

**VERİ MADENCİLİĞİ TEKNİKLERİNDEN C5.0  
ALGORİTMASI VE DESTEK VEKTÖR MAKİNELERİ İLE  
YAPAY SİNİR AĞLARININ SINIFLANDIRMA  
BAŞARILARININ KARŞILAŞTIRILMASI:  
İMALAT SEKTÖRÜNDE BİR UYGULAMA**

**Emre YAKUT**

**Doktora Tezi  
İşletme Anabilim Dalı  
Prof. Dr. Erkan OKTAY**

**2012**

**Her Hakkı Saklıdır**

**ATATÜRK ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
İŞLETME ANABİLİM DALI**

**Emre YAKUT**

**VERİ MADENCİLİĞİ TEKNİKLERİNDEN C5.0 ALGORİTMASI VE  
DESTEK VEKTÖR MAKİNELERİ İLE YAPAY SİNİR AĞLARININ  
SINIFLANDIRMA BAŞARILARININ KARŞILAŞTIRILMASI:  
İMALAT SEKTÖRÜNDE BİR UYGULAMA**

**DOKTORA TEZİ**

**TEZ YÖNETİCİSİ  
Prof. Dr. Erkan OKTAY**

**ERZURUM - 2012**



T.C.  
ATATÜRK ÜNİVERSİTESİ  
SOSYAL BİLİMLERİ ENSTİTÜSÜ



TEZ BEYAN FORMU

06/12./2012

SOSYAL BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

BİLDİRİM

Atatürk Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğine göre hazırlamış olduğum "Veri Madenciliği Tekniklerinden C5.0 Algoritması ve Destek Vektör Makineleri ile Yapay Sinir Ağlarının Sınıflandırma Başarılarının Karşılaştırılması: İmalat Sektöründe Bir Uygulama" adlı tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Atatürk Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

Lisansüstü Eğitim-Öğretim yönetmeliğinin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

Tezimin/Raporumun tamamı her yerden erişime açılabilir.

Tezimin/Raporum sadece Atatürk Üniversitesi yerleşkelerinden erişime açılabilir.

Tezimin/Raporumun ... yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

11. 12. 2012

[Tarih ve İmza]

[Emre YAKUT]



T.C.  
ATATÜRK ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ



TEZ KABUL TUTANAĞI

SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Prof. Dr. Erkan OKTAY danışmanlığında, Emre YAKUT tarafından hazırlanan bu çalışma 06 / 12 / 2012 tarihinde aşağıdaki jüri tarafından İşletme Anabilim Dalı'nda Doktora Tezi olarak kabul edilmiştir.

**Başkan** : Prof. Dr. Erkan OKTAY

İmza:

**Jüri Üyesi** : Prof. Dr. M. Vedat PAZARLIOĞLU

İmza:

**Jüri Üyesi** : Prof. Dr. Uğur YAVUZ

İmza:

**Jüri Üyesi** : Doç. Dr. M. Suphi ÖZÇOMAK

İmza:

**Jüri Üyesi** : Yrd. Doç. Dr. Emrah TALAŞ

İmza:

Yukarıdaki imzalar adı geçen öğretim üyelerine aittir. .... / ..... / .....

Prof. Dr. Mustafa YILDIRIM

Enstitü Müdürü

## İÇİNDEKİLER

<b>ÖZET</b> .....	<b>ix</b>
<b>ABSTRACT</b> .....	<b>x</b>
<b>KISALTMALAR DİZİNİ</b> .....	<b>xi</b>
<b>ÇİZELGELER DİZİNİ</b> .....	<b>xii</b>
<b>GRAFİKLER DİZİNİ</b> .....	<b>xiv</b>
<b>ŞEKİLLER DİZİNİ</b> .....	<b>xv</b>
<b>GİRİŞ</b> .....	<b>1</b>

## BİRİNCİ BÖLÜM

### VERİ MADENCİLİĞİ

<b>1.1. VERİ MADENCİLİĞİNE GİRİŞ</b> .....	<b>4</b>
<b>1.2. VERİ MADENCİLİĞİNİN TANIMI</b> .....	<b>5</b>
<b>1.3. VERİ MADENCİLİĞİ SÜRECİ</b> .....	<b>7</b>
1.3.1. Problemin Tanımlanması.....	8
1.3.2. Verilerin Hazırlanması .....	8
1.3.2.1. Veri Temizleme .....	9
1.3.2.2. Veri Normalleştirme .....	10
1.3.2.3. Veri İndirgeme.....	11
1.3.2.4. Veri Dönüştürme ve Entegrasyonu.....	11
1.3.3. Modelleme .....	12
1.3.3.1. Modelleme Tekniğinin Seçimi .....	12
1.3.3.2. Test Tasarımının Oluşturulması .....	13
1.3.3.3. Modelin Kurulması.....	13
1.3.4. Modelin Değerlendirilmesi.....	13
1.3.4.1. Sonuçların Değerlendirilmesi.....	13
1.3.4.2. Veri Madenciliği Sürecinin Gözden Geçirilmesi .....	14

1.3.4.3. Sonraki Adımların Belirlenmesi.....	14
1.3.5. Modelin Kullanılması.....	14
<b>1.4. VERİ MADENCİLİĞİNDE KARŞILAŞILAN SORUNLAR.....</b>	<b>15</b>
1.4.1. Veri Tabanı Boyutu.....	15
1.4.2. Gürültülü Veri.....	15
1.4.3. Eksik ve Kesin Olmayan Değerler.....	16
<b>1.5. VERİ MADENCİLİĞİ VE İLİŞKİDE OLDUĞU DİĞER DİSİPLİNLER ....</b>	<b>16</b>
<b>1.6. VERİ MADENCİLİĞİ YÖNTEMLERİ.....</b>	<b>18</b>
1.6.1. Bayesyen Sınıflandırma .....	18
1.6.2. Karar Ağaçları İle Sınıflandırma .....	20
1.6.2.1. ID3 Algoritması.....	23
1.6.2.2. C4.5 Algoritması.....	24
1.6.2.3. C.5 Algoritması.....	26
1.6.2.4. Sınıflandırma ve Regresyon Ağaçları (CART).....	27
1.6.3. Kümeleme Analizi .....	29
1.6.3.1