha çok iş yapmasını sağlamaktır. Bu arada elbette yönetime de aynı kolaylık fırsatını sunmaktır. Burada devlet, vatandaş ve şirket üçgeninde toplam yaşamı kolaylaştırıp, daha faydalı katkılar üretmek amaçlanmaktadır.¹⁹⁶

E-devlet çalışmaları, kamu hizmetlerinin sunumunun, bilgi işlem ve iletişim teknolojileri aracılığıyla daha düşük maliyetle, daha hızlı ve daha iyi bir şekilde vermek üzere, devletin bilgi temelli yeniden örgütlenmesi çalışmalarıdır. Bu çerçevede; vergi tahakkuku ve ödemeleri, kamu alımları, dış aktörlerle ilişkiler ve gümrükleri sosyal yardım ve transferler, yasaların uygulanması, kamunun sunduğu bilgilere erişim, kurumlar arası ve kurum içi iletişim, sağlık, eğitim, ortak kütüphaneler servisi gibi kamu ile olan ilişkilerin kolaylaştırılmasına yönelik düzenlemelerdir.¹⁹⁷

¹⁹⁴Braund vd., **a.g.e.**, p. 15.

¹⁹⁵Aktan ve Vural, **a.g.e.**, s. 17.

¹⁹⁶Kargı, **a.g.e.**, s. 201.

¹⁹⁷Özsağır, **a.g.e.**, s. 196.

E-devlet uygulaması ekonomik yönden düşünüldüğünde bireylerin ve firmaların bürokratik işlemlerini internet üzerinden online doldurdıkları formlar ile yapmaları devlete milyonlarca dolar tasarruf sağlayacaktır. Ayrıca bu yöndeki yapılanmalar ile devlette verimliliğin artması sağlanmaktadır. Bürokrasiyi ve giderleri azaltmak, hizmet kalitesini arttıracak, hizmetlere her an ulaşımı kolaylaştıracak ve yönetime katılımı arttıracaktır.¹⁹⁸

1.6.Bilgi Ekonomisi Değişkenleri ile İlgili Literatür Araştırması

Bu başlık altında bilgi ekonomisi değişkenleri ile ilgili yapılmış istatistiksel çalışmalara değinilmekte ve bu çalışmalarda yer alan değişkenlerin hangi yöntemlerle analiz edildikleri yer almaktadır.

Bilginin toplanması, işlenmesi ve dağıtımını kolaylaştıran unsurlar bilgi toplumunun altyapı parametreleri olarak adlandırılmaktadır. Bir ülkenin sahip olduğu bilgi altyapısı ilk olarak Japonya’da yapılan araştırmalarda ölçülmeye çalışılmıştır. Japon bilgi işleme ve geliştirme merkezi 1986 yılında bilgi toplumunun farklı parametrelerini içeren bir indeks geliştirmişlerdir. Bu indeks üç göstergeden oluşmaktadır. Bunlar:

- Donanım oranı: Bir sektördeki bilgisayar donanımının değeri ile ilgili endüstride çalışan işgücü sayısına bölünerek hesaplanmasıdır.
- Yazılım oranı: Bir endüstride son beş yıllık zaman süresi içinde kullanılan yazılım programlarının değeri ilgili endüstrideki işgücü sayısına bölünerek elde edilmesidir.
- İletişim oranı: Bir endüstrideki bilgi taşıma kapasitesinin çalışan sayısına bölünerek hesaplanmasıdır.¹⁹⁹

Bilgi ekonomisi, ortamın ekonomik gelişimi için etkin kullanılan bilgiye ihtiyaç olup olmadığını hesaba katmaktadır. Bilgi ekonomisi indeksi bir ülke ya da bir bölgenin bilgi ekonomisine doğru gelişim seviyesini yansıtan toplam indekstir.²⁰⁰

¹⁹⁸James N. Danziger, Kim Viborg Andersen, "The Impacts of Information Technology on Public Administration: An Analysis of Empirical Research from The "Golden Age" of Transformation", **International Journal of Public Administration** 25.5, 2002, p. 600.

¹⁹⁹Dordick ve Wang, **a.g.e.**, p. 34.

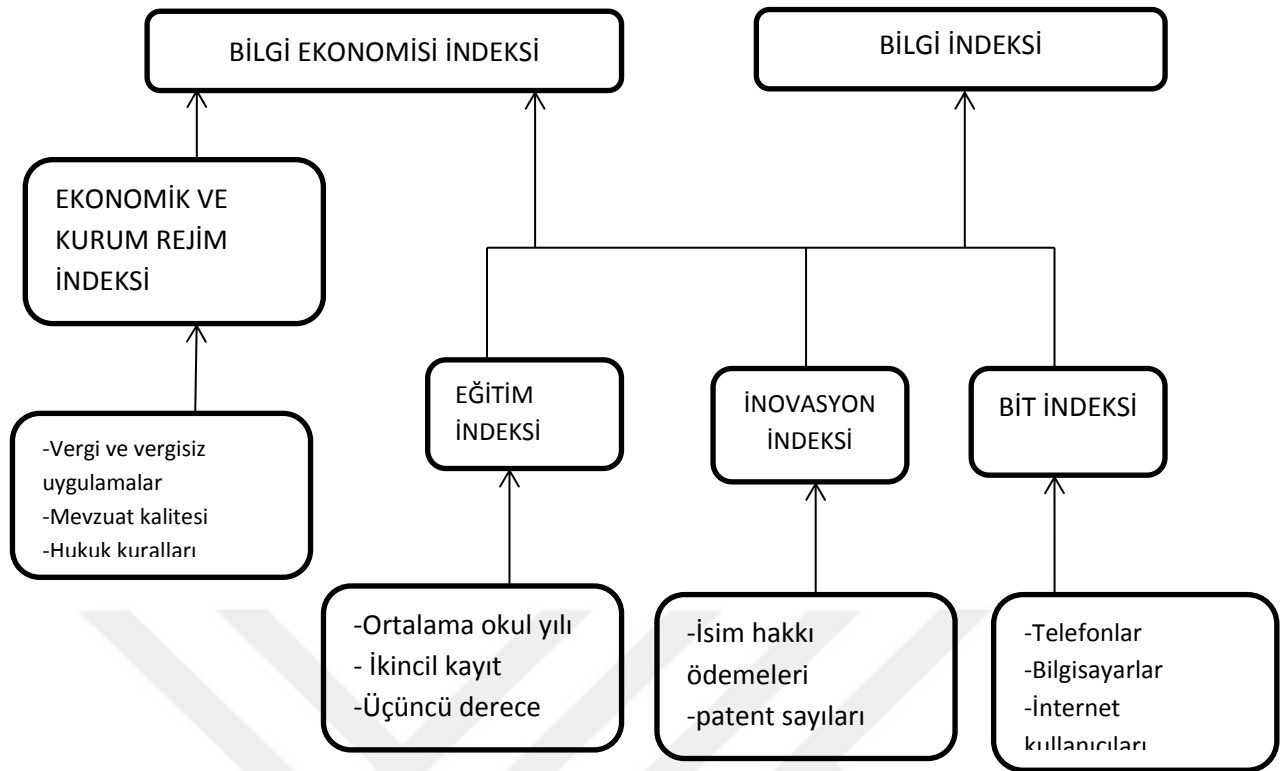
²⁰⁰OECD, (1996), **a.g.e.**, s. 12.

Bilgi ekonomisi ile ilgili yapılmış birçok çalışma bulunmaktadır. Bu çalışmalarda bilgi ekonomisinin özellikleri göz önüne alınarak birçok farklı değişken farklı istatistiki çalışmalarla analiz edilmiştir. Kullanılan değişkenler bilgi, iletişim ve teknoloji harcamaları, kamu Ar-Ge harcamaları, işletmelerin Ar-Ge harcamaları, fen ve mühendislik bölümü mezunları sayısı, kamu eğitim harcamalarının GSYİH içindeki yeri, insani gelişme endeksi, yüksek teknoloji ürünlerinin ihracat içerisindeki payı, Amerikan Patent Enstitüsü'ne kayıtlı patent sayısı, bilgi yoğun hizmetlerde çalışanların toplam istihdama oranı, cep telefonu aboneliği sayısı, internet erişimli işletmelerin yüzdesi, internet kullananların sayısı değişkenleridir.²⁰¹

OECD tarafından belirlenen bilgi ekonomisi indeksine göre, bir ülkenin bilgiyi üretmesi, benimseme ve yayma yeteneğini ölçmektedir ayrıca belli bir ülkede tüm bilgi gelişim potansiyelinin de belirlenmektedir. Bu indekste kullanılan başlıklar bilgi indeksinden farklı olarak ekonomik ve kurum rejimi, vergili ve vergisiz uygulamalar, mevzuat kalitesi ve hukuk kurallarından oluşmaktadır. Bilgi indeksi başlığı altında ise eğitim indeksi, inovasyon indeksi ve bilgi ve iletişim teknolojileri indeksi yer almaktadır. Eğitim indeksi ortalama okul yaşı, ikincil kayıt yani ikinci üniversite, üçüncül kayıt yani üçüncü üniversite başlıklarından oluşmaktadır. İnovasyon indeksi; isim ücreti ve geliri, patent hesabı ve makale başlıklarından oluşmaktadır. Bilgi ve iletişim teknolojileri indeksi ise telefonlar, bilgisayarlar ve internet kullanıcılarından oluşmaktadır. Bunlar aynı zamanda bilgi ekonomisi indeksine de girmektedir.²⁰²

²⁰¹Erkan Oktay, Sebahattin Kaynak, "Türkiye ve Avrupa Birliği Ülkelerinin Bilgi Ekonomisi Girdi ve Çıktı Değişkenleri Arasındaki Kanonik İlişkinin Araştırılması", **Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi** 10.2., 2007, s. 427.

²⁰²Csaba Makó, Brian Mitchell, "Knowledge Economy and Innovation: A European Comparative Perspective", **Psychosociological Issues in Human Resource Management** 1.2., 2013, p. 1.



Şekil. 1.4.Bilgi Ekonomisi İndeksi ve Bilgi İndeksi.

Avrupa Birliği ülkeleri ve Türkiye'nin bilgi ekonomisi değişkenleri ile ilgili yapılmış bir çalışmada girdi ve çıktı değişkenleri belirlenerek, bu değişkenler arasındaki kanonik ilişki analiz edilmiş olup, girdi ve çıktı değişkenleri arasında 0,98 oranında kanonik ilişki elde edilmiştir. Çalışmada kullanılan değişkenler bilgi iletişim ve teknoloji harcamaları, kamu Ar-Ge harcamaları, işletmelerin Ar-Ge harcamaları, fen ve mühendislik bölümü mezunları sayısı, kamu eğitim harcamalarının GSYİH içindeki yeri, insani gelişme endeksi, yüksek teknoloji ürünlerinin ihracat içerisindeki payı, Amerikan Patent Enstitüsü'ne kayıtlı patent sayısı, bilgi yoğun hizmetlerde çalışanların toplam istihdama oranı, cep telefonu aboneliği sayısı, internet erişimli işletmelerin yüzdesi, internet kullananların sayısı değişkenleridir.

Bilgi ekonomisini ölçmek için bilgi girdilerini ve bilgi çıktılarını ölçmek gerekmektedir. Bilgi girdilerini ölçmek bilimsel ve teknik bilgiye yapılan yatırımı ölçmektedir. Bilgi akışının gelişimi, sektörlerin ve ekonomilerin bilgi ve Ar-Ge yoğunluğu ölçülerinin daha iyi kazanç sağladığını göstermektedir. Bilgi çıktılarını ölçmek için ve bilgi ekonomisinin performansını değerlendirmek için ise öncelik Ar-Ge'ye, özel ve sosyal getiri oranlarının belirleyicilerine, diğer bilgi girdilerinin

gelişimine dikkat edilmelidir. Dolayısıyla istihdam, çıktı, verimlilik ve rekabet açısından topluma, firmalara ve bireylere bilginin getirisinin ölçülmesini kapsamaktadır. Hem firma seviyesindeki araştırmalara hem de makro seviyedeki ekonometrik analizlere dayalı olmaktadır.²⁰³

Ülkelerin ekonomik ilerlemesi üzerinde etkili olan bilgi ekonomisi faktörleri ile ilgili çok değişkenli regresyon analizi yardımıyla, 118 ülkenin GSYİH verileri üç gelir grubuna ayrılmıştır. Düşük gelir grubundaki ülkelerde ekonomik gelişmeyi etkileyen bilgi ekonomisi değişkenleri eğitim ve bilgi iletişim teknolojileri olarak tespit edilmiştir.²⁰⁴ OECD tarafından standartlaştırılan temel bilgi belirleyicileri; Ar-Ge üzerindeki harcamalar, teknik personel ve mühendis istihdamı, patentler, teknoloji için uluslararası ödemeler dengesi, inovasyon ve insan kaynaklarıdır.²⁰⁵

GSMH'daki Ar-Ge harcamalarının payı, Ar-Ge faaliyetlerinde görev alan nitelikli personel sayısı, yapılan bilimsel çalışmaların yayın sayısı, bilgisayar, internet teknolojilerinden, iletişim araçlarından faydalanan bireylerin sayısı, ihracat rakamları içerisindeki teknolojik malların payı gibi değişkenler ülkelerin teknolojik olarak ve inovasyon faaliyetleri açısından yeterliliğini göstermektedir. Teknolojik yetkinlik açısından gelişmiş olan ülkelerle gelişmekte olan ülkeler arasında büyük uçurumların olduğu anlaşılmaktadır. Yukarıda anlatılan değişkenler açısından gelişmekte olan ülkelerin kendilerini ilerletmeleri gerekmektedir.²⁰⁶

Özel sektörün kamu harcamaları içindeki payı ülkeler arasında farklılık göstermektedir. Gelişme düzeyi yüksek ülkelerde, özel sektörün Ar-Ge harcamalarındaki payı daha fazladır. Buna karşılık gelişme düzeyi düşük olan ülkelerde kamu sektörüyle yükseköğretimin payı artmaktadır. Bunun en önemli

²⁰³Batagan, **a.g.e.**, p. 62.

²⁰⁴Dragomir Sundać, Irena Fatur Krmpotić, "Knowledge Economy Factors and The Development of Knowledge-Based Economy, **Croatian Economic Survey 13**, 2011, p. 126.

²⁰⁵OECD,(1996), **a.g.e.**, s. 31.

²⁰⁶Hüseyin Ağır, "Türkiye İle Güney Kore'de Bilim ve Teknoloji Politikalarının Karşılaştırması", **Bilgi Ekonomisi ve Yönetimi Dergisi 5.2.**, 2010, s. 34.

nedeni ekonomik gelişme düzeyi düşük olan ülkelerde Ar-Ge harcaması için gerekli olan fonların sadece devlet tarafından karşılanabilmesidir.²⁰⁷

1.7.Bilgi Ekonomisinin Dünya ve Türkiye’deki Yeri

Tarım toplumundan sanayi toplumuna geçiş ve bunun ötesinde de bilgi toplumunun oluşturulmasıyla birlikte küresel düzende bilgi hızla gelişip şekillenmektedir. Bilginin gelişimiyle düzenlenen ekonomik sistemde üretim ve verimlilik giderek artmaktadır. Bilgiye dayalı toplumların giderek daha çok ilerlemesiyle işgücü verimliliği artmakta, ekonomik büyüme daha da hızlanmakta bununla birlikte teknolojik anlamda ve bilimsel anlamda yeni buluşlar meydana çıkmaktadır. Henüz gelişimini tamamlamış olan ülkelerin, gelişmiş ülkelerin eriştiği bilgiye dayalı topluma ulaşabilmesi için çaba içerisinde olmaları gerekmektedir.

Türkiye açısından durum ise Türkiye'nin henüz sanayileşme sürecini tamamlamamış olmasının sonucu olarak birçok yapısal sorun ile karşı karşıya bulunduğu görülmektedir. Türkiye'nin sanayileşme sürecini tamamlaması ekonomik, sosyal, siyasal ve kültürel alanlarda yapısal değişimi gerçekleştirmesi ve uluslararası alanda küreselleşme sürecine uyum sağlaması açısından bilgi toplumuna yönelik gelişmeleri yakından izlemesi ve bu alana öncelik veren politikaları oluşturması gerekmektedir.²⁰⁸

Bilgi toplumu olma yolu uzun bir süreçtir. Bu süreç içerisinde geçmişte farklı alanlarda, farklı kurumlar çalışmalar yürütmüş olmakla birlikte, Türkiye’de bilgi toplumuna dönüşüm çalışmaları 2000’li yılların başından itibaren yoğunluk kazanmıştır.²⁰⁹ Gelişmiş ülkeler son yıllarda erişmiş bulundukları ileri gelişmişlik seviyesini ve bilgiye dayalı toplum seviyesini, gelişimlerinin başlangıç evresinde beşeri sermaye yatırımlarına verdikleri değer sayesinde kazanmışlardır. Türkiye’nin de dahil olduğu gelişmekte olan ülkelerin bilgiye dayalı topluma uyum ağlama aşmasında en temel problemi yetenekli işgücünün bulunmamasıdır. Bilgiye dayalı

²⁰⁷ Atik ve Dura, **a.g.e.**, s. 218.

²⁰⁸ Ethem Soner Çelikkol, Sema Yılmaz Genç, Okan Şeneldir, “Küreselleşme Olgusu ve Bilgi Toplumunun Gelişme Dinamiklerinin Türkiye Üzerindeki Olası Etkileri”, **International 8th Knowledge & Management Congress Proceedings (October 28-31)**, İstanbul, 2010, s. 215.

²⁰⁹ ,Ulvi Saran, “E-Devlet Ve Türkiye Bilgi Toplumu Stratejisi,” Uluslararası Katılımlı Bilgi Toplumu Sempozyumu, 2006, s.211.

toplumda yetenekli personel yokluğunun telafi edilmesi okuldaki eğitimin yanında okuldan sonraki eğitime, mesleki anlamdaki eğitime, yükseköğretimdeki eğitime, eğitim faktörünün devamlılığının kazanılmasına ve bunlar gibi unsurlara da bağlı bir durumdur.²¹⁰

Bilgi ve iletişim teknolojilerine ait üretim miktarları, bilgisayar ve ilgili aktiviteler kısmına ait katma değer, istihdam değişkeni, bilgi ve iletişim sektöründe çalışan bireylere ödenen ücretlerde hesaplanan yüzde değişim, bilgi ve iletişim teknolojileri sektöründe kullanılan örgüt sayılarına ait değişim verilerinin kullanıldığı bir çalışmada OECD ülkeleri arasında kümeleme analizi yapılmıştır. Yapılan analizde gelişmiş ülkelerin benzer kümelerde toplandığı görülmüştür. Türkiye ise daha çok Kuzey Avrupa ülkeleri ile aynı benzer değişimler göstermiş ancak gelişmekte olan bir ülke olarak bu değişkenler açısından ilerlemesinin yavaş olduğu görülmüştür.²¹¹

Üretim faktörlerinin kaliteli ve nicelikli olmasıyla birlikte teknolojik ilerlemelerinde artmasıyla ekonomik gelişmişlik düzeyi daha da arttırılabilmektedir. Teknolojik anlamda ilerlemenin elde edilmesindeki en temel ve kuvvetli faktör ise bilginin üretilmesidir. Bilgi üretimini gerçekleştiren ülkeler, teknolojik anlamda faydalı ilerlemeler sağlayarak, rekabet ettikleri ülkelere nazaran daha da önde gelişme hızına sahiptirler. Türkiye birçok Avrupa ülkesine göre bilgi üretimi alanında geride bulunmaktadır. Bu alanda ilerleyebilmesi için bu alanda yetenekli personelin eğitilmesi, eğitimli personel istihdamının arttırılması gerekmektedir. Mevcut işgücünü bünyesinde barındırması gerekmektedir. Eğitim seviyesinin artması bilgi üretimini gerektiren mesleklerin kolaylıkla icra edilmesini sağlayıp istihdamı arttırırken, diğer yandan da bilginin üretilmesini hızlandırarak arttıracaktır ve bu sayede teknolojik ilerleme seviyesi daha da yükselecektir. Bu durumda ekonomik büyümeye olumlu yansıyacaktır.²¹²

²¹⁰Coşkun Can Aktan, Mehtap Tunç, **Bilgi Toplumu ve Türkiye, Moderniteden Postmoderniteye Değişim**, Ed. by. Coşkun Can Aktan, Çizgi Kitapevi, Konya, 2003, s. 119.

²¹¹Yeloğlu, **a.g.e.**, s. 249.

²¹²Hayriye Atik, Aytekin Altıparmak, "Avrupa Birliği ve Türkiye'de Bilgi Sektörünün İstihdama Katkısı", **Bilgi Ekonomisi ve Yönetimi Dergisi 6.1.**, 2011. s. 235.

İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

2.1. Veri Madenciliği Tanımı ve Gelişimi

Gelişen dünyada sürekli bilgi oluşumu ve akışı meydana gelmektedir. Son yıllarda kullanılabilir bilginin artması, kullanıcılar tarafından bu bilgilerin stoklanarak daha faydalı bilgiler oluşturmalarını sağlamaktadır. Günümüzde biriken bilgi stokları birçok veri tabanını oluşturmakla birlikte bu verilerin dijital ortamda saklanması veri tabanlarının oluşumunu daha da hızlandırmaktadır. Dijital ortamlarda oluşturulan bu veri tabanları içerisinde, işlendiğinde ortaya çıkacak birçok faydalı bilgi yer almaktadır. Veri tabanları arasındaki gizli ve faydalı bilgileri ortaya çıkarmak giderek daha da önem kazanmaktadır. Bu noktada veri madenciliği kavramı devreye girmektedir ve birçok farklı yöntemle gizli ve faydalı bilgileri ortaya çıkarmaktadır.

1990'ların ortasından beri veriden daha fazla bilgi elde etmeyi vurgulayan birbiriyle ilgili 3 yeni alan bilgi sistemlerinde ve bilgi teknolojisinde meydana gelmiştir. Bu alanlar veri ambarı, bilgi yönetimi ve veri madenciliğidir. Bu alanlardan veri madenciliği, verilerdeki kabul edilebilir, özgün, potansiyel fayda sağlayan ve açık olan korelasyonları ve modelleri tanımlamayı amaçlamaktadır. Bilgisayar donanımı ve yazılımındaki ilerlemelerle birlikte birçok veri madenciliği uygulamaları önceki zamanlardansa şimdi firmalara daha erişilebilir ve daha satın alınabilir özelliktedirler.²¹³

Veri madenciliği büyük miktarlardaki veriyi farklı metotlarla hem anlaşılabilir hem de faydalı hale getiren ve beklenmedik ilişkileri bulmak için gözlemsel veri setlerinin analiz edilmesini sağlayan bir yöntemdir.²¹⁴ Daha önceleri tespit edilememiş, geçerliliği olan ve uygulanabilen bilgilerin çoğunun büyük veri tabanlarından temin edilmesi ve bu bilgilerin organizasyon kararları alınırken kullanılmasıdır. Burada üzerinde durulması gereken olan önemli noktalardan biri temin edilecek bilgiler hakkında daha önceleri bilinmezlik durumunun olmasıdır.

²¹³Hian Chye Koh, Wei Chin Tan, Goh Chwee Peng, "Credit Scoring Using Data Mining Techniques", **Singapore Management Review** 26.2., 2004, p. 101.

²¹⁴David J. Hand, Heikki Mannila, Padhraic Smyth, **Principles of Data Mining**, MIT Press, 2001, p. 6.

Erişilecek bilginin önceden bilinmeme durumunun olmasıyla anlatılmak istenen söz konusu bilgilerle sağlanacak sonucun öngörülememesi anlamıdır.²¹⁵ Veri madenciliği sorunları çözmeye yardım eden tanımlayıcı ve tahmin edici modeller meydana getirerek ve veri analiz ederek işlemektedir.²¹⁶

Birçok veri bir araya getirilerek üzerlerinde çalışmalar ve çözümlerler gerçekleştirilmesiyle farklı istatistik ve matematik uygulamalar kullanılabilmektedir. Bu uygulamalarda kullanılan veri sayısının artmasıyla birçok sorunda meydana çıkmaktadır. Bağlantılı veri tabanları üzerinden bu uygulamaları hayata geçirmek oldukça zorlaşmaktadır. Birçok veriyi barındıran veri yığınları üzerinde çözümlerler yapmak için hem farklı veri tabanı uygulamalarına hem de yeni çözümler tekniklerine ihtiyaç duyulmaktadır. Verileri yönetebilmek için veri ambarları ve bu veriler üzerinde çözümlerler yapılmaktadır. Böylece faydalı bilgilere ulaşılmasının sağlanması için veri madenciliği kavramı meydana çıkmıştır. Veriler arasındaki ilişkileri açığa çıkarmak veri madenciliği ile mümkün görülmektedir. Veri madenciliği herhangi bir kuruluştaki elde edilen veriler içerisindeki mevcut ya da gelecekteki herhangi bir zamanda gün yüzüne çıkarılabilecek gizli bilgilerin su yüzüne getirilme süreci olarak değerlendirilmektedir. Bu nedenle veri madenciliği faaliyetlerinin kurum ve kuruluşların karar mekanizmalarına destek sağlamada kritik öneme sahip olduğu söylenebilmektedir.²¹⁷

Veri tabanı yazılımları bilgisayar destekli sistemlerin önemli bir unsuru olarak değerlendirilmektedir. Veri tabanlarının yönetilmesi için kurulan sistemler, bağlantılı veri ve programların bir araya gelmesinden ortaya çıkmaktadır. Veri tabanı herhangi bir kuruma ait olan bilgilerden oluşan bir ortamı kapsamaktadır. Veri tabanlarına ait sistemler çoğu veri kümesinin belli bir düzen içinde tutulabildiği ve bu verilerin farklı yazılımların bağlantısıyla yönetilebildiği bir ortam olarak

²¹⁵Gökhan Silahtaroglu, **Veri Madenciliği**, Papatya Yayınları, İstanbul, 2013, s. 12.

²¹⁶H. Ali Ata, İbrahim H. Seyrek, "The Use of Data Mining Techniques in Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi** 14.2.,2009, s. 2.

²¹⁷Yalçın Özkan, **Veri Madenciliği Yöntemleri**, Papatya Yayıncılık Eğitim, 2013, s. 38.

düşünülebilmektedir.²¹⁸ Veri madenciliği verideki modelleri ve ilişkileri keşfetmek için istatistiksel algoritmaları kullanan, komplike veri araştırma imkanı olarak tanımlanmaktadır. Veri madenciliği verideki gizli örüntüleri ve ilişkileri ortaya çıkarmaktadır. Anlamlı sonuçlar çıkarmak için ilerlemek zorunda olan adımları tanımlayan bilgi keşfi sürecinin bir bölümüdür.²¹⁹

İlk olarak bilgisayar kullanımı 1950’de sayım için gerçekleştirilmiştir. Daha sonraları verilerin birikmesiyle veri tabanı ve verilerin stoklanması kavramları teknolojik alanda anlam kazanmaya başlamıştır. 1960’lı yılların sonlarına doğru basit öğrenmeli bilgisayarlar bilim insanları tarafından hayata geçirilmiştir. İlişkisel Veri Tabanı Yönetim Sistemleri ile ilgili uygulamalardan 1970’li yıllarda faydalanılmaya başlanmıştır. Zaman ilerledikçe bilgisayar alanında uzmanlaşan kişiler aynı zamanda basit kurallı uzmanlaşmış sistemleride geliştirmişleridir. Basit anlamda makine öğrenimi böylece hayata geçmiştir.²²⁰ Bu sistemler 1980’li yılların başlarında daha çok yayılarak bilimsel kapsamda, mühendislikle gibi alanlarda uygulamaya konulmuştur. Söz konusu yıllarda firmalar, bu firmaların müşterileri, rekabet ettikleri firmalar ve kendi ürünleri ile ilgili verilerinin yer aldığı veri tabanlarını meydana getirmişlerdir. Bu veri tabanlarında kendileri ile ilgili devasa boyutlarda bilgileri muhafaza etmektedirler. Bu verilere ise SQL sorgulama dili ya da vb. gibi programlarla erişilebilmektedir. Veri tabanlarındaki veri miktarının sürekli artması, 1990’lı yıllarda bu biriken verilerden nasıl istifade edileceğini düşündürmeye başlamıştır. VTBK ile ilgili oluşturulan çalışma grupları ile 1989 yılında yapılan bir toplantıda ve 1991 yılında söz konusu toplantının sonuç raporu yerine geçen “Gerçek Veri Tabanlarında Bilgi Keşfi” makalesinde veri madenciliği ile ilgili kavramların ve tanımların açığa çıkarılmasıyla bu süreç hızlandırılmıştır. 1992’de ise ilk veri madenciliği yazılımı gerçekleştirilmiştir. 2000’li yılların başlarından itibaren ise bu gelişim devam ederek birçok farklı alanda yayılmaya başlamıştır. Veri madenciliği

²¹⁸ A.e., s. 14.

²¹⁹ Chris Rygielski, Jyun-Cheng Wang, David C. Yen, Data, "Mining Techniques for Customer Relationship Management", **Technology in Society** 24.4., 2002, s. 488.

²²⁰ Pieter Adrians, Dolf Zantinge, **Data Mining**, İngiltere, 1997, s. 15.

uygulamalarında alınan başarılı sonuçlarla bu alan daha çok ilgi görmeye başlamıştır.

221

Dönem	Teknolojiler	Özellikler
Veri Biriktirme (1960s)	Bilgisayarlar, kayıtlar, diskler	Geçmişe yönelik statik veri dağıtımı
Veri Erişimi (1980s)	İlişkisel veritabanları (RDBMS), SQL sorguları	Geçmişe yönelik kayıt seviyesinde statik veri dağıtımı
Veri Dolaşımı (1990s)	Online Analitik İşleme (OLAP), Çok Boyutlu Veritabanları, Veri Ambarları	Geçmişe yönelik çoklu seviyede dinamik veri dağıtımı
Veri Madenciliği (2000s)	İleri Algoritmalar, Çok işlemcili Bilgisayarlar, Büyük Boyutlu Bilgisayarlar	İleriye dönük, tahmin edici bilgi dağıtımı

Tablo 2.1. Veri Madenciliği Gelişimi.

Bugünkü teknoloji, verilerin aracılığı ile kullanıcılara hayata yönelik yol göstermektedir. İlk aşamada veri birikimi sağlanmaktadır. Daha sonra bireysel şekilde biriktirilen veriler toplam ya da ortalama gibi basit hesaplamalar yapmak için kullanılmaktadır. İkinci aşama veri erişimi olarak gösterilmektedir. Sonraki aşama ise yapılandırılmış bir formatta veri depolamak için veri tabanlarının kullanılmasıdır. Bu aşamada teknoloji, bilgi yönetimi ve büyük çaplı politikalar için de kullanılmaktadır. Son olarak geri bildirim, şirket birimleriyle etkileşimli bilgi değişimi sunan online analitik araçlar yer almaktadır.²²²

Bilgi keşfi verilerin içinden bilinmeyen, potansiyel olarak faydalı, kesin bilginin alınmasıdır. Mevcut veri tabanlarının sayısı ve büyüklüğündeki gelişimle birlikte bu verileri analiz etmek için insani yetenekler yetersiz kalmaktadır. Veri

²²¹Serkan Savaş, Nurettin Topaloğlu, Mithat Yılmaz, "Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri", **İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi Yıl:11 Sayı: 21**, 2012, s. 5.

²²²Rygielski, Wang, ve Yen, **a.g.e.**, s. 486.

madenciliği hem veri tabanı hem de makine öğrenme araştırmacıları tarafından 1990'lardan beri en fazla gelecek vadeden araştırma konuları içinde yer almaktadır.²²³ Geniş veri tabanlarından enformasyon ve bilgiyi ortaya çıkartmak veri tabanı sistemleri ve makine öğrenmesi alanlarındaki birçok kullanıcı tarafından ana araştırma konusu olarak tanımlanmaktadır. Birçok sektör firması tarafından ise yüksek gelirli fırsatla birlikte önemli bir alan olarak tanımlanmaktadır.²²⁴

Veri madenciliği veri tabanında önceden ortaya çıkarılmamış ve öngörülemeyen bağlantıların ikincil anlamda analizi olarak da tanımlanmaktadır. Aynı zamanda istatistik, matematik, veri tabanı teknolojileri, örüntü tanıma ve makine öğrenmesi ile ilişkili yeni ve farklı bir disiplin olarak açıklanmaktadır.²²⁵ Veri madenciliği yöntemleri uzun bir saha araştırması ile mal ve hizmet tasarlama sürecinin neticesidir. Veri madenciliğinin başlangıç noktası bilgisayarlarda ilk depolanan verilere kadar uzanmaktadır ve veri erişimindeki ilerlemelerle devam etmektedir.²²⁶

2.1.1. Veritabanlarında Bilgi Keşfi ve Veri Madenciliği

Veri tabanı sistemlerinin giderek kullanımının artması ve büyüklüklerindeki devasa artışın, örgütleri stoklanan veri yığınlarından ne şekilde yararlanılabileceği sorunu ile yüz yüze bırakmıştır. Geleneksel sorgular ya da raporlama araçları ile ilgili büyük veri kitlelerinin yetkinlik gösterememesi, veri tabanlarında bilgi keşfi (VTBK) adı ile yeni arayışların içine girilmesine sebep olmaktadır. VTBK süreci içinde, modelin kurulumu ve değerlendirme adımlarında oluşan veri madenciliği uygulaması kritik kısmı ifade etmektedir. Bu kritik önemden dolayı araştırmacılar VTBK ve veri madenciliği kavramlarının birbirlerine benzediğini düşündüklerinden eş anlamlı da kullanılmaktadırlar.²²⁷ Veri analizi, test edilmek istenen hipoteze karar verilmesi

²²³William J. Frawley, Gregory Piatetsky-Shapiro, Christopher J. Matheus, "Knowledge Discovery in Databases: An Overview", **AI Magazine** **13.3.**, 1992, p. 3.

²²⁴Ming-Syan Chen, Jiawei Han, Philip S. Yu, "Data Mining: An Overview from A Database Perspective", **IEEE Transactions on Knowledge and Data Engineering** **8.6.**, 1996, p. 870.

²²⁵, David J. Hand, "Data Mining: Statistics and more?.", **The American Statistician**, **Vol.52 No.2**, 1998, p.115.

²²⁶Rygielski, Wang, ve Yen, **a.g.e.**, s. 485.

²²⁷, Haldun Akpınar, "Veri Tabanlarında Bilgi Keşfi Ve Veri Madenciliği", **İÜ İşletme Fakültesi Dergisi**, **29..1**, 2000, p.19.

aşamasıdır ve VTBK sürecinin kalbini oluşturmaktadır. Bu aşamada istatistiksel paket programları kullanılmaktadır.²²⁸

Kuramsal seviyede VTBK alanı, veriyi anlamlandırmak için teknik ve metotların geliştirilmesiyle ilgilenmektedir. Sürecin merkezinde model çıkarımı ve keşfi için özel veri madenciliği yöntemlerinin uygulanması yer almaktadır. Biriktirebildiğimiz büyük miktarlardaki verileri kullanmak için beşeri analiz özelliklerini çoğaltma, yani hem ekonomik açıdan hem de bilimsel açıdan önemli olarak görülmektedir. Firmalar müşterilere daha değerli hizmetler sunmak, etkinliklerini arttırmak ve rekabet avantajı sağlamak için devasa büyüklükteki verileri kullanmaktadırlar. Çevreyle ilgili ele alınan veriler, yaşadığımız evrenin modelleri ve teorilerini kurmak için kullandığımız esas bulgulardır. Bilgisayarlar insanlara çok büyük miktarlarda veriyi biriktirme imkânı sağlamaktadır. VTBK büyük hacimlerdeki verilerden anlamlı yapılar ve modelleri keşfetmek için yardım etmektedir. VTBK hepimiz için dijital bilgi çağının hayatımızın bir gerçeği olduğu problemine yönelmek için bir girişimdir.²²⁹

VTBK verilerinden faydalı bilgiyi keşfetme sürecinin tamamını oluştururken, veri madenciliği bu sürecin sadece bir aşamasıdır. Veri madenciliği verilerden modeller çıkarmak için özel algoritma uygulamasıdır. Verilerin hazırlanması, verilerin seçilmesi, verilerin temizlenmesi, önceki uygun bilginin birleştirilmesi ve veri madenciliği sonuçlarının uygun yorumlanması gibi adımlar, veriden faydalı bilginin çıkarılmasını garantilemek için zorunlu olarak gösterilmektedir.²³⁰ VTBK'nın veri madenciliği girdisi verilerden örüntü bulmak için istatistikle, makine öğrenmesi ile ve model tanımlama ile ilgilenmektedir. VTBK verilerin nasıl stoklandığını ve verilere nasıl erişildiğini, algoritmaların devasa veri setlerini nasıl ölçekleyebildiğini ve nasıl etkin olarak çalıştığını, sonuçların nasıl yorumlandığını ve görselleştirildiğini, insan makine etkileşiminin nasıl faydalı olarak modellendiğini ve

²²⁸ Ronald J. Brachman, Tej Anand, "The Process of Knowledge Discovery in Databases. In: Advances in Knowledge Discovery and Data Mining", **American Association for Artificial Intelligence**, 1996, p. 40.

²²⁹ Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, "From Data Mining to knowledge Discovery in Databases", **AI Magazine** 17.3., 1996, p. 36.

²³⁰ A.e., p. 38.

desteklendiğini kapsayan veriden bilgi keşfetme sürecinin tamamına odaklanmaktadır.²³¹

Dil, kesinlik, ilginç sonuç, etkinlik gibi özellikler veri tabanlarında bilgi keşfinin temel karakteristiklerini oluşturmaktadır.

- Örüntüler ve diller: Bilginin birçok farklı çeşidi olmasına rağmen verilerde keşfedilebilmektedir. Bilgi keşfedilirken kullanılan dil doğrudan anlaşılabilir olmalıdır.
- Kesinlik: Kesinliğin derecesini temsil etmek ve yaymak, yatırım yapılan keşif sisteminin ne kadar güvenilir olduğunu belirlemek için zorunludur.
- İlginç sonuç: Birçok model herhangi bir veri tabanından çıkarıldığı için bir nevi ilginç sayılan modellerin sonuçları bilgi olarak düşünülebilmektedir. Modeller özgün olma, basit olmama ve kullanışlılık gibi özelliklere sahipse ilginç görülebilmektedir. Sistem bilgisinin faaliyet alanı ya da kullanıcı bilgisinin faaliyet alanı olabilen bir modelin özgün olup olmadığı, varsayılan referans sistemine bağlıdır.
- Etkinlik: Keşif süreci bilgisayarda etkin olarak uygulanmaktadır. Uygulama zamanı ve kullanılan alan, girdi uzunluğunun düşük dereceli polinomal bir fonksiyonu ise kullanılan algoritma etkin olarak düşünülmektedir.²³²

Veri tabanı sistemleri ile ilgili yazılımlar bilgisayar sistemlerinin kritik bir unsuru olarak değerlendirilmektedir. Verilerin ve programların bir araya getirilmesiyle, veri tabanlarının yönetilmesini sağlayan sistemler oluşturulmaktadır. Verilerin bir araya gelerek oluşturduğu topluluk "veri tabanı" olarak değerlendirilmektedir. Veri tabanı verilerle oluşturulan kümelerin düzenli olarak takip edildiği ve bu verilerin farklı yazılımlarla organize edildiği bir ortam olarak düşünülebilmektedir.²³³

VTBK süreci, seçimle birlikte, ön işleme, alt örnekleme, veri tabanı dönüşümleri ve modelleri düzenlemek için veri madenciliği metodlarını uygulamaktadır. VTBK, düzenlenmiş modeller kümesinin geçerli saydığı bilgiyi

²³¹ A.e., p. 40.

²³² Frawley, Shapiro ve Matheus, a.g.e., p. 50.

²³³ Özkan, a.g.e., s. 14.

tanımlama için veri madenciliği ürünlerini değerlendirmekle birlikte veri tabanını kullanmayı kapsamaktadır.²³⁴

Veritabalarında bilgi keşfi verilerdeki geçerli, ilginç, potansiyel olarak kullanışlı ve sonuç olarak anlaşılabilir modelleri tanımlayan anlaşılması zor bir süreçtir. Veri madenciliği ise veri tabanlarında bilgi keşfi sürecinin sadece bir adımındır. VTBK'nin diğer adımları ise; veri madenciliği sürecinin amaçlarını ve uygulamanın etkinlik alanının anlaşılmasını sağlamaktır. Bunlar:

- Hedef veri setini seçmek
- Veri setini bütünleştirme ve kontrol etmek
- Veri temizleme, önişleme ve dönüşüm
- Model geliştirme ve hipotez kurmak
- Uygun veri madenciliği algoritmasını seçmek
- Sonuçları yorumlama ve görselleştirmek
- Sonuçları test etmek ve doğrulamak
- Keşfedilen bilgiyi kullanmak ve devam ettirmektir.

Veri setinin seçilmesi ve oluşturulması: Bu aşamada veri yığınları birleştirilerek sorgulama yapmak amacıyla elverişli örneklem hacminin belirlenmesi amaçlanmaktadır.

Uygulama alanını anlama ve geliştirme: Bu adımda uygulamanın gerçekleştirileceği alan anlaşılmaya çalışılmaktadır ve çeşitli hazırlıklarla birlikte bilginin keşfindeki amaç geliştirilmektedir.

Veri temizleme: Verilerin hatalı olması durumunda meydana gelebilecek sıkıntıları gidermek için yapılan bir süreçtir. Veri durağan ve güncellenmeyecekse temizleme süreci başlayacaktır. İlk veri tabanı oluşturulduğu gibi kalacaksa, sürekli veri kalitesi sorunları meydana gelecektir. Sonuç olarak yüzeysel basit bir veri madenciliği çalışması gibi görünse de, gerçekte biriktirilebilen ve analiz edilebilen

²³⁴Fayyad, Shapiro ve Smyth, **a.g.e.**, p. 42.

veriyi üretmek amacıyla önemli örgütsel altyapı revizyonu için itici bir güç olabilmektedir.²³⁵

Verinin ön işlenmesi ve temizlenmesi: Bu aşama da seçilmiş örnekleme alan yanlış ve eksik özellikteki verilerin ayıklandığı ve değiştirilerek temizlendiği bir aşamadır ve bu işlemler açığa çıkarılan bilginin değer ve kalite boyutunu artırmaya yönelik yapılmaktadır.

Verinin işlenmesi: Seçilmiş olan örnekleme ile ilgili bulunmayan özellikler ve yinelenen veriler bu aşama da temizlenmektedir. Bu adımın hedefi veri madenciliğindeki sorgunun doğru çalışmasını sağlamaktır.

Modelin geliştirilmesi ve hipotezin kurulması: Tanımlanmış probleme en doğru modelin araştırılması, fazla sayıda model kurularak denenmesiyle mümkün olmaktadır. Bu nedenden dolayı verilerin hazırlanması ve modelin kurulması, denenmesi aşamaları en iyi sonucu verdiği düşünülen model elde edilinceye kadar tekrarlanan bir süreçtir.²³⁶

Uygun bir veri madenciliği programı ve algoritmasının seçilmesi: Veri madenciliği uygulamalarından hangisinin kullanılacağına bu adımda karar verilmektedir. Bu aşama en baştaki veri tabanlarında bilgi keşfi sürecinin hedefine ve önceki adımlarına bağlı olarak ilerlemektedir. Bu adımda ayrıca strateji belirlenmeli, uygulanacak tekniklerle ilgili karar alınmalı, örüntü araştırmak amacıyla kullanılacak olan kesin bir yöntemle karar verilmesi gerekmektedir.

Veri madenciliği algoritmasını çalıştırılması: Bu adımda bir önceki aşamada seçilen veri madenciliği algoritması uygulanmaktadır. Bu aşamada hedef, elde edilen sonuçlar tatmin edici oluncaya kadar algoritmanın çalıştırılmasıdır.²³⁷

Sonuçların yorumlanması ve görselleştirilmesi: Bu adımda elde edilen, uygunluğu kabul edilen model ya da modellerin kurulduktan sonra, veri madenciliği uygulamaları sonucu araştırmanın probleminin ne kadar çözüldüğü

²³⁵Brachman ve Anand, a.g.e.,p. 41.

²³⁶Zaki Mohammed, **Introduction to Data Mining**, Springer Verlag, 2003, p. 39.

²³⁷Bahadır Özdemir, Selim Aksoy, "Image Classification Using Subgraph Histogram Representation. In: Pattern Recognition (ICPR)", **20th International Conference on. IEEE**, 2010. p. 351.

değerlendirilmektedir. Bu aşamada elde edilen sonuçlar değerlendirilir, veri madenciliği uygulama süreci gözden geçirilir ve sonraki aşamalarda ne yapılacağı konuları incelenir. Bu adımın sonunda veri madenciliği uygulama sonuçları üzerinde düşünülerek kullanımla alakalı karar verilmektedir.²³⁸

Sonuçları test etmek ve doğrulamak: Elde edilen sonuçların geçerliliğinin test edilmesi ve onaylanması gerekmektedir. Bu aşamada sonuçlar doğrulandıktan sonra kabul edilmektedir.

Keşfedilen bilgiyi kullanmak ve devam ettirmek: Zaman içerisinde bütün sistemlerin özelliklerinde ve dolayısıyla ürettikleri verilerde ortaya çıkan değişiklikler, kurulan modellerin sürekli olarak izlenmesini ve gerekiyorsa yeniden düzenlenmesini gerektirecektir.²³⁹

2.1.2. Veri Ambarı ve Veri Madenciliği

Veri tabanı sistemlerinin avantajlarına rağmen, özellikle karar destek uygulamalarındaki gereksinimleri karşılamakta zorlandığı görülmüştür. Bu gereksinime paralel olarak verinin farklı biçimde saklanması ve veriye hızlı erişimin sağlanması gibi başka yollar aranmaya başlanmıştır. Bununda ötesinde kurumun ürettiği tüm bilgileri, boyutu ne olursa olsun işlemek söz konusu olacaktır. Söz konusu verinin boyutları bazı kurumlarda o denli fazladır ki, bu verileri geleneksel veri tabanı yöntemleriyle işlemek olanaksızdır. Veri ambarı adı verilen kavram bu tür gereksinimlerden dolayı ortaya çıkmıştır. Veriler üzerinde çözümlemeler yapmak amacıyla çeşitli istatistiksel ve matematiksel yöntemler kullanılabilmektedir. Ancak veri sayısı arttıkça sorunlar ortaya çıkacaktır. Özellikle ilişkisel veritabanları üzerinde bu çözümleri yapmak zorlaşacaktır. Bu tür veriler üzerinde çözümlemeleri yapabilmek için hem yeni veri tabanı kavramlarına hem de yeni çözümleme yöntemlerine gereksinim duyulmaktadır. Veriyi yönetmek amacıyla "veri ambarı" ve verileri çözümleyerek "yararlı bilgiye" erişilmesini sağlayan "veri madenciliği"

²³⁸ 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 14.1.**, 2009, s. 37.

²³⁹ Ziya Y. Ayık, Abdülkadir Özdemir, Uğur Yavuz, "Lise Türü ve Lise Mezuniyet Başarısının, Kazanılan Fakülte ile İlişisinin Veri Madenciliği Tekniği ile Analizi", **Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 10.2.**, 2007, s. 444.

kavramları ortaya atılmıştır. Veri ambarı, birbirleriyle bütünleşik olmayan uygulamaların bütünleştirilmesi açısından bir olanak sağlamaktadır. Veri ambarı, bir zaman boyutu içinde analitik işlemlerin yapılması için ihtiyaç duyulan bilgi temelini sağlamakta ve karar verme sürecinde yöneticilere destek vermek üzere hazırlanmaktadır.²⁴⁰

Veri ambarı, firmaların karar destek işlemleri ve online analiz faaliyetleri için hazır hale gelen işlemsel verinin temizlenmesi ve biriktirilmesi eğiliminden oluşan ve ilgili alanlara ayrılan veri tabanlarından meydana gelmektedir.²⁴¹

2.1.2.1. Veri Ambarı Temel Nitelikleri

Veri ambarı, temelinde konu olan, bütünleşik, farklı zamanlara ait, kalıcılık özelliğine sahip veri topluluklarının, örgütsel anlamda kara alma da öncelikli olarak kullanılmasını ifade etmektedir.²⁴² Kurumların belli başlı amaçlarına ya da konularına yönelik olması veri ambarında yer alan verilerin tasarımının ve uygulamasının belli konulara göre düşünülmesini gerektirmektedir. Veri ambarının kapsamındaki verilerin en göze çarpan özelliği entegre durumda bulunmalarıdır. Bu bütünleşme farklı biçimlerde olabilmektedir. Verilerin kodlanmasındaki ortak görüş, ölçü birimlerindeki nicel özelliklerin fiziksel görünümündeki tutarlılığa dikkat edilmesidir. Veri ambarındaki tüm veriler zamanın belli bir anına aittir. Veri ambarında yer alan veri sadece okunabilir bir yapıda olabilmektedir. Veri yönetiminin ihtiyaçlarına cevap vermek amaçlı düzenlendiği için günlük olarak işleme tabi tutulmamaktadır yani silinmemekte ve güncelleştirilmemektedir.²⁴³

Veri ambarı, veri kümelerine VTBK adımı için kritik yönden yardım sağlamaktadır: bunlar verinin temizlenmesi ve veriye erişimdir.

- Verinin Temizlenmesi: Örgütler ellerinde bulundurdıkları geniş kapsamdaki veriler ve veri tabanlarının mantıksal görünümünü anlamakta zorlanmaktadırlar. Bu nedenle

²⁴⁰Özkan, a.g.e., s. 20-22.

²⁴¹Fayyad, Shapiro ve Smyth, a.g.e., p. 42.

²⁴²William H. Inmon, "Information architecture for the 90's: legacy systems, operational data store, data warehouse". **PRISM TechTopic**, Vol.1, No.13, 1993, p.120.

²⁴³Özkan, a.g.e., s. 26.

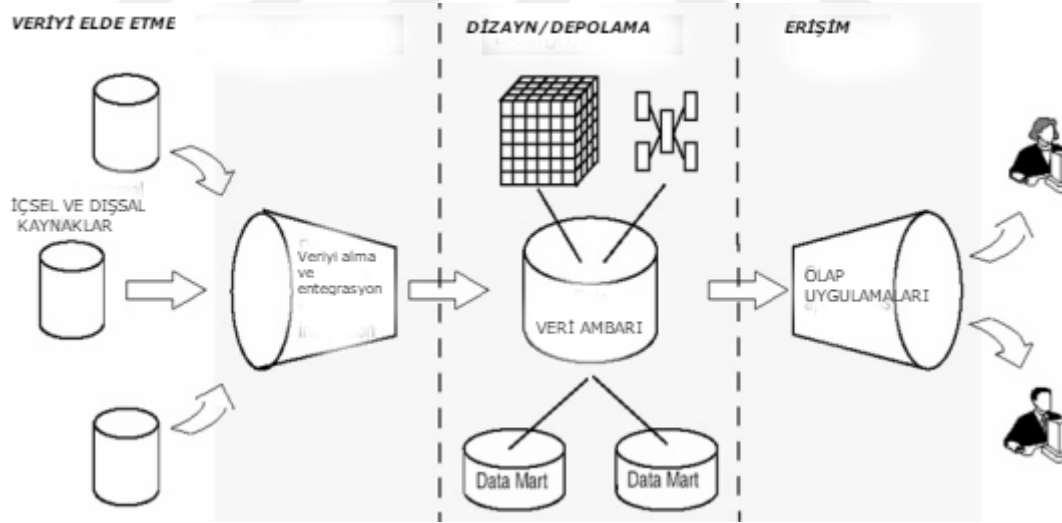
verilerin haritalanması, sonuçlarının tek bir isimle adlandırılmasının yanı sıra kayıp verinin temizlenmesi ve gürültü ile hataların tespit edilmesi gerekmektedir.

- Veriye Erişim: Tarihi olarak sıralanmış verileri elde etmek zor olduğu için bu verilere ulaşım yollarının bulunması için iyi tanımlanmış teknikler meydana getirilmelidir. İlk olarak örgütler ve örgüt elemanları verilerinin depolanma ve ulaşım sorunlarına odaklanmalıdırlar.²⁴⁴

2.1.2.2. Veri Ambarı Mimarisi

Veri ambarı mimarisindeki genel özellikler aşağıdaki gibidir:

- a) Çeşitli kaynaklardan elde edilen veriler üzerinde dönüştürme işlemi uygulanmaktadır.
- b) Veri ambarı meydana getirilmektedir.
- c) Kullanmak isteyen kişiler için veri ambarına erişim sağlanmaktadır.²⁴⁵



Şekil 2.1 .Veri Ambarı Mimarisi

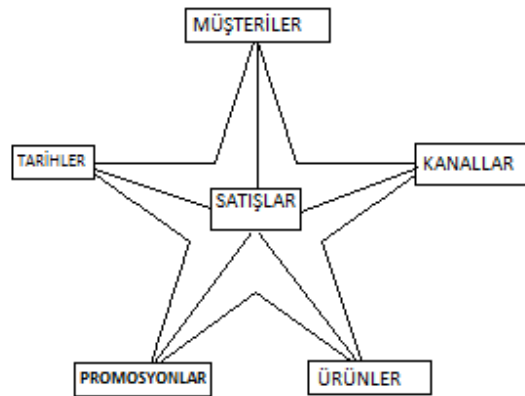
Veri ambarı ve OLAP(Online Analytical (Processing) teknolojileri karar alma desteği olarak 90'lardan beri geniş bir kabul alanı bulmaktadır. Veri ambarı OLAP araçlarını manipüle etmektedir. OLAP araçları genellikle veri küpü olarak

²⁴⁴ Ali Serhat Koyuncugil, Nermin Özgülbaş, "Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları", *International Journal of Informatics Technologies* 2.2.,2009, s. 23.

²⁴⁵ Özkan, a.g.e., s. 30.

adlandırılan çok boyutlu veri görüşlerinin gezinme ve görselleştirme mekanizmalarını sunmaktadır. Veri küpü ambardaki veriyi incelenmesi için çok boyutlu olarak temsil etmektedir. Veri küpü ölçümlere sahip olan gerçeklerden oluşmaktadır. Bu ölçümler nitelik olarak adlandırılan ve kategorik tanımlamalar kümesinden oluşturulan her bir boyuta dayalı değerlerdir. Örneğin müşteri, ürün, bölge ve zaman gibi ilgili boyutların dahil edildiği satışlar uygulaması düşünülebilir. Bu uygulamadaki ilgili ölçü satış miktarıysa, OLAP unsuru dikkate alınan boyutlardaki tek bir üyeye uyum sağlayan satış ölçüsünü yansıtmaktadır. Bir boyut ve hiyerarşiye göre organize edilebilmektedir. Örneğin lokasyon boyutu şehir, eyalet, bölge hiyerarşisini düzenleyebilmektedir.²⁴⁶

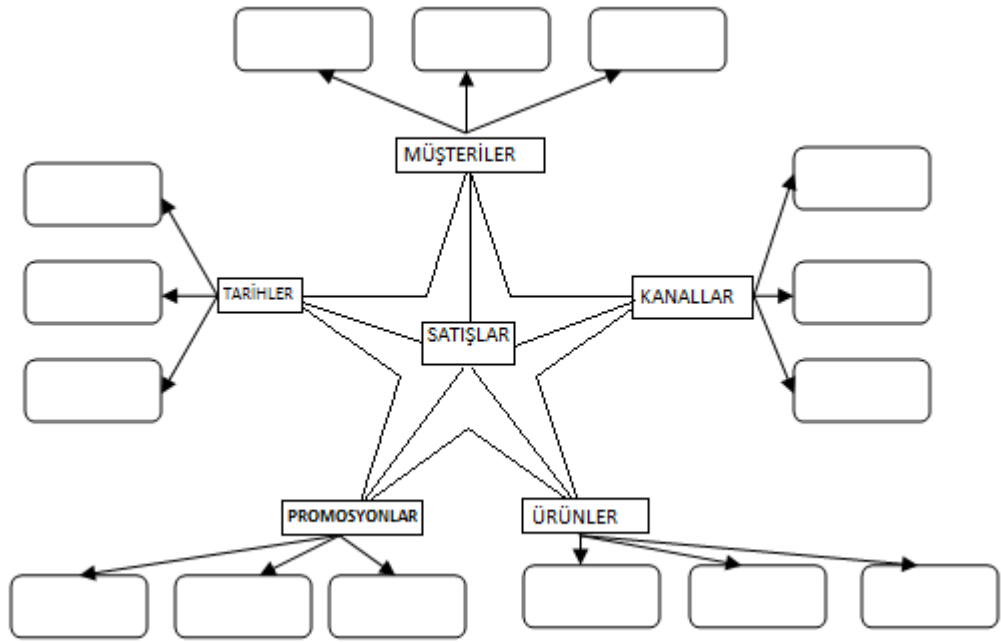
Veri ambarlarının mimarileri temelde üç farklı şema kullanılarak tasarlanmaktadır. Bu şemalar yıldız şema, kar tanesi şema ve ana tablo birliği şema şeklindedir. Yıldız şema da ortada yer alan bir ana tablo etrafında veri ambarı mimarisini oluşturan tablolar bulunmaktadır. Ana tablo işerin temel ölçümlerini içeren özelliklerden oluşmaktadır. Bir ana tablo ve ona ait olan özellikler ayrıca boyutların tablolarıyla alakalı yabancı anahtarları kapsamaktadır. Boyutları oluşturan tablo, ana tabloda gizlenen verinin indeksleri ve düzenlenen özelliklerinden meydana gelmektedir. Bu tablo söz konusu boyutları açıklayan özellikleri kapsamaktadır.



Şekil 2.2 Veri Ambarı Mimarisi Yıldız Şema.

²⁴⁶Messaoud vd., "OLEMAR: An Online Environment for Mining Association Rules in Multidimensional Data", **Data Mining and Knowledge Discovery Technologies** 35, 2007, p. 30.

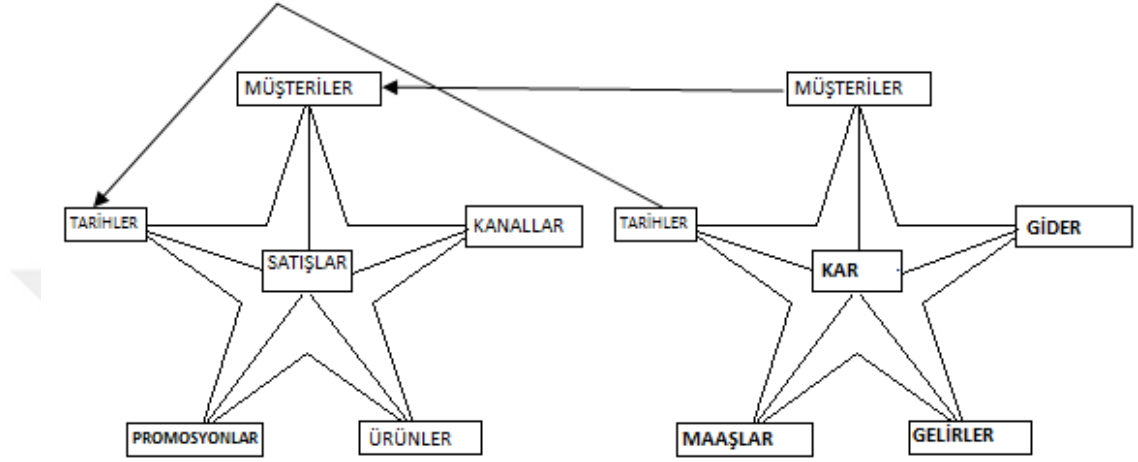
Kar tanesi şema türünde yıldız şema görünümünden oldukça farklıdır. Buradaki boyut tabloları, normalize edilen diğer boyut tablolarına bağlı kalmaktadır. Bu model yıldız şemanın geliştirilmiş şeklidir. Ana tabloların sahip olduğu her bir boyut tablosu, farklı işlemlere tabi tutulmuş boyut tablolarını da kapsamaktadır. Buradaki boyut tabloları bir çok özelliğe sahip iseler, normalize edilmek durumunda kalmaktadır. Yıldız şema bu özelliği desteklemediğinde kar tanesi şema tercih edilmektedir.



Şekil 2.3. Veri Ambarı Mimarisi Kar tanesi Şema.

Ana tabloların bir araya getirilmesiyle oluşan bu şema da ise mevcut durumdaki boyut tabloları ortak kullanılmaktadır ve görünümünde birkaç yıldız şeması iç içe geçmiş gibi görülmektedir. Bazı veri tabanlarının dizaynların da, boyut tablolarındaki veriler çok fazla yer kaplayarak farklılık gösterebilmektedir. Bu gibi durumlar söz konusu olduğunda, genel tasarım için yıldız şema ya da kar tanesi şema mimarisi kullanılamamaktadır. Her iki şemanın bileşimine ihtiyaç duyulmaktadır.

Böyle durumlarda tüm tasarımda ne yıldız ne de kar tanesi modeli kullanılamaz. Bu bileşimde karma model olarak adlandırılmaktadır.²⁴⁷



Şekil 2.4. Veri Ambarı Mimarisi Ana Tablolar Şema.

2.1.2.3. Veri Ambarının İçerdiği Veri

Veri ambarı, içerdiği veri açısından da göz önüne alındığında farklı bir yapıya sahip olduğu görülmektedir. Veri ambarının içerdiği veri

- Metadata
- Ayrıntı veri
- Eski ayrıntı veri
- Düşük düzeyde özetlenmiş veri
- Yüksek düzeyde özetlenmiş veri olarak sınıflandırılmaktadır.

Metadata, karar mekanizmalarıyla ilgili destek sistemlerinin uzmanlarına yardımcı olmak amacıyla meydana getirilen bir dizindir. Ayrıca veri ambarının içeriğinde neler bulunduğunu göstermektedir. Kullanılan verinin yapısını ortaya koyan bilgileri içermektedir. İşlemsel çevreden veri ambarına dönüştürülen verinin

²⁴⁷ Silahtaroglu, a.g.e., s. 25.

konumları hakkında bilgileri içeren bir kılavuzdur. İşlemsel ortamdan elde edilen verilerin, algoritmalarından hangisine göre düşük veya yüksek düzeyde özetlendiği hakkındaki bilgileri içeren bir kılavuzdur. Ayrıntı veri denilince, sadece şu andaki en son ayrıntı veri kastedilmemektedir. Bu ayrıntılar belirli bir dönemi kapsayabilir. Örneğin satışlara ilişkin veri son beş yılın ayrıntılarını içerebilir. Eski ayrıntı veri, şu andaki ayrıntı veriye göre daha düşük bir ayrıntı düzeyine indirilerek saklanmaktadır. Bu tür veriye çok sık biçimde erişilmediği için, disk üzerinde saklanabileceği gibi genellikle başka saklama birimlerine de yerleştirilebilmektedir. Veri ambarlarının düzenlenmesi sırasında hangi verilerin nasıl özetlenmesi gerektiği ve özetleme işinin hangi düzeyde gerçekleşeceği belirlenmelidir. Yüksek seviyede özetlenen veriler, ayrıntı veriden daha yüksek seviyede özetlenerek daha kolay ulaşılabilir hale getirilen verilerdir. Bu türdeki veri de veri ambarların bir unsuru olarak yer almaktadır.²⁴⁸

Metadata, direkt olarak işlemsel çevrede yer alan verileri içermemektedir. Karar Destek Mekanizmaları çalışmalarına yardım amacıyla meydana getirilen dizin olarak tanımlanmaktadır. Veri ambarlarına işlemsel alanlardan dönüştürülerek getirilen verilerin, yerleri ile ilgili bilgi vermektedir. İşlemsel alandan gelen verilerin algoritma çeşidine göre yüksek ya da düşük düzeyde özetlenmesiyle ilgili bilgi vermektedir.²⁴⁹

Ayrıntı Veri: Bu veri türü en son gerçekleşen olayları kapsamaktadır. Aynı zamanda hemen işlenemediği için diğer verilere oranla daha büyük hacimlidir.

Eski Ayrıntı Veri: Ayrıntı veri türünün haricinde kalan veriler olarak tanımlanmaktadır. Çok eski tarihlere ait verilerdir.

Düşük Düzeyde Özetlenmiş Veri: Ayrıntı veriden filtrelenerek meydana çıkarılan düşük düzeyde özetlenen verilerdir.

Yüksek Seviyede Özetlenmiş Veri: Ayrıntı verinin daha yüksek seviyede özetlenmesiyle kolay bir şekilde ulaşılabilir hale getirilmesiyle oluşturulan veridir.

²⁴⁸ Özkan, a.g.e., s. 29.

²⁴⁹ (Çevrimiçi), <http://ab.org.tr/ab13/bildiri/175.pdf>, 20.11.2016.

Veri ambarında yer alan bir başka veri tabanı da data marttır. Data mart pazarlama alanına yönlendirilmiş olan tematik veri tabanıdır. Pazarlama veri tabanının küçültülmesidir. Örnek olarak yeni ya da potansiyel müşterilerle iletişim kurmak için gerekli bilgileri içeren bir şirket arşivi düşünülebilir. Başka bir deyişle müşteri ilişkilerini yönetmeye odaklı bir veri tabanından bahsetmektedir. Data mart gibi tematik veri yapılarının meydana getirilmesi veri madenciliği faaliyeti için bilgi verici bir çevreye doğru ilk ve temel adımı temsil etmektedir.²⁵⁰

2.2. Veri Madenciliği Amaçları

Veri madenciliğinin amaçları aşağıdaki kategorilerden oluşmaktadır. Bunlar;

Veri işleme: VTBK süreçlerinin amaçları ve gereklerine bağlı olan analistle veriyi dönüştürebilmekte, temizleyebilmekte, deneyebilmekte, birleştirebilmekte ve seçebilmektedirler.

Tahmin etme: Tahmin edici bir model, belli bir veri biriminin spesifik bir özelliği için değer tahmin edebilmektir.

İlişki: Bir modelin otomatik tahmini ile aynı birimin aynı özellik değerleri üzerindeki başka özellik değerlerinin bağımsızlığının analiz edilmesidir. Yeni kayıtlar için bu özellik değerlerini tahmin edebilen bir modelin otomatik üretilmesidir.

Sınıflandırma: Önceden tanımlanmış kategorik sınıf kümesi veri birimlerine ait bu sınıfları belirlemektedir. Sınıflandırma kavramı, basit bir veri kümesiyle birlikte tanımlanan farklı sınıflar arasındaki verinin dağıtılmasıdır.

Kümeleme: Belli veri birimleri kümesinin, benzer özellikteki birimlerin katılımı ile birlikte gruplandığı sınıfların işleme alınmasıdır.

Bağlantı analizi: Bir modelin mevcudiyeti gibi, birimlerin başka bir modelin varlığını ifade etmesidir. Bu ilişkiler aynı veri birimi içindeki özellikler arasındaki birliktelikler olabilmektedir.

²⁵⁰Johannes Grabmeier, Andreas Rudolph, "Techniques of Cluster Algorithms in Data Mining", **Data Mining and Knowledge Discovery** 6.4., 2002, p. 350.

Model görselleştirme: Görselleştirme, keşfedilen bilginin insanlar tarafından anlaşılabilir ve yorumlanabilir hale gelmesinde önemli rol oynamaktadır.

Açıklayıcı veri analizi: Ön yargılı varsayım ve modellere bağımlılığı olmayan yani ilginç örüntüleri tanımlamayı sağlayan etkileşimli bir incelemedir.²⁵¹

Veri madenciliğinin görevleri altı kategori olarak ayrılmıştır. Bu kategoriler tanımlama, tahmin etme, öngörü, sınıflandırma, kümeleme ve ilişkilendirmedir.²⁵²

Tahmin etme, sayısal bir bağımlı değişkenin değerini bir ya da daha fazla bağımsız değişkene dayalı olarak tahmin etmeye çalışırken, tanımlama verideki modelleri ve eğilimleri tahmin etmeye çalışmaktadır. Tahmine benzer olarak, sınıflama teknikleri belli manipüle edilmiş değişkenlere dayalı bir bağımlı değişkenin değerini bulmak için kullanılmaktadır ama bu durumda nümerik değişkenlerdense kategorik değişkenler tercih edilmektedir. Dolandırıcılıkları saptama, sınıflamanın yaygın bir örneğidir. Öngöründe gelecekte olması gereken değerler bulunmaya çalışılmaktadır. Örnek olarak ise hisse senedi fiyatlarının gelecek aylardaki fiyatı ya da hangi takımın şampiyon olacağı ile ilgili öngöründe bulunmaktır. Kümelemede ise oluşumlar arasındaki benzerlikleri bulmak için çalışmalar yapılmaktadır ve bu oluşumlar gruplandırılmaktadır. Son olarak ilişkilendirmede ise belli oluşumların birkaç özelliği arasındaki ilişkiyi belirlemek amaçlanmaktadır.²⁵³

Veri madenciliğini etkin yönetmek amacıyla ilk olarak uygulanan sistemin hangi özelliklere sahip olması gerektiği incelenmelidir. Bununla birlikte veri madenciliği tekniklerinin gelişmesiyle hangi zorluklarla karşılaşılacağı da araştırılmalıdır.

1) Farklı uygulamalarda kullanılan birçok farklı veri çeşidi bulunduğundan bilgi keşfinden beklenen, sistemde yer alan farklı veri türleri durumunda veri madenciliğini etkin olarak yerine getirmesidir.

²⁵¹ Michael Goebel, Le Gruenwald, "A Survey of Data Mining and Knowledge Discovery Software Tools", **ACM SIGKDD Explorations Newsletter 1.1.**, 1999, p. 21.

²⁵² Daniel T. Larose, **Discovering Knowledge in Data: An Introduction to Data Mining**, John Wiley & Sons, 2014, p. 10.

²⁵³ Seyrek ve Ata, **a.g.e.**, s. 14.

2)Veri madenciliği tekniklerinin ölçülebilirliği ve etkinliği: Veri tabanlarındaki yüksek miktardaki veriden bilgiyi etkin olarak seçip çıkartmak amacıyla bilgi keşfi algoritmaları geniş veri tabanları için uygun, ölçülebilir ve etkin olmak zorundadır.

3) Veri madenciliği sonuçlarının anlamlılığı, kesinliği ve kullanışlılığı: Ortaya çıkarılan bilgi veri tabanı içeriğini tanımlamalıdır ve belli uygulamalar için kullanışlı olmalıdır.

4) Farklı veri madenciliği istek ve sonuç türlerinin ifade edilmesi: Farklı bilgi türleri, büyük miktarda veriden ortaya çıkarılmaktadır. Bununla birlikte ortaya çıkarılan bilgi farklı açılardan incelenmektedir ve farklı formlarda sunulmaktadır.

5) Çoklu soyutlama seviyelerinde çıkarılan interaktif bilgi: Bir veri tabanından tümüyle ortaya çıkarılacak olan bilgiyi, tahmin etmek zor olduğu için yüksek seviye veri madenciliği sorgusu ileriki araştırmalarda belli başlı ilginç sonuçları keşfedebilen bir araştırma olarak hareket etmelidir.

6) Farklı veri kaynaklarından bilgi çıkarmak: İnternetin de dâhil olduğu geniş alanlı bilgisayar ağları çoğu veri kaynağıyla iletişime geçebilir ve büyük miktarda dağıtılmış, heterojen veri tabanlarını düzenleyebilmektedir.

7) Gizlilik ve veri güvenliğinin korunması: Veri farklı soyutlama seviyelerinde ve çoğu farklı açıdan incelendiğinde, veri güvenliğini ve gizlilik gizliliğin ihlaline karşı koruma amacını tehdit etmektedir.²⁵⁴

2.3.Veri Madenciliği Uygulama Süreci

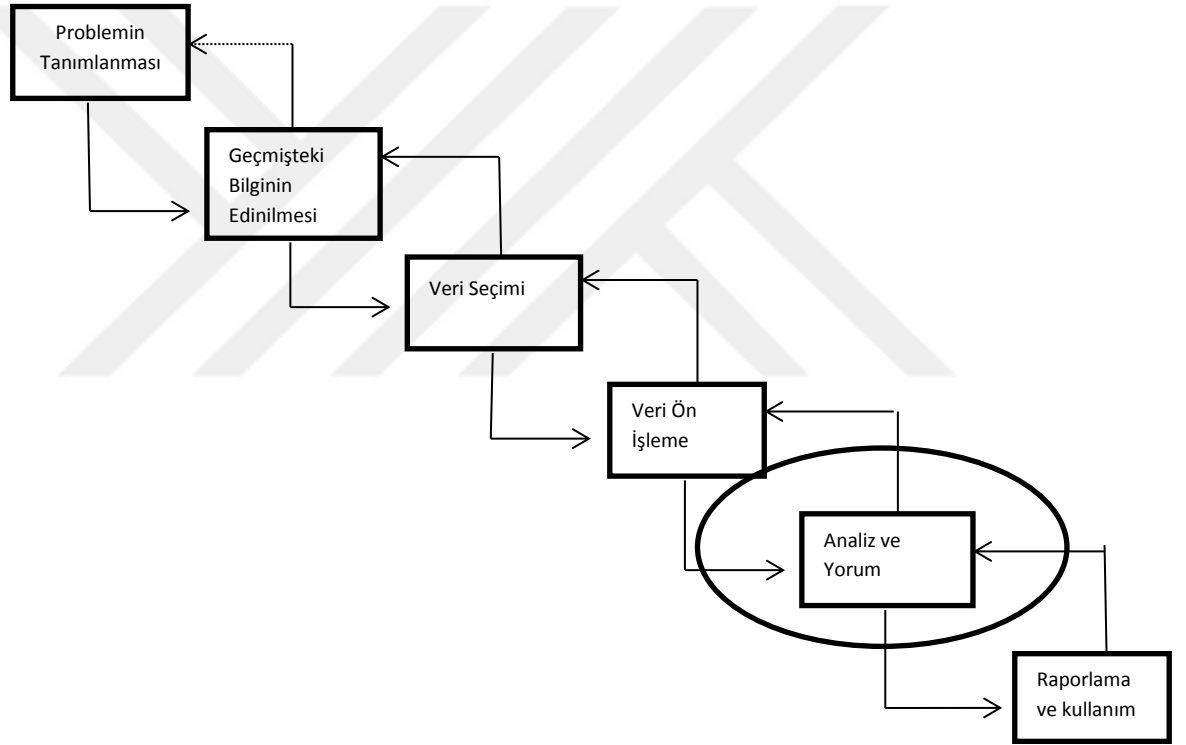
Veri madenciliği, veri yığınlarından önceleri bilinmeyen, gizli, büyük olasılıkla yararlı bilginin belli bir süreçte meydana getirilmesi faaliyeti olarak tanımlanmaktadır.²⁵⁵ Veri madenciliği veriden bilgiyi seçip çıkartmak için algoritma uygulamasını ve yapısını ele almaktadır. Veri madenciliği araştırmacı ve tekrarlı bir

²⁵⁴ Chen, Han ve Yu, a.g.e., p. 880.

²⁵⁵ G.Piatetsky-Shapiro, "The Data-Mining Industry Coming of Age", **IEEE Intelligent Systems** 14(6), 1999, p. 33.

süreçtir. Veri analizi boyunca, yeni bilgi keşfedilmektedir ve yeni hipotezler formüle edilmektedir. Süreç boyunca farklı aşamalar arasında hareket edilmektedir.²⁵⁶

Veri madenciliği bilgi keşfi sürecinin bir adımı olduğundan veri ile ilgili hazırlık süreçlerinden sonra uygulama yöntemi belirlenmektedir. Belirlenen yönetime göre uygun algoritmalarla analizler yapılmaktadır. Her veri madenciliği tekniğinin farklı algoritmaları bulunduğundan bu süreçte adımlar kullanılan tekniğe göre farklılık gösterebilmektedir.



Şekil 2.5. Veri Madenciliğindeki Önemli Adımlar.

Veri tabanlarında bilgi keşfi sürecinin bir aşaması olan veri madenciliğinin çekirdeği model kurmaktır. Modelin doğru bir şekilde kurulabilmesi amacıyla yapılacak projenin hedefinin iyice anlaşılması gerekmektedir. Her hedef ile alakalı birçok algoritma bulunmaktadır. Böyle bir durumda mevcut verilere göre uygun olan algoritmaların tümü çalıştırılmaktadır. Eldeki verilere uygun algoritma çalıştırılınca,

²⁵⁶ Ad J. Feelders, "Prior Knowledge in Economic Applications of Data Mining", In **European Conference on Principles of Data Mining and Knowledge Discovery**, Springer Berlin Heidelberg, 2000, p. 273.

en mantıklı sonucu vereni bulmak amacıyla değerlendirmelerde bulunmaktadır.²⁵⁷ Her veri madenciliği modelinin belli bir hayat çevrimi bulunmaktadır. Bazı çözümlemelerde özellikler statik özellik taşımaktadır ve modelin tekrardan düzenlenmesine ihtiyaç duyulmamaktadır. Ancak çoğu işin özellikleri sıklıkla değişim göstermektedir. Yeni ve farklı verilerin dâhil edilmesiyle mevcut modelin tekrar eğitilmesine ihtiyaç duymaktadır. Dolayısıyla modelin kurulmasından sonra eğer ki sıklıkla veri tabanına veri ekleniyorsa modeli oldukça sık güncellemek gerekmektedir.²⁵⁸

Veri madenciliğinde aşağıdaki adımlar takip edilmektedir.

- 1)Uygulama anlayışını geliştirmek
- 2)Keşfetmek için kullanılan hedef veri setini oluşturmak
- 3)Veri ön işlemlerini yapmak, veriyi temizlemek
- 4)Eğer mümkünse verinin sabit gösterimlerini bulmak ve değişken sayılarını azaltmak
- 5)Veri madenciliği tekniğini seçmek
- 6)Veri madenciliği algoritmasını seçmek
- 7)İlgili modeli araştırmak/inceleme
- 8)Çıkarılan modeli yorumlamak. Eğer gerekliyse 1'den 7'ye kadar olan adımları tekrarlamak
- 9)Elde edilen bilgiyi birleştirmek ve bir rapor hazırlamak.

İlk iki adım veri analizinin bütün formlarına aittir. Bu adımlar ilgili ilk bilginin meydana geldiğini, proje amaçlarını ve verinin oluşmasını kapsamaktadır. Eğer veri ambarı hazır, veri ambarının metadatası verinin meydana geldiği yeri tam

²⁵⁷ Adem Tekerek, "Veri Madenciliği Süreçleri ve Açık Kaynak Kodlu Veri Madenciliği Araçları", **Akademik Bilişim 11**, 2011, s. 3.

²⁵⁸ Zhao Hui Tang, Jamine MacLennan, **Data Mining with SQL Server 2005**, John Wiley&Sons, 2005, p. 180.

olarak saptamaya yardım etmektedir. 3. adım ambarlamak için verinin hazırlanması aşamasıdır.

Örgütsel veri, ambarlamaya konulmadan önce temizlenir ve onaylanır. Bilgi keşfi için ön işleme verideki gürültüyü ve aykırılıkları kaldırmak için kullanılmaktadır. 4. ve 5. adımlar veri sınıflama tekniğini seçmeyi ve değişkenlerin sayısını azaltmayı kapsamaktadır. Çok fazla değişkenle birlikte sonuçları anlamlandırmak zor olmaktadır. Bu nedenle dönüşüm teknikleri uygulanmaktadır ve sabit olan veri simgelerin bulunması amacıyla denemeler yapmaktadır.

6. ve 7. adımlar VTBK sürecinin kalbidir. Seçilen algoritmalar gerçekleştirilecek olan tekniğe bağlıdır. Gerçek madenciliğin yapıldığı adımlardır. Uygun modelleri kapsamaktadır. Ayrıca veri ambarına uygun modelleri belirlemeyi de kapsamaktadır. Modeller sınıflandırma kurallarını, regresyon fonksiyonlarını yansıtmaktadır. Özel modeller çeşitli alanlardan gelmektedir. Bu alanlar istatistik veri analizi, yapay sinir ağları, uzman sistemler, bulanık mantık, akıllı üstlenici, çok boyutlu analiz, veri görselleştirme ve karar ağaçlarıdır.

8. adım genel işleyişe beşeri geri getirmektedir. Adım 7'nin veri madenciliği sonuçların içsel olarak tutarlı ve makul olup olmadığını yargılayan bir analist tarafından incelenmesi gerekmektedir. Analist sonuçlarla tatmin olmazsa, düzeltilmiş sorguları ve durumları birlikte yeniden çalıştırabilmektedir. Yani süreç iteratiftir. İlk olarak analist çıktıyla tatmin olmalıdır ve sonuçları yorumlamalıdır.

9. adımda bulgular raporlanmaktadır. Bu adımlar karar alıcılara daha anlaşılabilir bir formda sonuçları sunmayı kapsamaktadır. Grafiksel olarak modelleri yansıtmak, modellere dayalı yapılanma kuralları, önceki sonuçlarla birlikte fazlalıkları ve çelişkileri tekrar çözmek, doğal dildeki bulguları açıklamak gibi yaklaşımların hepsi bu aşama kapsamındadır. Son olarak bulgulara dayalı fiiller gerçekleştirilmektedir.²⁵⁹

²⁵⁹H. Michael Chung, Paul Gray, "Special Section: Data Mining", **Journal of Management Information Systems** 16.1, 1999, p. 15.

Veri madenciliği konu alanı uzmanlığı, geniş veri tabanları için veri madenciliği algoritmalarının uygulama tecrübesini gerektirmektedir. Bu bağlamda veri madenciliği adımları aşağıdaki maddelerden meydana gelmektedir.

- **Verinin Temizlenmesi:** Veri tabanlarında bulunan tutarlı olmayan ve yanlış veriler gürültü olarak adlandırılmaktadır. Verilerdeki gürültünün temizlenmesi amacıyla; eksik değerli kayıtlar kaldırılmaktadır, kaybolan değerlerin yerlerine ise sabit değerler verilebilmektedir, kayıp verilerin yerine diğer verilerin ortalamaları da alınıp yazılabilmektedir. Son olarak ise uygun bir tahmin yöntemiyle de eksik veriler tamamlanabilmektedir.
- **Verinin Bütünleştirilmesi:** Çeşitli veri kaynaklarıyla veya veri tabanlarıyla ulaşılan verinin tümünün bir arada değerlendirilebilmesi amacıyla farklı çeşitlerdeki verinin tek bir çeşide döndürülmesi işlemi olarak ifade edilmektedir. Bu durumun en yaygın örneğiyle cinsiyet değişkeninde karşılaşılmaktadır. Birçok çeşitte bulundurulabilen bir veri olup herhangi bir veri tabanında 0/1 diye kaydedilirken diğer veri tabanında E/K veya Erkek/Kadın diye kaydedilebilmektedir. Bilginin keşfedilmesinde ki başarı, verilerin uyumlu olmasına bağlı olmaktadır.
- **Verinin İndirgenmesi:** Uygulanacak olan veri madenciliği çözümlemesinden elde edilen sonucun değiştirilemeyeceği düşünülüyorsa, verilerin miktarı ya da değişken sayısı azaltılabilmektedir. Verinin indirgenebilmesi için kullanılan teknikler, veriyi sıkıştırma, örnekleme yapma, genelleme, birleştirme veya veri küpüyle boyut indirgeme olarak gösterilmektedir.
- **Verinin Dönüştürülmesi:** Verilerin seçilmiş modele uygun bir şekilde içeriğinin de aynı kalması sağlanarak dönüştürülmesidir. Bu işlem uygulanacak modele göre şekillendirilmektedir. Kullanılacak olan değişkenlerin ortalamaları ya da varyansları birbirlerinden oldukça farklıysa en büyük ortalama ve varyansa sahip olan değişkenlerin diğer değişkenler üzerinde etkisi görülebilmektedir ve bu durum değişkenlerin görevlerini de negatif yönde etkileyebilmektedir. Bu nedenle veriler üzerinde normalizasyon yapılmaktadır.
- **Veri Madenciliği Algoritmasının Uygulanması:** Veriler uygun hale getirilmekte ve daha sonra verilere uygun veri madenciliği algoritması uygulanmaktadır.

- Sonuçların Sunulması ve Değerlendirilmesi: Algoritma uygulanırken sonuçlar elde edildikten sonra söz konusu sonuçlar düzene koyularak, ilgili birimlere sunulmaktadır. Örnek olarak hiyerarşik kümeleme tekniği uygulandığında sonuçların dendogramla sunulmasıdır.²⁶⁰

2.4. Veri Madenciliği Modelleri

Veri madenciliği faydalı bilgi sağlayan önceden bilinmeyen ilişkileri ortaya çıkarmak için veri analizi olarak tanımlanmaktadır. Veri madenciliği temel olarak veri analiz ederek çalışmaktadır. Sorunları çözmeye yardım eden tanımlayıcı ve tahmin edici modellerden oluşmaktadır.²⁶¹

2.4.1. Sınıflandırma

Veri madenciliği içerisinde iki ana gruptan biri olan sınıflandırma, veri dizisinin istatistik ve/veya makine öğrenimi yöntemleri kullanılarak önceden belirlenen sınıflara atanması işlemidir. Sınıflama yöntemleri makine öğrenimi terminolojisinde denetimli öğrenme (supervised learning) başlığı altında toplanmıştır. Veri madenciliği çok büyük yığınlardaki detayların dışında bilinmeyen yapıları ve ilişkileri ortaya çıkartmak için çok büyük veri setlerini analiz etmeyi sağlayan tüm yöntem ve teknikleri kapsamaktadır. Elde edilen bilgi filtrelenmekte, hazırlanmakta ve aynı zamanda faydalı kararlar ve stratejiler için sınıflandırılmaktadır.²⁶²

Gerek istatistik gerekse makine öğrenimi temelli çeşitli sınıflandırma yöntemleri geliştirilmiştir. İstatistik yöntemler içerisinde lineer ve lineer olmayan başlıca yöntemler şunlardır: Lineer regresyon analizi, lojistik regresyon analizi, diskriminant analizi, bayes sınıflandırma yöntemleridir. Makine öğrenimi yöntemleri içerisinde ise karar ağaçları, en yakın komşu yöntemi, yapay sinir ağları ve destek vektör makineleri başlıca yöntemlerdir.²⁶³

²⁶⁰ (Çevrimiçi), <http://ab.org.tr/ab13/bildiri/175.pdf>, 20.11.2016.

²⁶¹ Hand, **a.g.e.**, p. 1.

²⁶² Grabmeier ve Rudolph, **a.g.e.**, p. 320.

²⁶³ Haldun Akpınar, Data Veri Madenciliği-Veri Analizi, İstanbul, Papatya Yayıncılık, 2014, s.77.

Sınıflandırma veri madenciliğinin en çok kullanılan uygulamalarından bir tanesidir. Her gün günlük yaşamda sık sık meydana gelen bir işe benzemektedir. Örneğin bir hastane tıbbi hastalardan belli bir hastalığa yakalanmada yüksek, düşük ve orta risk grubunda olanları sınıflandırabilmektedir. Seçim anket firması, herhangi bir siyasi partiye oy vermeye niyeti olan ve kararsız olan insanlarla mülakat yapabilmektedirler ve bu insanları sınıflandırabilmektedirler. Sınıflandırma bir tahmin çeşididir. Tahmin edilen değerlerin bir etiket olması sınıflandırma olarak adlandırılırken, sayısal tahmin ise regresyon olarak adlandırılmaktadır.²⁶⁴

Sınıflandırma algoritmaları, elde edilen eğitim kümesinden belli bir dağılım şekli öğrenmektedirler ve daha sonra sınıfı belli olmayan test verileri getirildiğinde doğru olarak sınıflara ayırmaya çalışmaktadırlar. Sınıflandırma, en genel tanımıyla 2 evreden oluşmaktadır:

- Eğitim verileriyle öğrenme
- Öğrenilen değerlerle test verileri üzerinde sınıflandırmadır.²⁶⁵

2.4.2. Kümeleme

Küme analizi nesnelerin belirlenen özniteliklerine ve dolayısı ile öznitelik değerlerine göre alt gruplara, kümelere ayrılabilmesi için istatistik ve makine öğrenimi alanlarında geliştirilen çeşitli yöntem ve süreçlerdir. Algılama ve öğrenmenin temel ögesi, nesnelerin benzerliklerine göre sınıflandırmasıdır. Makine öğreniminde denetimsiz öğrenmenin temel araçlarından biri olan küme analizi, nesnelerin benzerlik ilişkilerine göre gruplandırması ile insan beyninin tipik bir akıl yürütme işlevini taklit etmektedir. Bu bakış açısına göre küme analizi gizli örüntülerin denetimsiz öğrenme yaklaşımı ile aranmasıdır. Örneğin bitkilerin hayvanlardan, köpeklerin kedilerden ayırt edilmesi çocukluğun erken dönemlerinde geliştirilen bilinçaltı öğrenme mekanizmaları ile gerçekleştirilen kümeleme süreçleridir.²⁶⁶

²⁶⁴Max Bramer, **Principles of Data Mining**, Springer, London, 2007, s. 6.

²⁶⁵Şadi Evren Şeker, **İş Zekası ve Veri Madenciliği**, 2013, s. 77.

²⁶⁶Akpınar, **a.g.e.**, s.56.

Kümeleme problemleri, genel olarak çıkarılmış olan özellikler kullanılarak bir veri setindeki değerlerin farklı kümeler arasında paylaşılmasıdır. Sınıflandırma ile arasındaki fark, bir veri kümesindeki verilerin arasında ilişki olup olmadığının sorgulanmasıdır.²⁶⁷

Kümeleme analizi çoğu disiplinle alakalı bir uygulamadır. Bu disiplinlerin en başında istatistik, bilgisayar, matematik gelmektedir. Bilgisayar alanında kümeleme analiziyle alakalı çalışmalara ses, resim tanıma ve makine öğrenmesi örnek olarak verilebilir. Kümeleme analizi istatistik alanında çok değişkenli tahminlerde ve örüntü tanıma konularında yardım sağlamaktadır.²⁶⁸ Veri kümeleme, seyrek ve kalabalık alanları tanımlamakta ve veri kümesinin dağılım modellerini keşfetmektedir. Bununla birlikte türeyen kümeler orijinal veri kümesinden daha etkin ve etkili görselleştirebilmektedir.²⁶⁹

Kümeleme benzer amaçlara göre verilerin gruplara ayrılmasıdır. Küme diye adlandırılan her bir grup, diğer grupların birimlerine benzemeyen ancak birbirine benzeyen birimlerden oluşmaktadır. Makine öğrenmesi açısından kümeler gizli örüntülerle uyumlu olmaktadır. Kümelerin araştırılması denetimsiz öğrenmedir. Bununla birlikte neticelenen sistem bir veri anlayışını temsil etmektedir. Kümeleme bir sıklık tahmin problemi olarak görülebilmektedir. Bu da geleneksel çok değişkenli istatistiksel tahminlerin konusu olmaktadır. Kümeleme hem de görüntü işlemedeki veri sıkıştırma içinde oldukça geniş bir şekilde kullanılmaktadır.²⁷⁰

2.4.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler

Birliktelik kuralları belli başlı türlerdeki veri bağlantılarını açıklayan bir model olarak tanımlanmaktadır. Herhangi bir ürünün satın alınması durumunda bu ürün yanında başka bir ürününde satın alınmasıyla bir birliktelik kuralı oluşturulabilmektedir. Bu açıdan tanımlayıcı özelliklere sahip bir model olduğu

²⁶⁷ Şeker, a.g.e., s. 82.

²⁶⁸ P. Arthur Dempster, Nan M. Laird, Donald B. Rubin, "Maximum Likelihood from Incomplete Data Via the EM Algorithm", **Journal of The Royal Statistical Society Series B (Methodological)**, 1977, p. 35.

²⁶⁹ Tian Zhang, Raghu Ramakrishnan, Miron Livny, "BIRCH: An Efficient Data Clustering Method for Very Large Databases", In: **ACM Sigmod Record**, 1996, p. 5.

²⁷⁰ Pavel Berkhin, **A Survey of Clustering Data Mining Techniques**, In: **Grouping Multidimensional Data**, Springer Berlin, 2006, p. 30.

görülmektedir. Belli bir ürün ve bu ürünle birlikte başka ürünlerinde alınması durumunda birliktelik kuralları gündeme gelmektedir. Dolayısıyla perakendecilik sektöründe faaliyet gösteren firmalarda daha fazla uygulanmaktadır. Örneğin bir süpermarkette yapılan alışverişin incelenip hangi ürünün hangi ürün ile satın alındığının belirlenmesi birliktelik kurallarını ilgilendirmektedir. Bunun dışında örneğin iletişim ağlarında meydana gelen hataların belirlenmesinde de kullanılabilir. ²⁷¹

Birliktelik kuralları bilgisayar bilimi alanında geliştirilmiştir ve pazar sepet analizi gibi önemli uygulamalarda kullanılmaktadır. Belli bir müşteri tarafından satın alınan ürünler arasındaki birliktelikleri ölçmek için kullanılmaktadır. Bir web kullanıcısı tarafından sıralı olarak bakılan sayfalar arasındaki birliktelikleri ölçmek için kullanılmaktadır. Amaç bütün işlemler kümesindeki birlikte meydana gelen eleman gruplarını vurgulamaktır. Veri tabanı her bir işlem için gerçekleşen öge listesini içermektedir. Her bir bireyin veri kümesinde bir defadan daha fazla görülüp görülmediğine dikkat edilmelidir. Birliktelik kuralları özel model tipleri arasındaki kuralları dikkate almaktadır. Bu model tipleri öge kümeleri olarak adlandırılmaktadır. Öge kümesindeki her bir değişken çift değerlidir. Yani özel durum o değişken için doğruysa 1 değeri verilir yanlışsa 0 değeri verilmektedir. Birliktelik kuralları değişkenler arasındaki ilişkileri belirlemek için kullanılan en yaygın metottur. Oldukça kompleks ve dinamik olan küresel bir analiz için çok geniş veri kümelerini araştırıp çıkartma amacıyla kullanılmaktadır. ²⁷²

Birliktelik kuralları öge kümeleri arasında meydana gelen bağımlılıklar üzerinde faydalı bilgi sunan örüntülerdir. Apriori gibi mevcut birliktelik kuralları oldukça büyük kurallar kümesini ortaya çıkarmaktadır. Bu kuralları anlamlandırmak için faydalı örüntüleri vurgulayan belli biçimdeki grup kurallarına ihtiyaç duymaktadır. ²⁷³

²⁷¹ Margaret H. Dunham, **Data Mining-Introductory and Advanced Topics**, Person Education, 2003, p. 8.

²⁷² Grabmeier ve Rudolph, **a.g.e.**, p. 332.

²⁷³ Koh, Tan ve Peng, **a.g.e.**, p. 182.

Birliktelik kuralları ile oluşturulan modeller aşağıdaki örneklerden de anlaşıldığı gibi aynı zamanda meydana gelen bağlantıların açıklanmasında kullanılmaktadır. Aşağıda sunulan örneklerde görüldüğü gibi eş zamanlı olarak gerçekleşen ilişkilerin tanımlanmasında kullanılır.

- Tüketiciler kola satın alırken, % 90 olasılıkla çekirdekte satın almaktadır.
- Az yağlı peynir ve diyet yoğurt alan tüketiciler, % 80 olasılıkla diyet bisküvi de satın almaktadır.²⁷⁴

Birliktelik kuralları aslında olaylar arasında gerçekleşen probabilistik ilişkiyi açıklamaktadır. Gerçekleşen olaylar arasındaki bağlantı ise sıklıkla bir arada gözlemlenen olaylardır. Veri tabanlarının ilk adımı veri tabanındaki birliktelik durumlarının açığa çıkarılmasıdır. Veri tabanındaki herhangi bir X'in aynı zamanda Y'yi içermesi bir birlikteliktir. Bu durum ile alakalı yaygın bir örnek olarak : "Mama içeren %70 alışverişin, %10'u aynı zamanda çocuk bezi de içermektedir." Burada %70 güven düzeyini %10 ise bu güven seviyesinde bulunan desteği işaret etmektedir.²⁷⁵

Birliktelik kuralları için veri madenciliği 2 adımda gerçekleşmektedir. İlk adım bütün sık öge kümelerini bulmakla meydana gelmektedir. Bu öge kümeleri asgari destek olarak adlandırılan ve veri tabanlarında yer alan kullanıcıya ait belli ve özellikli sıklıktan oluşmaktadır. İkinci adım sık öge kümeleri arasındaki kuralları düzenlemekle meydana gelmektedir.²⁷⁶

Örüntü tanıma, herhangi bir yazılı metni tanımak ya da o metnin çok benzerini bulmak örüntü tanımanın konusuna girer. Bunun dışında parmak izi, ses, yüz tanıma, kan hücrelerinin kıyaslanması, el yazılarının tespit edilmesi gibi alanlarda da uygulanmaktadır. Dolayısı ile örüntüden kasıt el, yüz, resim, çizim ve ses gibi varlıkların sayısal ortamda sergiledikleri şekillerdir. Aslında örüntü tanımada yapılan işlem bir çeşit sınıflandırmadır. Elimizde, ulaşılmaması gereken bir örnek vardır ve biz bu örneğin benzerini veya aynısını aramaktayız. Bir algoritmayla bu örneğe

²⁷⁴ Akpınar, a.g.e., s.46

²⁷⁵ Silahtaroglu, a.g.e., s. 50.

²⁷⁶ Zaki vd., "New Algorithms for Fast Discovery of Association Rules", **KDD Proceedings Vol. 97**, 1997, p. 284.

benzeyenleri bir araya topluyor, daha sonra bunları en çok benzerden, en az benzere doğru sıralıyoruz. Örüntü tanıma amacıyla geliştirilen birden fazla algoritma bulunmaktadır. "K-en yakın komşu algoritması", "doğrusal sınıflayıcı", "üstel sınıflayıcı" bunlardan sadece birkaç tanesidir.²⁷⁷

Ardışık zaman örüntüleri, veri madenciliği ve perakendecilik sektöründen gelen yoğun ilgiyle gelişimini sürdürmektedir. Bir dizi olayın gelmesiyle birlikte o olaydaki alt dizilerden oluşan örüntüler ardışık zamanlı olarak isimlendirilmektedir.²⁷⁸ Barkod teknolojisindeki gelişme, perakendecilik sektöründe faaliyette bulunan firmaların yüksek hacimlerde ve sıklıkla denilebilecek aralıklarla meydana gelen bilgilerin, hızla stoklanmasına imkan sağlamıştır. Satış miktarlarını gösteren bu verilere sepet verisi denilebilmektedir.²⁷⁹ Bu veriler genel anlamda işlem tarihi, satılmış olan malları, işlem sıra numarası gibi bilgileri kapsamaktadır. İşlemin gerçekleştiği yerde ki burası tipik olarak bir süpermarkettir; üyelik kartları ve kulüp kartları diyebileceğimiz kartlar kullanılmaktaysa, bu bilgilerin içinde müşterilerin numaraları da bulunabilmektedir. Bu kayıtların incelenmesiyle, ardışık zaman örüntüleri tespit edilebilmektedir. Eğer müşterilerin bir kısmı, önce bir CD oynatıcısı, daha sonraki bir zamanda ABC isimindeki bir filmi, ardından XYZ isimindeki bir filmi satın alıyorsa bu bir ardışık örüntü oluşturacaktır. Başka müşteriler ABC filmi ile XYZ arasındaki zaman diliminde, farklı filmler ya da farklı ürünler alsalar bile onlarda bu örüntüye dâhildirler.²⁸⁰

2.4.3.1. Apriori Algoritması

Apriori algoritması, büyük ölçekteki veri tabanlarının üzerinde yer alan veri madenciliği çözümlmeleri için geliştirilmiştir. Genellikle ilişki kurallarının çıkarımlarında kullanılan algoritmadır. Algoritmanın hedefi veri tabanlarında yer alan satır değerleri arasındaki ilişkiyi meydana çıkarmaktır. Algoritmanın adında yer alan prior kelimesi latince de önce anlamına gelmektedir. Algoritma kendinden

²⁷⁷ Silahtaroglu, a.g.e., s. 54.

²⁷⁸ Marek Wojciechowski, "Interactive Constraint-Based Sequential Pattern Mining", **East European Conference on Advances in Databases and Information Systems**, Springer Berlin, 2001, p. 2.

²⁷⁹ Rakesh Agrawal, Ramakrishnan Srikant, "Fast Algorithms for Mining Association Rules", **Proc. 20th Int. Conf. VeryLarge Data Bases, VLDB. Vol. 1215**, 1994, p. 490.

²⁸⁰ Silahtaroglu, a.g.e., s. 55.

önceki çıkarımlarla bağlantılı olduğu için bu isim verilmiştir. Algoritma yapısal açıdan aşağıdan yukarıya yaklaşımına sahip olduğu için her defasında bir eleman araştırmakta ve bu elemanların diğer elemanlarla ilgisini ortaya koymaya çalışmaktadır.²⁸¹

Tipik apriori algoritması apriori k elemanlı kümelerin $k + 1$ elemanlı kümeleri ortaya çıkarması için kullanılan bir araştırma olarak bilinmektedir ve iteratif bir yaklaşımla çalışmaktadır. İlk olarak, sık sık 1 elemanlı kümelerin sıklık grupları minimum destek sağlayan bu elemanların biriktirilmesi ve her bir eleman için sayının toplanması amacıyla veri tabanı taranarak elde edilmektedir. Sonuçlanan küme L_1 diye belirtilmektedir. Daha sonra L_1, L_2 'yi bulmak için kullanılmaktadır. L_2 ise iki elemanlı kümelerin sıklık grubudur. L_2 ise L_3 'ü bulmak için kullanılmaktadır. Bu şekilde k elemanlı kümelerin sıklığı bulunabilmektedir. Her bir L_k sonucu veri tabanının tam olarak taranmasını gerektirmektedir.²⁸²

Bu algoritma aşağıda belirtilen aşamalara sahiptir:

- a) Birliktelik çözümlemesini gerçekleştirmek için ilk olarak destek ve güven kıstaslarını karşılaştırmak amacıyla sınır değerler belirlenmektedir. Uygulamada bulunan sonuçların bu sınır değerlerine eşit veya büyük olması beklenmektedir.
- b) Veri tabanı taratılarak çözümlemeye katılacak ürünlerin her biri için tekrarlanmış sayılar hesaplanmaktadır. Bu sayılara destek sayıları denmektedir. Bu tekrarlanan sayılar eşik destek sayıları ile kıyaslanmaktadır. Bu sayılardan küçük olan değerleri bulunan satırlar çözümlemenin içinden çıkarılmakta ve şartlara uygun olan kayıtlar ele alınmaktadır.
- c) Üstteki aşamada ele alınan ürünler bu defa ikili şekilde gruplandırılmaktadır. Böylece bu gruplara ait tekrar sayıları elde edilmektedir. Elde edilen bu sayılarda eşik destek sayıları ile kıyaslanmaktadır. Eşik destek sayılarından küçük değerleri olan satırlar çözümlemenin içerisinden çıkarılmaktadır.

²⁸¹ Şeker, a.g.e., s. 86.

²⁸² Jaishree Singh, Hari Ram, JS Sodhi, "Improving Efficiency of Apriori Algorithm Using Transaction Reduction", **International Journal of Scientific and Research Publications** 3.1, 2013, s. 2.

d) Bu defa üçlü, dörtlü vb. gruplandırmalarla bu gruplara ait destek sayılarının elde edilmesiyle ve eşik değerleriyle kıyaslanmaktadır. Eşik değerlerine uygun görüldüğü sürece işlemler devam etmektedir.

e) Ürün grupları tespit edildikten sonra destek ölçütü göz önüne alınarak birliktelik kuralları bitirilmektedir. Bu kurallara ait her bir kural ile ilgili olan güven kıstasları tespit edilmektedir.²⁸³

2.5. Veri Madenciliği Teknikleri

Veri tabanlarının sayısı ve büyüklüğündeki gelişme, bu verileri analiz etmekte birçok yöntemin yetersiz kalmasını sağlamaktadır. Böylece veriden bilgiyi çekip çıkarmak için bir ihtiyaç ve fırsat oluşmaktadır. Bilgi keşfi veriden, önceden bilinmeyen potansiyel olarak faydalı, kesin bilginin alınmasıdır. Son yıllarda veri madenciliği hem veri tabanı hem de makine öğrenme araştırmacıları tarafından 1990'lardan beri en fazla gelecek vadeden araştırma konuları içinde yer almaktadır ve birçok yöntemi bulunmaktadır.²⁸⁴

Veri madenciliği süreci kümeleme analizi, veri özetleme, sınıflandırma kurallarını öğrenme, bağımlılık ağlarını bulma, değişikliklerin analiz edilmesi ve anomalileri tespit etme gibi farklı birçok yöntemi kapsamaktadır.²⁸⁵

Veri madenciliği modelleri işlev görme durumlarına göre ,

- Sınıflama (Classification) ve Regresyon (Regression),
- Kümeleme (Clustering),
- Birliktelik Kuralları (Association Rules) ve Ardışık Zamanlı Örüntüler (Sequential Patterns) olarak sınıflandırılmaktadır.²⁸⁶

Veri madenciliği araçları veri almakta ve bir model formunda realitenin temsilini meydana getirmektedir. Sonuçlanan model verideki mevcut örüntü ve

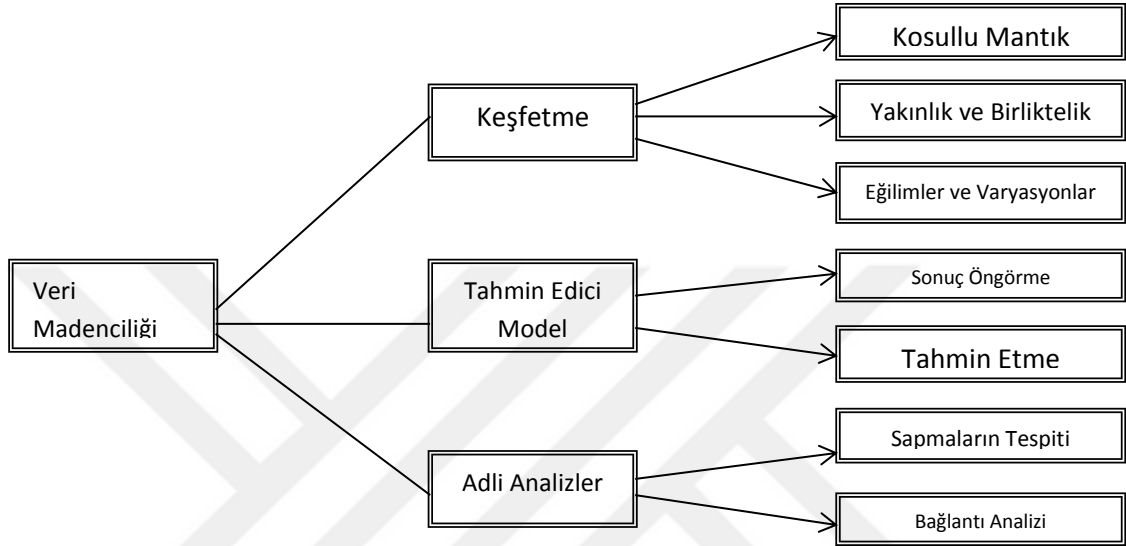
²⁸³ Özkan, a.g.e., s. 159.

²⁸⁴ Jiawei Han, Yandong Cai, Nick Cercone, "Knowledge Discovery in Databases: An Attribute-Oriented Approach", In: VLDB, 1992, p. 1.

²⁸⁵ Ruth Dilly, Data Mining: An Introduction, Student Notes, 1995, p. 2.

²⁸⁶ Akpınar, a.g.e., s.5

ilişkileri tanımlamaktadır. Veri madenciliği aktiviteleri üç genel kategoriye ayrılmaktadır.



Şekil 2.6. Süreç Eğiliminden Veri Madenciliğinin Analizi.

Keşfetme: Önceden tahmin edilmiş bir fikir ya da model olabilecek bir hipotez dışında gizli modelleri bulmak için veri tabanını arama sürecidir.

Tahmin edici modelleme: Veri tabanından keşfedilen modelleri oluşturma ve geleceği tahmin etmek için kullanma sürecidir.

Adli analiz: Anomalileri bulmak için ortaya çıkarılan örüntüleri uygulama sürecidir.

Veri madenciliği firma problemlerini çözmede altı tip model kullanmaktadır. Bu modeller sınıflama, regresyon, zaman serileri, kümeleme, birliktelik analizi ve sırlamadır.²⁸⁷ Sınıflama ve regresyon modelleri tahmin etmek için kullanılırken, birliktelik ve sıralama modelleri davranış tanımlamak için kullanılmaktadır.

²⁸⁷Herb Edelstein, "Data Mining: Exploiting the Hidden Trends in Your Data", **DB2 Online Magazine**, (Çevirimiçi) <http://www.psy.gla.ac.uk/~steve/pr/edel.html>, 27.11.2016

Kümeleme modelleri ise tahmin etmek için ya da tanımlama yapmak için kullanılmaktadır.²⁸⁸

VTBK'nın veri madenciliği aşaması belirli yöntemleri tekrarlayan iteratif bir uygulamadan meydana gelmektedir. Bilgi keşfinin amaçları sistemin kullanım amaçlarına göre tanımlanmaktadır. Bu amaçlar keşfetme ve doğrulama olarak ikiye ayrılmaktadır. Doğrulama amacı sistemin kullanılmasıyla ve hipotezin doğrulanmasıyla sınırlıdır. Keşfetme amacıyla ise sistem otomatik olarak yeni modeller bulmaktadır. Bu amacın tahmin etme ve tanımlama olarak iki alt amacı bulunmaktadır. Tahmin etme amacıyla sistem belirli birimlerin gelecekteki davranışlarını keşfetmek için modeller bulmaktadır. Tanımlamada ise insani olarak anlaşılabilir bir formda bir kullanıcıya tanıtım için modeller bulunmaktadır. Tahmin etme ve tanımlama amaçlarına belli veri madenciliği teknikleri uygulanarak ulaşılabilmektedir. Bu metotlar sınıflama, regresyon, kümeleme yöntemleridir. Sınıflama önceden belirlenmiş birkaç sınıfın birine veri birimini eşleyen fonksiyondur. Finansal pazarlardaki eğilimlerin sınıflandırılması bu yöntemle örnek olarak gösterilebilmektedir. Regresyon ise gerçek değerli tahmin değişkenine bir veri birimini eşleyen fonksiyondur. Reklam gider fonksiyonu olarak yeni bir ürün için müşteri talebini tahmin etmek örnek olarak gösterilebilmektedir. Kümeleme yöntemi ise veriyi ve sonlu bir kategori setini tanımlamak için araştırma yapan yaygın bir gruplama işi olarak açıklanmaktadır. Pazarlama veri tabanlarında homojen alt popülasyonu keşfetmek örnek olarak gösterilebilir.²⁸⁹ Özetlemede ise veri, alt küme için kısa bir tanım bulma amacını taşıyan yöntemleri kapsamaktadır. Bağımlı modellemede ise değişkenler arasındaki önemli bağımlılıkları tanımlamak amacıyla bir model kurulmaktadır. Değişimlerin ve saptamaların belirlenmesi ise önceden ölçülmüş değerlerden verideki en önemli değişimleri keşfetmeye odaklanmaktadır.²⁹⁰

Denetimli ve denetimsiz ifadeleri birbirilerinin zıttı anlamları taşımaktadırlar. Denetimsiz teknikler daha fazla verinin anlaşılması, tanınması, keşfedilmesi, amacıyla kullanılmaktadır. Bir sonraki uygulanması gereken teknikler için fikir

²⁸⁸ Rygielski, Wang ve Yen, **a.g.e.**, p. 488.

²⁸⁹ Fayyad, Shapiro ve Smyth, **a.g.e.**, p. 43.

²⁹⁰ **A.e.**, p. 43.

vermeyi hedeflemektedir. Denetimli teknikler de ise verilerden faydalı bilgilerin ve sonuçların elde edilmesi amaçlanmaktadır. Denetimsiz bir teknikle ulaşılan bilginin ya da sonucun mümkün olduğu kadar denetimli bir teknikle doğrulanması eldeki bulguların doğrulukları ve geçerlilikleri yönünden çok önemli görülmektedir. Kümeleme analizi denetimli ve denetimsiz tekniklerin ayrımı için güzel bir örnek olarak gösterilebilir. Mesela aşamalı kümeleme analizi göz önüne alındığında birimlerin ve değişkenlerin birbirleri arasındaki farklı benzerlik ölçütlerine uygun kümeleme yapılırken küme sayısı en başta belirtilmemektedir. Bu nedenden dolayı kümeleme analizi denetimsiz bir tekniktir. Aşamalı olmayan kümeleme analizi tekniklerinden olan k ortalamalar yönteminde ise analize giren birimlerin uygun bir şekilde k kümeye gruplanması amaçlanmaktadır. Analizin başında küme sayısı belirlenmektedir. Bu nedenden dolayı k ortalamalar kümeleme analizi denetimli bir tekniktir.²⁹¹

Denetimli ve denetimsiz teknikleri bütün süreç yönünden ele almak gerekirse;

- Denetimsiz yöntemler daha fazla verinin anlaşılmasına, keşfedilmesine, tanımlanmasına yönelik uygulanan ve sonraki teknikler için fikir sunması beklenen yöntemlerdir.
- Denetimli teknikler ise verilerden bilgi ve netice çıkarımında bulunmaya yönelik uygulamalardır.²⁹²

Denetimli (Supervised) veri madenciliği yöntemler:

- En yakın k komşu (k-Nearest-Neighbor)
- K-ortalamalar kümeleme (K-means clustering)
- Regresyon modelleri (Regression models)
- Kural çıkarımı (Rule induction)
- Karar ağaçları (Decision trees)
- Sinir ağları (Neural networks)

²⁹¹ Ali Serhan Koyuncu, "Veri Madenciliği ve Sermaye Piyasalarına Uygulanması", **Sermaye Piyasası Kurulu Araştırma Raporu**, 2007, s. 89.

²⁹² (Çevrimiçi), http://www.theartling.com/dmintro/dmintro_2.htm, 15.10.2016.

Denetimsiz (Unsupervised) Veri Madenciliği yöntemleri:

- Aşamalı kümeleme (Hierarchical clustering)
- Kendi kendini düzenleyen haritalar (Self organized maps) olarak gruplandırılabilir.

Yeni jenerasyon veri madenciliği tekniklerinden olan yapay sinir ağları, insanların beyin yapılarının hesaplama şekli göz önüne alınarak oluşturulmaktadır. Girdiler ve çıktılar arasındaki küçük hesaplama birimleriyle oluşturulan sonuçların birleştirilmesiyle neticelendirilmeye çalışılan bir tekniktir. Uygulanma, anlaşılma ve yorumlanma yönünden karar ağacı ne kadar basitse, yapay sinir ağları da o seviye de zor olarak gerçekleştirilmektedir. Yalnızca model oluşturma, sonuçları yorumlama aşamasının ötesinde, doğru bir model kurabilmek için ağın eğitimindeki dengenin önemi oldukça büyüktür. Fazla eğitilmiş bir ağ, önceden gözlenmemiş bir gözleme yönelik tahmin kabiliyetini yitirirken, az eğitilmiş bir ağ ise yanlış tahmin verebilmektedir.²⁹³

Verilerle ilgili herhangi bir grup oluşturulmamaktadır. Sınıflandırma modellerinde ise modele giren verilerin sınıfı bilinmemektedir. Ayrıca modele yeni bir veri katıldığında, katılan verinin sınıfının hangisi olacağı öngörülebilmektedir. Ancak kümeleme modellerinde sınıfı olmayan veriler grup şeklinde verilere bölünmektedir.²⁹⁴

2.5.1.Karar Ağaçları

Karar ağacı uygulaması veri madenciliği alanında ucuz olma, yorumlanma kolaylığı, veri tabanı sistemleri ile basitçe bütünleşebilme, yüksek güvenilirlikte olma gibi özelliklerinden dolayı diğer sınıflandırma modellerine göre daha yaygın kullanılmaktadır.²⁹⁵

Karar ağaçları yöntemiyle verilerin sınıflandırılması iki adımdan oluşan bir işlemdir. Bu süreçte ilk adımı öğrenme aşaması oluşturmaktadır. Bu adımda daha

²⁹³ Koyuncuoglu ve Özgülbaş, **a.g.e.**, s. 27.

²⁹⁴ Gurumurthy D. Ramkumar, Arun N. Swami, Hitachi America, "Clustering Data without Distance Functions", **IEEE Data Eng. Bull.** **21.1.**, 1998, p. 12.

²⁹⁵ Serhat Özkes, **Veri Madenciliği Modelleri ve Uygulama Alanları**, 2003, s. 68.

önce elde edilen eğitim verisiyle model geliştirmek için sınıflandırma algoritması kullanılarak analiz yapılmaktadır. Sınıflandırma kuralları ya da karar ağaçları, öğrenilen modeli ifade etmektedir. Sınıflandırma işlemi ikinci adımı temsil etmektedir. Sınıflandırma adımında test verileriyle kuralların ya da karar ağaçlarının doğruluk düzeyi belirlenmeye çalışılmaktadır. Eğer doğrulama düzeyi onaylanabilir seviye de ise kurallar yeni verilerin sınıflandırılması için kullanılabilir.

Belli bir sınıfa olası üyeler olacak elemanları belirlemek, farklı pozisyonları kategorilere ayırmak, gelecek olayları öngörmek, için kural oluşturmak, belli başlı gruplara ait olan bağlantıları tanımlamak, kategorileri birleştirmek gibi faaliyetlerde karar ağaçları uygulanabilmektedir.²⁹⁶

Sınıflama işlevlerinden en çok faydalanılan algoritma karar ağaçları algoritmalarıdır. Başka tekniklerle karşılaştırıldığında karar ağaçlarını yapılandırmanın ve anlamının daha basit olduğu söylenebilmektedir. Sonraki aşamada ise veriler bu ağaca uygulanmaya başlanır. Elde edilen sonuca göre söz konusu kayıtlar sınıflandırılmaktadır.²⁹⁷ Temel olarak iki adımdan oluştuğu söylenebilir: Birincisi ağacın kurulması, ikincisi de verilerin teker teker ağaca uygulanarak sınıflandırmanın gerçekleştirilmesi şeklindedir. Karar ağaçları meydana getirilirken hangi algoritmanın kullanılacağı önemlidir. Seçilen algoritmaya göre ağaç şekil değiştirebilmektedir. Farklı yapıya sahip ağaçlar farklı sınıflama sonucu verebilmektedir.²⁹⁸

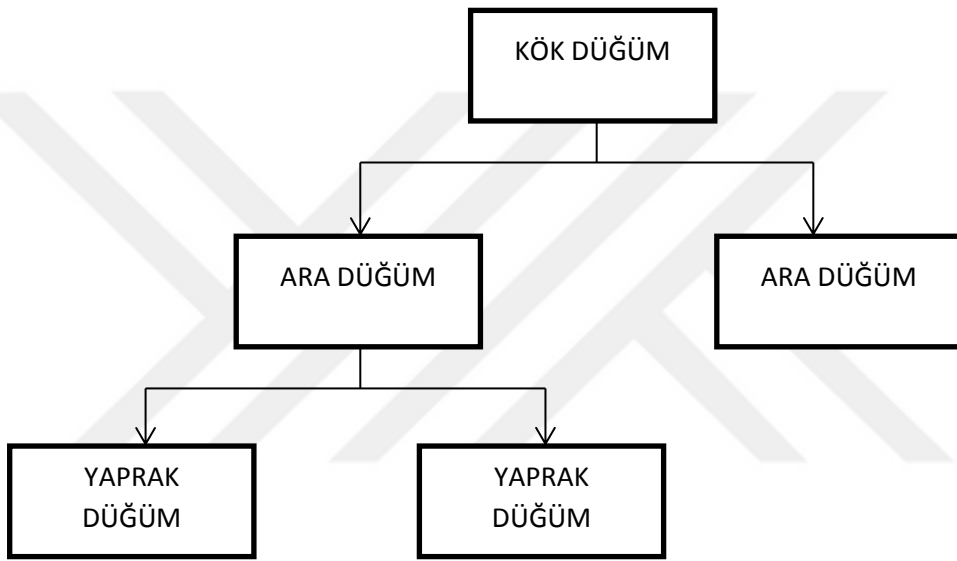
Karar ağaçlarında bazı algoritmalar geleceği tahmin özellikleri içerse de genel olarak bir sınıflandırma aracıdır. Karar ağaçları öğreniminde en önemli tekniklerin başında veri dizisini homojenleştirecek özneliğin bulunması gelmektedir. Verilen bir kararın önceden belirlenmiş olasılıklar dahilinde, başka karar noktalarını tetiklediği ardışık karar verme süreci, karar ağaçları yönetimi ile kolayca görselleştirilip hesaplanabilmektedir. Bir karar ağacı, doğal bir ağaçta olduğu gibi kök, dal ve yapraklardan meydana gelmektedir. Bir karar ağacında bu oluşum, kök düğüm, ara düğüm ve bir karar ağacının sona erdiği noktalar olan yaprak düğüm

²⁹⁶ Akpınar, **a.g.e.**, s.10.

²⁹⁷ Agrawal, **a.g.e.**, p. 488

²⁹⁸ Silahtaroglu, **a.g.e.**, s. 68.

kavramları ile ifade edilmektedir. Aşağıdaki şekilde tipik bir karar ağacı çıktısı görülmektedir. Şekilden izlenebileceği gibi kök düğümden başlayarak, her hiyerarşide veri dizisinin belirli kriterlere göre bölünmesi gerçekleşmektedir. Seçilen karar ağacı algoritmasına göre bölünme işlemi her aşamada ikili, üçlü veya çoklu olabilmektedir.²⁹⁹



Şekil 2.7. Karar Ağacı Genel Görünümü.

Karar ağacı öğrenme aşamasında öncelikle ağacın yapısı oluşturulmaktadır. Bu sayede ağaç yaprakları düzeyinde, sınıflara ait etiketler ve bu yapraklara ait olan baştan yukarı doğru uzanan kollarda, nitelikler üzerindeki işlemler ifade edilmektedir. Veri madenciliği alanında karar ağacına ait öğrenme iki esas hedef için kullanılmaktadır. Karar ağaçlarına ait öğrenmenin söz konusu amaçlara yönelik adları aşağıda yer almaktadır.

- Sınıflama problemi: Bir bireyin giderlerinden o bireyin almış olduğu eğitimin öngörülmesi gibi hedef kümenin birkaç sınıftan birine yerleştirilmesinin amaçlandığı ve sınıfların tanımlarının yapıldığı problemlerdir.

²⁹⁹ Akpınar, a.g.e.,s.200.

- İlkeleme problemi: Veri madenciliği probleminde yer alan sınıfın yerine sayısal bir değer tespit edilip, söz konusu sınıfın o değere dönüştürülmesidir.

300

2.5.1.1.ID3 Algoritması

ID3 algoritmasının en önemli özelliği entropiye bağlı olmasıdır. Özellikler birbirlerinden bağımsızdırlar. Bu algoritmanın amacı, oluşturulma aşamasında öğrenme kümesinde yer alan verilerin oldukça birbirine benzer bir şekle getirilmeye çalışılmasıdır. Bu sayede ağacın derinliğinin minimum seviyede kalması sağlanmaktadır. Karmaşık bir durumun oluşması böylece engellenmiş olmaktadır. Kazanç ise maksimum düzeyde gerçekleşmektedir. Karmaşıklığın tespit edilmesi için entropi adı verilen bilgiye dayalı bir teori kullanılmaktadır. Entropi değerlerin aralıklarına, sınıfların etiket sayılarına göre değişebilmektedir.³⁰¹

Tüm değişkenler arasında sınıflamada en iyi ayırma özelliğine sahip değişkenin bulunmasında entropi kavramı fayda sağlamaktadır. Entropi elde bulunan bilgilerin sayısallaştırılmasını ifade etmektedir yani beklenti içerisinde olmamanın maksimum hale getirilmesini ifade etmektedir. Entropi veri kümeleri içerisindeki belirsizliklerin, ilginçliklerin ve rastgeleliklerin ölçülmesinde kullanılmaktadır. Eğer tüm veriler tek bir sınıfa ait ise örneğin bütün kişiler tek bir futbol takımını destekleyse, herhangi birisine desteklediği takım sorulduğunda alınan yanıt şaşırtıcı olmayacaktı. Böyle bir halde entropi 0 olarak belirlenecekti. Entropi 0-1 aralığında yer almaktadır. Tüm olasılık değerleri eşit hesaplandığında entropi maksimum değere sahip olacaktır. Matematik alanında entropi aşağıdaki gibi ifade edilmektedir.

- $\langle p_1, p_2, \dots, p_n \rangle$ olasılıkları ifade ederse tüm olasılıkların toplamı 1 olmalıdır.
- $\sum_{i=1}^n p_i = 1$, bu durumda entropi aşağıdaki gibi olacaktır:

$$H(p_1, p_2, \dots, p_n) = \sum \left(p_i \log\left(\frac{1}{p_i}\right) \right)$$

Ham haldeki verilerin entropisi ile yani en baştaki entropi değeri ile her alt kısmın entropi ağırlıkları toplamı arasında yer alan farklar hesaplanmaktadır.

³⁰⁰ Şeker, a.g.e., s. 89.

³⁰¹ (Çevirimiçi) <http://bolubeyi.net/karar-agaclari-ve-algoritma-turleri.html>, 27.11.2016

Hesaplanan farkın ait olduğu bölüm hangisi için büyükse o alt bölüme doğru dallanma yapılmaktadır.

$$\text{Kazanç}(D;S) = H(D) - \sum_{i=1}^n P(D_i)H(D_i)$$

En son aşamada ise daha önce denildiği gibi ağaç budanmaktadır. Budama işlemlerinde kullanıcı tarafından verilen minimum destek ve güven seviyeleri esas alınarak yukarıya doğru dallar birleştirilir.³⁰²

2.5.1.2.C 4.5 ve C.5 Algoritmaları

Sayısal değerleri olan özelliklerin karar ağaçlarında kullanılmasına imkan sağlamaktadır ve önceden bilinmeyen özellik değerlerine sahip olan örneğe ait kümeler için karar ağacının ne şekilde tasarlanacağı ile ilgili bir yol göstermektedir.³⁰³

C4.5 algoritması ID3 algoritmasına göre belli avantajlar sunmaktadır. Bu avantajlarda ilki ID3 algoritmasında karar ağacı meydana getirilirken kayıp verilerin göz önüne alınmamasıdır yani kazanç seviyesinin hesaplanması durumunda sadece eksik verisi bulunmayan kayıtlardan faydalanılmaktadır. C4.5 algoritması ise eksik olan verileri, uygulamadaki diğer değişkenlerin yardımıyla tahmin ederek kazanç oranının belirlenmesini sağlamaktadır.³⁰⁴ Böylece duyarlılık ve anlamlılık açısından daha ileri düzeyde kuralların çıkartıldığı bir ağaç üretilmektedir.

2.5.1.3.CART Algoritması

Karar ağacı uygulamasında yer alan algoritmalarından birisi de CART algoritmasıdır. Bu teknikte ID3 algoritması tekniğinde ki gibi dalları en iyi şekilde ayırmada entropiden faydalanmaktadır. Ayrılma işleminde en iyi koşulu tespit etmek için ID3 ve C4.5 algoritmalarından daha farklı özelliğe sahip bir teknik kullanmaktadır. CART algoritmasında kullanılan en iyi dallara ayırma kriteri şu şekilde hesaplanır³⁰⁵ CART ya da C&RT olarak ifade edilen bu tekniğe ait

³⁰² Silahtaroglu, a.g.e., s. 78.

³⁰³ Özkan, a.g.e.,s. 37.

³⁰⁴ Dunham, a.g.e.,p. 33.

³⁰⁵ Yisehac Yohannes, Patrick Webb, **Classification and Regression Trees, CART: A User Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity**, Intl Food Policy Res. Inst., 1999, p. 15.

algoritma ikiyeşerli araçlar üretmektedir. Dolayısı ile düğümlerden hangilerinin kök olacağı veya düğüm olarak kalacağı ile ilgili karar alınmaktadır. Düğüm belirlendikten sonra ikiye ayrımın hangi noktadan gerçekleşeceği de hesaplanmalıdır yani söz konusu algoritma dalların ayrılması işleminde en uygun değişkeni tespit etmektedir. Ayrıca belirlediği değişken ikiden çok farklı çeşitte değerler içeriyorsa nasıl iki gruba bölüneceğini de tespit etmektedir.³⁰⁶

2.5.2. Yapay Sinir Ağları

Yapay Sinir Ağı (YSA) modelleri, bağlantı kuran modelleyici (connectionist modeller), paralel bir şekilde dağıtılan işleme modeli ya da nöronorfik (neuronorphic) gibi farklı ifadelerle isimlendirilmektedir.³⁰⁷ Ancak hangi şekilde isimlendirilse isimlendirilsin sistemde yer alan sinir yapısının ele alınarak iyi bir performans sağlamasıyla ortaya çıkmıştır. Bu şekilde hesaplama elemanları arasındaki yoğun ilişkilerden oluşmaktadır. Özellikle konuşma görüntü algılama gibi alanlarda geniş bir potansiyeli bulunmaktadır. Hesaplama hızı YSA modellerinde paralel ve yüksek hızda olmalıdır.³⁰⁸ İnsan sinirinin bir modelini esas alan teknik, sinir ağları yöntemini meydana getirmektedir. Sinir ağı bir ya da daha fazla çıktı tahmin etmek için kullanılmaktadır. Aynı zamanda bir girdi kümesi olarak verilmektedir.³⁰⁹

Yapay sinir ağları örneklerden öğrenmeye bağlı öğrenme tekniklerine göre uygulanmaktadır. Yani bir olay ile alakalı gerçekleşmiş örneklerin kullanılmasıyla, girdiler ve çıktılar arasında bulunan bağlantılar öğrenilmeye çalışılmakta ve bu bağlantılara göre yeni örnek çıktıları belirlenmektedir. YSA'da öğrenmeyi yapacak sistem ve algoritma öğrenme stratejilerine göre değişmektedir. Öğrenme stratejilerinden ilki denetimli öğrenmedir. Bu stratejide olayın öğrenilmesine bir yol gösteren yardımcı olmaktadır. Çok katmanlı algılayıcılar bu öğrenme türüne örnek olarak verilebilir. Destekleyici öğrenme stratejisinde de bir yardımcı öğretmen vardır ancak işlevi farklıdır. LVQ ağı bu tür öğrenmeye örnek olarak verilebilir. Denetimsiz

³⁰⁶ Silahtaroglu, a.g.e., s. 82.

³⁰⁷ Richard Lippmann, "An Introduction to Computing with Neural Nets", **IEEE Assp Magazine** 4. 2., 1987, s. 43.

³⁰⁸ Akpınar, a.g.e., s.237.

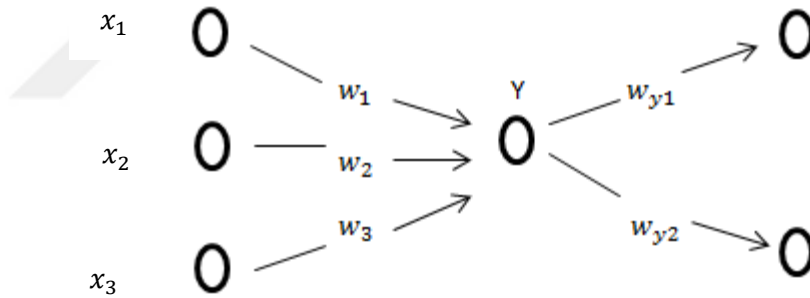
³⁰⁹ Bramer, a.g.e., s. 6.

stratejide herhangi bir yardımcı öğretmen bulunamamaktadır. İlişkilerin, sistemin kendisi tarafından öğrenilmesi beklenmektedir. ART ağları bu stratejiye örnek olarak verilebilir.³¹⁰

Genel olarak bir YSA modeli aşağıdaki şekilde yer alan bileşenlerden oluşmaktadır.

- N tane katmandan
- Her katmanda nöronlara benzeyen ve aynı işlevi gerçekleştiren, farklı sayılarla ifade edilebilen hesaplama elemanlarından
- Hesaplama elemanları arasında yer alan yoğun ilişkilerden meydana gelmektedir.

Farklı YSA modelleri içerisinde yer alan hesaplama elemanları düğüm, birim ya da işlem elemanı vb. ifadelerle adlandırılmaktadır.³¹¹



Şekil 2. 8. Yapay Sinir Ağları Genel Görünümü.

Buradaki Y nöronu ele alınırsa, bu nöron X_1 , X_2 , X_3 nöronlarından işaret alır (x_1 , x_2 ve x_3 işaretler). X_1 , X_2 , X_3 nöronlarını Y nöronuna bağlayan ağırlıklar ise sırasıyla w_1 , w_2 , w_3 olarak gösterilmektedir. Bu durumda gizli nöron durumundaki Y nöronunun girdisi, y -girdi, gelen işaretlerin ağırlıklarla çarpımının toplamıdır.

$$y\text{-girdi} = w_1x_1 + w_2x_2 + w_3x_3$$

³¹⁰ Ercan Özetmel, Yapay Sinir Ağları, Papatya Yayıncılık, 2006, s.24

³¹¹ Akpınar, a.g.e., s.240.

İleri sürümlü yapay sinir ağlarında kullanılan w ağırlıkları her seferinde ΔW kadar düzeltilerek yenilenir.

$$w_1^{yeni} = w_1^{eski} + \Delta w_1$$

Dolayısıyla algoritmanın en ince noktası, ΔW değerlerini bularak en uygun w ağırlıklarına ulaşmaktır. Bunun için her defasında ortaya çıkan hatayı en aza indirecek bir yapı kullanılmaktadır. Yapay sinir ağlarına dayalı olarak gerçekleştirilen algoritmalar sınıflandırma işlemlerinde olduğu gibi kümeleme ve örüntü tanıma işlemleri için de kullanılabilmektedirler. Sınıflandırma için kullanıldıkları zaman en basit ifadesiyle algoritma, bir öğrenme sürecini sürdürmektedir. Bu öğrenme sürecinde çıktı katmanına ulaşabilmek için w ağırlıklarını hesaplamaya çalışır. Öğrenme için ayrılmış veri kümesi üzerinde bu ağırlıklar hesapladıktan sonra, eldeki verilerin diğer kısmını kullanarak öğrenmenin ne kadar gerçekleştiğini bulmak için ağırlıkları test eder. Eğer test sonucunda bulunan ağırlıkların etkinliği doğrulanırsa, algoritma öğrenme işlemini tamamlamış olmaktadır. Aksi durumda w ağırlıkları üzerinde düzeltme ya da yeniden değer hesaplama işlemleri yapılmaktadır. Öğrenme işlemi tamamlandıktan sonra verilen herhangi bir yeni verinin, eldeki ağırlıklar yardımı ile bağlı olduğu sınıf hesaplanabilmektedir. Yapay sinir ağlarının öğrenmesi uzun zaman alır, ancak bir kez öğrendikten sonra çok duyarlı bir şekilde sınıflandırma yapılabilmektedir.³¹²

Sinirsel ağların nöronları tabakalar olarak organize edilmiştir. Bu tabakalar girdi, çıktı ve gizli tabakalardır. Girdi seviyesi dışsal çevreden bilgi almaktadır. Her bir nöron açıklayıcı bir değişkenle uyumludur. Girdi tabakası herhangi bir hesaplama ile gerçekleşmemektedir. Bir sonraki seviyeye bilgi geçişini sağlamaktadır. Çıktı tabakası sistem dışındaki ağlar tarafından gönderilen son sonuçları üretmektedir. Sinir ağlarında 2 ya da daha fazla tepki değişkeni bulunmaktadır. Girdi ve çıktı değişkeni arasında bir ya da daha fazla orta tabaka bulunabilmektedir. Bunlar ise gizli tabakadır ve dışsal çevreyle doğrudan bağlantılı değildirler. Bu tabakalar sadece analiz içindir. İşlevleri ise girdi değişkenleri ve çıktı

³¹² Silahtaroglu, a.g.e., s. 126.

değişkenleri arasındaki ilişkiyi sağlamakla birlikte geriye daha yakın adapte olmalarını da sağlamaktadır.³¹³

2.5.3. Genetik Algoritmalar

Algoritma alanı ağırlıklı olarak genomla ilgili terminoloji terimlerini kullanmaktadır. Doğal dünyada belli bir çevre gücünün sınırları ve zorlukları en uygun dölü üretmek için farklı türlerle savaşılmaktadır. Genetik algoritma dünyasında çeşitli potansiyel sonuçların uygunluğu karşılaştırılmaktadır ve en uygun potansiyel çözümler her zamankinden daha çok optimal sonuçları üretmek için yavaş yavaş gelişmektedir. Genetik algoritmalar doğal seçimin kullandığı, ayrıca şirket ve araştırma problemlerini çözmek için kullanılan süreçleri sayısal olarak taklit etmeye de çalışmaktadır. 60-70'li yıllarda geliştirilen genetik algoritmalar; biyolojik olarak genetik bilginin eş seçimi tekrar üretilmesi, mutasyonu ve çapraz tedavisi gibi esinlenen faktörlerin etkilerini çalışmak için bir çerçeve sağlamaktadır.

Vücuttaki her bir hücre aynı kromozom kümesini içermektedir. Aynı zamanda tasarım olarak işleyen DNA dizilerini de içermektedir. Her bir kromozom genlere ayrılmaktadır. Genler ise göz rengi gibi özel bir nitelik kodlamak için tasarlanan DNA bloklarından oluşmaktadır. Belir bir özelliği belirleyen bir genin değişik (alternatif) hallerinden her biri aleldir. Her bir gen kromozom üzerindeki belli bir mahalde bütün olarak bulunmaktadır. Gen oluşumu yeni bir kromozomun her iki ebeveynin kromozomlarının özelliklerinin birleştirilmesiyle düzenlenmektedir. Dölün bir kromozomunda tek bir genin değişmesi olan mutasyon, rastgele nispeten ve nadir meydana gelebilmektedir. Genetik algoritma üç işlevle kullanılmaktadır:

1) Seçim: Seçim işlevi hangi kromozomların yeniden üretileceğini seçmek için kullanılan bir metottan bahsetmektedir.

2) Çapraz Geçiş: Bu işlev 2 yeni dölü meydana getiren gen oluşumunu gerçekleştirmektedir. Bunu yerine getirirken rastgele bir yer seçerek yapmaktadır.

³¹³ Grabmeier ve Rudolph, a.g.e., p. 341.

Aynı zamanda seçim boyunca seçilen 2 kromozom arasındaki yerin sağ ve soluna müteakip değiş tokuş yaparak gerçekleşmektedir.

3) Mutasyon: Bir kromozomdaki belli bir mahaldeki parçaların rastgele değişmesiyle meydana gelmektedir.

Çoğu genetik algoritma potansiyel sonuçların koleksiyonunu iteratif olarak güncelleyerek çalışmaktadır. Potansiyel sonuçların derlenmesi ise popülasyon olarak adlandırılmaktadır. Popülasyonun her bir üyesinin her bir döngüdeki uygunluğu değerlendirilmektedir. Yeni bir popülasyon işleci kullanan eski popülasyonla yer değiştirmektedir. Yeni popülasyon üreme için seçilen en uygun üyelerden oluşmaktadır. Uygunluk fonksiyonu kromozom üzerinde işlev gören gerçek değerli bir fonksiyondur yani $f(x)$ 'deki x 'in uygunluğu değerlendirme sırasında kromozomun aldığı sayısal değerdir.

Adım 1: Verinin parça dizilerinde kodlandığı farz edilmektedir. Çapraz geçiş oranı p_0 ve mutasyon oran p_m belirlenir.

Adım 2: L uzunluğundaki n kromozomun her bir kümesinden oluşan popülasyon seçilmektedir.

Adım 3: Uygunluk fonksiyonu popülasyondaki her bir kromozom için hesaplanmaktadır.

Adım 4: n döl üretilene kadar seçim, çapraz geçiş ve mutasyon aşamaları yenilenir.

Adım 5: Kromozomların yeni popülasyonu mevcut popülasyonla yer değiştirmektedir.

Adım 6: Son verme kriteriyle karşılaşıp karşılaşılmadığı kontrol edilmektedir.³¹⁴

³¹⁴ Daniel T. Larose, **Data Mining Methods and Models**, A John Wiley&Sons, Inc Publication, Canada, 2005, p.240.

2.5.4. Destek Vektör Makineleri

Destek Vektör Makineleri (DVM) son zamanlarda uzaktan algılama sahasında başarılı bir şekilde uygulanan sınıflama tekniğidir. Bu sınıflama yöntemi verinin farklı bir boyutta yer almasıyla gerçekleşmektedir yani verinin olduğundan yüksek olan bir boyuta dönüştürülmesiyle meydana getireceği hiper düzlem aracılığıyla iki grubun birbirlerinden ayrılması ilkesini esas almaktadır. Yüksek boyutta dönüşüm yapılması sırasında farklı nitelikler alan fonksiyonlar kullanılmaktadır³¹⁵ Sınıflama problemini kareli optimizasyonla oluşan probleme dönüştürmesi DVM'nin sağladığı en önemli fayda olarak ifade edilmektedir. Bu şekilde problem çözmeye yönelik öğrenme düzeyindeki işlem kalabalığı ortadan kalkmaktadır. Diğer yöntemlere göre daha hızlı sonuç sağlamaktadır.³¹⁶

Sınıflama alanında uygulanan etkisi büyük kullanımı kolay tekinlerdendir. Sınıflama amacıyla düzlemde yer alan gruplar arasındaki sınırın çizilmesiyle grupları ayırma olanağı sağlamaktadır. Söz konusu sınır, çizildiği yer açısından iki farklı gruba ait birimlere en uzak mesafede yer almalıdır. Dolayısıyla DVM'leri sınırın ne şekilde çizilmesi gerektiğini belirlemektedir. Verilerin doğrusal bir şekilde ayrılabilmesi ya da ayrılamaması Destek Vektör Makineleri'nde karşılaşılan durumları ifade etmektedir.³¹⁷ Destek Vektör Makinelerinde karşılaşılabilecek iki durum, verilerin lineer olarak ayrılacakları bir yapıda olması veya lineer olarak ayrılamayan yapıda olmasıdır.

2.5.4.1. Doğrusal Destek Vektör Makineleri

Eğitimde kullanılan N elemanla oluşturulan verilerin $\theta = \{x_i, y_i\}$ $i=1,2,\dots,N$ olduğu farzedilsin. İki değere sahip olan veriler doğrusal olarak ayrılacak yapıda ise doğrudan normal bir düzlemle ayrılabilir. Bu düzlem

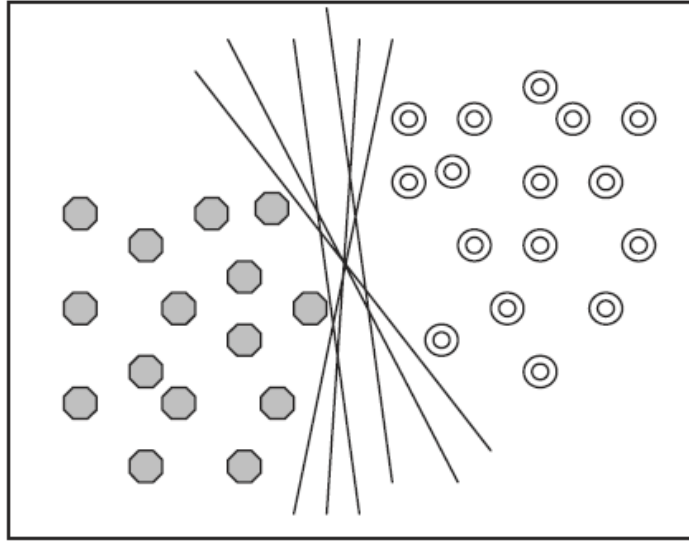
³¹⁵ Taşkın Kavzoğlu, İsmail Çölkesen, "Destek Vektör Makineleri ile Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi", **Harita Dergisi** **144.7.**, 2010, s. 2.

³¹⁶ Stanislaw Osowski, Linh Tran Hoai, Tomasz Markiewicz, "Support Vector Machine-Based Expert System for Reliable Heartbeat Recognition", **IEEE Transactions on Biomedical Engineering** **51.4.**, 2004, p. 149.

³¹⁷ Şeker, **a.g.e.**, s. 83.

ayırıcı aşırı düzlem olarak isimlendirilmektedir. DVM’de hedef söz konusu düzlemde uzaklığın sınıflara eşit uzaklıkta belirlenmesidir.³¹⁸

Şekil2.9’da gösterildiği gibi iki sınıfa ait veriler birbirlerinden ayrılarak birden fazla hiper düzlem çizilebilmektedir ama Destek Vektör Makinelerinde amaç kendine en yakın mesafedeki noktaların belirlenerek aradaki mesafeyi maksimum seviyeye ulaştıran hiper düzlemi elde etmektir.

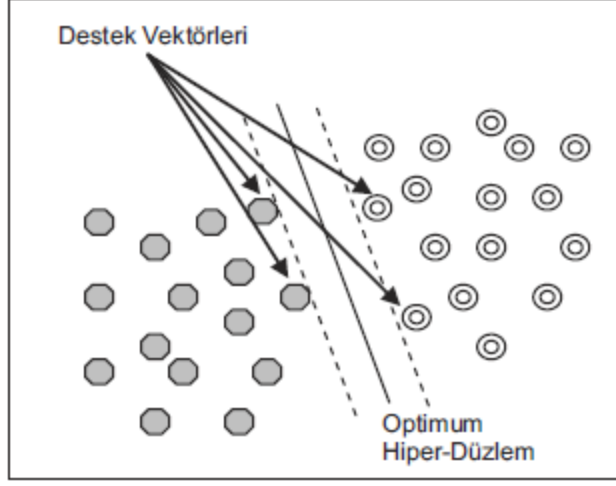


Şekil 2.9. Optimum Hiper-Düzlem ve Destek Vektörleri

Şekil 2.10’da görüldüğü gibi sınırın maksimum hale getirilmesiyle en elverişli ayırmanın yapıldığı düzlem optimum hiper düzlem olarak adlandırılmaktadır. Bu sınırın genişlik ölçüsünü kısıtlayan noktalar da Destek Vektörleri olarak adlandırılmaktadır.³¹⁹

³¹⁸ Emre Yakut, Bekir Elmas, Selahattin Yavuz, "Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi** 19.1., 2014, s. 143.

³¹⁹ Kavzoğlu ve Çölkesen, **a.g.e.**, s. 8.



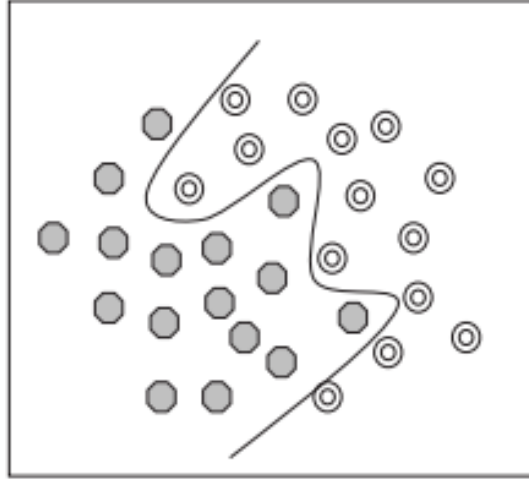
Şekil 2.10. Optimum Hiper-Düzlem ve Destek Vektörleri

2.5.4.2. Doğrusal Olmayan Destek Vektör Makineleri

Destek vektör makinelerinde ayırım bir aşırı düzlem üzerinde iki gruba ayıran bir doğruyla ve bu doğrunun bağlı olduğu bir tolerans ile sağlanmaktadır. Böyle bir durum her defa da beklenen neticeyi vermeyebilmektedir.³²⁰ Doğrusal olmayan Destek Vektör Makineleri yöntemi algoritması, veri setinin doğrusal bir fonksiyon aracılığıyla tam ya da belli bir hatayla ayıramadığı durumlarda kullanılmaktadır. Gerçek yaşama ait problemlerde de veri setinin hiper düzlem aracılığıyla doğrusal bir şekilde ayırımının yapılması oldukça zordur. Bu nedenden dolayı sınıfların ayrılması ayırma eğrisiyle öngörülmeye çalışılmaktadır. Uygulamaya ait eğrinin ön görülmesi zor bir işlemdir.³²¹ Doğrusal ayırımın yapılamaması durumuna ait geometrik gösterim Şekil 2.11'de yer almaktadır.

³²⁰Şeker, a.g.e., s.83

³²¹Sevgi Ayhan, Şenol Erdoğan, "Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi", *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9.1., 2014, s.185



Şekil 2.11. Doğrusal Olarak Ayrılmayan Hiper-Düzlemin Belirlenmesi

Bu durumdan dolayı eğitim verilerine ait bir bölümün optimum hiper düzlemde belirli bir tarafta bulunmasından kaynaklı problem pozitif yapay değişken tanımlanarak çözüme ulaştırılmaktadır. C ile ifade edilen pozitif değerler ($0 < C < \infty$) Şekil 2.12’de yer alan sınırı maksimum değere ulaştırmakta ve hatalı sınıflamayı minimum seviyeye çekmektedir. ³²²Düzenleme parametresiyle ve yapay değişkenle direkt ayrımı yapılamayan veriler için optimizasyon amaçlı problem;

$$\min \left[\frac{\|w\|^2}{2} + C \cdot \sum_{i=1}^R \varepsilon_i \right] \text{şeklini alır.}$$

Buna dayalı sınırlamalar ise;

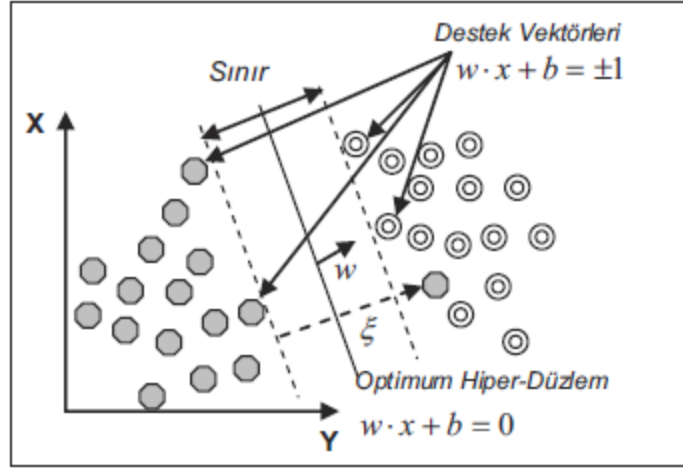
$$y_i(w \cdot \varphi(x_i) + b) - 1 \geq 1 - \varepsilon_i$$

$$\varepsilon_i \geq 0 \text{ ve } i = 1, 2, \dots, N$$

şeklinde ifade edilmektedir. ³²³

³²²Corinna Cortes, Vladimir Vapnik, “Support-Vector Networks”, **Machine Learning, Vol. 20, No.3**, 1995, p.280

³²³Kavzoğlu ve Çölkesen, **a.g.e.**,s.9



Şekil 2.12. Optimum Hiper- Düzlem Ve Destek Vektörleri

2.5.5. Bayes Sınıflandırma Tekniği

Naive Bayes ise veri madenciliği işleminde kullanılabilecek sınıflandırma algoritmalarının en basitlerinden bir tanesidir. Naive Bayes algoritması her koşulun neticede meydana getirdiği etkilerin olasılık anlamda tespit edilmesi temeline dayanmaktadır. NaiveBayes algoritmasında formül aşağıdaki gibi hesaplanmaktadır:

$$p(C|F_1, \dots, F_n) = \frac{p(C)p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)}$$

Herhangi bir sonucun gerçekleşme olasılığı söz konusu sonucu etkileyen bütün kriterlerin onu sağlayabilme olasılıklarının çarpımından elde edilmektedir.

NaiveBayes, bayes ağ türlerinin içinde, bütün özelliklerin bağımsız belli bir sınıf değeri olduğu en basit formudur. Bu durum bağımsızlık koşulu olarak açıklanmaktadır. Koşullu bağımsızlık varsayımının nadiren gerçek dünyay uygulamalarında geçerli olduğu açıktır.³²⁴

³²⁴ (Çevrimiçi) <http://www.aai.org/Papers/FLAIRS/2004/Flairs04-097.pdf>, 26.11.2016.

2.5.6.K-En Yakın Komşu

Yaygın bir şekilde kullanılan algoritmalarından birisi olarak ifade edilmektedir. İngilizce kaynaklar "K-Nearest Neighbor" ya da "KNN" diye tanımlamaktadır. Sınıflama yapılırken veri tabanındaki her bir kaydın diğer kayıtlarla olan uzaklığı hesaplanır. Ancak, herhangi bir kayıt için başka kayıtlardan yalnızca k tanesi göz önüne alınmaktadır. Algoritmanın adından da görüldüğü gibi k sayıdaki kayıt yani veri tabanındaki nokta, uzaklığı hesaplanmış noktaya diğer kayıtlara göre daha yakın olanlardır. Bu teknik, özellikle coğrafik alanda bilgi tabanlı sistemler de yaygın olarak kullanılmaktadır. Belirlenmiş noktaya en yakın şehir, istasyon vs. belirlenmesi aslında k-en yakın komşu algoritmasının temelini oluşturmaktadır.³²⁵ Sınıflamada yararlanılan bu algoritmanın çalışmasıyla sınıflama esnasında meydana getirilen niteliklerden, sınıflanması beklenen kişinin önceki kişilerden k adetne yakınlık mesafesine bakılmaktadır.³²⁶

K-en yakın komşu algoritması makine öğrenme yöntemleri içerisinde en basit sınıflandırma yaklaşımlarından biridir. Bu algorithmada sınıflandırılmak istenilen nesne, öznelilik değerlerine göre kendisine en yakın komşu veya komşuların sınıfına atanır. Sınıflandırmada aktif rol alan komşuların sayısı k ile gösterildiği için, algoritma k-en yakın komşu algoritması olarak isimlendirilmektedir. Algoritma son derece basit olmakla birlikte k değerinin belirlenmesi algoritmanın etkinliği açısından önemlidir.³²⁷

Bu algoritmanın adı "tek bağlantı kümeleme yöntemi" olarak da isimlendirilmektedir. İlk başta bütün gözlemlerin sahip olduğu değerler farklı birer küme gibi değerlendirilmektedir. Söz konusu kümelerin sıra sıra birleştirilmesiyle yeni kümeler oluşturulmaktadır. Bu teknikte ilk olarak gözlemler arasında yer alan mesafeler tespit edilmektedir. İ. Ve j. gözlemleri arasında yer alan mesafelerin belirlenmesinde aşağıdaki Öklid uzaklık bağıntısı kullanılabilir:

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

³²⁵Kevin Beyer vd, "When is "Nearest Neighbor" Meaningful?.In: International Conference On Database Theory", **Springer Berlin Heidelberg**, 1999, p. 230.

³²⁶Şeker, **a.g.e.**, s.90

³²⁷Akpınar, **a.g.e.**, s.320.

Ele alınan uzaklık ölçüleri $\text{Min } d(i,j)$ seçilmektedir. Söz konusu mesafeyle alakalı satırlar bir araya getirilerek yeni bir küme ortaya çıkarılmaktadır. Oluşan duruma göre uzaklıkların tekrardan hesaplanması gerekmektedir.

Tek bir gözlemle meydana gelen kümelerin arasındaki mesafeler yukarıdaki formül ile doğrudan hesaplanabilmektedir. Ancak iki kümenin birden çok gözleme sahip olması durumunda ve kümeler arasında yer alan uzaklıkların tespit edilmesi gerekli görüldüğünde başka bir yol takip edilmektedir. İki kümeye dâhil olan gözlemlerin arasındaki “birbirlerine en yakın mesafeye sahip olanlar” tespit edilmektedir. Belirlenen bu mesafe iki kümenin birbirlerine olan mesafesi olarak kabul edilmektedir.³²⁸

Algoritma çalışırken k değeri ilk başta seçilmektedir. Bu değerin yüksekliği birbirleriyle benzerlikleri bulunmayan noktaların bir arada toplanmasına sebep olmaktadır. Bu değerin küçüklüğü de aynı sınıfa ait noktaların ayrı sınıflara atanmasına neden olmaktadır. Genel olarak k değerleri 3,5,7 olarak seçilmektedir.³²⁹

K-en yakın komşu tekniğiyle oluşturulan algoritma, gözlem değerleriyle oluşturulan herhangi bir küme için aşağıda aşamaların uygulanmasıyla meydana gelmektedir:

a) K parametresi tespit edilir. Belirlenen parametre önceden tespit edilmiş bir noktaya en yakın mesafede bulunan komşuların sayısını ifade etmektedir. Öklid mesafe ölçümü en çok kullanılmaktadır.

b) Algoritmada belirlenmiş noktanın en yakın komşuları belirleneceği için tespit edilen nokta ile diğer noktalar arasında yer alan mesafeler teker teker hesaplanmaktadır.

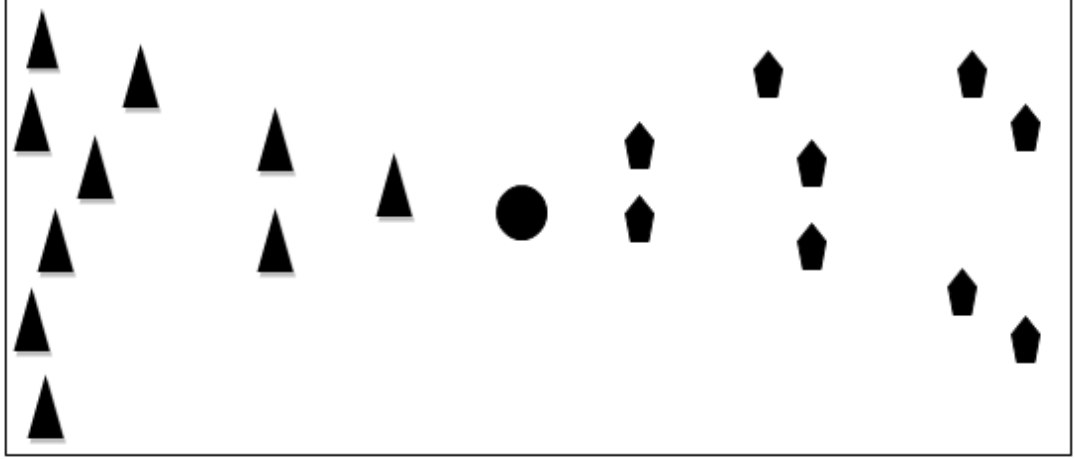
c) Yukarıda hesaplanmış olan mesafelere göre satırlar dizilir ve bu satırlar arasında en küçük değerli olan k adedi seçilmektedir.

³²⁸ Silahtaroglu, a.g.e.,s.110

³²⁹ Khan, Maleq, Ding, Qin ve Perrizo,William “K-Nearest Neighbor Classification on Spatial Data Streams Using P Trees”, In: **Pacific-Asia Conference on Knowledge Discovery and Data Mining**, Springer Berlin Heidelberg, 2002, p. 517.

d) Seçilmiş olan satırların hangi sınıfa ait olduğu belirlenir ve en fazla yinelenen kategori değeriyle belirlenmektedir.

e)Seçilmiş olan kategori, ön görülmesi beklenen gözlem değerlerinin sınıfı olarak kabul edilir.³³⁰



Şekil 2.13. K En Yakın Komşu Örnek Şekli

Yukarıdaki Şekil 2.13'de daire şekli $k=3$ için üçgen sınıfındayken, $k=5$ için beşgenlerle aynı sınıftadır. $k=3$ için en yakın ilk üç simge üçgendir. $k=5$ için en yakın 5 simge ise beşgendir.

Bugüne kadar, k-en yakın komşu algoritmasının daha etkin ve daha hızlı çalışan birçok uyarlaması tasarlanmıştır. Sınıflandırmanın dışında kümeleme işaretlerinde de kullanılabilmektedir.³³¹

³³⁰ Özkan, a.g.e., s. 118.

³³¹ Silahtaroglu, a.g.e.,s.119.

ÜÇÜNCÜ BÖLÜM

KÜMELEME ANALİZİ

3.1.Kümeleme Analizi Nedir?

Günümüzde bireyler kritik kararlar vermeden önce geriye dönüp bakarak sorunun genelini anlamaya çalışmaktadırlar. Ancak çoğu zaman sorunun genelini çözebilmek çok karmaşık olmaktadır. Devasa büyüklükte veri tabanları fazla hacimde alanı kapsayabilmekte ve aşırı karmaşık bir yapısı bulunduğundan en iyi sonuç veren yöntemler bile bu veri yığınları arasından anlamlı neticeler çıkaramamaktadırlar. Fazla karmaşık olan problemlerin çözümünü bulmakta takip edilen teknik genel olarak problemi daha kolay çözülebilecek hale getirmekle gerçekleşmektedir yani problem alt sorunlara bölünerek ve her biri ayrı ayrı çözüldükten sonra elde edilen çözümler birleştirilerek sonuç elde edilmeye çalışılır. Ancak bazı koşullarda verilerin dağılımları, nereden bölünmeleri gerektiği konusunda kestirimde bulunmayı engellemektedir. Bu nedenle otomatik olarak küme tespit etme teknikleri bulunmuştur.³³²

İstatistiksel yöntemlerden bazıları değişkenler ve bu değişkenlere ait veriler arasında bulunan benzerlikler ya da farklılıkları bulmaya imkân sağlamaktadır. Özellikle çok değişkenli istatistik teknikler çeşitli niteliklere ait değişimlerin aynı zamanda değerlendirilmesine olanak sağlamaktadır. Verilerin içerisinde meydana gelen gruplamaları meydana çıkaran yöntem çok değişkenli istatistik tekniklerden olan kümeleme analizidir. Karışık veri setleri içinde var olan yapıların ortaya çıkarılmasını sağlayan kümeleme analizinin, kendi içerisinde farklı yöntemleri bulunmaktadır. Kümeleme analizinde hedef birbirlerine göre farklılıklar veya yüksek seviyede benzerlikler taşıyan verilerin veya değişkenlerin kümelerde gruplanmasını sağlamaktır. Her kümede yer alan birimlerin aynı hassasiyeti taşıması ve tüm veride daha çok homojen kümelerin oluşması beklenmektedir.³³³ Sınıflamanın tersine,

³³² Özgür M. Dolgun, Tülin Özdemir, Doruk Oğuz, "Veri Madenciliğinde Yapısal Olmayan Verinin Analizi: Metin Ve Web Madenciliği", *İstatistikçiler Dergisi: İstatistik Ve Aktüerya*, **2.2.**,2009, s.50

³³³ Carme Hervada Sala, Eusebi Jarauta Bragulat, "A Program To Perform Ward's Clustering Method On Several Regionalized Variables", *Computers & Geosciences*, **Vol.30 No.8** ,2004,p.883.

tekrardan tanımlanan sınıflara bağlı olmamaktadır. Kümeleme yöntemi, denetimsiz öğrenmeyle (unsupervised learning) gerçekleşmektedir.³³⁴

Kümeleme analizi çok değişkenli istatistiksel yöntemlerden olup, grup sayıları önceden bilinmeyen verilerin birbirlerine olan benzerlik durumlarına göre gruplandırılması için uygulanmaktadır. Birimler ve değişkenler açısından benzer olan veriler kümeleme analiziyle ayrı kümelerle toplanmaktadır.³³⁵ Kümeleme analizi belli bir gözlem setinin ya da kümelerin kısıtlı bir sayıya ayrılmasını hedefleyen çok değişkenli istatistik tekniklerdendir. Söz konusu ayırım aynı grupta yer alan gözlemlerin birbirlerine benzer durumları varken farklı grupta yer alan gözlemlerin birbirleri arasında farklı durumları olması kaydıyla yapılmaktadır.³³⁶

Kümeleme analizi, iki gözlem arasındaki benzerlikler ya da farklılıklar dikkate alınarak yapılmaktadır.³³⁷ Kümeleme teknikleri; uzaklık matrisi veya benzerlik matrisi kullanarak birimlerin veya değişkenlerin birbirleri içerisinde homojen ve birbirleri aralarında heterojen gruplamalar oluşturmaya imkan sağlayan tekniklerdir.³³⁸

Kümeleme süreci heterojen bir yapıda olan kitlenin daha homojen alt gruplara veya kümelerle bölünmesiyle gerçekleşmektedir. Kümeleme işleminde önceden belirlenen sınıflara göre bölünme yapılmamaktadır. Bu da sınıflama ile kümelemeyi ayıran en önemli farktır. Sınıflama işleminde ise veriler gerçekleştirilen bir eğitim sonucunda meydana gelen bir modele uygun olarak önceden belirlenen bir sınıfa atamayla gerçekleşmektedir. Kümelemede ise başlangıçta tanımlanan sınıf ya da

³³⁴ Şebnem Koltan Yılmaz, Said Patır, "Kümeleme Analizi Ve Pazarlamada Kullanımı", **Akademik Yaklaşımlar Dergisi**, 2.1.,2012, s.94.

³³⁵ Zeki Çakmak, Nevin Uzgören, Gülnur Keçek, " Kümeleme Analizi Teknikleri İle İllerin Kültürel Yapılarına Göre Sınıflandırılması Ve Değişimlerinin İncelenmesi", **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, 12.3.,2005, s.20

³³⁶ Tim H. NEIL, **Applied Multivariate Analysis**, New York, 2002, p.515

³³⁷ Richard Johnson, Dean Wichern, **Applied Multivariate Statistical Methods**, New Jersey, 1992, p.65

³³⁸ Kazım Özdamar, **Paket Programlar ile İstatistiksel Veri Analizi**, İstanbul, Kaan Kitapevi, 5. Baskı, 2004, s.528.

örnekler yer almaktadır. Elde edilen sınıfların hangi anlamlara geldiği ise uygulamayı yapan kişiye bağlı kalmaktadır.³³⁹

3.2.Kümeleme Analizinin Özellikleri ve Amacı

Kümeleme analizi öncelikli hedef nesnelerin özelliklerini göz önüne alarak gruplama yapmak olan çok değişkenli istatistik yöntemlerden birisidir. Kümeleme analizinde söz konusu bu gruplama başlangıçta belirlenen kritere uygun olarak gerçekleştirilmektedir. Kümelerin kendi içerisinde fazla homojenlik, kendi aralarında ise fazla heterojenlik taşımasına dikkat edilmektedir. Kümeleme analizinin genel yapısı içerisinde rastlantı değişkeni önemli bir yer kapsamaktadır. Söz konusu değişken kümeleme analizinde nesnelerin karşılaştırılması amacıyla kullanılmakta olan nitelikleri temsil eden değişkenler kümesinden oluşmaktadır. Küme analizinde yer alan küme rastlantı değişkeni sadece nesnelerin karşılaştırılması amacıyla kullanılmakta olan değişkenleri kapsadığı için söz konusu nesnelere ait özellikleri belirlemektedir. Yalnızca nesneleri karşılaştırmak için kullanılan değişkenleri içerdiği için nesnelerin özelliğini belirlemektedir. Kümeleme analizinde, rastlantı değişkeni deneysel anlamda tahmin edilmektedir. Bunun yanı sıra bu değişkenin araştırmayı yapan tarafından belirlenmesi bu tekniği diğer çok değişkenli istatistik yöntemlerden ayırmaktadır. Kümeleme analizinde hedef rastlantı değişkenini tahmin etmek değildir ancak bu değişkene bağlı olarak nesneleri kıyaslamaktır.³⁴⁰ Kümeleme analizinin amaçları arasında radikal tipleri belirlemek, grupların içerisinde ön tahmin yapmak, hipotez testini uygulamak, verilerin yerine kümeleri değerlendirmek ve aykırı değerleri bulmakta yer almaktadır.³⁴¹

İyi bir kümeleme analizinde aşağıda bulunan özellikler yer almalıdır:

³³⁹Metin Vatansever, "Görsel Veri Madenciliği Tekniklerinin Kümeleme Analizlerinde Kullanımı ve Uygulanması", **Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Yayınlanmamış Yüksek Lisans Tezi**, 2008, s.53

³⁴⁰Joseph F. Hair, Rolph E. Anderson, Ronald L. Tatham Ve C. William Black,**Multivariate Data Analysis With Readings**, USA, 1995, 4th Edition, p. 423.

³⁴¹Ömer Yılmaz, Sinan M. Temurlenk, Türkiye'deki İstatistik Bölgelerin Kışı Başına Düşen Gelir Açısından Hiyerarşik Ve Hiyerarşik Olmayan Kümeleme Analizi İle Değerlendirilmesi: 1965-2001, **Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi**, **19.2.**,2005, s.80.

- Ölçeklenebilme özelliği olmalıdır. Az sayıda kayıt bulunan veri tabanına da yüzbinlerce kaydı olan kümelere de uygulanma özelliği bulunmalıdır.
- Verilerin farklı çeşitleriyle kullanılabilir olmalıdır. Veri tabanının hem kategorik hem de sayısal verilerine uygulanabilmelidir.
- Düzgün şekle sahip olmayan kümelerde de uygulanabilir olmalıdır.
- Giriş değişkenlerine minimum sayıda gereksinim duyulmalıdır. Bir teknikte giriş değişkenine ne kadar az gereksinim duyuluyorsa, o kadar bağımsız sonuçlar elde edilir.
- Verilerde yer alan gürültülerle de kullanılabilme özelliği bulunmalıdır.
- Kümedeki veri kayıtlarının sıralaması analizden bağımsız olmalıdır. Analize kümedeki elemanlardan hangisiyle başlanırsa başlansın sonuç değişime uğramamalıdır.
- Devasa boyutlardaki veri tabanlarında da uygulanma özelliği olmalıdır.
- Sonuçlar kolay yorumlanabilmeli ve işlevsel özelliği bulunmalıdır.³⁴²

Normallik, sabit varyans, doğrusallık varsayımlarının kümeleme analizi içinde önemi çok az olmakla birlikte, asıl dikkat edilecek durumlar; örneklemin popülasyonu temsil etme başarısı, değişkenler arasındaki çoklu bağlantı ve gözlemlerdeki aşırı değerlerdir. Kümeleme analizinin temel varsayımı, veri matrislerinin analizden önce tahmin ve koşul değişkenlerinin alt matrislerine bölüştürülmemesidir. Diğer varsayımı ise verilerin bir kısmının homojen bir kısmının heterojen olmasıdır. Çok değişkenli istatistiksel analizlerin en önemli varsayımlarından normallik kümeleme analizinde önemli görülmemektedir. Sadece uzaklık değerlerinin normal dağılması yeterlidir.³⁴³

³⁴² Esra Dinçer, Veri Madenciliğinde K-Means Algoritması ve Tıp Alanında Uygulanması, **Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi**, 2006, s.35.

³⁴³ Hüseyin Tatlıdil, **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Ankara, Engin Yayınları, 1992, s.252.

Birimler p değişkene göre hesaplanmaktadır. Benzerliği ortaya çıkarma için kullanılan bazı ölçülerle kümeleme analizinde homojen gruplar oluşturulmaktadır. Kümeleme analizinin hedefler dört gruba ayrılmaktadır:

1) n sayıda birimin, nesnenin, oluşumun, p tane değişken değerine göre belirlenen niteliklere uygun mümkün olduğunca kendi içlerinde homojen ve kendi aralarında heterojen alt gruplara ayrılması

2) p tane değişken, n tane birimle belirlenen değerlere uygun ortak özelliklerin açıklandığı farz edilen alt kümeler ayrılması ve ortak faktörlerin meydana çıkarılması

3) Hem birimlerin hem değişkenlerin bir arada ele alınmasıyla, ortak n birimin p değişkene uygun ortak özelliklerle alt kümeler ayrılması,

4) Birimlerin, p adet değişkene uygun belirlenen değerleri için, takip ettikleri biyolojik ve tipolojik sınıflandırmayı ortaya çıkarmaktır.³⁴⁴

3.3.Kümeleme Analizi Adımları

Kümeleme analizinin uygulama adımları veri matrisini belirleme, benzerlik ya da farklılık matrisini belirleme, kümeler ayırma ve yorumlama gibi adımlardan oluşmaktadır.

- Veri matrisini belirleme: Birimlerin veya değişkenlerin doğal gruplanmalarıyla ilgili kesin bilgiler olmadığı popülasyonlardan elde edilen n tane birimin ve p adet değişkenin gözlemlerinin belirlenmesidir.
- Benzerlik ya da farklılık matrislerini belirleme: Birimler/değişkenler birbirleri arasındaki benzerlikleri veya farklılıkları ortaya çıkaran makul bir benzerlik ölçümüyle birimlerin/değişkenlerin birbirleriyle aralarında olan mesafelerin tespit edilmesidir.
- Kümelerin ayrılması: Uygun kümeleme tekniğinin yardımı sonucu benzerliklerle ya da farklılıklarla oluşturulan matrislere göre birimlerin veya değişkenlerin makul sayıda kümeler bölünmesidir.

³⁴⁴Yetiş Şazi Murat, Alper Şekerler, “ Trafik Kaza Verilerinin Kümeleme Analizi Yöntemi İle Modellenmesi”, **Teknik Dergi**, **20.98.**,2009, s.4761.

- Sonuçların yorumlanması: Elde edilen kümeleri yorumlamak ve kümeleme yapısına bağlı kalarak ortaya çıkarılan hipotezleri onaylamak amacıyla gereken analitik tekniklerin uygulanmasıdır.³⁴⁵ Tekniğin kullanılması sonucunda nesnelerin kümelere ayrımı yapılmış olmaktadır. Analizin son aşaması ise kümeleme analiziyle elde edilen sonuçların anlamlılık değerlerinin yorumlanmasıdır. Kümeleme analizinin neticesinde kümeleri meydana getiren elemanların birbirlerine benzerlikleri varken, diğer küme elemanlarıyla farklılık göstermektedirler. Kümeleme süreci başarılı olarak sonuçlanırsa geometrik bir çizimle birimlerin küme içindeki konumları birbirine yakinken, kümeler birbirinden uzaklık gösterecektir.³⁴⁶

3.4. Uzaklık veya Benzerlik Ölçüleri

Genellikle kullanılmakta olan kümeleme teknikleri birimlerin arasında yer alan mesafelere dayalı benzerlik veya farklılık matrisine uygun işlemler gerçekleştirmektedir. Çeşitli kümeleme teknikleri farklı mesafe ölçümlerine göre farklı neticeler verebilmektedir. Ayırmaya bağlı kümeleme teknikleri her veri setiyle her bir birimi sadece bir kümeye ayırmaktadır. Bu sayede aşamalı veya aşamalı olmayan kümeleme teknikleri her bir birim için karar vermekteler ve o birimi belli bir kümenin elemanı yapmaktadırlar. Sonuçlar açısından benzerlikler taşıyan yöntemler de belli başlı birimlerin farklı kümelerde bulunduğu da izlenmiştir.³⁴⁷

Küme, birbirleri arasında yakın mesafe bulunan birey ya da nesnelerin meydana getirdiği beraberlik olarak tanımlanmaktadır. Bu birliktelik çok boyutlu uzayda yer almaktadır. Dolayısıyla kümenin tanımlanması benzerlik ve farklılık kavramlarını da karşımıza çıkarmaktadır. Küme noktalarının geometrik gösterimi ikiden fazla boyutta gerçekleştirildiğinde, noktaların arasında yer alan mesafelerin birden fazla boyutta hesaplanmasına ihtiyaç duyulmaktadır. Veri matrisi X'de bulunan n birimle p değişkenin yer aldığı uzaklık matrisi D ile gösterilmektedir. Bu matrise ait elemanlar ise d_{ij} şeklinde gösterilmektedir. Diğer taraftan birimler arasındaki benzeme seviyeleri, benzerlik matrisi S, matrise ait elemanlar da s_{ij} biçiminde ifade edilir. Birimler arasındaki benzerlikler $s_{ij} = 100(1 - d_{ij}/\max d_{ij})$

³⁴⁵ Murat ve Şekerler, a.g.e., s.4765.

³⁴⁶ Yılmaz ve Patır, a.g.e., s.93.

³⁴⁷ Murat ve Şekerler, a.g.e., 90.

şeklinde hesaplanmaktadır. P adet değişkene ait olduğunda birimler arasında yer alan mesafeler aşağıda yer alan tekniklere ilişkin formüller kullanılarak hesaplanmaktadır.³⁴⁸

Sayısal veriler için kullanılmakta olan uzaklık ölçütlerinin formülleri aşağıda verilmiştir.

Minskowski Uzaklık Ölçüsü

$$d_{\gamma}(x_i, x_j) = \sum_{k=1}^p [|x_{ik} - x_{jk}|^{\gamma}]^{1/\gamma}; \quad \gamma \geq 1$$

Manhattan City – Block Uzaklık Ölçüsü

$$d_1(x_i, x_j) = \sum_{k=1}^p |x_{ik} - x_{jk}|$$

Öklid Uzaklık Ölçüsü

$$d_2(x_i, x_j) = \sum_{k=1}^p [|x_{ik} - x_{jk}|^2]^{1/2}$$

Mahalanobis Uzaklık Ölçüsü

$$d_2(x_i, x_j) = \left[\sum_{k=1}^p w_k^2 (x_{ik} - x_{jk})^2 \right]^{1/2}$$

İki kümenin ortalama vektörlerinin karşılaştırılmasında kullanılmaktadır.³⁴⁹

³⁴⁸Münevver Turanlı, Ünal H. Özden, Serdar Türedi, “Avrupa Birliği’ne Aday ve Üye Ülkelerin Ekonomik Benzerliklerinin Kümeleme Analiziyle İncelenmesi”, **İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi**, 5.9., 2006, s.99

³⁴⁹A.e.,s.100

3.5.Benzerlik Matrisinin Oluşturulması

Kümeleme tekniğinde genellikle kullanılmakta olan N tane nesneye ait olan p tane ölçüm sonucunun elde edildiği N x p boyutlu X veri matrisi aşağıda yer almaktadır.

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \cdots & x_{1p} \\ x_{21} & x_{22} & x_{23} \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} \cdots & x_{np} \end{bmatrix}$$

Matrisin içinde yer alan x_{ij} ifadeleri j'inci değişkene ait i'inci birey veya nesnenin sahip olduğu değeri ifade etmektedir. Kümeleme analizinin genel amacı birbiriyle benzerlik taşıyan birey ya da nesnelerin aynı gruplarda bir araya gelmesidir. Kümelerin oluşturulması sırasında yapılan ayırım işlemi iki gözlem arasındaki benzerlik ya da uzaklık ölçümlerine göre gerçekleştirilmektedir. Bu açıdan bireyler arasında yer alan benzerlik ölçümünde yakınlık ya da uzaklık ölçülerinden hangisinin kullanılacağı bu tekniğin en önemli problemlerinden birini oluşturmaktadır. Uygulamada sıklıkla kullanılmakta olan temel benzerlik ölçüleri:³⁵⁰

- 1) korelasyon katsayıları,
- 2) uzaklık ölçüleri,
- 3) birliktelik katsayıları ve
- 4) olasılığa dayalı benzerlik katsayılarıdır.

Korelasyon katsayısı ile uzaklık ölçüleri genel olarak sosyal bilimler alanında ölçü olarak kullanılmaktadır. Korelasyon katsayısı genel olarak iki değişken arasında yer alan ilişkiyi ifade eden sayıdır. Ancak, iki gözleme ait vektör arasında yer alan ilişkinin belirlenmesinde de kullanılabilir. x_i ve x_j gözlemleri arasında bulunan ilişkiye ait ölçü olan Kari Pearson'un Çarpım Moment Korelasyon Katsayısı:

³⁵⁰ Mark S. Aldenderfer, Roger K. Blashfield, **Cluster Analysis. Sage University Paper Series On Quantitative Applications In The Social Sciences**, Beverly Hills, 1984, Sage publications , p.22.

$$r_{jk} = \frac{\sum (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum (x_{ij} - \bar{x}_j)^2 \sum (x_{ik} - \bar{x}_k)^2}}$$

şeklinde olup , x_{ij} ; j gözlem için i. değişkenin değerini, x_j ise j. gözlem için bütün değişken değerlerinin ortalamasıdır. Bu teknik değişken değerlerinin oransal ya da aralıklı ölçekle ölçülmesi durumunda kullanılmaktadır. Katsayıya ait değerler [-1,1] aralığında olmakla birlikte, 0 değeri j. ve k. gözlemler arasında benzerlik bulunmadığını göstermektedir. $r_{jk}=1$ olması j. ve k. gözlemlerin arasındaki bütün ilişkiyi gösteriyor olmasına karşın x_j ve x_k gözlemlerinin aynı olduğunu göstermemektedir.

x_i ve x_j gözlem vektörleri arasında yer alan $d(x_i, x_j) = d_{ij}$ uzaklık mesafesinin ifade edilmesi için geliştirilen ve bir çok istatistik paket programlarında bulunan uzaklık ölçüleri vardır. Söz konusu ölçülere ait değerlerin üst sınırı bulunmamakla birlikte ölçeğe bağımlılıkları bulunmaktadır. Ayrıca aşağıda yer alan kriterlere uymaları gerekmektedir.³⁵¹

$d(x_i, x_j) > 0$; eğer $i = j$ ise $d(x_i, x_j) = 0$

$d(x_i, x_j) = d(x_j, x_i)$ simetri özelliği

$d(x_i, x_j) < d(x_i, x_k) + d(x_k, x_j)$ üçgen eşitsizlik özelliklerini sağlaması gerekmektedir.

Yukarıdaki özellikleri bünyesinde bulunduran N x N boyutu olan simetrik D uzaklıklar matrisi aşağıdaki şekildedir.

$$D = \begin{bmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1N} \\ & 0 & d_{23} & \cdots & d_{2N} \\ & & 0 & \cdots & d_{3N} \\ & & & 0 & d_{N-1,N} \\ & & & \cdots & 0 \end{bmatrix}$$

³⁵¹Brian Everitt, **Cluster Analysis**, London, Heinemann Educational Books, 1974, p.56.

Birimlerin birbirleri arasında olan benzerlik seviyeleri benzerlik (S) matrisi ile gösterilmektedir. Benzerlik matrisine ait elemanlar D matrisi elemanlarına göre belirlenmektedir. $S = 100*[1-d_{ij} / \max(d_{ij})]$ matrisinin elemanları s_{ij} biçiminde gösterilir. S benzerlik matrisinin gösterimi ise aşağıdaki şekildedir:³⁵²

$$S = \begin{bmatrix} S_{11} & & & & \\ S_{21} & S_{22} & & & \\ S_{31} & S_{32} & S_{33} & \dots & \\ \dots & \dots & \dots & \dots & \\ S_{n1} & S_{n2} & S_{n3} & \dots & S_{nn} \end{bmatrix}$$

3.6. Kümeleme Analizi Yöntemleri

Bireyleri ya da nesneleri kümelemede birçok farklı yöntem kullanılabilir. Bu farklı teknikler i hiyerarşik ve hiyerarşik olmayan olarak ayrılmaktadır.

3.6.1. Hiyerarşik Olmayan Yöntemler

Hiyerarşik olmayan kümeleme yöntemi g gruptaki n gözlemin ($g < n$) bir bölümünü gözlemek için yol göstermektedir. Hiyerarşik kümelemenin tersine süreç belirlenen en iyilik ölçütünü sağlayan bir kısmı sonuç olarak vermektedir. En iyilik ölçütü belli bir grup sayısı için maksimum içsel tutarlılığı sağlayan grup kararıdır. Hiyerarşik olmayan kümeleme grup sayılarını tanımlamakla başlamaktadır. Bu da g'nin farklı değerleriyle analizin yürütülmesiyle ve kümelemenin iyi yapılması için uygun indeksleri karşılaştırarak en uygun sonucu belirlemeyle yapılmaktadır. En çok

³⁵²Veli Yılandıcı, “ Bulanık Kümeleme Analizi İle Türkiye'deki İllerin Sosyoekonomik Açidan Sınıflandırılması”, **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 15.3., 2010, s.120.

kullanılan hiyerarşik olmayan kümeleme k ortalamaları yöntemidir.³⁵³ Küme sayısı ile ilgili bilgi veya araştırmacı tarafından kanaat getirilen anlamlılık seviyesi ve bir küme sayısı bulunuyorsa, hiyerarşik olmayan kümeleme teknikleri kullanılabilir.³⁵⁴

3.6.1.1. K Ortalamalar Yöntemi

K-means algoritmasına ait genel mantık n adet veri nesnesiyle meydana gelen bir veri setinin, giriş parametresi olan k tane kümeye bölümlenebilmesidir. Bölümlemenin sonucunda oluşan kümeler içerisindeki benzerlikleri maksimum yapmak ve kümeler arasındaki benzerlikleri minimum yapmak amaçlanmaktadır. Kümelerdeki benzerlik kümeyle ait orta bir değer olarak öngörülen bir nesne ve küme içindeki başka nesnelerin arasında yer alan uzaklık ortalamaları ile hesaplanmaktadır.³⁵⁵ En çok kullanıma sahip olan ve denetimsiz tekniklerden olan k-means algoritmasında atama süreci, her bir verinin yalnızca tek bir küme içerisinde olmasına izin vermektedir. Dolayısıyla bu algortima keskin bir yapıya sahiptir. Kümedeki merkez nokta söz konusu kümeyi temsil etmektedir. Eşit ölçüde büyüklüğe sahip kümeler tespit etmeye meyilli bir yöntemdir.³⁵⁶

K-means algoritmasının bilgisayardaki uygulanma süreci aşağıda yer almaktadır:

1. Kümenin sayısı (k) tespit edilir. Söz konusu bu değer algoritmanın dışından gelmektedir.
2. k tane rastgele ya da belli bir teknikle kümeye ait merkezler belirlenmektedir.
3. Küme merkezlerine bütün elemanların uzaklık değerleri hesaplanmaktadır.

³⁵³ Chandrasekaran Mohan, Inderpal S. Narang, "Method For Non-Hierarchical Lock Management In A Multi-System Shared Data Environment", 1996, s.39.

³⁵⁴ Turanlı, Özden ve Türedi, **a.g.e.**, s.100.

³⁵⁵ HanJiawei, Micheline Kamber, **Data Mining: Concepts And Techniques**, San Francisco, Morgan Kaufmann, 2001, s.77.

³⁵⁶ Meltem Işık, AliY. Çamurcu, "K-Means, K-Medoids Ve Bulanık C-Means Algoritmalarının Uygulamalı Olarak Performanslarının Tespiti", **İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi**, **6.11.**, 2007, s.56.

4. Elemanlar yakın bulundukları merkezlere uygun olarak kümelenmektedir.

5. Önceki aşamada meydana gelen kümelere ait ortalamalar hesaplanır ve bu kümelere ait yeni merkezler tespit edilir.

6. En sonda elde edilen kümelere ait merkezler bir önceki kümelerin merkez değerlerine eşitleninceye veya tespit edilen döngü sayısına erişilene kadar yinelenmektedir.³⁵⁷

Kümeleme analizinde faydalı neticeye ulaşabilmek amacıyla değişken seçimi ve küme sayısını belirlemek önemli bir işlemdir. Küçük örneklemeler içindeki kümeleri belirleyebilmek amacıyla aşağıda yer alan eşitlikten sıklıkla yararlanılmaktadır:

$$k = \left(\frac{n}{2} \right)^{\frac{1}{2}}$$

Mariot tarafından sunulan teknikte ise;

$$M = k^2 |W|$$

En az M değerini gösteren küme adedi gerçek küme sayısını ifade etmektedir.³⁵⁸

3.6.1.2. K Medoids Yöntemi

Medoid kümeleme yönteminde, küme içi gözlemlere ait n birimin kendi içinde benzer olmasıyla ve kümelerin arasındaki gözlem değerlerinin ise farklı olmasıyla k kümenin ayrımını yapmak amaçlanmaktadır. Bu amaç medoid ismi verilen k küme ve bu kümeleri tanıttacak olan çekirdek noktalar yardımıyla gerçekleştirilmektedir. Medoid bir kümedeki başka olgular arasında yer alan farklılıkların minimum hesaplandığı birimlerden oluşmaktadır.³⁵⁹ Bu algoritmanın esasını, verinin farklı yapıdaki özellikleriyle meydana gelen k adet temsilci nesnenin

³⁵⁷ Dinçer, a.g.e., s.8.

³⁵⁸ Belgin Atamer, Kümeleme Analizi (Cluster Analysis) ve Kümeleme Analizinin İlaç Sektörüne Uygulanması, **Marmara Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi**, İstanbul, 1992, s.45.

³⁵⁹ Özdamar, a.g.e., s.336.

bulunması oluşturmaktadır. 1987’de Kaufman ve Rousseeuw tarafından geliştirilen algoritma yaygın olarak kullanılmaktadır. Temsil yeteneği olan nesnenin kümedeki diğer nesnelerle olan ortalama mesafeyi minimum hale getirmesi onu kümedeki en merkezi nesne yapmaktadır. Dolayısıyla bu ayırma yönteminde her bir nesnenin ve referans noktasının arasında var olan benzersizliklerin toplam değerini minimum yapma mantığı kullanılmaktadır. Kümeleme analizinde temsil yeteneği olan nesneler genellikle merkezi tipler olarak adlandırılmaktadır.³⁶⁰

3.6.1.3. Bulanık Kümeleme

Kümeleme analizinin bölüm başında tanımı, veri matrisine ait ve gruplama hesapları önceden yapılmayan birimlerin birbiriyle alt kümelere ayırımı sağlayan teknikler olarak açıklanmıştı. Kümeleme analizinin temelini veri setine ait değişkenler dikkate alınarak birimler arasında yer alan mesafenin göz önüne alınması oluşturmaktadır. Ayrıca yeni bir birimin gruplardan hangisine dahil olacağının tahmin edilmesi de önemli bir yer tutmaktadır. Kümeleme analizi olasılıklı, katı ve bulanık diye üye kayıtları açısından incelenmektedir. Bütün gözlemler ya bir kümeye aittirler ya da ait değildirler. Bulanık kümeleme de üyelik kayıtları 0 ve 1 arasında değişmektedir. Verilerin eş zamanlı olarak birden çok kümeye üyeliği söz konusu olmaktadır. Bulanık kümeleme tekniğinde veri kayıtlarının kümeye üye olma olasılıkları toplamı 1’dir. Bu teknikte herhangi bir birimin bir kümede bulunma olasılık değeri bütün olası kümeler içinde 0 ve 1 arasında değişmektedir. Böylece verilere ait noktalar eş zamanlı olarak 1’den çok kümeye ait olabilmektedir. Kümelere bağlı üyelerin durumu bulanık olduğundan veriye ait noktanın hangi kümeye dâhil bulunduğunu gösteren bir tek değer bulunmaktadır. Bu tek bir değer yerine değerler kümesi bulunmaktadır. Herhangi bir birime ait belli bir bölüm bir kümeye üye iken, başka bir bölümüyle söz konusu kümenin dışında olabilmektedir. Hangi kümede olma olasılığı yüksekse o kümeye atanmaktadır. Bu nedenle klasik kümeleme tekniklerini bulanık kümelemenin alt konusu gibi incelemek uygun görülmektedir.³⁶¹

³⁶⁰ Işık ve Çamurcu, **a.g.e.**, s.76

³⁶¹ Yılcı, **a.g.e.**, s.456

3.6.2.Hiyerarşik Yöntemler

Hiyerarşik kümeleme dendogram olarak da bilinen bir küme hiyerarşisi meydana getirmektedir. Her küme nodu küçük kümeler içermektedir. Küçük kümeler ise ortak büyük kümelerle kaplanan noktalarla ayrılmış kardeş kümelerdir. Bu yaklaşım öge boyunun farklı seviyelerindeki verinin keşfedilmesini sağlamaktadır. Hiyerarşik kümeleme yöntemleri agglomeratif (aşağıdan yukarıya) ve ayırıcı (yukarıdan aşağıya) yöntemler olarak ayrılmaktadır.

Hiyerarşik kümelemenin avantajları;

- 1) Öge boyunun seviyesine bağlı esneklik
- 2) Benzerlik ya da uzaklığın herhangi bir formunu işleme kolaylığı
- 3)Herhangi bir nitelik tipine uygulanabilirliktir.

Dezavantajları ise;

- 1) Son verme kriterinin belirsizliği
- 2)Hiyerarşik algoritmalar bir kez yapılandırıldıktan sonra kümelerin tekrar değerlendirilmemesidir.³⁶²

Hiyerarşik kümeleme tekniğinde ilk olarak n adet kümenin var olduğu kabul edilmektedir. İlk aşamada en yakın mesafeye sahip olan iki küme birleştirilmektedir. İkinci aşamada ise küme adedi bir azaltılarak yenilenen uzaklıklar matrisi hesaplanır. Bu iki aşama $n-1$ kere yinelenir.³⁶³

3.6.2.1. Merkezi Yöntem

Her gruptaki birimlerin ortalama değeri grubun merkezi olarak kabul edilmektedir. Öncelikle öklid uzaklıklarının kareleri hesaplanarak benzerlik ya da uzaklık matrisi oluşturulduktan sonra i . birim ile j . birim arasındaki uzaklık

³⁶²Pavel Berkhin," A Survey Of Clustering Data Mining Techniques. In:Grouping Multidimensional Data", **Springer Berlin Heidelberg**, 2006, p. 30.

³⁶³Turanlı, Özden ve Türedi, **a.g.e.**, s.99.

$$D_{ij}^2 = \sum_{k=1}^p (x_{ik} - x_{jk})^2 \text{ dir.}$$

Buradaki D_{ij}^2 i ve j birimleri arasındaki uzaklıkların karesi toplamını, x_{ik} k. değişkeninin i. birimine ait değerini, x_{jk} k. değişkeninin j. birimine ait değeri ve p ise değişken sayısını ifade etmektedir. Oluşturulan bu matrisin incelenmesiyle hangi birimler arasındaki yakınlığın daha az olduğu saptanır. Yakınlığı az olan birimler aynı küme içinde olacak şekilde kümeler oluşturulmakta ve küme ortalamaları hesaplanmaktadır. Daha sonra yine öklid uzaklıklarının kareleri toplamı hesaplanarak aynı işlemler tüm birimlere ayrılmasına kadar sürdürülmektedir. Her aşamada yeni kümelerin elde edilmesi ve yeni birimlerin bu kümeye dâhil olmasıyla ortalamaların değişmesi karışık sonuçlar verebilmektedir. Çünkü bir kümedeki yakınlıkların ortalaması diğer bir kümedekinden küçük çıkabilmektedir. Yöntemin diğer yöntemlere üstünlüğü ise aşırı değerlerden az etkilenmemesidir.³⁶⁴

3.6.2.2. Tek Bağlantı Metodu

Tek bağlantı kümeleme yöntemi hiyerarşik kümeleme teknikleri içinde en kolay olan yöntemdir. Bu yöntemde iki küme arasında yer alan mesafe minimum ölçüye sahiptir. Dolayısıyla bir kümedeki herhangi bir birim değeri ile bir diğer kümedeki başka bir birim değeri arasında yer alan uzaklık minimum seviyededir. Oysaki tam bağlantı yöntemindeki uzaklıklar maksimum düzeydedir. Bu durum tek bağlantı yönteminin tercih edilmesini sağlamaktadır.³⁶⁵ Her aşamada p kümesi ile q kümesi birleştirildikten sonra oluşan yeni küme (k kümesi) ile başka bir küme (m kümesi) arasında yer alan benzerlik şu şekilde gösterilmektedir: Mesafe ölçüsü d_{ij} ile gösterildiğinde, k kümesi ve m kümesinin birbirlerine en yakın mesafede bulunan iki değeri arasında yer alan mesafe:

$$d_{km} = \min(d_{pm}, d_{qm}) \text{ şeklindedir.}$$

³⁶⁴ Neyran Orhunbilge, **Çok Değişkenli İstatistik Yöntemler**, İstanbul Üniversitesi İşletme Fakültesi Yayını, 2010, s.474.

³⁶⁵ John A. Hartigan, "Consistency Of Single Linkage For High-Density Clusters", **Journal Of The American Statistical Association**, vol. 76 No.374,1981, p.390.

K kümesi ve m kümesi birleştirilmişse, oluşan kümede yer alan bir değerin en yakın komşusuna olan uzaklığı en çok d_{km} kadar olacaktır.³⁶⁶

3.6.2.3. Tam Bağlantı Metodu

Tam bağlantı yönteminde bir birimin hangi birimlerle veya kümelerle bir araya getirileceği ve m. küme olarak gösterileceği, birimlerin yeni meydana gelen kümelerle hesaplanan mesafeleri göz önüne alınarak belirlenmektedir. m. kümenin hangi kümeye birleşim sağlayacağını tespiti, diğer kümeler olan k. ve l. kümelerin j. kümeyle olan uzaklıkları dikkate alınarak belirlenmektedir.

m. kümenin j. kümeyle olan mesafesi(d_{mj});

$d_{mj} = \max(d_{kj}, d_{lj})$ şeklinde belirlenmektedir.

Bu yöntem en uzak komşu niteliğine sahip birimleri birbirleriyle birleştirerek küme oluşturmayı içermektedir.³⁶⁷

3.6.2.4. Ortalama Bağlantı Metodu

Bu yöntemde süreç tek bağlantı ve tam bağlantı yöntemlerindeki gibi olmaktadır. Ancak kümeleme koşulu olarak bir küme içerisindeki değer ile başka küme içerisindeki değerler arasında yer alan ortalama mesafeler kullanılmaktadır. Bu yöntemde yer alan kümeler küçük varyans değerleriyle birbirine bağlı konumdadırlar. Bu yöntem genellikle tek bağlantı ve tam bağlantı arasında neticeler vermektedir. Bu nedenle alt bir teknik gibi kullanılmaktadır.³⁶⁸ Bu yöntemin değişime uğramış çeşitleri bulunmaktadır. Yaygın olarak uygulanan türde gözlem çiftlerinin arasında yer alan uzaklıkların aritmetik ortalamasıyla hesap yapılmaktadır. Bu teknikten genellikle biyoloji alanında yararlanılmaktadır. Sosyal bilimler alanında da gün geçtikçe kullanımı artmaktadır. Genel olarak bu yöntemde tam bağlantı yöntemiyle aynı dendogramlar meydana getirilmektedir. Ancak yöntemlerin her

³⁶⁶Michael R. Anderberg, **Cluster Analysis For Applications**, New York, 1988, s.137.

³⁶⁷Özdamar, a.g.e., s.298.

³⁶⁸Azize Celile Günay Atbaş, " Kümeleme Analizinde Küme Sayısının Belirlenmesi Üzerine Bir Çalışma", **Ankara Üniversitesi Fen Bilimleri Enstitüsü Yayınlanmamış Yüksek Lisans Tezi**, 2008, s.17.

birinde farklı uzaklıklar tanımladığından bir araya getirmeler farklı düzeylerde meydana çıkabilmektedir.³⁶⁹

M. kümenin hangi birim veya kümeyle bir araya getirileceği D matrisinde yer alan uzaklık hesapları göz önüne alınarak tespit edilmektedir. Daha önceden oluşturulan k ve l kümelerinden hangisiyle birleşim sağlanacağını tespit etmek için j. kümeyle k. ve l. kümelerin mesafe hesaplamaları aşağıdaki gibi yapılmaktadır:

M. kümenin j. küme ile olan uzaklığı ;

$$d_{mj} = (N_k d_{kj} + N_l d_{lj}) / N_m$$

şeklinde hesaplanmaktadır. Bütün eleman ve kümeler birbiri içinde bir araya getirilene kadar aynı adımlar yinelenmektedir. Meydana getirilen kümeler mesafe hesaplamalarına göre dendogram grafiğiyle görselleştirilerek aşamalı birleştirmeler görüntülenmektedir.³⁷⁰

3.6.2.5. Ward Metodu

Ward, “minimum varyans yöntemi” adıyla hiyerarşik bir tekniği oluşturmuştur. Bu teknik kümeler arasında yer alan mesafeyi değerlendirmek amacıyla varyans analizi yöntemini kullandığı için diğer tekniklerden farklı görülmektedir. Kümeleme analizinde, her aşamada bir araya getirilen kümelerde;

$$ESS = \sum_{i=1}^n (x_i - \bar{x})^2 \text{ ifadesinin en az artması hedeflenmektedir.}^{371}$$

Buradaki n, kümede yer alan nesne adedini, x_i , i'inci nesnenin değerini, x nesne değerlerine ait ortalamayı göstermektedir. Ward algoritması, kümelerin kütle merkezleri arasında yer alan öklid uzaklığına bağlı hareket etmektedir. Bu tekniğin diğer hiyerarşik yöntemlerle test edilmesi önerilmektedir. Ward tekniğinde benzerlik kriteri için öklid uzaklığının karesi hesaplanarak, iki küme arasındaki uzaklık;

³⁶⁹Seniye Ümit Fırat, “Kümeleme Analizi: İstihdamın Sektörel Yapısı Açısından Avrupa Ülkelerinin Karşılaştırılması”, *İstanbul Üniversitesi Sosyal Bilimler Dergisi*, **3.2.**,1995, s.46.

³⁷⁰ Özdamar, **a.g.e.**, s.349.

³⁷¹ Everitt, **a.g.e.**, s.79.

$$D_{KL} = \frac{\|\bar{X}_K - \bar{X}_L\|^2}{\frac{1}{N_K} + \frac{1}{N_L}}$$

eşitliği ile belirlenmektedir.

Ward metodu, genel anlamda küçük ölçekli kümeleri ortaya çıkarmasına rağmen etkisi büyük bir yöntemdir. Bu teknik aynı sayıda nesneyi kapsayan kümeler meydana getirme eğiliminde bir yöntemdir.³⁷²



³⁷²Bahattin Yanık, İlhan Avcı, Bölgesel Debi Süreklilik Eğrilerinin Elde Edilmesi, **İTÜ Dergisi**, 4.5., 2010, s.22.

DÖRDÜNCÜ BÖLÜM

UYGULAMA

Bu doktora çalışmasının uygulama kısmında kullanılan veriler Avrupa Birliği üye ülkeleri ve Türkiye'nin 2013-2016 yılları arasındaki ekonomik değişkenleridir. Uygulamada söz konusu ülkelerin 34 bilgi ekonomisi değişkeniyle birlikte GSYİH, istihdam ve ihracat ithalat oranı değişkenleri kullanılmıştır.

İlk olarak ülkelerin her yıl için ayrı ayrı GSYİH, istihdam ve ihracat ithalat oranı değişkenleri kullanılarak kümeleme analizi yapılmıştır. Ekonomisi güçlü olan ülkeler 3.grupta, 3.gruba göre nispeten orta olan ülkeler 1.grupta, 1.gruptaki ülkelere nispeten ekonomisi küçük olan ülkeler 2.grupta yer almaktadır. Her yıl için bu gruplama ayrı ayrı yapılmıştır.

İkinci aşamada veri madenciliği tekniklerinden karar ağacı uygulaması yapılmıştır. Karar ağacı uygulamasında söz konusu ülkelerin 2013-2016 yılları arasında 34 tane bilgi ekonomisi değişkeni ve önceki kümeleme analizi uygulamasında gruplanan ülkelerin genel ekonomik büyüklük seviyesini gösteren değişkenler kullanılmıştır. Analizde hedef değişken ekonomik büyüklük değişkeni olarak alınmıştır. Tahminleyici değişken olarak ise bilgi ekonomisi değişkenleri kullanılmıştır.

Uygulamanın üçüncü aşamasında, ilk aşamadaki ülkelerin kümeleme analizi sonucu elde edilen ekonomik seviyeleri, bu ekonomik büyüklük seviyesini en çok etkileyen bilgi ekonomisi değişkenleri, ülkelerin Avrupa Birliği'ne üyelik yılları, refah ülkesi olma özellikleri ve coğrafi bölgeleri göz önüne alınarak veri madenciliği algoritmalarından Kohonen ağırları ile ülkeler gruplandırılmıştır. Meydana gelen kümelerin üzerinde hangi değişkenlerin ne kadar etkili ve ayırıcı özelliğe sahip olduğu belirlenmiştir.

Kümeleme analizi SPSS programı ile yapılmıştır. Veri madenciliği uygulaması ise yine SPSS programının uygulaması olan Clementine programı ile yapılmıştır. SPSS veri madenciliği çözümü, Clementine istatistik ve yapay zeka kökenli algoritmaları bir arada sunan, veri ambarlarına yaptığınız yatırımların geri dönüşü noktasında kritik bir çözümdür. Clementine, uygulama kolaylıkları, açık yapısı ve açık metodolojisi ile ülkemizde ve dünyada veri madenciliği uygulamalarında tercih edilirliliği en yüksek olan çözümdür. Bu özelliklerinden dolayı Clementine programı uygulama bölümünde kullanılmıştır.

Aşağıdaki bölümde verinin nasıl toplandığı, düzenleme aşamasında hangi işlemlere tabi tutulduğu hakkındaki detaylı bilgiler, yapılan uygulamalar ve varılan sonuçlar yer almaktadır.

4.1.Ülkelerin Ekonomik Büyüklüklerine Göre Kümeleme Analizi ile Gruplanması

Tezin uygulama bölümünde ilk olarak kümeleme analizi yapılmıştır. Yapılan kümeleme analizinde kullanılan değişkenler Avrupa Birliği üye ülkelerinin GSYİH, işsizlik ve ihracat/ithalat oranı değişkenleri göz önüne alınarak yapılmıştır. Analizde 35 tane ülke verisi kullanılmıştır. Kümeleme analizi ülkelerin söz konusu verileri ele alınarak 2013-2016 arasındaki yıllar için ayrı ayrı yapılmıştır. Uygulamanın ilk bölümünde yapılan kümeleme analizinde SPSS programı kullanılmıştır.

Çalışmada yer alan kümeleme analizi hiyerarşik olmayan yöntemlerden k-ortalamar tekniği ile yapılmıştır. Ancak hiyerarşik yöntemlerle de aynı sonuçların elde edildiği gözlenmiştir. Tüm verilerin metrik olmasıyla birlikte yöntem olarak k – ortalamar metodu yöntemi kullanılmıştır. K ortalamar yönteminde küme sayısı öncelikle araştırmacı tarafından belirlenmektedir. Ülkelerin ekonomik durumları önceden aşağı yukarı tahmin edilebildiğinden dolayı k ortalamar yöntemi tercih edilmiştir. K ortalamar yöntemi tezin üçüncü bölümünde ayrıntılı olarak ele alınmıştır.

Kümeleme analizi istatistiksel olarak herhangi bir varsayıma dayanmamaktadır. Ancak değişkenler arasında çoklu doğrusal bağlantı olması bu değişkenlerin birimleri ayırt edici özellik taşıması nedeniyle kümeleme analizinde

önemli olduğu bilinmektedir. Bu nedenle her yıl için değişkenler arasındaki Pearson korelasyon katsayıları ve anlamlılık dereceleri hesaplanmıştır.

Tablo 4.1.'de her yıl için hesaplanan korelasyon katsayı değerleri ve anlamlılık dereceleri yer almaktadır. Tabloya bakıldığında GSYİH ve ihracat ithalat oranı değişkenleri arasındaki korelasyonların her yıl için %70'in üzerinde olduğu görülmektedir. Bu değerlerin %1 ve %5 'e göre anlamlı olduğu söylenebilir. GSYİH ve işsizlik rakamları değişkenlerine bakıldığında ise korelasyon katsayılarının yıllara göre %40 ile %56 arasında değiştiği görülmektedir. Bu değerler ise %5 anlamlılık seviyesine göre anlamlıdır. İhracat ithalat oranı ve işsizlik değişkenleri arasındaki korelasyonlar ise %40 ile %50 arasında değişmektedir. 2016 ve 2015 yıllarındaki katsayılar %1 ve %5'e göre anlamlıyken; 2013 ve 2014 yıllarındaki değerler ise %5'e göre anlamlıdır.

Genel olarak değişkenler arasında ortalamanın üzerinde ilişki olduğu söylenebilir. Her yıl için katsayıların anlamlı olduğu söylenebilir. Dolayısıyla bu değişkenler kullanılarak kümeleme analizi yapılabilir.

Yıllar	GDP/EXPTÖİMP		GDP/UNEMP		EXPTÖİMP/UNEMP	
	Pearson Korelasyon Katsayısı	Anlamlılık	Pearson Korelasyon Katsayısı	Anlamlılık	Pearson Korelasyon Katsayısı	Anlamlılık
2016	0,743	0,000	-0,465	0,036	-0,5	0,005
2015	0,766	0,000	-0,456	0,02	-0,496	0,002
2014	0,7	0,000	-0,4	0,04	-0,450	0,04
2013	0,732	0,000	-0,45	0,02	- 0,4	0,02

Tablo 4.1. Kümeleme Analizi Değişkenlerinin Arasındaki Korelasyon Katsayılarının Anlamlılık Testleri.

Ülkelerin ekonomik büyüklük düzeyini belirlemek için kullanılan GSYİH, ihracat ithalat oranı ve işsizlik değişkenlerinin anlamlı olup olmadığı sınanmıştır. Tablo 4.2.'de her yıl için ve her değişken için hesaplanan F istatistikleri ve anlamlılık seviyeleri yer almaktadır. Tablodan da görüldüğü gibi, %5 ve %1 hata payı ile tüm değişkenler istatistiksel olarak anlamlıdır. Bu doğrultuda, seçilen

değişkenlerin ülkeleri ekonomik gelişmişlik açısından sınıflandırmak için uygun olduğu söylenebilir.

Yıllar	GDP		UNEMP		EXPTOİMP	
	F İstatistiği	Anlamlılık	F İstatistiği	Anlamlılık	F İstatistiği	Anlamlılık
2016	141,176	0,000	4,298	0,022	15,529	0,000
2015	164,831	0,000	4,613	0,017	17,032	0,000
2014	173,056	0,000	6,692	0,004	9,745	0,000
2013	178,819	0,000	7,505	0,002	12,161	0,000

Tablo 4.2. Kümeleme Analizi Değişkenlerinin Anlamlılık Testleri.

Tablo 4.3.'de değişkenlerin ortalama değerleri yani küme merkezleri yer almaktadır. Böylece kümelerin içindeki gözlemlerin özelliklerini ortaya koymak mümkün olabilmektedir. Tüm yıllar için GSYİH değişkeni en iyi olan ülkeler 3.kümede yer almaktadır. Aynı şekilde ihracat ve ithalat oranı en iyi olan ve işsizliği en düşük olan ülkelerde 3. Kümede yer almaktadır. Bu nedenle 3.küme de ekonomik büyüklük seviyesi yüksek olan ülkeler yer almaktadır diyebiliriz. Tablodaki değerlere bakıldığında 1.kümede 2.kümedeki ülkelere göre daha iyi değerler olduğu için ekonomik büyüklük seviyesi ortalamanın üzerinde olan ülkeler olarak adlandırabiliriz. 2.kümedeki ülkeler ise diğer ülkelere göre ekonomik olarak seviyesi düşük olan ülkelerdir.

Yıllar	GDP			EXPTOİMP			UNEMP		
	1.Küme	2.Küme	3.Küme	1.Küme	2.Küme	3.Küme	1.Küme	2.Küme	3.Küme
2013	36341	12350	75433	1,08	0,93	1,31	7,99	12,78	5,03
2014	36825	12450	75336	1,10	0,96	1,28	8,18	13,11	5,17
2015	37891	12730	75400	1,06	0,89	1,32	7,93	11,66	4,47
2016	41627	13823	76666	1,07	0,9	1,3	7,16	10,73	4,73

Tablo 4.3 Kümeleme Analizi Değişkenlerinin Küme Merkezleri.

Tablo 4.4'de her yıl için yapılan kümeleme analizinde hangi ülkelerin hangi gruplara dâhil olduğu yer almaktadır.

Ülkeler	2013	2014	2015	2016
Belçika	1	1	1	1
Bulgaristan	2	2	2	2
Çek Cumhuriyeti	2	2	2	2
Danimarka	1	1	1	1
Almanya	1	1	1	1
Estonya	2	2	2	2
İrlanda	1	1	1	1
Yunanistan	2	2	2	2
İspanya	1	1	1	2
Fransa	1	1	1	1
Hırvatistan	2	2	2	2
İtalya	1	1	1	2
Kıbrıs	2	2	2	2
Letonya	2	2	2	2
Litvanya	2	2	2	2
Lüksemburg	3	3	3	3
Macaristan	2	2	2	2
Malta	2	2	2	2
Hollanda	1	1	1	1
Avusturya	1	1	1	1
Polonya	2	2	2	2
Portekiz	2	2	2	2
Romanya	2	2	2	2
Slovenya	2	2	2	2
Slovakya	2	2	2	2
Finlandiya	1	1	1	1
İsveç	1	1	1	1
İngiltere	1	1	1	1
İzlanda	1	1	1	1
Norveç	3	3	3	3
İsviçre	3	3	3	3

Karadağ	2	2	2	2
Makedonya	2	2	2	2
Sırbistan	2	2	2	2
Türkiye	2	2	2	2

Tablo 4.4. Kümeleme Analizi Gruplama Sonuçları.

Her yıl için genel olarak aynı kümeleme analizi sonuçlarının elde edildiği görülmektedir. Sadece önceki yıllarda 1.kümede yer alan İtalya ve İspanya 2016 yılında 2.kümeye geçmiştir. Yani İtalya'nın ve İspanya'nın ekonomik olarak gerilediği söylenebilir. 1. grupta ekonomisi ortalamanın üzerinde olan ülkeler yer almaktadır. Bu ülkeler Belçika, Danimarka, Hollanda, İrlanda, Fransa, Avusturya, İsveç, İngiltere, İzlanda'dır. Ekonomik seviyesi yüksek olan ülkeler ise 3.grupta yer almaktadır. Bu ülkeler sadece 3 tanedir. Lüksemburg, Norveç ve İsviçre ülkeleri bu grubu oluşturmaktadır. Aralarında Türkiye' nin de bulunduğu 2.grupta ekonomik olarak az gelişmiş ülkeler yer almaktadır. Bulgaristan, Çek Cumhuriyeti, Estonya Yunanistan, ispanya, Hırvatistan, İtalya, Kıbrıs, Letonya, Litvanya, Makedonya, Malta, Polonya, Portekiz, Romanya, Slovenya, Slovakya, Karadağ, Yugoslavya ve Sırbistan bu grupta yer almaktadır.

4.3. Veri Madenciliği Uygulama Süreci

Veri madenciliği uygulama süreci önceki bölümlerde de açıklandığı gibi küçük ve birbirleriyle bağlantılı adımlardan oluşan bir süreçtir. Bu süreçte önemli olan verilerin kendisi ve doğruluğudur. Verilere bağlı olarak seçilen modeller ve algoritmalarda titiz davranmak gerekmektedir. Doğru ve güvenilir sonuçlara ulaşmak için veri madenciliği sürecinde yer alan adımların doğru ve eksiksiz bir biçimde yerine getirilmesi gerekmektedir.

Veri madenciliğinin her bir aşaması sonraki diğer aşamasıyla bağlantılıdır. Veri madenciliği sürecinin birbirine bağımlı süreçlerden oluşması , veri yığınları arasında, soyut kazılar yapılarak veriyi ortaya çıkarmanın yanı sıra, bilgi keşfi sürecinde örüntülerin ayrıştırılarak süzülmesi ve bir sonraki adıma hazır hale getirilmesini ifade etmektedir.

Veri madenciliği uygulama aşamasında SPSS clementine programı kullanılmıştır. Tez çalışmasının uygulama süreci ayrıntılı olarak alt başlıklarda ele alınmıştır.

4.3.1. Problemin Tanımlanması

Veri madenciliği uygulamalarında ilk olarak ele alınacak problemin tüm çerçevesiyle ele alınması gerekmektedir. Burada amaç veri madenciliği uygulamasının ne amaçla yapılacağı ve bu amaca ulaşmak isteyen birimin tanımlanmasıdır. Bu nedenle amaç doğru ve eksiksiz bir biçimde açıklanmalıdır. Uygulama sonunda elde edilen sonuçların amaca ne derecede katkı sağlayacağı öngörülebilmelidir.

Tez çalışmasının uygulama kısmında Avrupa Birliği üye ülkelerinin ve Türkiye'nin 2013-2016 yılları arasındaki 34 bilgi ekonomisi değişkeni, her yıl için GSYİH, işsizlik ve ihracat ithalat oranı değişkenleri ile hesaplanan kümeleme analizi uygulama sonuçları göz önüne alınarak veri madenciliği tekniklerinden karar ağacı uygulaması gerçekleştirilmiştir. Yapılan karar ağacı analizi ile genel ekonomik seviyesini etkileyen en önemli bilgi ekonomisi değişkenlerinin belirlenmesi hedeflenmiştir. Sonraki aşamada ise ülkelerin ekonomik büyüklük seviyeleri, coğrafi konumları, refah seviyeleri ve karar ağaçları ile belirlenen bilgi ekonomisi değişkenleri göz önüne alınarak Kohonen ağları ile kümeleme analizi yapılmıştır.

4.3.2. Veri Setinin Hazırlanması

Uygulamada kullanılan veri setinin hazırlık aşaması bu bölümü oluşturmaktadır. Bu bölümde verilerin hangi kaynaktan nasıl toplanıp bir araya getirildiği, hangi değişkenlerin dönüşümü yapıldığı alt başlıklar halinde anlatılmıştır.

4.3.2.1. Verilerin Toplanması

Uygulama bölümünün ilk aşamasında yapılan kümeleme analizinde kullanılan GSYİH, ihracat ithalat oranı, işsizlik değişkenleri Eurostat isimli internet sitesinden alınmıştır. Veri madenciliği tekniklerinden karar ağacı uygulamasında kullanılmak üzere yapılan kümeleme analizi sonuçlarına ek olarak 34 bilgi ekonomisi değişkeni verileri de adı geçen aynı siteden sağlanmıştır.

Veri madenciliği tekniklerinden Kohonen ağırları ile yapılacak uygulamada refah ülkesi olan ve olmayan ülkeler tespit edilmiştir. Bu veriler ise Legatum Institute isimli web sitesinden alınmıştır ve bu uygulamaya değişken olarak eklenmiştir. Üyelik yılları ile coğrafi bölge değişkenleri ile ilgili bilgiler ise wikipedia isimli internet sitesinden alınmıştır.

4.3.2.2.Verinin Önişlenmesi

Değişkenlerde yer alan eksik veriler doktora çalışmasında kullanılmak üzere web sitesinden istenmiş olup bir kısmı temin edilmiştir. Ancak bir kısmı temin edilememiştir. Eksik olan verileri olan değişkenler ve ülkeler analize dahil edilmemiştir. Bu nedenle Türkiye ile birlikte 35 Avrupa Birliği üye ülkesi çalışmaya dâhil edilebilmiştir. Toplamda 34 tane bilgi ekonomisi değişkeni ve kümeleme analizi sonucu elde edilen ekonomik seviye değişkeni veriler arasında yer almaktadır. Veriler adı geçen internet sitesinden Excel dosyası halinde her değişken için farklı bir dosyada alınmıştır. SPSS dosyasında alınan bütün veriler bir araya getirilmiştir.

Ayrıca Kohonen ağırları ile yapılacak olan analizde kullanılmak üzere ülkelerin refah verileri, üyelik yılı ve coğrafi bölge verileri de yer almaktadır. Refah seviyesi verileri refah ülkesi olan ve olmayan ülkeler olarak ayrılmıştır. Üyelik yılı değişkeni ise ülkelerin 2000 yılından önce üye olan ve 2000 yılından sonra üye olan ülkeleri ifade etmektedir. Coğrafi bölge değişkeninde ise ülkeler bulundukları konuma göre kuzey, güney, doğu ve batı olarak gruplara ayrılmıştır.

Kümeleme analizinde kullanılan değişkenler 2013-2016 yılları arasında ülkelerin GSYİH, istihdam ve ihracat ithalat oranı verileridir.

Karar ağacı uygulamasında ise 34 tane bilgi ekonomisi değişkeni kullanılmıştır. Bu değişkenler uygulamada d1,d2....olarak kodlanmıştır. Bu değişkenlerin açıklamaları ise;

D1: Düzenli olarak internet kullanan bireylerin yüzdesi

D2: E-bankacılık ve e-ticaret kullanım yüzdesi

D3: Ağ bağlantısına sahip bilgisayarlar üzerinden sipariş alan girişimlerin yüzdesi

D4: Evde internet erişimi olan hane halkı yüzdesi

D5: Genişband internet erişimine sahip hane halkı

D6: Genişband internet erişimine sahip girişimler

D7: Genişband internet erişimine sahip çalışanlar

D8: Fonksiyonel departmanlar arasındaki bilgi paylaşımını sağlamak için kurumsal kaynak planlama yazılımlarına (ERP) sahip olan girişimlerin yüzdesi

D9: Bilgisayar kullanan girişimlerin yüzdesi

D10: İnternet ya da internet dışında başka ağlar aracılığıyla satış yapan girişimlerin yüzdesi

D11: İstihdam edilen bireylerin internete erişim yüzdesi

D12: Web site aracılığıyla kamu kuruluşlarından bilgi alma yüzdesi

D13: İnternet üzerinden form doldurma ve gönderme yüzdesi

D14: Kamu kuruluşlarıyla iletişim kurabilmek için interneti kullanan bireylerin yüzdesi

D15: Girişimlerin e-ticaret iş hacmi

D16: Online satın alım yapan girişimlerin yüzdesi

D17: Web siteler aracılığıyla yapılan toplam satışlar

D18: İnternet ya da internet dışında başka ağlar aracılığıyla yapılan satışlar ve

satın alımların değeri

D19: Web siteler aracılığıyla yapılan toplam satışlar

D20: İnternet üzerinden satın alım yapan bireylerin yüzdesi

D21: İleri teknoloji ihracat yüzdesi

D22: Toplam istihdam içerisinde bilgi ve iletişim teknolojileri personeli yüzdesi

D23: Mobil internet kullanım yüzdesi

D24: Toplam istihdam içerisinde AR-GE personeli yüzdesi

D25: BİT eğitilmiş çalışanları olan girişimlerin yüzdesi

D26: Toplam AR-GE personel yüzdesi

D27: Girişimlerin toplam AR-GE gideri

D28: Patent uygulamalarının sayısı

D29: Telekomünikasyon sektörü toplam gelirleri

D30: Telekomünikasyon sektörü toplam yatırımları

D31: Toplam istihdam içerisinde BİT personeli yüzdesi

D32: Yüksek teknoloji ticaret hacmi

D33: Çalışanlarına BİT eğitimi veren girişimlerin yüzdesi

D34: GSYİH'da AR-GE gideri yüzdesi şeklinde ifade edilmektedir.

Tablo 4.5.'de verilerin gösterimi ayrıntı şekilde yer almaktadır.

Table (58 fields, 140 records) #2

File Edit Generate

	ulkeler	yil	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10	d1
1	Belcika	01-Oca-...	7...	56.00	24...	78...	7...	94...	5...	33...	34...	24...	50...
2	Bulgari...	01-Oca-...	5...	4.00	7.00	51...	5...	79...	2...	20...	22...	7.00	22...
3	Cek Cu...	01-Oca-...	6...	34.00	26...	65...	6...	92...	3...	24...	25...	26...	34...
4	Danima...	01-Oca-...	8...	79.00	29...	92...	8...	93...	6...	33...	33...	29...	64...
5	Almanya	01-Oca-...	7...	45.00	24...	85...	8...	91...	5...	24...	25...	24...	52...
6	Estonya	01-Oca-...	7...	68.00	13...	74...	7...	96...	4...	10...	10...	13...	44...
7	Irlanda	01-Oca-...	7...	43.00	23...	81...	6...	90...	4...	19...	20...	23...	46...
8	Yunani...	01-Oca-...	5...	9.00	9.00	54...	5...	80...	3...	12...	14...	9.00	33...
9	Ispanya	01-Oca-...	6...	31.00	14...	67...	6...	96...	4...	22...	22...	14...	47...
10	Fransa	01-Oca-...	7...	54.00	14...	80...	7...	98...	4...	33...	34...	14...	45...
11	Hirvatist...	01-Oca-...	5...	21.00	27...	66...	6...	90...	3...	19...	20...	27...	38...
12	Italya	01-Oca-...	5...	21.00	6.00	63...	5...	94...	3...	21...	22...	6.00	37...
13	Kibris	01-Oca-...	5...	21.00	8.00	62...	6...	95...	3...	21...	22...	8.00	36...
14	Letonya	01-Oca-...	7...	47.00	9.00	69...	6...	88...	3...	10...	11...	9.00	39...
15	Litvanya	01-Oca-...	6...	43.00	17...	60...	6...	99...	3...	23...	23...	17...	39...
16	Lukse...	01-Oca-...	9...	63.00	16...	93...	6...	96...	5...	23...	23...	16...	55...
17	Macaris...	01-Oca-...	6...	26.00	12...	67...	6...	87...	3...	9.00	10...	12...	33...
18	Malta	01-Oca-...	6...	41.00	18...	77...	7...	95...	3...	24...	25...	18...	37...
19	Hollanda	01-Oca-...	9...	80.00	20...	94...	8...	98...	5...	26...	26...	20...	57...
20	Avusturya	01-Oca-...	7...	45.00	16...	79...	7...	91...	4...	26...	26...	16...	43...

Table Annotations

Tablo 4.5 .Bilgi ekonomisi değişkenlerinin gösterimi.

Analize dâhil edilen değişkenler SPSS programında bir araya getirilirken eksik verili değişken ve ülkeler alınmamıştır. Tablo 4.6.'da eksik verinin bulunmadığı görülmektedir.

Quality of [39 fields] #6		
Field	% Complete	Valid Records
ulkeler	100	140
yil	100	140
d1	100	140
d2	100	140
d3	100	140
d4	100	140
d5	100	140
d6	100	140
d7	100	140
d8	100	140
d9	100	140
d10	100	140
d11	100	140
d12	100	140
d13	100	140
d14	100	140
d15	100	140
d16	100	140
d17	100	140
d18	100	140
d19	100	140
d20	100	140
d21	100	140
d22	100	140
d23	100	140
d24	100	140
d25	100	140
d26	100	140
d27	100	140
d28	100	140
d29	100	140
d30	100	140
d31	100	140
d32	100	140
d33	100	140
d34	100	140
ekonomikbyklk	100	140
cografibolge	100	140
refahsvy	100	140

Tablo 4.6. Değişkenlere Ait Veri Sayıları.

4.3.2.3.Verilerin Dönüştürülmesi

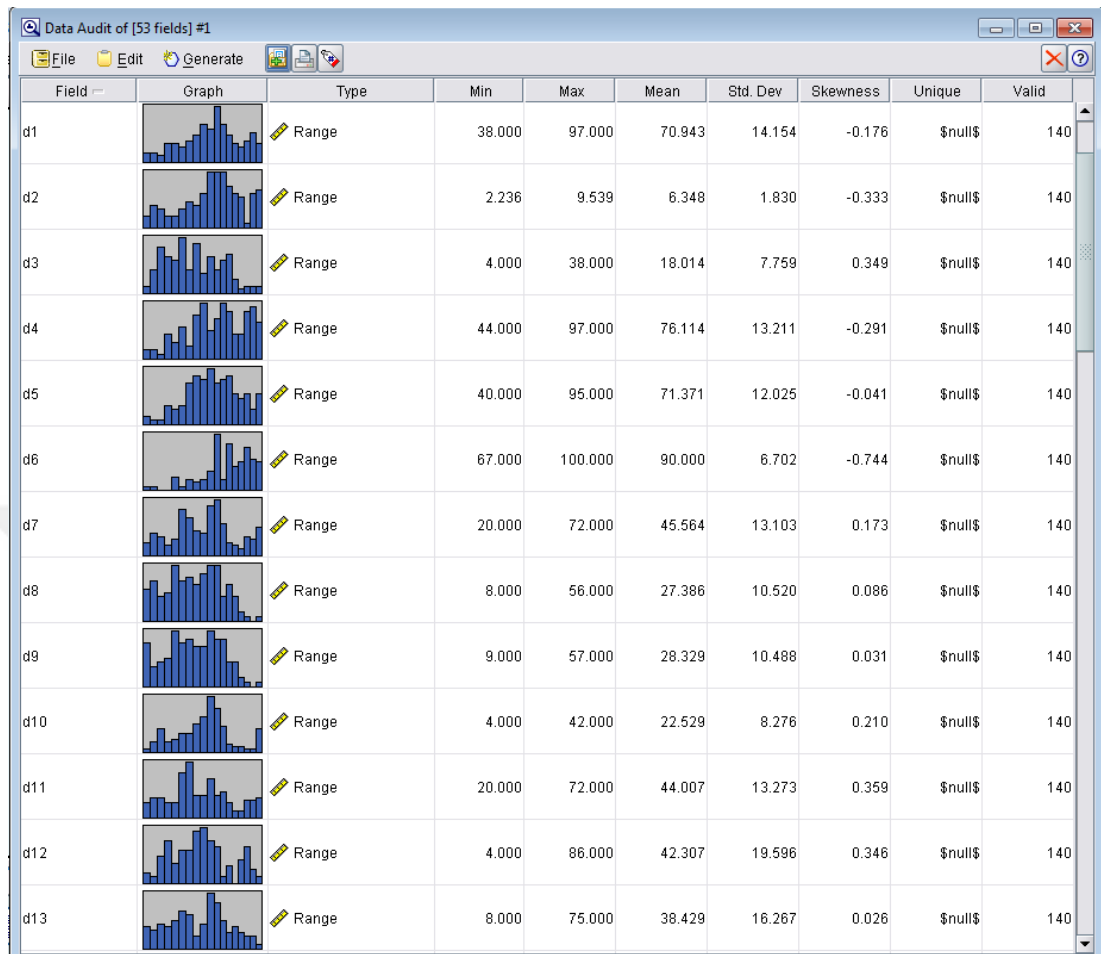
Uygulama bölümünde ilk olarak kümeleme analizi yöntemi kullanılmıştır. Bu amaçla, ilk aşamada ülkeler 1, 2 ve 3 olarak gruplandırılmıştır. 3. grupta yer alan ülkeler ekonomik seviye olarak gelişmiş olan ülkeler, 1. grupta yer alan ülkeler ekonomik seviye açısından ortalama değerlere sahip olan ülkeler, 2. grupta ise nispeten daha zayıf ekonomik seviyeye sahip olan ülkeler yer almaktadır. Bu şekilde gruplandırılan ülkeler karar ağacı analizinde 34 tane bilgi ekonomisi değişkeniyle birlikte kullanılmıştır.

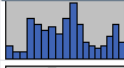
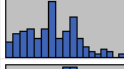
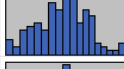
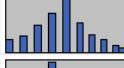
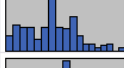
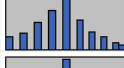
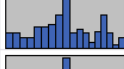
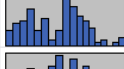
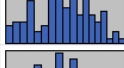
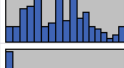
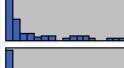
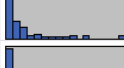
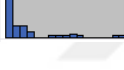
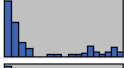
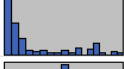
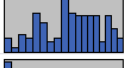
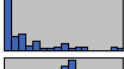
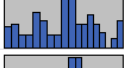
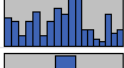
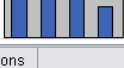
Yapılan Kohonen ağırları ile kümeleme analizinde refah seviyesi değişkeni refah ülkesi olan ülkeler 1 ile olmayan ülkeler ise 0 ile gösterilmiştir. Ayrıca coğrafi bölge değişkeni de 4 grup halinde yer almıştır. 1.grupta batı Avrupa ülkeleri, 2.grupta doğu Avrupa ülkeleri, 3.grupta kuzey Avrupa ülkeleri son olarak 4. Grupta ise güney Avrupa ülkeleri yer almaktadır. Üyelik yılı değişkeni ise ülkelerin 2000 yılı öncesi ve sonrası üye olan ülkelere göre belirlenmiştir. 2000 yılı öncesinde üye olan ülkeler 1 olarak kodlanırken, 2000 sonrasındaki ülkeler ise 0 ile kodlanmıştır.

	ulkeler	d20	d33	d24	d34	refahse...	cogrfibo...	uyelikyil
1	Belcika	42.00	34.00	1.76	1.83	1.00	1	1.00
2	Bulgari...	12.00	8.00	0.85	0.72	0.00	2	0.00
3	Cek Cu...	26.00	22.00	1.31	1.45	0.00	2	0.00
4	Danima...	67.00	28.00	2.19	2.26	1.00	4	1.00
5	Almanya	64.00	29.00	1.64	2.14	1.00	1	1.00
6	Estonya	46.00	13.00	0.94	1.12	0.00	4	0.00
7	Irlanda	44.00	30.00	1.85	1.16	1.00	4	1.00
8	Yunani...	24.00	15.00	1.14	0.72	0.00	2	1.00
9	Ispanya	32.00	23.00	1.13	0.91	1.00	3	1.00
10	Fransa	49.00	20.00	1.30	1.66	1.00	1	1.00
11	Hirvatist...	26.00	22.00	0.68	0.63	0.00	2	0.00
12	Italya	18.00	12.00	1.16	0.99	0.00	3	1.00
13	Kibris	19.00	22.00	0.36	0.34	0.00	2	0.00
14	Letonya	27.00	12.00	0.59	0.47	0.00	4	0.00
15	Litvanya	22.00	10.00	0.81	0.77	0.00	4	0.00
16	Lukse...	63.00	29.00	2.01	2.09	1.00	1	1.00
17	Macaris...	23.00	16.00	0.83	1.03	0.00	2	0.00
18	Malta	43.00	23.00	0.76	0.57	1.00	3	0.00
19	Hollanda	59.00	22.00	1.62	1.50	1.00	1	1.00
20	Avusturya	46.00	37.00	1.78	2.29	1.00	1	1.00

Tablo 4.7.Kohonen Ağları ile Kümeleme Uygulaması İçin Düzenlenen Veriler.

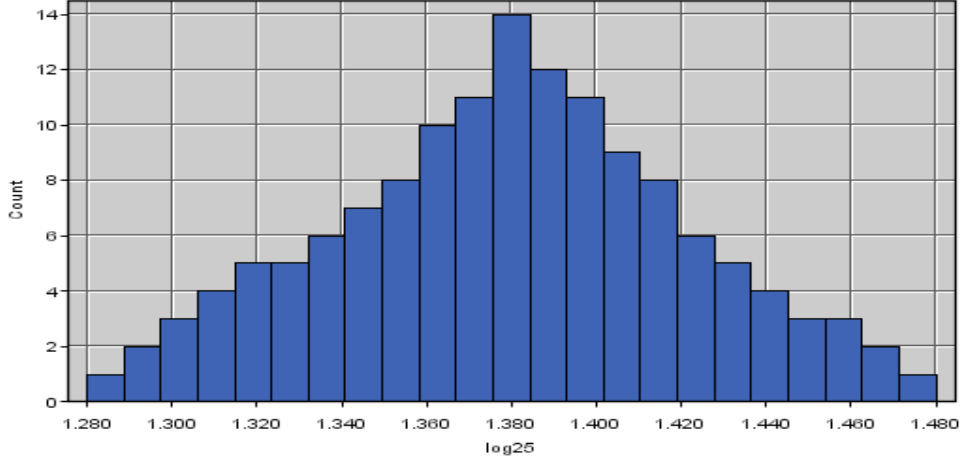
Tablo 4.8.'de verilerin histogramları ve temel istatistik değerleri yer almaktadır. Tablodan da görüldüğü gibi d25,d26,d27,d28,d29,d30 ve d32 değişkenlerinin normal dağılmadığı görülmektedir. Bu değişkenlerin normal dağılıma yaklaştırılması gerekmektedir. Bu nedenle adı geçen değişkenlere logritmik dönüşüm uygulanmıştır. Tabloda verilerin türleri de yer almaktadır. Bilgi ekonomisi değişkenleri sayısal verilerden oluştuğu için "Range" olarak işaretlenmiştir. Kohonen ağları analizinde kullanılacak olan refah ve üyelik yılı değişkenleri ise mantıksal değerleri ifade ettikleri için "Flag" olarak işaretlenmiştir. Karar ağacı analizinde kullanılacak olan ekonomik büyüklük ve Kohonen ağları analizinde kullanılacak olan coğrafi bölge değişkeni ise kategorik değerleri ifade ettiği için "Set" olarak işaretlenmiştir.



d14		 Range	5.000	88.000	46.507	21.017	0.222	\$null\$	140
d15		 Range	1.000	37.000	13.664	7.296	0.438	\$null\$	140
d16		 Range	1.000	36.000	16.914	7.396	0.086	\$null\$	140
d17		 Range	1.000	10.000	4.786	1.996	0.223	\$null\$	140
d18		 Range	1.000	37.000	13.529	7.485	0.414	\$null\$	140
d19		 Range	1.000	10.000	4.786	1.996	0.223	\$null\$	140
d20		 Range	3.000	75.000	37.086	17.662	0.081	\$null\$	140
d21		 Range	1.000	29.600	12.642	6.395	0.160	\$null\$	140
d22		 Range	5.000	55.000	28.300	11.483	-0.024	\$null\$	140
d24		 Range	5.000	45.000	21.836	9.517	0.387	\$null\$	140
d25		 Range	1.700	453.900	74.918	103.096	1.805	\$null\$	140
d26		 Range	321.000	379140.000	45781.864	76435.823	2.825	\$null\$	140
d27		 Range	104.068	59058.000	5907.448	11204.989	3.138	\$null\$	140
d28		 Range	101.560	22963.170	1887.531	3899.224	4.051	\$null\$	140
d29		 Range	235.000	60400.000	14791.493	19686.365	1.263	\$null\$	140
d30		 Range	32.000	8190.000	1753.214	2232.225	1.320	\$null\$	140
d31		 Range	2.390	4.880	3.742	0.614	-0.185	\$null\$	140
d32		 Range	244.000	76236.000	12364.164	16347.254	1.950	\$null\$	140
d33		 Range	0.100	2.100	1.069	0.512	0.048	\$null\$	140
d34		 Range	0.100	3.420	1.647	0.846	0.101	\$null\$	140
ekonomikbyklk		 Set	1.000	3.000	--	--	--	3	140
refahsvy		 Flag	0.000	1.000	--	--	--	2	140
uyelikyil		 Flag	0.000	1.000	--	--	--	2	140
cgrfblg		 Set	1	4	--	--	--	4	140

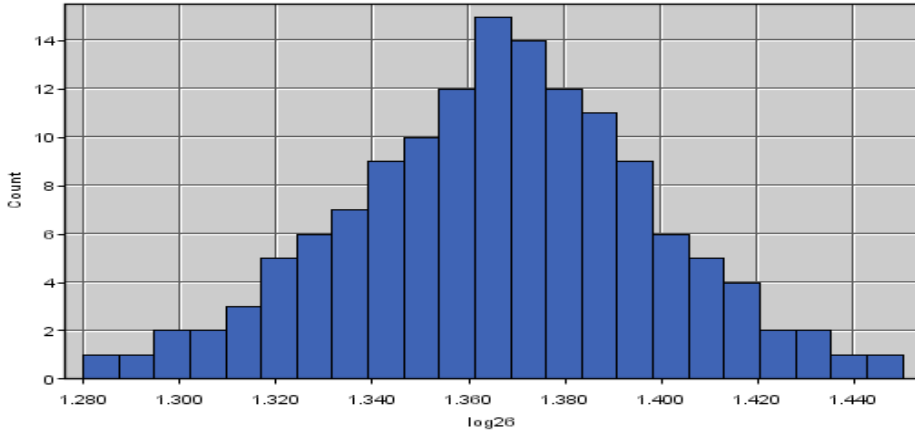
Tablo 4.8. Verilere ait Temel İstatistikler.

Bilgi ekonomisi değişkenlerinden d25 değişkeni çarpık dağılım gösterdiği için logaritmik dönüşümle normal dağılıma yaklaştırılmıştır.



Tablo 4.9. Logaritmik Dönüşüm Sonrası D25 Değişkenine ait Dağılım Grafiği.

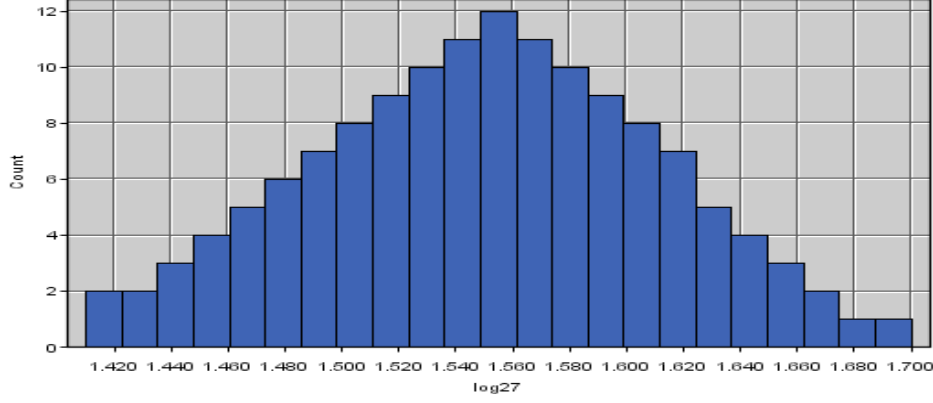
Yapılan tanımlayıcı istatistikler sonucu d26 değişkeninin de çarpık dağılıma sahip olduğu görülmektedir. Yapılan logaritmik dönüşümle aşağıdaki tablodan da görüldüğü gibi normal dağılıma yaklaştırılmıştır.



Tablo4.10. Logaritmik Dönüşüm Sonrası D26 Değişkenine ait Dağılım Grafiği.

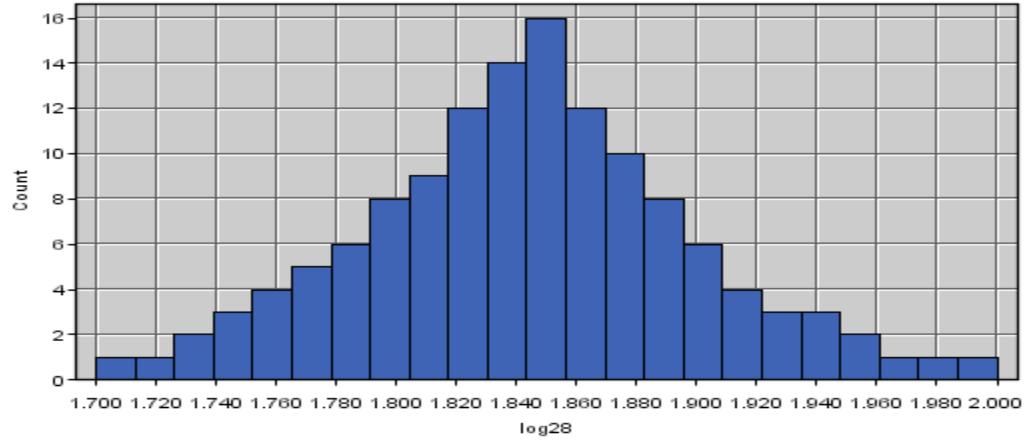
Çarpık dağılıma sahip değişkenlerden biri de d27 değişkeni olarak görülmektedir. Yine d27 değişkenini de yapılan logaritmik dönüşümle değişkenin

dağılımı normal dağılıma yaklaştırılmıştır. Aşağıdaki tabloda dağılımın grafiği yer almaktadır.



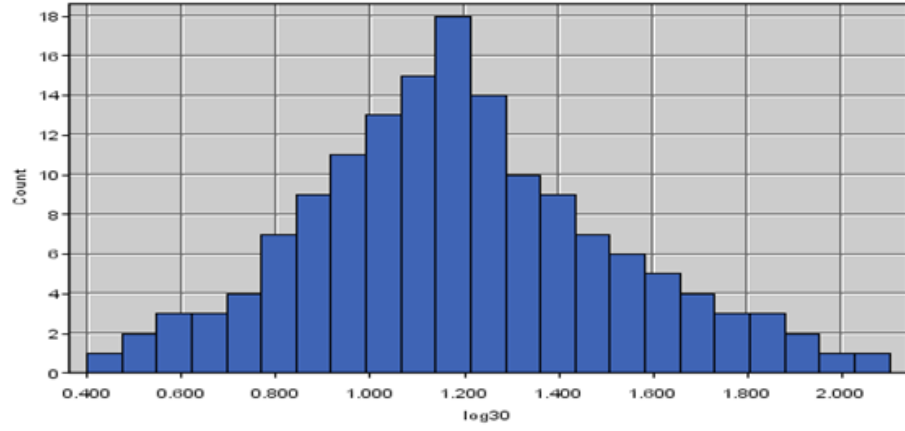
Tablo 4.11. Logaritmik Dönüşüm Sonrası D27 Değişkenine ait Dağılım Grafiği.

Dönüşüm yapılan değişkenlerden biri de d28 değişkenidir. Bu değişkeninde dağılımının çarpık olmasından dolayı logaritmik dönüşümle, aşağıdaki grafikten de görüldüğü gibi normal dağılıma yaklaştırılması sağlanmıştır.



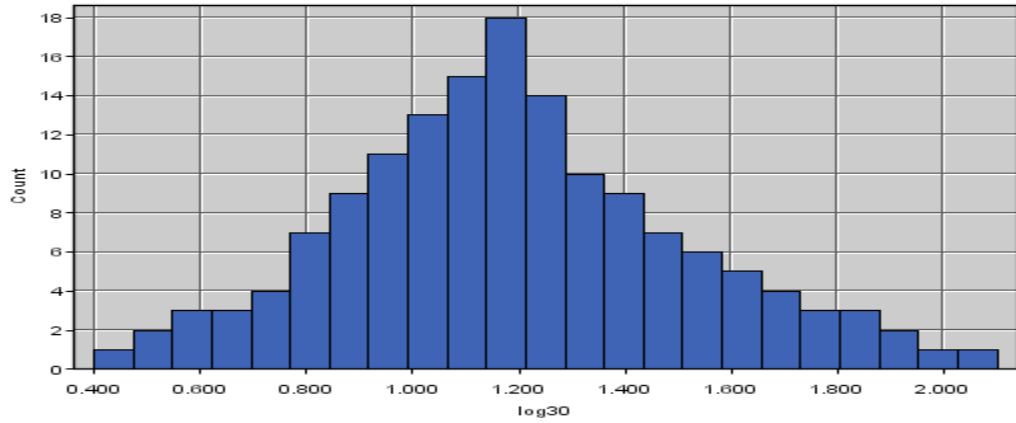
Tablo 4.12. Logaritmik Dönüşüm Sonrası D28 Değişkenine ait Dağılım Grafiği.

Şekil 4.13.'te ise çarpık dağılıma sahip olan ve logaritmik dönüşümle normal dağılıma yaklaştırılan d29 değişkeninin dağılım grafiği yer almaktadır.



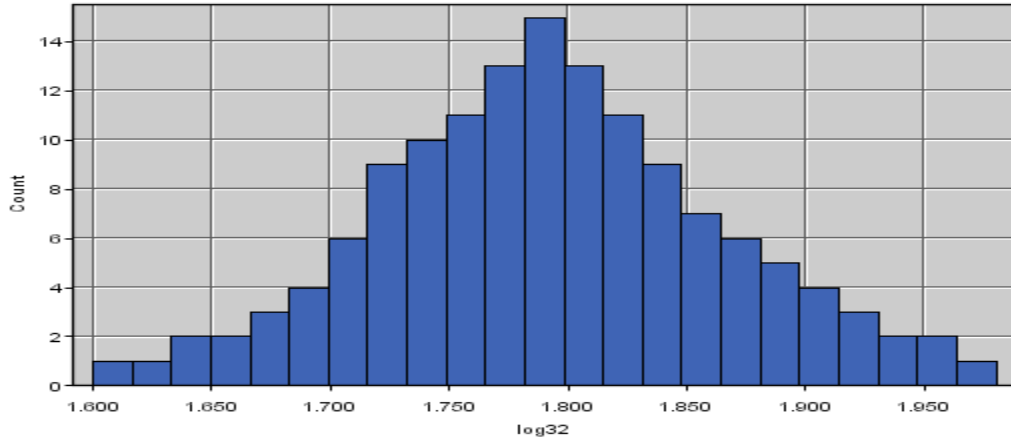
Tablo 4.13. Logaritmik Dönüşüm Sonrası D29 Değişkenine ait Dağılım Grafiği.

Aşağıdaki şekilde ise çarpık dağılıma sahip olan ve logaritmik dönüşümle normal dağılıma yaklaştırılan d30 değişkeninin dağılım grafiği yer almaktadır.



Tablo 4.14. Logaritmik Dönüşüm Sonrası D30 Değişkenine ait Dağılım Grafiği.

Son olarak d32 değişkeni de logaritmik dönüşümle birlikte normal dağılıma yaklaştırılmıştır. Aşağıdaki şekilde bu değişkenin dağılım grafiği yer almaktadır.



Tablo 4.15. Logaritmik Dönüşüm Sonrası D32 Değişkenine ait Dağılım Grafiği.

4.3.3. Modelin Kurulması

Doğru ve tutarlı sonuçlar elde etmek açısından kullanılan modeller çok önemlidir. Bu aşama önceki bölümlerde de anlatıldığı gibi veri tabanlarında bilgi keşfi sürecinin en önemli adımıdır. Bu aşamada veri madenciliği uygulamalarından karar ağaçları ve Kohonen ağları kullanılmıştır. Sonuçlar ayrıntılı olarak alt başlıklarda yer almaktadır.

4.3.3.1. Bilgi Ekonomisi ve Ekonomik Büyüklük İle İlgili Karar Ağacı Uygulaması

Ülkelerin ekonomik seviyelerine göre hangi bilgi ekonomisi değişkenlerinin ön plana çıktığını ve bu açıdan profillerini belirlemek için karar ağacı kullanılmıştır. Önceki bölümlerde açıklandığı gibi 34 adet bilgi ekonomisi değişkeni ve ekonomik seviye değişkenleri karar ağaçları uygulamasında kullanılacaktır. Ekonomik seviye değişkeni ülkelerin ekonomik yapısını düşük, orta ve yüksek olmak üzere 3 kategoriye ayırmaktadır.

Çalışmanın karar ağaçları uygulamasında C5.0 algoritması kullanılmıştır. Sayısal niteliklerin de karar ağacı oluşturmasına olanak sağlayan bir algoritmadır. C5.0 algoritması denetimli bir sınıflandırma algoritması olup, hedef değişkenin bulunduğu bir veri kümesi ile karar ağacı eğitilip bir model oluşturulmaktadır. C5.0 algoritması sınıflandırma işlemini gerçekleştirdiği için ilk önce karar ağacı eğitilmektedir. Elde edilen model yeni bir veri setiyle de test edilmekte ve modelin

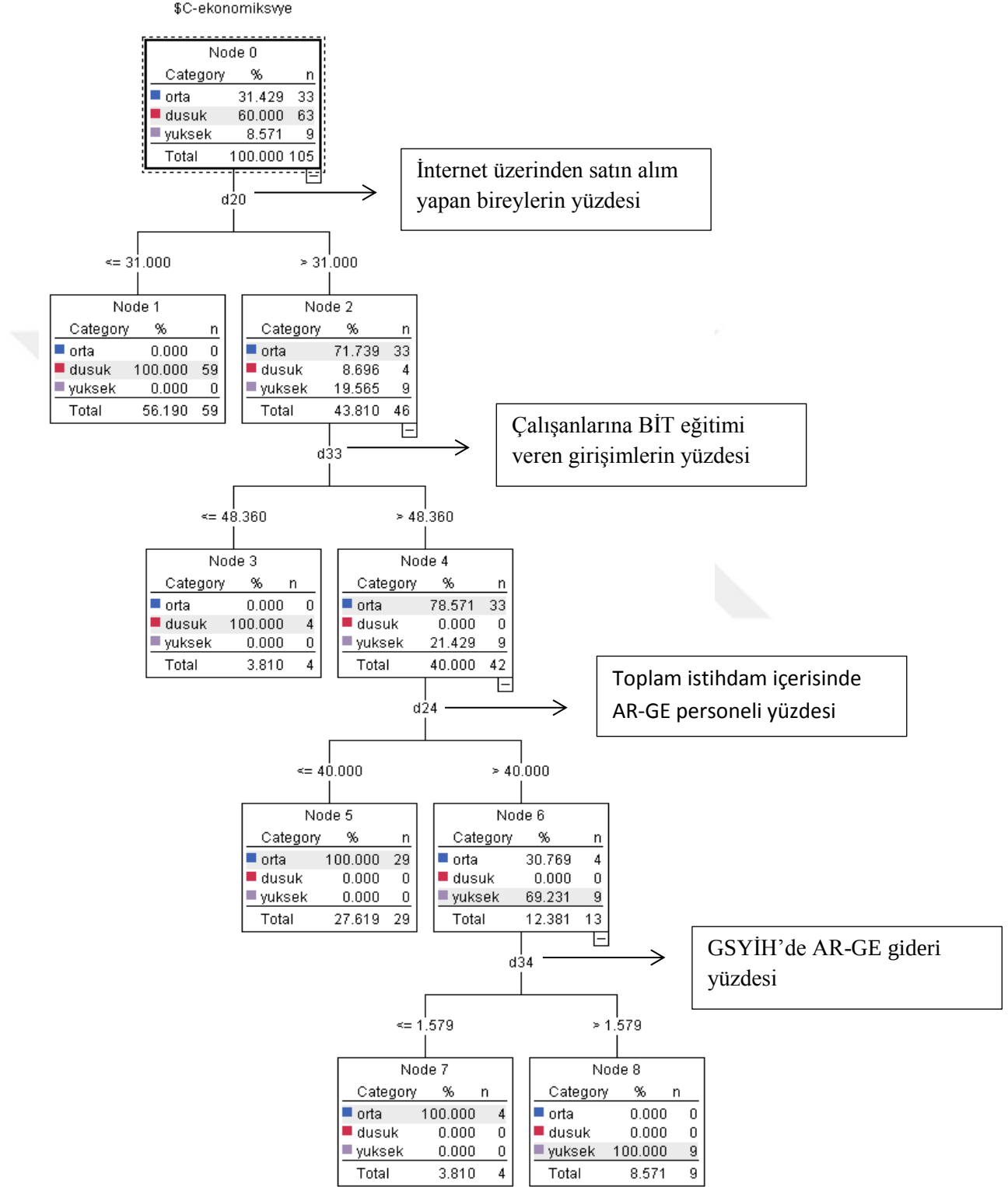
performansı gözlenmektedir. Bu nedenle eğitim veri seti olarak 140 adet veri içinden 105 tanesi rastgele seçilmiştir ve model oluşturulmuştur.

Önceki uygulama adımında yapılan kümeleme analizi sonucu ekonomik büyüklük seviyelerine göre gruplandırılan ülkeler bu aşamada hedef değişken olarak kullanılmıştır. Dolayısıyla hedef değişken ekonomik seviye olarak belirlenmiş olup 1, 2 ve 3 olarak gruplandırılmıştır. 1. grupta ekonomik açıdan orta seviyede olan ülkeler, 2. grupta ekonomik ortalamanın altında olan gelişmemiş ülkeler ve 3. grupta ekonomik açıdan gelişmiş olan ve ortalama seviyenin üzerindeki ülkeler yer almaktadır.

Tahminleyici alanlar olarak ise bilgi ekonomisi değişkenleri kullanılmıştır. Girdi değişkeni olarak kullanılan bilgi ekonomisi değişkenlerinin ayrıntılı açıklamaları önceki başlıklarda açıklanmıştır. Oluşturulan karar ağacı modeli aşağıdaki şekilde yer almaktadır.

Karar ağacı modeline göre bilgi ekonomisi değişkenlerinden internet üzerinden satın alım yapan bireylerin yüzdesi değişkenini ifade eden D20 değişkeni değeri %32'den küçük olanlar ekonomik açıdan düşük seviyede olan ülkeler olarak sınıflandırılmaktadır. %32'den büyük olan ülkeler ise ekonomik açıdan ortalama seviyede olan ve ekonomik açıdan yüksek seviyede olan ülkeler olarak sınıflandırılmıştır. Aynı şekilde D33 değişkeni yani çalışanlarına BİT eğitimi veren girişimlerin yüzdesi değişkeni değeri %48'in üstünde olan ülkeler de ekonomik açıdan ortalama değerlere sahip olan ve bu gruba göre ekonomik açıdan daha iyi seviyedeki ülkeler olarak elde edilmiştir.

Ekonomik açıdan orta seviyedeki ve yüksek seviyedeki ülkeleri ayıran değişken ise D24 değişkeni olarak belirlenmiştir. D24 değişkeni ise toplam istihdam içerisinde AR-GE personeli yüzdesidir. Bu değişkenin değeri %40'ın altında olan ülkeler ekonomik açıdan orta seviyedeki ülkeler olarak sınıflanırken, %40'ın üstünde olan ülkeler ise ekonomik açıdan yüksek seviyedeki ülkeler olarak sınıflanmıştır. Benzer şekilde D34 değişkeni de yani GSYH'de AR-GE gideri yüzdesi değişkeni 1,579'nın altında olan ülkeler ekonomik açıdan orta seviyedeki ülkeler iken bu değer üstündeki ülkeler ekonomik açıdan yüksek seviyedeki ülkelerdir.



Tablo 4.16. Karar Ağacı Analizine Ait Dendogram.

Elde edilen modelin sınıflandırma performansı ile ilgili tablo ise aşağıda yer almaktadır. Geriye kalan 35 veri ile model test edilmiştir. Her iki veri seti için sınıflandırma başarısı Tablo 4.16’da yer almaktadır. Modelin doğruluğu analiz edildiğinde %96,67 oranında doğruluk payı olduğu ve yanlışma payının %3,33 olduğu görülmektedir. Modeldeki doğru sınıflama başarısı, test veri setiyle ölçülerek %94,32 olarak elde edilmiştir.

SINIFLANDIRMA	EĞİTİM VERİ KÜMESİ		TEST VERİ KÜMESİ	
	SAYI	YÜZDE	SAYI	YÜZDE
DOĞRU	101	96,67	33	94,32
YANLIŞ	4	3,33	2	5,78

Tablo 4.17. Karar Ağacı Analizi Sınıflandırma Performansı.

Karar ağaçları ile ilgili kurulan bu model göz önüne alındığında ekonomik açıdan düşük düzeydeki olan ülkelerin BİT ile ilgili olarak eğitim seviyesini arttırmaları, internetin alışverişlerde daha etkin olarak kullanmaları ekonomilerini bir üst seviyeye ulaştırabilecektir. Ekonomik açıdan orta seviyedeki ülkelerin ise AR-GE ve inovasyona daha çok önem vermeleri halinde ekonomik olarak daha da ilerleme sağlayabilecekleri görülmektedir.

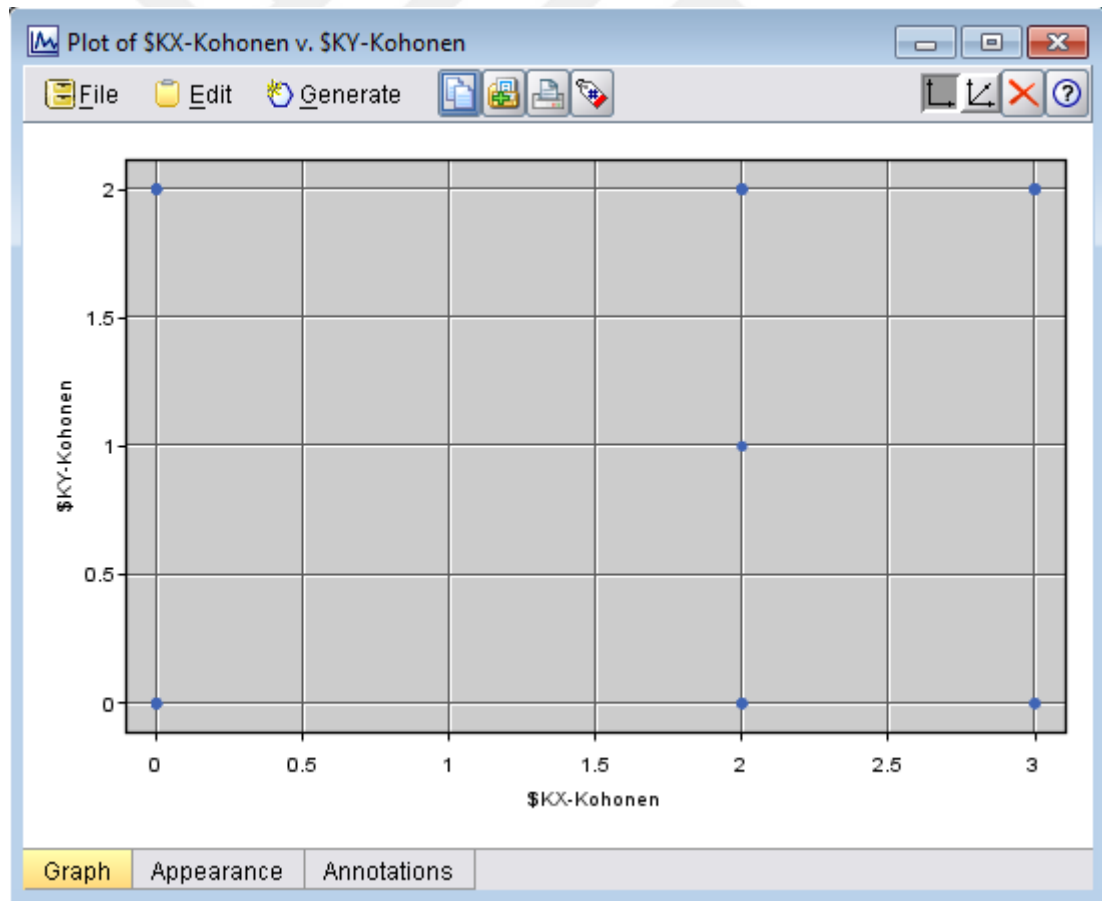
4.3.3.2. Kohonen Ağları İle Kümeleme Analizi

Kohonen ağları, girdi (bağımsız) değişken kümesindeki örüntülerin açığa çıkarılması amacıyla kullanılmaktadır. Bu ağın çıktısında, gözlemler gruplandırılmış olarak elde edilmektedir. Kohonen ağlarında, kestirim yapılacak bir çıktı (bağılı) değişken bulunmadığından, denetimsiz öğrenim gerçekleştiren bir sinir ağı türüdür.³⁷³

³⁷³Oğuzlar, A. (2005). Kümeleme Analizinde Yeni Bir Yaklaşım: Kendini Düzenleyen Haritalar (Kohonen Ağları). *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 19(2).

Uygulamanın bu aşamasında ülkelerin 2016 yılına ait olan bilgi ekonomisi değişkenleri kullanılmıştır. Bu değişkenler internet üzerinden satın alım yapan bireylerin yüzdesi (d20) , çalışanlarına BİT eğitimi veren girişimlerin yüzdesi (d33) , toplam istihdam içerisinde AR-GE personeli yüzdesi (d24) ve GSYİH’da AR-GE gideri yüzdesi(d34) değişkenleridir. Bu değişkenler bir önceki karar ağaçları uygulamasında tespit edilen bilgi ekonomisi değişkenleridir.

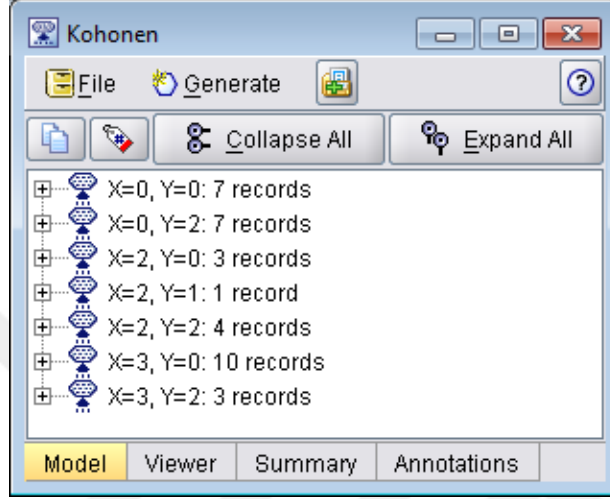
Ekonomik büyüklük değişkeni, refah seviyesi değişkeni, üyelik yılı değişkeni ve coğrafi bölge değişkeni bu aşamada kullanılan diğer değişkenlerdir. Tablo 4.17’de yer alan grafikte Kohonen Ağları ile yapılan analizde elde edilen kümeler ve koordinatları yer almaktadır. Bu kümelerin koordinatları $(X=0,Y=0)$, $(X=0,Y=2)$, $(X=2,Y=0)$, $(X=2,Y=1)$, $(X=2,Y=2)$, $(X=3,Y=0)$, $(X=3,Y=2)$ ’dir.



Tablo 4.18. Küme Koordinatlarına Ait Grafik Gösterimi.

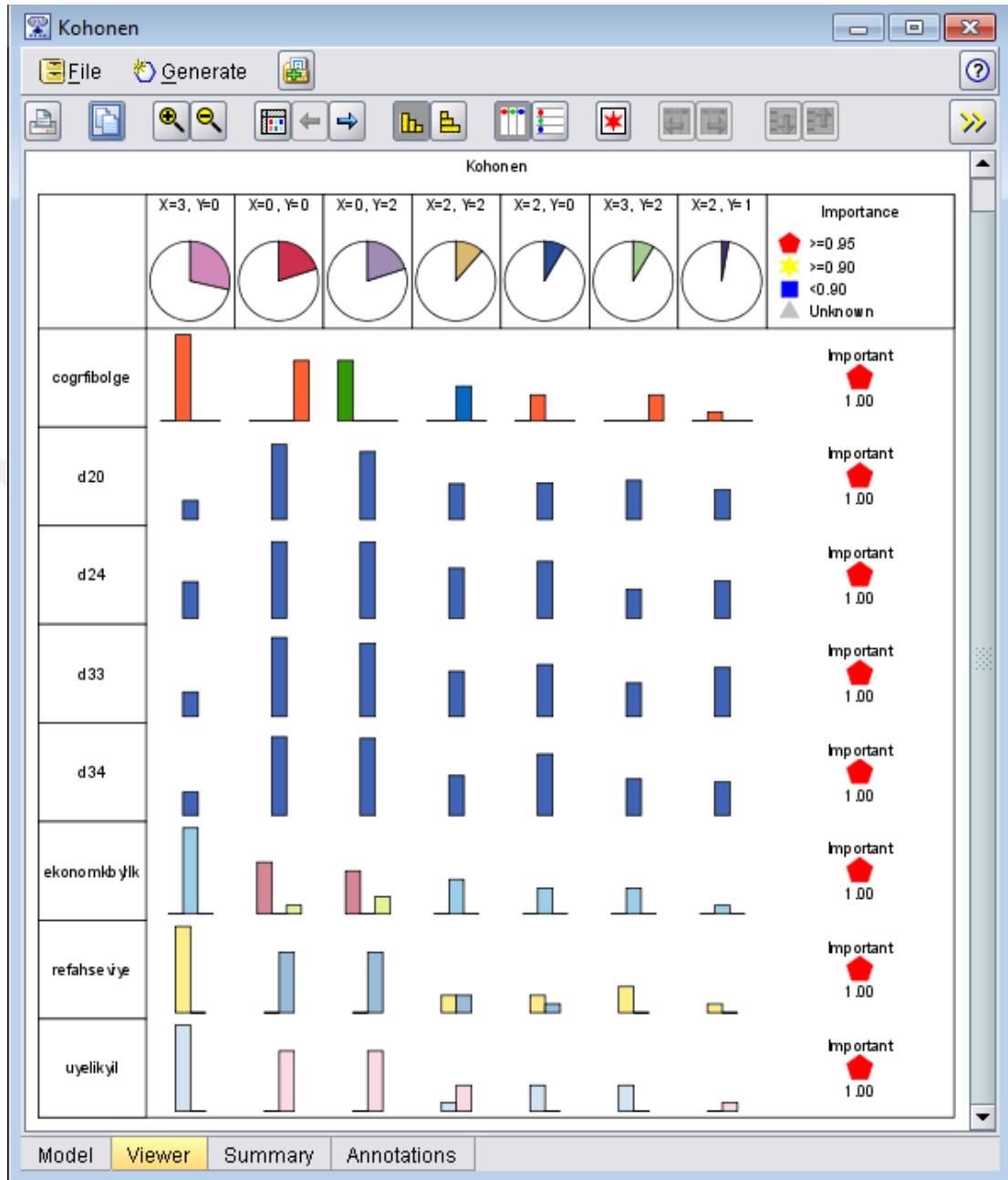
Tablo 4.18’de Kohonen Ağları ile yapılan analizde elde edilen küme koordinatlarında yer alan kayıt sayıları bulunmaktadır. $(X=3,Y=0)$ kümesinde 10

kayıt , $(X=0,Y=0)$ kümesinde 7 kayıt, $(X=2,Y=0)$ kümesinde 7 kayıt , $(X=2,Y=2)$ kümesinde 4 kayıt, $(X=2,Y=0)$ kümesinde 3 kayıt, $(X=3,Y=2)$ 3 kayıt ve $(X=2,Y=1)$ kümesinde 8 kayıt bulunmaktadır. En çok birimin yer aldığı kümeler $(X=3, Y=0)$ kümesindeki 10 kayıt ile $(X=0,Y=0)$ ve $(X=0,Y=2)$ kümesindeki 7 kayıttır.



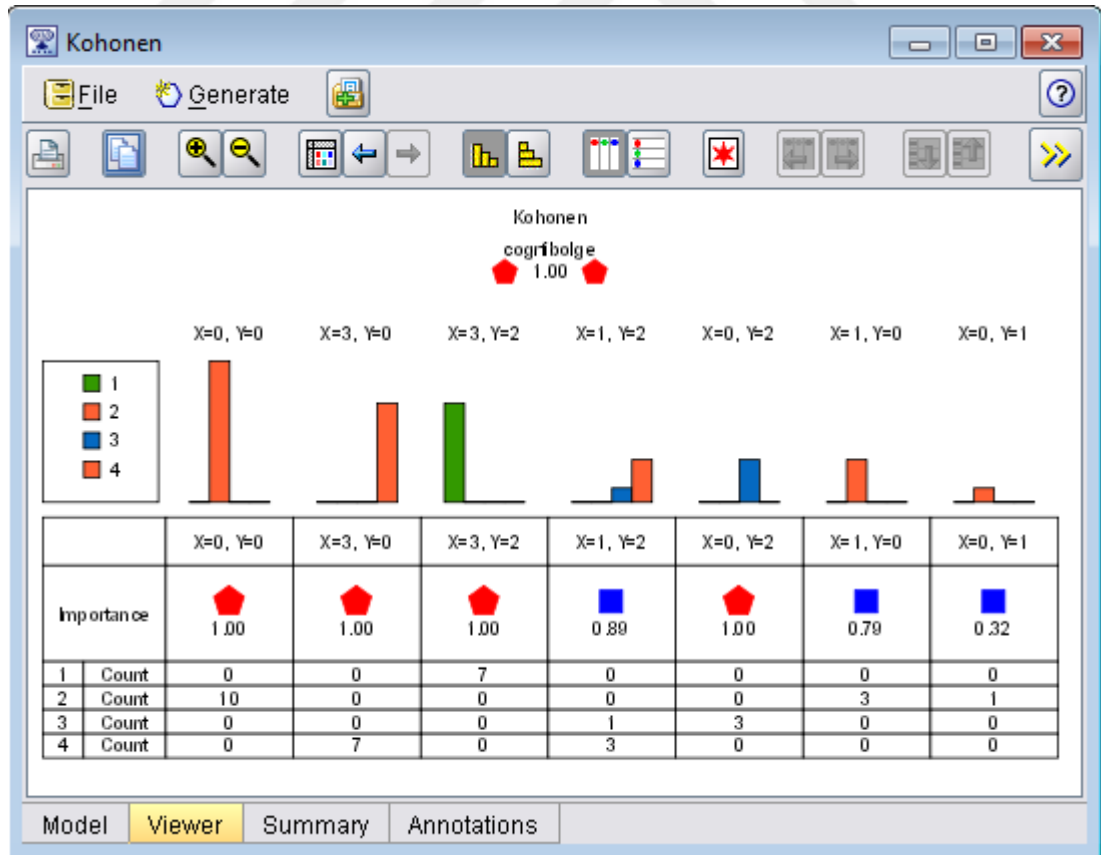
Tablo 4.19. Küme Koordinatlarına ait Kayıtlar.

Tablo 4.19’da oluşturulan kümelerin ve değişkenlerin etkilerinin görselleştirilmiş hali yer almaktadır. Ekonomik büyüklük değişkeni, refah seviyesi değişkeni, üyelik yılı değişkeni, coğrafi bölge değişkeni, internet üzerinden satın alım yapan bireylerin yüzdesi değişkeni (d20), çalışanlarına BİT eğitimi veren girişimlerin yüzdesi değişkeni (d33), toplam istihdam içerisinde AR-GE personeli yüzdesi değişkeni (d24) ve GSYH’de AR-GE gideri yüzdesi değişkeni(d34) ayırt edici özelliğe sahiptir. Değişkenlerin kümeler üzerindeki etkisi anlamlı bulunmuştur.



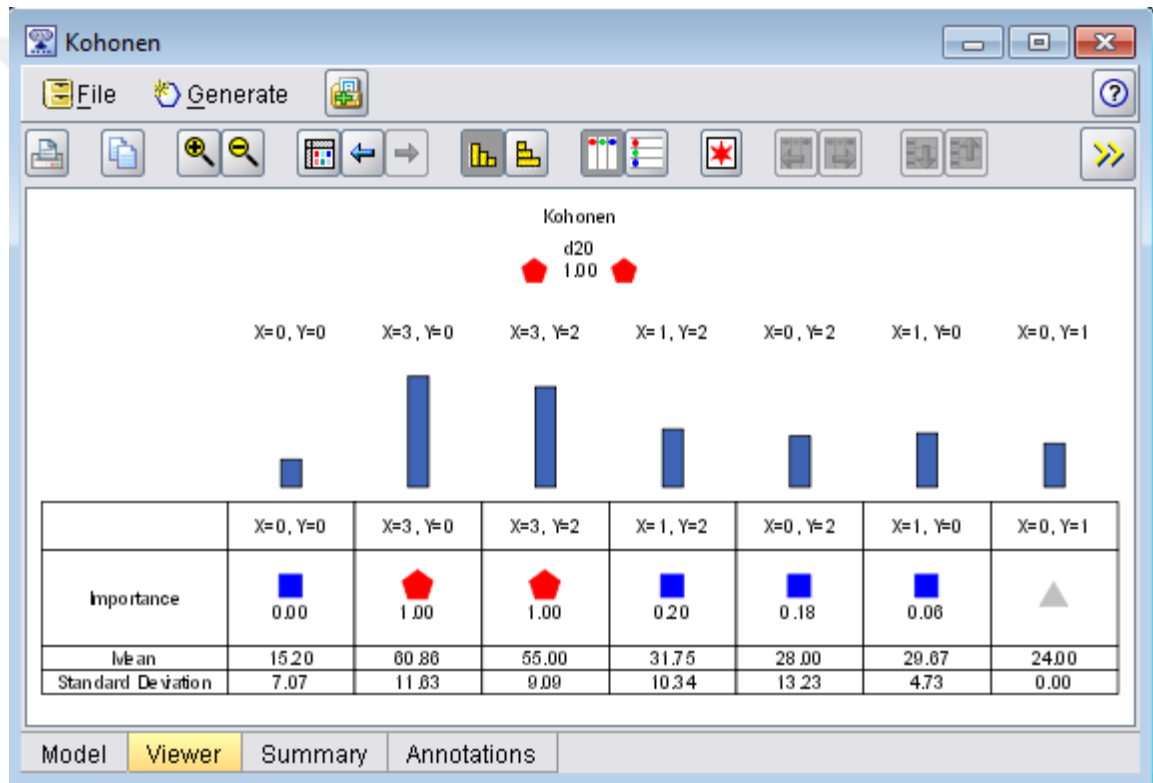
Tablo 4.20..Kohonen Ağları ile Kümeleme Analizi Sonuçları.

Coğrafi bölge değişkeni önemli bir özellik olarak görülmektedir. Bu değişken (0)(0), (3)(0), (3)(2), (0)(2), (0)(2) kümeleri üzerinde etkilidir. (0)(0) kümesinde doğu Avrupa ülkeleri (Karadağ, Makedonya, Sırbistan, Türkiye, Romanya, Bulgaristan, Kıbrıs, Yunanistan, Macaristan ve Polonya), (3)(0) kümesinde kuzey Avrupa ülkeleri (İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya), (3)(2) kümesinde batı Avrupa ülkeleri (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) ve (0)(2) kümesinde ise güney Avrupa ülkeleri (Malta, İspanya ve Portekiz) çoğunlukla yer almaktadır. (1)(2) kümesinde 1 tane güney ülkesi (İtalya) 3 tane kuzey ülkesi (İrlanda, Letonya ve Litvanya) bulunmaktadır. (1)(0) kümesinde 3 tane Doğu Avrupa ülkesi (Slovenya, Çek Cumhuriyeti Hırvatistan) bulunmaktadır. (0)(1) kümesinde 1 tane doğu Avrupa ülkesi (Slovakya) yer almaktadır.



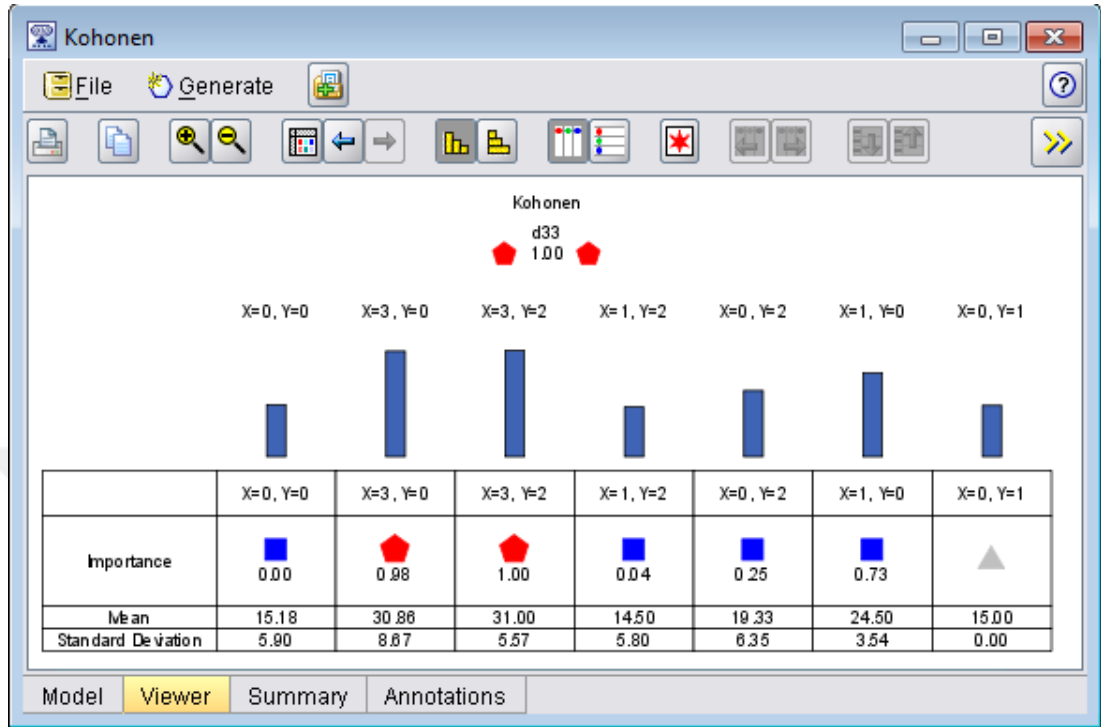
Tablo 4.21. Coğrafi Bölge Değişkeninin Kümeler Üzerindeki Etkisi.

Tablo 4.21’de internet üzerinden satın alım yapan bireylerin yüzdesi (d20) bilgi ekonomisi değişkeni (3)(0) ve (3)(2) kümelerinde etkili olduğu görülmektedir. Ortalama değeri ise 60,86 ve 55 olarak hesaplanmıştır. Aynı zamanda bu kümeler batı ve kuzey ülkelerinin yer aldığı kümelerdir. Bu ülkelerin hangileri olduğuna ortalama değerler göz önüne alınarak bakıldığında; (3)(0) kümesindeki 7 kuzey ülkesi İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya’dır. (3)(2) kümesindeki 7 batı ülkesi de grupta yer alan batı ülkeleri Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika’dır.



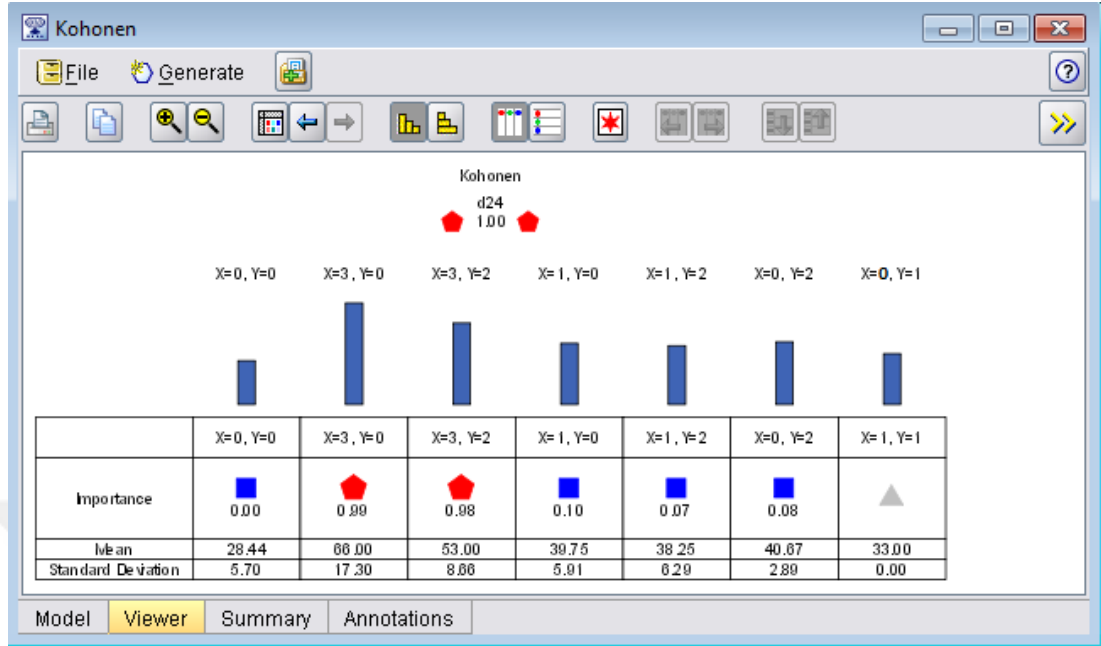
Tablo 4.22. İnternet Üzerinden Satın Alım Yapan Bireylerin Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.

Tablo 4.22’de çalışanlarına BİT eğitimi veren girişimlerin yüzdesi (d33) bilgi ekonomisi değişkeni (3)(0) ve (3)(2) kümelerinde etkili olduğu görülmektedir. Bu bilgi ekonomisi değişkeninin ortalama değeri 31 ve 30,86 olarak hesaplanmıştır. Aynı zamanda bu kümeler kuzey(İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya) ve batı ülkelerinin (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) yer aldığı kümelerdir.



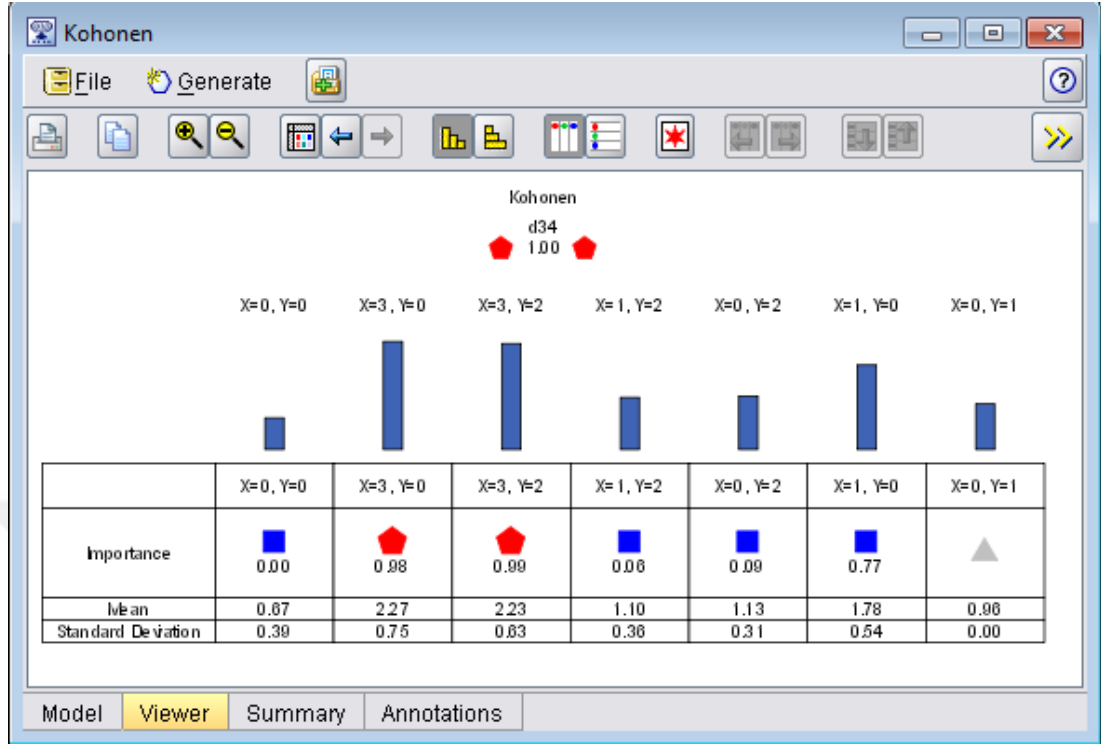
Tablo 4.23. Çalışanlarına BİT Eğitimi Veren Girişimlerin Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.

Toplam istihdam içerisinde AR-GE personeli yüzdesi (d24) bilgi ekonomisi değişkeni (3)(0) ve (3)(2) kümelerinde etkili bulunmuştur. Tablo 4.23’de görüldüğü gibi ortalama değerler ise 66 ve 53 olarak hesaplanmıştır. Bu kümeler kuzey(İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya) ve batı ülkelerinin (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) yer aldığı kümelerdir.



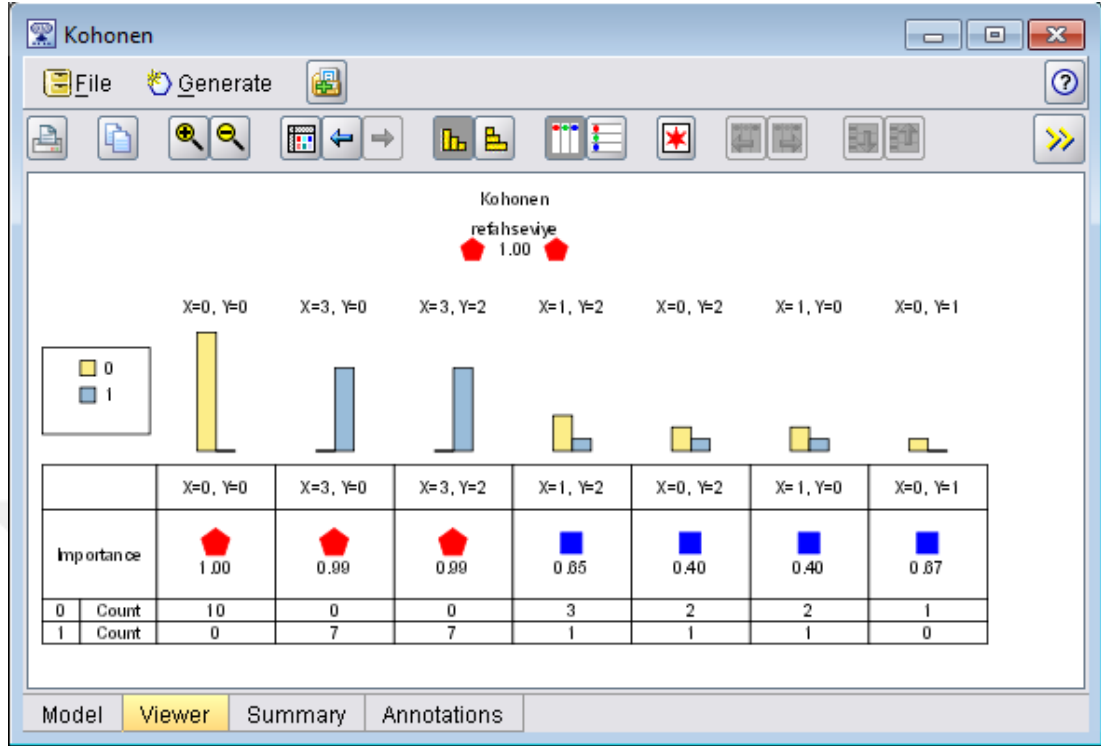
Tablo 4.24. Ar-Ge Personeli Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.

Tablo 4.24’de GSYH’de AR-GE gideri yüzdesi (d34) bilgi ekonomisi değişkeninin de (3)(0) ve (3)(2) kümelerinde etkili olduğu anlaşılmaktadır. Ortalama değerler ise 2,27 ve 2,23 olarak hesaplanmıştır. Aynı zamanda bu kümeler kuzey(İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya) ve batı ülkelerinin ülkeleri (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) yer aldığı kümelerdir.



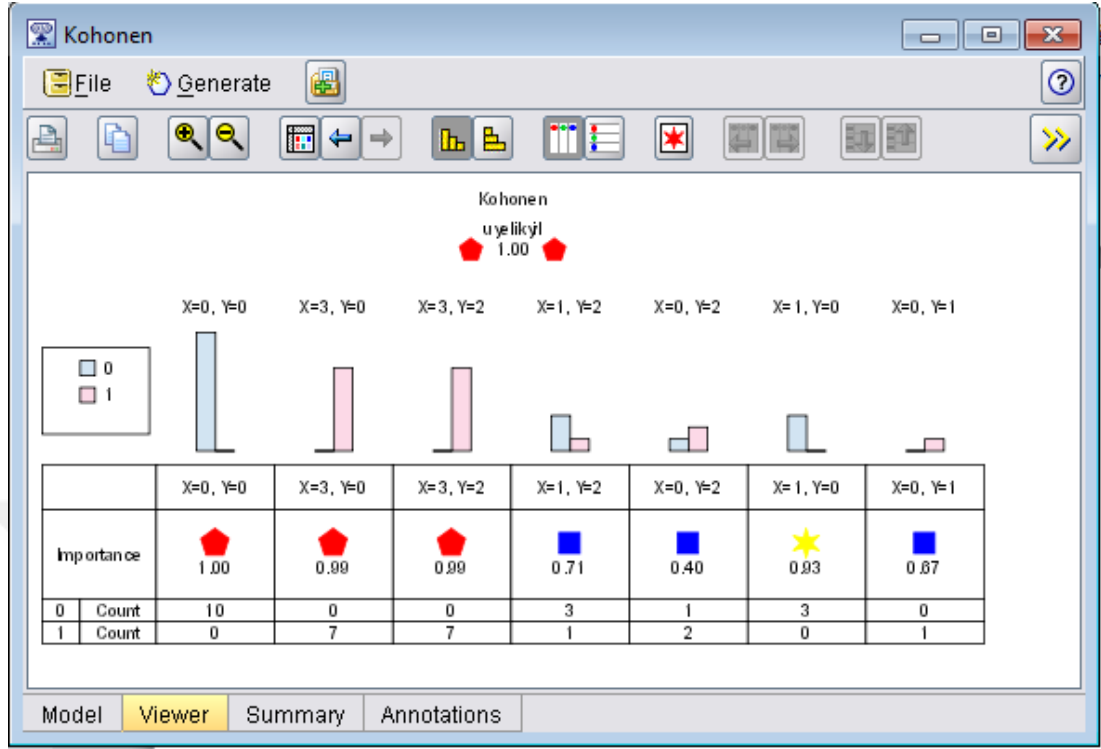
Tablo 4.25. GSYİH’de Ar-Ge Gideri Yüzdesi Değişkeninin Kümeler Üzerindeki Etkisi.

Tablo 4.25’de refah seviyesi değişkeninin (3)(2), (1)(0) ve (0)(0) kümelerinin üzerinde etkili olduğu görülmektedir. (0)(0) kümesinde yani doğu Avrupa ülkelerinin (Karadağ, Makedonya, Sırbistan, Türkiye, Romanya, Bulgaristan, Kıbrıs, Yunanistan, Macaristan ve Polonya) yer aldığı ve bilgi ekonomisi değişkenlerinin etkili olmadığı kümede refah ülkesi olmayan ülkeler yer almaktadır. (3) (0) ve (3)(2) kümelerinde yani batı ülkeleri (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) ve kuzey Avrupa ülkelerinin (İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya) yer aldığı ve bilgi ekonomisi değişkenlerinin etkili olduğu bu kümelerde ise refah ülkesi olan birimler yer almaktadır. (1)(2) kümesinde yer alan İrlanda, (0)(2) kümesinde yer alan Malta ve (1)(0) kümesinde yer alan Slovenya refah ülkeleri olarak yer almaktadır.



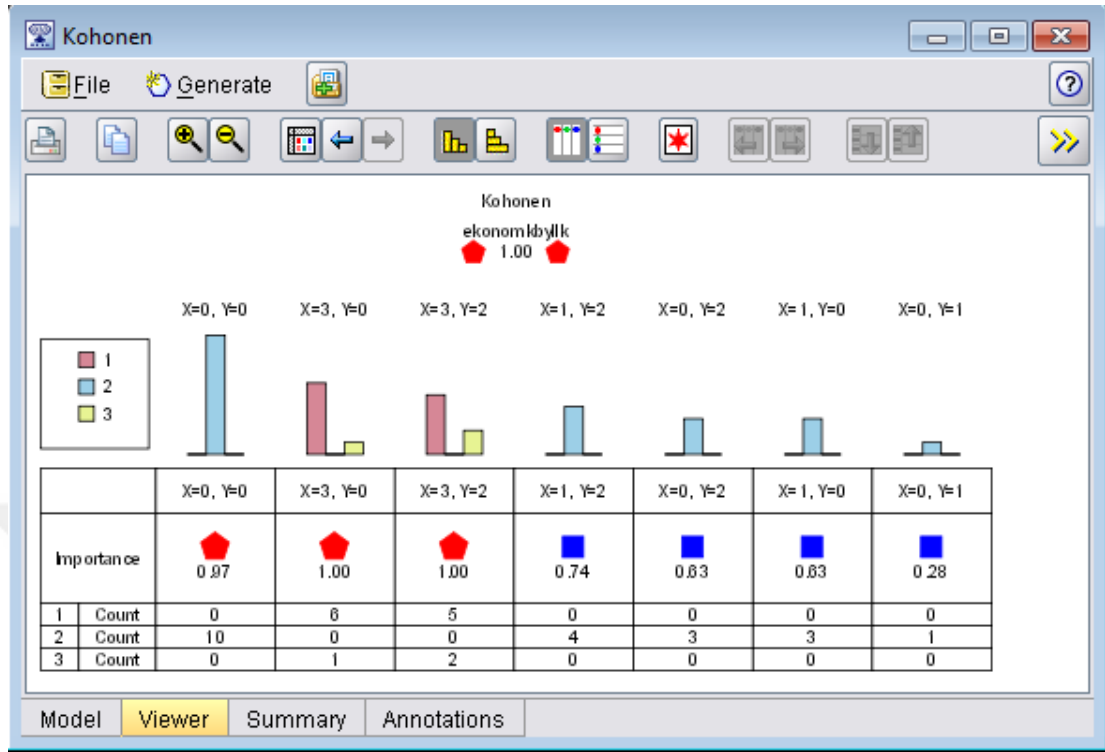
Tablo 4.26. Refah Seviyesi Değişkeninin Kümeler Üzerindeki Etkisi.

Tablo 4.26’da üyelik yılı değişkeninin (0)(0), (3)(0) ve (3)(2) kümelerinin üzerinde etkili olduğu görülmektedir. (0)(0) kümesinde yani doğu Avrupa ülkelerinin (Karadağ, Makedonya, Sırbistan, Türkiye, Romanya, Bulgaristan, Kıbrıs, Yunanistan, Macaristan ve Polonya) yer aldığı, bilgi ekonomisi değişkenlerinin etkili olmadığı ve refah ülkesi olmayan ülkelerin yer aldığı kümede 2000 yılı sonrası üye olan ülkeler bulunmaktadır. (3)(0) ve (3)(2) kümelerinde yani batı ülkeleri (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) ve kuzey Avrupa ülkelerinin (İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya) yer aldığı, bilgi ekonomisi değişkenlerinin etkili olduğu ve refah ülkesi olan birimlerin yer aldığı bu kümelerde ise 2000 yılı öncesi üye olan ülkeler bulunmaktadır. (1)(2) kümesinde yer alan İtalya, (0)(2) kümesinde yer alan Portekiz ve İspanya, (0)(1) kümesinde yer alan Slovakya 2000 yılı öncesi üye olan ülkelerdir.



Tablo 4.27.29 Üyelik Yılı Değişkeninin Kümeler Üzerindeki Etkisi.

Ekonomik büyüklük seviyesi değişkeni (0)(0), (3)(0) ve (3)(2) kümeleri üzerinde etkili olan bir değişkendir. (0)(0) kümesi doğu Avrupa ülkelerinin (Karadağ, Makedonya, Sırbistan, Türkiye, Romanya, Bulgaristan, Kıbrıs, Yunanistan, Macaristan ve Polonya) yer aldığı kümedir. Bu küme bilgi ekonomisi değişkenlerinin etkili olmadığı ve refah ülkesi olmayan ülkelerin yer aldığı ve 2000 yılı sonrası üye olan ülkelerin bulunduğu kümedir. Bu küme aynı zamanda ekonomik seviye olarak gelişmemiş ülkelerin yer aldığı küme olarak da hesaplanmıştır. Ekonomik seviye değişkeninin etkili olduğu (3)(0) ve (3)(2) kümelerinde ise kuzey ülkeleri (İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya ve Estonya) ve batı ülkeleri (Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika) bulunmaktadır. Aynı zamanda bu ülkeler refah ülkesi ve 2000 yılı öncesi üye olan ülkelerdir. Bu kümelerde ekonomik seviye olarak gelişmiş olan ve gelişmiş ülkelere nispeten daha iyi durumda olan yer almaktadır. Sözü edilen bilgiler Tablo 4.27’de yer almaktadır. (3)(0) kümesinde yer alan Norveç ve (3)(2) kümesinde yer alan Lüksemburg ve İsviçre tablodan da görüldüğü gibi ekonomik açıdan gelişmiş olan ülkelere nispeten daha iyi durumda olan ülkeler arasında yer almaktadır.



Tablo 4.28. 30 Ekonomik Büyüklük Değişkeninin Kümeler Üzerindeki Etkisi.

4.3.4. Modelin Değerlendirilmesi ve Kullanılması

Veri madenciliğinin bu aşamasında kabul edilen modelin katkıları ile ilgili değerlendirmeler yapılmaktadır. Ayrıca geçerliliği kabul edilen modelin hedefler açısından değerlendirmesi de bu aşamada gerçekleştirilmektedir. Elde edilen modeller farklı alanlarda ve uygulamalarda direk kullanılabileceği gibi, bu uygulamaların içine katılarak sürecin bir parçası gibi entegre edilebilmektedir. Uygulama aşamasında elde edilen sonuçlar ülkelerin ekonomik politikalarına katkı sağlayabilir. Özellikle bilgi çağında ülkeler ekonomilerinin geliştirilebilmesi için eğitim ve Ar-Ge faaliyetlerine öncelik verebilirler.

4.3.5. Modelin İzlenmesi

Bu çalışmada geçerliliği kabul edilen model, ekonomik verileri içerdiği için ülke verilerinde zaman içerisinde meydana gelen değişiklikler dikkate alınarak yeniden düzenlenebilir. Sürekli gelişen ve değişen dünyada daha farklı faktörler ekonominin gelişimine yön verebilmektedir. Elde edilen yeni sonuçlar mevcut sonuçlarla karşılaştırılarak izlenebilir.

SONUÇ

Son yıllarda teknolojinin geldiği noktada, biriktirilen yığınlarla veri ve bu verilerdeki anlamlı bilgilere ulaşmak için teknolojinin bütün imkânlarının kullanılması birçok alanda kullanılan veri madenciliğini ortaya çıkarmıştır. Genellikle bankacılık, pazarlama, finans alanında ticari amaçlı olarak kullanılan veri madenciliği teknikleri bu tez çalışmasında farklı olarak ekonomi alanında uygulanmıştır. Bu çalışmada veri madenciliğinin genel olarak kullanıldığı alanlar dışında farklı amaçlar içinde kullanım alanının olabileceği gösterilmiştir. Elde edilen bilgilerin ülkeler açısından ekonomi faaliyetlerinin yönlendirilmesinde kullanılabileceği ifade edilmiştir.

Ülkelerin geçmiş yıllarda gerçekleşen ekonomi rakamları, bu alanda bulundukları konumları ve değişkenlerin birbirleri ile ilgili ilişkileri açısından önemli bilgiler verebilmektedir. Veri yığınları üzerindeki bu önemli bilgilerin tespit edilmesi ülkeler açısından faydalı olacaktır. Ekonomi alanındaki veriler üzerinde yapılan veri madenciliği uygulaması ile ülkeler bilgi ekonomisi açısından değerlendirilmiştir.

Veri madenciliği yönetimlerinin bilgi ekonomisi değişkenlerine uygulanmasıyla birlikte ülkeler bazında ekonomi politikalarına yön verecek değerlendirmeler yapılabileceği düşünülmektedir.

Veri madenciliği teknikleri bu çalışmada Türkiye'nin de dahil olduğu 35 Avrupa ülkesinin 2013-2016 yılları arasındaki ekonomik değişkenlerine uygulanmıştır. Verilerin düzenlenmesi, toplanması ve ne gibi işlemlerin uygulandığı ile ilgili bilgiler bir önceki bölümde açıklanmıştır.

Tezin uygulama bölümünde Avrupa Birliği üye ülkeleri ve Türkiye'nin bilgi ekonomisi alanında yer alan yıllık verileri, elde edilecek sonuçların ekonomi politikası kararlarına destek sağlaması amacıyla değerlendirilebilmeleri için veri madenciliği sürecine tabi tutulmuştur.

Uygulamada veri madenciliği modellerinden sınıflandırma ve kümeleme modelleri seçilerek ekonomik seviyeyi en çok etkileyen bilgi ekonomisi değişkenleri

tespit edilmeye çalışılmıştır. Bu nedenle sınıflandırma modeli için karar ağacı uygulaması, kümeleme modeli için ise Kohonen ağırları uygulaması seçilmiştir.

Çalışmanın ilk bölümünde bilgi ekonomisi genel çerçevesi detaylı bir şekilde incelenmiştir. Bu bölümde bilgi ekonomisi ile ilgili değişkenlerin yansırı genel ekonomik çerçeve ile ilgili alanlara da yer verilmiştir. İkinci bölümde veri madenciliği modelleri ve teknikleri ayrıntılı olarak incelenmiştir. Üçüncü bölümde uygulamada kullanılan çok değişkenli istatistik tekniklerden kümeleme analizi ayrıntılı bir şekilde incelenmiştir.

Uygulamada öncelikle her bir değişkenin genel istatistik değerleri belirlenmiştir. Söz konusu değişkenlere ait dağılım grafikleri, ortalamalar, minimum ve maksimum değerler, standart sapma değerleri ve veri türleri tablo şeklinde gösterilmiştir. Normal dağılım göstermediği tespit edilen değişkenler logaritmik olarak dönüştürülerek normal dağılıma yaklaştırılmıştır. Bu değişkenler, BİT eğitilmiş çalışanları olan girişimlerin yüzdesi, Toplam AR-GE personel yüzdesi, girişimlerin toplam AR-GE gideri, patent uygulamalarının sayısı, telekomünikasyon sektörü toplam yatırımları, yüksek teknoloji ticaret hacmi değişkenleridir.

Uygulamanın sonraki aşamasında ülkelerin GSYİH, İşsizlik ve ihracat-ithalat oranı değişkenleri göz önüne alınarak kümeleme analizi yapılmıştır. Böylece ülkelerin her yıl için genel anlamda ekonomik gruplanmaları tespit edilmiştir. Sadece 2016 yılında İtalya ve İspanya ortalama seviyede olan ülkeler grubundan gelişmekte olan ülkeler grubuna geçmiştir. Bu ayrıntı dışında analiz sonucunda her yıl için aynı gruplanmalar elde edilmiştir. Kümeleme tekniği olarak hiyerarşik olmayan yöntemlerden k-ortalamalar kümeleme analizi kullanılmıştır. Analiz sonucu, ekonomik seviye değerleri ortalama değerlere sahip olan gelişmiş ülkeler 1.grubu oluşturmuştur. Bu ülkeler Belçika, Danimarka, Hollanda, İrlanda, Fransa, Avusturya, İsveç, İngiltere, İzlanda'dır. Ekonomik değerleri ortalamanın üzerinde olan ülkeler ise 3.grupta oluşturmuştur. Bu ülkeler Lüksemburg, Norveç ve İsviçre'dir. Aralarında Türkiye'nin de bulunduğu 2.grupta ekonomik seviye olarak ortalamanın altında olan az gelişmiş ülkeler yer almaktadır. Bulgaristan, Çek Cumhuriyeti, Estonya, Yunanistan, İspanya, Hırvatistan, İtalya, Kıbrıs, Letonya,

Litvanya, Makedonya, Malta, Polonya, Portekiz, Romanya, Slovenya, Slovakya, Karadağ, Yugoslavya ve Sırbistan bu grupta yer almaktadır. Türkiye yapılan kümeleme analizi sonucunda gelişmekte olan ülkeler grubuna dahil olmuştur.

Bir sonraki adımda C5.0 algoritması kullanılarak karar ağacı uygulaması gerçekleştirilmiştir. Hedef değişken olarak önceki aşamada kümeleme analizi ile belirlenen genel ekonomik seviye grupları, tahminleyici değişken olarak 35 tane bilgi ekonomisi değişkeni kullanılmıştır. Ekonomik seviye açısından en ayırıcı bilgi ekonomisi değişkenleri tespit edilmiştir. Bu analiz sonucunda ekonomik olarak gelişmekte olan ülkeleri (Bulgaristan, Çek Cumhuriyeti, Estonya Yunanistan, İspanya, Hırvatistan, İtalya, Kıbrıs, Letonya, Litvanya, Makedonya, Malta, Polonya, Portekiz, Romanya, Slovenya, Slovakya, Karadağ, Yugoslavya ve Sırbistan) diğer iki gruptan ayıran değişkenler internet üzerinden satın alım yapan bireylerin yüzdesi değişkeni ve çalışanlarına BİT eğitimi veren girişimlerin yüzdesi değişkeni olarak görülmüştür. Ekonomik olarak ortalama değerlere sahip ülkeler (Belçika, Danimarka, Hollanda, İrlanda, Fransa, Avusturya, İsveç, İngiltere, İzlanda) ve bu ülkelere nispeten daha iyi durumda olan ülkeleri (Lüksemburg, Norveç ve İsviçre) ayıran değişkenler ise toplam istihdam içerisinde Ar-Ge personeli yüzdesi değişkeni ile GSYİH’da yer alan Ar-Ge gideri yüzdesi değişkenleri olarak tespit edilmiştir.

En son aşamada ise ülkelerin ekonomik seviye, refah düzeyi, coğrafi konumları, Avrupa Birliği’ne üyelik yılları kategorik değişkenleri ve karar ağacı analiziyle tespit edilen bilgi ekonomisi değişkenleri dikkate alınarak Kohonen ağırları ile kümeleme analizi yapılmıştır. Yapılan analiz sonucunda 7 adet küme oluştuğu gözlenmiştir. Oluşan bu kümelerde kuzey ve batı bölgelerde yer alan, 2000 yılı öncesi AB’ye üye olan, refah seviyesi yüksek ve bilgi ekonomisi değişken değerleri ortalamanı üzerinde olan ülkeler bir arada yer almaktadır. Bu ülkeler İngiltere, İzlanda, Danimarka, Norveç, İsveç, Finlandiya, Estonya Almanya, Lüksemburg, İsviçre, Hollanda, Fransa, Avusturya ve Belçika’dır. Doğu Avrupa’da yer alan 2000 yılı sonrası üye olan, refah seviyesi düşük, ekonomik seviye olarak gelişmekte olan, bilgi ekonomisi değişken değerleri ortalamadan düşük olan ülkeler ise kendi aralarında grup oluşturmuşlardır. Bu ülkeler Karadağ, Makedonya, Sırbistan, Türkiye, Romanya, Bulgaristan, Kıbrıs, Yunanistan, Macaristan ve Polonya’dır.

Uygulama sonucunda Türkiye'nin ekonomik düzeyi düşük olan ve bilgi ekonomisi değişkenleri açısından doğu Avrupa ülkeleri ile aynı kategoride yer aldığı görülmektedir. Ülkelerin Kohonen analizinde İtalya, İrlanda, Letonya ve Litvanya ülkeleri, Slovenya Çek Cumhuriyeti ve Hırvatistan ülkeleri, Malta, İspanya ve Portekiz ülkeleri kendi aralarında ayrı gruplar halindeyken, Slovakya ise tek başına bir kümede yer almıştır.

Yapılan tüm analiz ve uygulamalar sonucunda ekonomik olarak gelişmekte olan ülkelerin gelişmiş ülkeler seviyesine ulaşabilmeleri için internet teknolojileri kullanımına ve bilgi ve iletişim teknolojileri eğitime önem vermeleri gerektiği söylenebilir. Gelişmiş olan ülkeler ise daha iyi konuma gelebilmek için Ar-Ge ve inovasyon faaliyetlerine daha çok yoğunlaşmalıdırlar. Ülkeler özellikle ilköğretim çağından itibaren bilgisayar ve teknoloji eğitimlerine ağırlık verip kalifiye Ar-Ge personeli yetiştirmeyi hedeflerlerse ekonomik anlamda ciddi adımlar atabileceklerdir.

KAYNAKÇA

- Adaçay, Funda Rana : “Bilgi Ekonomisine İlişkin Temel Göstergeler Açısından Avrupa Birliği ve Türkiye’nin Karşılaştırılması”, **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi 19**, 2007, s.189.
- Adrians, Pieter ve Zantinge, Dolf : **Data Mining**, İngiltere, 1997.
- Agrawal, Rakesh ve Srikant, Ramakrishnan : “Fast Algorithms for Mining Association Rules”, **Proc. 20th Int. Conf. Very Large Data Bases, VLDB. Vol. 1215**, 1994, p. 490.
- Ağır, Hüseyin : “Türkiye ile Güney Kore’de Bilim ve Teknoloji Politikalarının Karşılaştırması”, **Bilgi Ekonomisi ve Yönetimi Dergisi 5.2.**, 2010, s. 34.
- Akolaş, Arzu : “Bilişim Sistemleri ve Bilişim Teknolojisinin Küreselleşme Olgusu ve Girişimcilik Üzerine Yansımaları”, 2000, s. 32-33.
- Akpınar, Haldun : “Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği”, **İ.Ü. İşletme Fakültesi Dergisi, 29.1.**, 2000, p. 19.
- Akpınar, Haldun : **Data Veri Madenciliği-Veri Analizi**, İstanbul, Papatya Yayıncılık, 2014.
- Aktan, Coşkun Can ve Tunç, Mehtap : **Bilgi Toplumu ve Türkiye, Moderniteden Postmoderniteye Değişim**, Ed. by. Coşkun Can Aktan, Çizgi Kitapevi, Konya, 2003.
- Aktan, Coşkun Can ve Vural, İstiklal Y. : **Yeni Ekonomi ve Yeni Rekabet**, Türkiye İşveren Sendikaları Konfederasyonu, 2004.
- Albayrak, Ali Sait ve Yılmaz, Şebnem Koltan : “Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri Üzerine Bir Uygulama”, **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi 14.1.**, 2009, s. 37.
- Aldenderfer, Mark S. ve Blashfield, Roger K. : **Cluster Analysis, Sage University Paper Series On Quantitative Applications in The Social Sciences**, Beverly Hills, Sage Publications, 1984.
- Anderberg, Michael R. : **Cluster Analysis for Applications**, New York, 1988, p. 137.
- Andriessen, Daniel ve Tissen, Rene : **Weightless Wealth: Find Your Real Value in a Future of Intangible Assets**, Financial Times Prentice, 2000.
- Aslan, Özgür ve Öner, Selcen : “İnternet Ekonomisi”, **İstanbul Üniversitesi İletişim Fakültesi Hakemli Dergisi 26.**, 2006, s. 17.
- Ata, H. Ali ve Seyrek, İbrahim H. : “The Use of Data Mining Techniques in Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms”, **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi 14.2.**, 2009, s. 2.
- Atamer, Belgin : **Kümeleme Analizi (Cluster Analysis) ve Kümeleme Analizinin İlaç Sektörüne Uygulanması**, **Marmara Üniversitesi Sosyal Bilimler Enstitüsü Yüksek lisans Tezi**, İstanbul, 1992.

- Atbaş, Azize Celile Günay : “Kümeleme Analizinde Küme Sayısının Belirlenmesi Üzerine Bir Çalışma”, **Ankara Üniversitesi Fen Bilimleri Enstitüsü Yayınlanmamış Yüksek Lisans Tezi**, 2008, s. 17.
- Atik, Hayriye ve Altıparmak, Aytekin : “Avrupa Birliği ve Türkiye’de Bilgi Sektörünün İstihdama Katkısı”, **Bilgi Ekonomisi ve Yönetimi Dergisi** 6.1., 2011. s. 235.
- Ayhan, Sevgi ve Erdoğan, Şenol : “Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi”, **Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi**, 9.1., 2014, s.185
- Ayık, Ziya Y., Özdemir, Abdülkadir ve Yavuz, Uğur : “Lise Türev ve Lise Mezuniyet Başarısının, Kazanılan Fakülte ile İlişkisinin Veri Madenciliği Tekniği ile Analizi”, **Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi** 10.2., 2007, s. 444.
- Barışık, Salih ve Yirmibeşçik, Oya : “Türkiye’de Yeni Ekonominin Oluşum Sürecini Hızlandırmaya Yönelik Uyum Çabaları”, **Uluslararası Yönetim İktisat ve İşletme Dergisi** 2.4., 2012, s. 41.
- Batagan, Lorena : “Indicators for Knowledge Economy”, **Informatica Economică** 4, 2007, s. 44.
- Bell, Daniel : **The Coming of The Post Industrial Society: A Venture in Social Forecasting** Newyork, 1973.
- Berberoğlu, Bahar : “Yaşam Boyu Öğrenme İle Bilgi ve İletişim Teknolojilerin Açısından Türkiye’nin Avrupa Birliği’ndeki Konumu”, **Bilgi Ekonomisi ve Yönetimi Dergisi** 5.2., 2010, s. 15.
- Berkhin, Pavel : “A Survey Of Clustering Data Mining Techniques”, **In: Grouping Multidimensional Data**, Springer Berlin Heidelberg, 2006, p. 30.
- Berkhin, Pavel : **A Survey of Clustering Data Mining Techniques, In: Grouping Multidimensional Data**, Springer Berlin, 2006.
- Beyer, Kevin, Goldstein, Jonathan, Ramakrishnan, Raghu ve Shaft, Uri : “When Is Nearest Neighbor Meaningful?”, **In: International Conference On Database Theory**, Springer Berlin Heidelberg, 1999, p. 230.
- Bilge, Neyran Orhun : **Çok Değişkenli İstatistik Yöntemler**, İstanbul Üniversitesi İşletme Fakültesi Yayını, 2010.
- Birkinshaw, Julian ve Sölvell, Örjan : “Preface”, **International Studies of Management & Organization** 30.2, 2000, p. 84.
- Black, Sandra E. ve Lynch, Lisa M. : “How to Compete: The Impact of Workplace Practices and Information Technologyon Productivity”, **Review of Economics and Statistics** 83.3, 2001, p. 440.
- Blakeley, Nic, Lewis, Geoff ve Mills, Duncan : “The Economics of Knowledge: What Makes Ideas Special for Economic Growth?”, **New Zealand Treasury, Polic Perspective Paper**, 2005, p. 2.
- Brachman, Ronald J. ve Anand, Tej : “The Process of Knowledge Discovery in Databases. In: Advances in Knowledge Discovery and Data Mining”, **American Association for Artificial Intelligence**, 1996, p. 40.

- Bramer, Max : **Principles of Data Mining**, Springer, London, 2007, s. 6.
- Braund Paul, v.d. : “Information and Communications Technology for Economic Development”, **A Report on the Global E-Discussion, Technical Report, World Bank**, 2006, p. 9.
- Bresnahan, Timothy F., Brynjolfsson, Erik ve Hitt, Lorin M. : “Information Technology, Workplace Organization and The Demand for Skilled Labor: Firm-Level Evidence”, **National Bureau of Economic Research**, 1999, p. 52.
- Brinkley, Ian : “The Knowledge Economy: How Knowledge is Reshaping The Economic Life of Nations”, **The Work Foundation's Three Year Knowledge Economy Programme**, London, 2008, p. 15.
- Chandrasekaran, Mohan ve Narang, Inderpal S. : “Method for Non-Hierarchical Lock Management In A Multi-System Shared Data Environment”, 1996, s. 39.
- Chen, Derek ve Dahlman, Carl J. : “The Knowledge Economy, The KAM Methodology and World Bank Operations”, **World Bank Institute Working Paper**, 2005, p. 5.
- Chen, Ming-Syan, Han, Jiawei ve Yu, Philip S.: "Data Mining : An Overview from A Database Perspective", **IEEE Transactions on Knowledge and Data Engineering** 8.6., 1996, p. 870.
- Choi, Soon-Yong, Stahl, Dale O. ve Whinston, Andrew B. : “Intermediation, Contracts Contracts and Micropayments in Electronic Commerce”, **Electronic Markets** 8.1., 1998, p. 201.
- Chung, H. Michael ve Gray, Paul : “Special Section: Data Mining”, **Journal of Management Information Systems** 16.1, 1999, p. 15.
- Cohen, Wesley M. ve Levinthal, Daniel A. : “Innovation and Learning: The Two Faces of R&D”, **The Economic Journal**, 1989, p. 70.
- Cooke, Philip : “Regional Innovation Systems, Clusters, and The Knowledge Economy”, **Industrial and Corporate Change** 10.4, 2001, p. 949.
- Cornu, Bernard : **Collective Intelligence and Capacity Building in a Network Society, Education and The Knowledge Society: Information Technology Supporting Human Development**, Ed. by. Tom J. van Weert, New York, Springer, 2004,
- Cortes, Corinna ve Vapnik, Vladimir : “Support-Vector Networks”, **Machine Learning, Vol. 20, No.3**, 1995, p. 280.
- Coşlu, Eda : **Veri Madenciliği**, Akademik Bilişim 2013 – XV. Akademik Bilişim Konferansı Bildirileri 23-25 Ocak 2013 – Akdeniz Üniversitesi, Antalya
- Cox, Buket Öztuna : **Avrupa Birliği Hukukunda Elektronik Ticaret ve Türkiye’deki Gelişmeler**, Pusula Yayıncılık, İstanbul, 2002.
- Çakmak, Zeki, Uzgören, Nevin ve Keçek, Gülnur : “Kümeleme Analizi Teknikleri İle İllerin Kültürel Yapılarına Göre Sınıflandırılması ve Değişimlerinin İncelenmesi”, **Dumlupınar Üniversitesi Sosyal Bilimler Dergisi**, 12.3., 2005, s.20.

- Çalık, Deniz ve Çınar, Özge Pelin : “Geçmişten Günümüze Bilgi Yaklaşımları Bilgi Toplumu ve İnternet”, **XIV. Türkiye'de İnternet Konferansı**, İstanbul Bilgi Üniversitesi, 2009, s. 8.
- Çelikkol, Ethem Soner, Genç, Sema Yılmaz ve Şeneldir, Okan : “Küreselleşme Olgusu ve Bilgi Toplumu'nun Gelişme Dinamiklerinin Türkiye Üzerindeki Olası Etkileri”, **International 8th Knowledge & Management Congress Proceedings (October 28-31)**, İstanbul, 2010, s. 215.
- Çukurçayır, M. Akif ve Çelebi, Esra : “Bilgi Toplumu ve E-Devletleşme Sürecinde Türkiye”, **Uluslararası Yönetim İktisat ve İşletme Dergisi**, 2012, s. 61.
- Dahlman, Carl ve Andersson, Thomas : **Korea and The Knowledge Based Economy: Making The Transition**, World Bank Institute, 2000.
- Dahlman, Carl, Xue, Lan ve Gürüz, Kemal : “Lifelong Learning and The Knowledge Economy”, **Summary of The Global Conference on Lifelong Learning**, Stuttgart, Germany, 2002, p. 17-18.
- Danziger, James N. ve Andersen, Kim Viborg : “The Impacts of Information Technology on Public Administration: An Analysis of Empirical Research from The 'Golden Age' of Transformation”, **International Journal of Public Administration** 25.5, 2002, p. 600.
- Dempster, P. Arthur, Laird, Nan M. ve Rubin, Donald B. : “Maximum Likelihood from Incomplete Data Via The EM Algorithm”, **Journal of The Royal Statistical Society Series B (Methodological)**, 1977, p. 35.
- Dilly, Ruth : **Data Mining: An Introduction**, Student Notes, 1995.
- Dinçer, Esra : “Veri Madenciliğinde K-Means Algoritması ve Tıp Alanında Uygulanması”, **Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi**, 2006, s. 35.
- Dolgun, Özgür M., Özdemir, Tülin ve Oğuz, Doruk : “Veri Madenciliğinde Yapısal Olmayan Verinin Analizi: Metin ve Web Madenciliği”, **İstatistikçiler Dergisi: İstatistik Ve Aktüerya**, 2.2., 2009, s. 50.
- Dordick, Herbert S. Ve Wang, Georgette : **Information Society: A Retrospective View**, Sage Publications Inc., 1993.
- Doz, Yves, Santos, Jose ve Williamson, Peter : **From Global to Metanational: How Companies Win in The Knowledge Economy**, Boston, Harvard Business Press, 2001.
- Driouchi, Ahmed, Azelmad, El Mustapha ve Anders, Gary C. : “An Econometric Analysis Of The Role Of Knowledge In Economic Performance”, **The Journal of Technology Transfer**, 31.2, 2006, p. 250.
- Drucker, Peter F. : “Knowledge-Worker Productivity: The Biggest Challenge”, **California Management Review** Vol. 41. No.2, 1999, p. 10.
- DTI Competitiveness White Paper : “Our Competitive Future: Building The Knowledge Driven Economy”, London: Department of Trade and Industry, 1998, p. 23.
- Dunham, Margaret H. : **Data Mining-Introductory and Advanced Topics**, Person Education, 2003.

- Dunning, John H. : **Regions, Globalization, and The Knowledge-Based Economy**, Oxford University Press, 2002.
- Dura, Cihan ve Atik, Hayriye : “Bilgi Toplumu”, **TC Kültür Bakanlığı Yayınları**, Ankara, 1990, s. 4.
- Dura, Cihan ve Atik, Hayriye : **Bilgi Toplumu Bilgi Ekonomisi ve Türkiye**, İstanbul, 2002.
- Durdu, Ömer F. : “Avrupa Birliği Yolunda Bilgi Teknolojisinin Türkiye Ekonomisi İçin Önemi”, **Akademik Bilişim Konferansları**, (Çevrimiçi) <http://ab.org.tr/ab03/tammetin/171.doc>, 13.05.2007
- Edelstein, Herb : “Data Mining: Exploiting the Hidden Trends in Your Data”, **DB2 Online Magazine**, (Çevrimiçi) <http://www.psy.gla.ac.uk/~steve/pr/edel.html>, 27.11.2016
- Ekin, Nusret : **Küresel Bilgi Çağında Eğitim-Verimlilik, İstihdam**, İTO Yayınları, İstanbul, 1999.
- Erkan, Hüsnü : “Türkiye’nin Stratejik Önceliği: AB ya da Bilgi Toplumu”, **Stratejik Araştırmalar Dergisi**, Sayı.13, 2009, s. 2.
- Ersel, Birsal : “Bilgi Çağında Çalışma İlişkileri ve Beyin Göçü”, **2. Ulusal Bilgi, Ekonomi ve Yönetim Kongresi, Bildiriler Kitabı**, İzmit, 2003, s. 715.
- Everitt, Brian : **Cluster Analysis**, London, Heinemann Educational Books, 1974.
- Fayyad, Usama, Piatetsky-Shapiro, Gregory ve Smyth, Padhraic : “From Data Mining to Knowledge Discovery in Databases”, **AI Magazine** 17.3., 1996, p. 36.
- Feelders, Ad J. : “Prior Knowledge in Economic Applications of Data Mining”, **In European Conference on Principles of Data Mining and Knowledge Discovery**, Springer Berlin Heidelberg, 2000, p. 273.
- Fırat, Seniye Ümit : “Kümeleme Analizi: İstihdamın Sektörel Yapısı Açısından Avrupa Ülkelerinin Karşılaştırılması”, **İstanbul Üniversitesi Sosyal Bilimler Dergisi**, 3.2., 1995, s. 46.
- Fiske, John : **Introduction to Communication Studies**, Routledge, 2010.
- Foray, Dominique : **Economics of Knowledge**, MIT Press, 2004.
- Frascati, Manual : **Proposed Standard Practice for Surveys on Research and Experimental Development**, 2014.
- Frawley, William J., Piatetsky-Shapiro, Gregory ve Matheus, Christopher J. : “Knowledge Discovery in Databases: An Overview”, **AI Magazine** 13.3., 1992, p. 3.
- Garrick, John ve Clegg, Stewart : “Knowledge Work and The New Demands of Learning”, **Journal of Knowledge Management** 4.4, 2000, p. 283.

- Godin, Benoit : “The Knowledge Economy: Fritz Machlup’s Construction of a Synthetic Concept”, **Project on the History and Sociology of S&T Statistics Working Paper 37**, 2008, p. 4.
- Goebel, Michael ve Gruenwald, Le : “A Survey of Data Mining and Knowledge Discovery Software Tools”, **ACM SIGKDD Explorations Newsletter 1.1.**, 1999, p. 21.
- Gottschalk, Petter : **Strategic Knowledge Management Technology**, Idea Group Publishing, London, 2005.
- Göksan, Turan Selçuk, Uzundurukan, Soner ve Keskin, S. Nilay : “Yaşam Boyu Öğrenme ve Avrupa Birliği'nin Yaşam Boyu Öğrenme Programları”, **İnşaat Mühendisliği Eğitimi Sempozyumu**, 2009, s.144.
- Grabmeier, Johannes ve Rudolph, Andreas : “Techniques of Cluster Algorithms in Data Mining”, **Data Mining and Knowledge Discovery 6.4.**, 2002, p. 350.
- Green, Andy : “Models of Lifelong Learning and The Knowledge Society”, **Compare 36.3**, 2006, p. 312.
- Haır, Joseph F., Anderson, Rolph E., Tatham, Ronald L. ve Black, C. William : **Multivariate Data Analysis With Readings**, 4th Edition, USA, 1995.
- Halal, William E. : **The Logic of Knowledge: KM Principles Support Agile System, Agile Information System: Conceptualization, Construction, and Managent**, Ed. Kevin C. Desouza, Oxford, 2006.
- Han, Jiawei, Cai, Yandong ve Cercone, Nick : “Knowledge Discovery in Databases: An Attribute-Oriented Approach”, **In: VLDB**, 1992, p. 1.
- Hand, David J. : “Data Mining: Statistics and More?”, **The American Statistician, Vol.52 No.2**, 1998, p. 115.
- Hand, David J., Mannila, Heikki ve Smyth, Padhraic : **Principles of Data Mining**, MIT Press, 2001.
- Harald, Eggerstedt, Hall, Rebecca Brideau ve Wijnbergen, Sweder van : **Turkey Informatics and Economic Modernization**, The World Bank, 1993.
- Hargreaves, Andy : **Teaching in The Knowledge Society: Education in The Age of Insecurity**, Teachers College Press, 2003.
- Hargreaves, David : “Knowledge Management in The Learning Society”, **Proceeding of Forum of OECD Education Ministers**, 2000, p. 5.
- Hartigan, John A. : “Consistency of Single Linkage For High-Density Clusters”, **Journal of The American Statistical Association, Vol. 76 No. 374**, 1981, p.390.

- Houghton, John ve Sheehan, Peter : “A Primer on The Knowledge Economy, Centre of Strategic Economic Studies”, Victorya University, 2002, p. 9.
- Inmon, William H. : “Information Architecture for The 90’s: Legacy Systems, Operational Data Store, Data Warehouse”, **PRISM Tech Topic, Vol.1, No.13**, 1993, p. 120.
- International Labour Office : **World Employment Report 2001: Life at Work in the Information Economy**, Geneva, 2001., p. 1.
- Işık, Meltem ve Çamurcu, Ali Y. : “K-Means, K-Medoids ve Bulanık C-Means Algoritmalarının Uygulamalı Olarak Performanslarının Tespiti”, **İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi**, 6.11., 2007, s. 56.
- Işık, Nihat ve Kılıç, Efe Can : “Bilgi Ekonomisi ve İktisadi Büyüme: OECD Ülkeleri Üzerine Bir Uygulama”, **Akdeniz İİBF Dergisi** 26, 2013, s. 23.
- Jiawei, Han ve Kamber, Micheline : **Data Mining: Concepts and Techniques**, San Francisco, Morgan Kaufmann, 2001.
- Johnson, Richard ve Wichern, Dean : **Applied Multivariate Statistical Methods**, New Jersey, 1992.
- Jolfsson, Erik B. ve Hitt, Lorin : “Information Technology as a Factor of Production: The Role of Differences Among Firms”, **Economics of Innovation and New Technology**, 1995, p. 17.
- Kargı, Nihal : **Bilgi Ekonomisi**, Ekin Kitabevi, Bursa, 2006.
- Karlsson, Charlie ve Johansson, Börje : “Towards a Dynamic Theory for The Spatial Knowledge Economy”, **Royal Institute of Technology Centre of Excellence for Studies in Science and Innovation** 20, 2004, p. 3.
- Kavzoğlu, Taşkın ve Çölkesen, İsmail : “Destek Vektör Makineleri ile Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi”, **Harita Dergisi** 144.7., 2010, s. 2.
- Kaynak, Selahattin ve Arslan, İbrahim : “In The Context of The Knowledge Economy of Parameters: Econometrics Analysis of the Organisation for Economic Co Operation and Development (OECD) Economies”, **African Journal of Business Management** 6.12, 2012, p. 4459.
- Kefela, Ghirmai T. : “Knowledge-Based Economy and Society Has Become a Vital Commodity to Countries”, **International Journal of Educational Research and Technology**, Vol.1, No.2, 2010, p. 69.
- Kelleci, Mehmet Ali : “Bilgi Ekonomisi, İşgücü Piyasasının Temel Aktörleri ve Eşitsizlik: Eğilimler, Roller, Fırsatlar ve Riskler”, Devlet Planlama Teşkilatı, 2003, s. 26.
- Kendrick, John W. : “Productivity Trends: Capital and Labour”, **National Bureau of Economic Research**, 1957, p. 15.
- Khan, Maleq, Ding, Qin ve Perrizo, William : “K-Nearest Neighbor Classification on Spatial Data Streams Using P Trees”, **In: Pacific-Asia Conference on Knowledge Discovery and Data Mining**, Springer Berlin Heidelberg, 2002, p. 517.

- Kıracı, Arzdar ve Tiryakioğlu, Murad : “Contribution of The New Economy to The Profitability of Turkish Firms: A Sectoral Analysis”, Ed. by Sedef Murat, İbrahim Güran Yumuşak, **Bilgi Ekonomisi ve Yönetimi Dergisi 5.2**, İstanbul, 2010, s. 3.
- Kibritçioğlu, Aykut : “İktisadi Büyümenin Belirleyicileri ve Yeni Büyüme Modellerinde Beşeri Sermayenin Yeri”, **A.Ü. Siyasal Bilgiler Fakültesi Dergisi**, Cilt:53, Sayı:1-4, Ankara, 1998, s. 226.
- Kitagawa, Fumi : “Universities and Regional Advantage: Higher Education and Innovation Policies in English Regions”, **European Planning Studies 12.6**, 2004, p. 105.
- Koh, Hian Chye, Tan, Wei Chin ve Peng, Goh Chwee : “Credit Scoring Using Data Mining Techniques”, **Singapore Management Review 26.2.**, 2004, p. 101.
- Koski, Heli, Rouvinen, Petri ve Ylä-Anttila, Pekka : “ICT Clusters in Europe The Great Central Banana and The Small Nordic Potato”, **Information Economics and Policy**, 2002, p. 146.
- Koyuncugil, Ali Serhan : “Veri Madenciliği ve Sermaye Piyasalarına Uygulanması”, **Sermaye Piyasası Kurulu Araştırma Raporu**, 2007, s. 89.
- Koyuncugil, Ali Serhat ve Özgülbaş, Nermin : “Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları”, **International Journal of Informatics Technologies 2.2.**, 2009, s. 23.
- Kurtulmuş, Numan : **Sanayi Ötesi Dönüşüm**, İstanbul, İz Yayınları, 1996.
- Kuznets, Simon : “Croissance et Structure Economiques”, Paris, 1972, s. 152.
- Larose, Daniel T. : **Data Mining Methods and Models**, A John Wiley & Sons, Inc Publication, Canada, 2005.
- Larose, Daniel T. : **Discovering Knowledge in Data: An Introduction to Data Mining**, John Wiley & Sons, 2014.
- Lehr, Bill ve Lichtenberg, Frank : “Information Technology and Its Impact On Firm-Level Productivity: Evidence from Government and Private Data Sources. In: Centre for Study of Living Standards”, **Conference On Service Sector Productivity and The Productivity Paradox.**, 1997, p. 2.
- Lengel, Laura : “The Information Economy and The Internet”, **Journal Mass Commun 2.**, 2009, p.2.
- Lippmann, Richard : “An Introduction to Computing with Neural Nets”, **IEEE Assp Magazine 4. 2.**, 1987, s. 43.
- Lundvall, Bengt-Åke ve Foray, Dominique : **The Knowledge Based Economy: From The Economics of Knowledge to The Learning Economy in Employment and Growth in The Knowledge Based Economy**, Paris, 1996.
- Macleod, Christine : “The Gifts of Athena: Historical Origins of The Knowledge Economy”, Ed. by Joel Mokyr, **The Journal of Economic History**, Princeton NJ and Oxford: Princeton University Press, 2002, s. 288.

- Makó, Csaba ve Mitchell, Brian : “Knowledge Economy and Innovation: A European Comparative Perspective”, **Psychosociological Issues in Human Resource Management 1.2.**, 2013, p. 1.
- Marti, J. M. V. ve Cabrita, M. do R. : **Entrepreneurial Excellence in the Knowledge Economy Intellectual Capital Benchmarking Systems**, Palgrave Macmillan, London, 2012.
- Messaoud, Riadh Ben, Rabaséda, Sabine Loudcher, Rokia, Missaoui ve Boussaid, Omar : “OLEMAR: An Online Environment for Mining Association Rules in Multidimensional Data”, **Data Mining and Knowledge Discovery Technologies 35**, 2007, p. 30.
- Mohammed, Zaki : **Introduction to Data Mining**, Springer Verlag, 2003.
- Murat, Yetiş Şazi ve Şekerler, Alper : “Trafik Kaza Verilerinin Kümeleme Analizi Yöntemi İle Modellenmesi”, **Teknik Dergi, 20.98.**, 2009, s. 4761.
- Neil, Tim H. : **Applied Multivariate Analysis**, New York, 2002.
- OECD, Brief Policy : “Science, Technology and Innovation in The New Economy”, **OECD Observer**, 2000, p. 2.
- OECD, EUROSTAT : **Guidelines for Collecting and Interpreting Innovation Data**, Oslo, 2005.
- OECD : **Knowledge Management in The Learning Society, Organisation for Economic Co-operation and Development**, , OECD Publishing and Centre for Educational Research and Innovation, 2000.
- OECD : **Organisation for Economic Co-operation and Development, Knowledge Management in The Learning Society**, OECD Publishing and Centre for Educational Research and Innovation, 2000.
- OECD : **The Knowledge- Based Economy, Organisation for Economic Co-operation and Development**, Paris, 1996, p. 15.
- OECD : **The Impact of the Crisis on ICTs and Their Role in The Recovery**, 2009.
- Oktay, Erkan ve Kaynak, Sebahattin : “Türkiye ve Avrupa Birliği Ülkelerinin Bilgi Ekonomisi Girdi ve Çıktı Değişkenleri Arasındaki Kanonik İlişkinin Araştırılması”, **Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 10.2.**, 2007, s. 427.
- Oliner, Stephen D. ve Sichel, Daniel E. : “The Resurgence of Growth in The Late 1990s: Is Information Technology The Story?”, 2000, p. 15.
- Osowski, Stanislaw, Hoai, Linh Tran ve Markiewicz, Tomasz : “Support Vector Machine-Based Expert System for Reliable Heartbeat Recognition”, **IEEE Transactions on Biomedical Engineering 51.4.**, 2004.
- Özdamar, Kazım : **Paket Programlar ile İstatistiksel Veri Analizi**, İstanbul, Kaan Kitapevi, 5. Baskı, 2004.

- Özdemir, Bahadır ve Aksoy, Selim : “Image Classification Using Subgraph Histogram Representation”, **In: Pattern Recognition (ICPR), 20th International Conference on. IEEE**, 2010. p. 351.
- Özekes, Serhat : **Veri Madenciliği Modelleri ve Uygulama Alanları**, 2003.
- Özkan, Yalçın : **Veri Madenciliği Yöntemleri**, Papatya Yayıncılık Eğitim, 2013.
- Özmen, Şule : **Ağ Ekonomisinde Yeni Ticaret Yolu: E-Ticaret**, İstanbul Bilgi Üniversitesi Yayınları, 2003.
- Özsağır, Arif : **Bilgi Ekonomisi**, Seçkin Yayıncılık, 2013.
- Piatetsky-Shapiro, Gregory : “The Data-Mining Industry Coming of Age”, **IEEE Intelligent Systems** **14(6)**, 1999, p. 33.
- Piva, Mariacristina ve Vivarelli, Marco : “Innovation and Employment: Evidence from Italian Microdata”, **Journal of Economics**, 2005, p. 66.
- Porat, Marc Uri : **The Information Economy: Definition and Measurement**, 1977.
- Powell, Walter W. ve Snellman, Kaisa : “The Knowledge Economy”, **Annual Review of Sociology** **Vol:30**, 2004, p. 200, 214.
- Ramkumar, Gurumurthy D., Swami, Arun N. ve America, Hitachi : “Clustering Data without Distance Functions”, **IEEE Data Eng. Bull.** **21.1.**, 1998, p. 12.
- Rukancı, Fatih ve Anameriç, Hakan : “Bilgi Toplumu ve Toplumun Bilgilenmesinde Kütüphanelerin Rolü”, **Kütüphaneciliğin Destanı Uluslararası Sempozyumu**, Ekim 2004, s. 3.
- Rygielski, Chris, Wang, Jyun-Cheng ve Yen, David C. : “Data Mining Techniques for Customer Relationship Management”, **Technology in Society** **24.4.**, 2002, s. 488.
- Sala, Carme Hervada ve Bragulat, Eusebi Jarauta : “A Program To Perform Ward's Clustering Method On Several Regionalized Variables”, **Computers & Geosciences**, **Vol.30 No.8**, 2004, p. 883.
- Saran, Ulvi : “E-Devlet ve Türkiye Bilgi Toplumu Stratejisi”, **Uluslararası Katılımlı Bilgi Toplumu Sempozyumu**, 2006, s. 211.
- Savaş, Serkan, Topaloğlu, Nurettin ve Yılmaz, Mithat : “Veri Madenciliği ve Türkiye’deki Uygulama Örnekleri”, **İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi Yıl:11 Sayı: 21**, 2012, s. 5.
- Saygılı, Şeref : “Bilgi Ekonomisine Geçiş Sürecinde Türkiye Ekonomisini Dünyadaki Konumu”, **Ekonomik Modeller ve Stratejik Araştırmalar Genel Müdürlüğü, Stratejik Araştırmalar Dairesi, Yayın No: 2675**, 2003, s. 5.
- Schleicher, Andreas : **Measuring Student Knowledge and Skills: A New Framework for Assessment**, Organisation for Economic Co-Operation and Development (OECD), Paris, 1999.

- Senn, James A. : **Information Technology in Business**, Prentice Hall Inc., New Jersey, 1995.
- Shapiro, Carl ve Varian, Hal R. : **Information Rules: A Strategic Guide to The Network Economy**, Harvard Business School Press, Cambridge, 1999.
- Silahtaroglu, Gökhan : **Veri Madenciliği**, Papatya Yayınları, İstanbul, 2013.
- Singh, Jaishree, Ram, Hari ve Sodhi, J.S. : “Improving Efficiency of Apriori Algorithm Using Transaction Reduction”, **International Journal of Scientific and Research Publications** 3.1, 2013, s. 2.
- Singh, R. P. : **Teaching Teachers for A in Knowledge Society**, Gyan Publishing House, New Delhi, 2008, p. 15.
- Soete, Luc : “Macroeconomic and Structural Policy in The Knowledge-Based Economy; National Policy Challenges”, **OECD Proceedings Industrial Competitiveness in the Knowledge-Based Economy; The New Role of Governments**, Paris, 1997, p. 136.
- Stewart, Thomas A. : “Intellectual Capital: The New Wealth of Nations”, New York, 1997, p. 6.
- Sundać, Dragomir ve Krmpotić, Irena Fatur : “Knowledge Economy Factors and The Development of Knowledge Based Economy”, **Croatian Economic Survey** 13, 2011, p. 126.
- Şeker, Şadi Evren : **İş Zekâsı ve Veri Madenciliği**, 2013.
- Şentürk, Ünal : “Bilgi Toplumu Parametreleriyle Türkiye’nin Sorgulanışı”, **I. Ulusal Bilgi, Ekonomi ve Yönetim Kongresi**, 2002, s. 195.
- Tang, Zhao Hui ve MacLennan, Jamine : **Data Mining with SQL Server 2005**, John Wiley & Sons, 2005, p. 180.
- Taşcı, Kamil ve Telli, Çağatay : “Bilgi Ekonomisinin Kuramsal Çerçevesi”, **XII. Türkiye’de İnternet Konferansı**, 2007, s. 19.
- Tatlıdil, Hüseyin : **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Ankara, Engin Yayınları, 1992.
- Tekerek, Adem : “Veri Madenciliği Süreçleri ve Açık Kaynak Kodlu Veri Madenciliği Araçları”, **Akademik Bilişim** 11, 2011, s. 3.
- Ternouth, Philip : “Universities in The Knowledge Economy: Higher Education Organisation and Global Change”, **Universities as Knowledge Producers**, London, 2012, p. 41.
- Ternouthi, Philip : **Universities in The Knowledge Economy Higher Education Organization and Global Change**, Ed. by. Paul Temple, Newyork, 2012.
- TISK : “Küresel Eğilimler ve TÜİK Çalışma Kaydı”, Yayın No: 164, Ankara, 1997, s. 76.
- Turanlı, Münevver, Özden, Ünal H. ve Türedi, Serdar : “Avrupa Birliği’ne Aday ve Üye Ülkelerin Ekonomik Benzerliklerinin Kümeleme Analiziyle İncelenmesi”, **İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi**, 5.9., 2006, s. 99.

- Turban, Efraim, King, Dave, Lee, Jae Kyu ve Viehland, Dennis : “Electronic Commerce: A Managerial Perspective”, 2002, p. 3.
- Tübitak : “Bilgi Toplumu Politikaları Üzerine Bir Değerlendirme (Dünya ve Türkiye)”, Eylül, 2002, s. 4.
- Uçkan, Özgür : **E-Devlet, E-Demokrasi ve Türkiye**, Literatür Yayıncılık, İstanbul, 2003.
- Vatansever, Metin : “Görsel Veri Madenciliği Tekniklerinin Kümeleme Analizlerinde Kullanımı ve Uygulanması”, **Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Yayınlanmamış Yüksek Lisans Tezi**, 2008, s . 53.
- Wislaw, Charles D. ve Bramer, William L. : **Future Work Putting Knowledge to Work in The Knowledge Economy**, The Free Press, Newyork, 1994.
- Wojciechowski, Marek : “Interactive Constraint-Based Sequential Pattern Mining”, **East European Conference on Advances in Databases and Information Systems**, Springer Berlin, 2001, p. 2.
- World Bank : **Building Knowledge Economies: Advanced Strategies for Development**, World Bank Institute Development Studies, Washington, D.C., 2007.
- World Bank : “Lifelong Learning and The Knowledge Economy”, **Summary of the Global Conference on Lifelong Learning Organized by The World Bank**, Stuttgart, Germany, 2003, p. 1.
- World Bank : “Measuring Knowledge in The World's Economies: Knowledge Assessment Methodology and Knowledge Economy”, **Knowledge for Development Program**, 2009, p. 1.
- Yakut, Emre, Elmas, Bekir ve Yavuz, Selahattin : “Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini”, **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi 19.1.**, 2014, s. 143.
- Yamak, Rahmi ve Koçak, Necmettin Alpay : “Bilgi Teknolojisi Harcamalarının Ekonomik Büyüme Üzerine Etkileri: 1993-2005”, **Bilgi Ekonomisi ve Yönetimi Dergisi**, 2007, s. 8.
- Yanık, Bahattin ve Avcı, İlhan : “Bölgesel Debi Süreklilik Eğrilerinin Elde Edilmesi”, **İTÜ Dergisi 4.5.**, 2010, s. 22.
- Yapraklı, Sevda ve Sağlam, Tuncay : “Türkiye’de Bilgi İletişim Teknolojileri ve Ekonomik Büyüme: Ekonometrik Bir Analiz (1980-2008)”, 2010, s. 580.
- Yaylalı, Muammer, Oktay, Erkan, Akan, Yusuf ve Kaynak, Selahattin : “Türkiye ve Avrupa Birliğine Üye Ülkelerin Bilgi Ekonomisi Performanslarının Veri Zarflama Analizi Metoduyla Karşılaştırılması”, **Marmara Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi C.21 S.1**, 2007, s. 10.
- Yeloğlu, Hakkı Okan : “Bilgi Ekonomisi Değişkenlerine Yönelik İlk İzlenimler: Türkiye-OECD Ülkeleri Karşılaştırmaları (1995-1999)”, **Bilgi Dünyası 10.2**, 2009, s. 246.

- Yeo, Benjamin JK : “Driving The Knowledge Economy: Explaining The Impact of Regional Innovation Capacity on Economic Performance”, **Contemporary Management Research** 6.1, 2010, p. 83.
- Yıllancı, Veli : “Bulanık Kümeleme Analizi İle Türkiye'deki İllerin Sosyoekonomik Açından Sınıflandırılması”, **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 15.3., 2010, s. 120.
- Yıldız, Sinan : “Patent Sisteminin Ekonomik Mantığı ve Tarihsel Analizi”, **İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi**, 2010, s. 55.
- Yılmaz, Ömer ve Temurlenk, Sinan M. : “Türkiye'deki İstatistik Bölgelerin Kişi Başına Düşen Gelir Açısından Hiyerarşik ve Hiyerarşik Olmayan Kümeleme Analizi İle Değerlendirilmesi: 1965-2001”, **Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi**, 19.2.,2005, s. 80.
- Yılmaz, Şebnem Koltan ve Patır, Said : “Kümeleme Analizi ve Pazarlamada Kullanımı”, **Akademik Yaklaşımlar Dergisi**, 2.1., 2012, s. 94.
- Yohannes, Yisehac ve Webb, Patrick : **Classification and Regression Trees, CART: A User Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity**, Intl Food Policy Res. Inst., 1999.
- Yumuşak, İbrahim Güran ve Bilen, Mahmut : “Türkiye Küresel Ağa Hazır Mı? Bilgi Ekonomisi İndeksi, Beşeri Kalkınma İndeksi ve Ağa Hazırlık İndeksi Göstergeleri Üzerine Bir Değerlendirme”, **Bilgi Ekonomisi ve Yönetimi Dergisi** 5.2, 2010, s.101.
- Yücel, İsmail Hakkı : “Türkiye'de Bilim Teknoloji Politikaları ve İktisadi Gelişmenin Yönü”, **T.C. Devlet Planlama Teşkilatı (DPT) Yayın No: 2690**, Ankara, 2006, s. 10.
- Zaim, Halil ve Koçak, Orhan : “Bilgi Çalışanının Memnuniyeti”, **Jounal of Yaşar University**, 2010, s. 1.
- Zaki, Mohammad. J., Parthasarathy, Srinivasan, Ogihara, Mitsunori ve Li, Wenmin : “New Algorithms for Fast Discovery of Association Rules”, **KDD Proceedings Vol. 97**, 1997, p. 284.
- Zhang, Tian, Ramakrishnan, Raghu ve Livny, Miron : “BIRCH: An Efficient Data Clustering Method for Very Large Databases”, **In: ACM Sigmod Record**, 1996, p. 5.
- : Karar Ağaçları ve Algoritma Türleri, (Çevrimiçi) <http://bolubeyi.net/karar-agaclari-ve-algoritma-turleri.html>, 27.11.2016
- : The Optimality of Naive Bayes, (Çevrimiçi) <http://www.aaai.org/Papers/FLAIRS/2004/Flairs04-097.pdf>, 26.11.2016.
- : An Introduction to Data Mining, (Çevrimiçi) http://www.theartling.com/dmintro/dmintro_2.htm, 15.10.2016.

ÖZGEÇMİŞ

Gonca YÜZBAŞI KÜNÇ 03.08.1986 yılında Gaziantep’te doğmuştur. İlk ve orta öğretimi aynı şehirde tamamlamıştır. 2003 yılında Abdülkadir Konukoğlu Lisesinden mezun olmuştur.

2003 yılında Çukurova Üniversitesi iktisadi ve İdari Bilimler Fakültesi İşletme (İngilizce) lisans programına başlamış, 2008 yılında mezun olmuştur. aynı yılda Gaziantep Üniversitesi İşletme Anabilim Dalı’nda Yüksek Lisans’a başlamış, 2009 yılında Adıyaman Üniversitesi İşletme bölümüne Araştırma Görevlisi olarak atanmıştır. 2010 yılında Yüksek Lisansını tamamlamış ve 2011 yılında İstanbul Üniversitesi Sayısal Yöntemler Bilim Dalı’nda doktora Programına başlamıştır.

YÜZBAŞI, halen Adıyaman Üniversitesi İşletme Bölümünde Araştırma Görevlisi olarak çalışmaya devam etmektedir.

İstanbul, 2018.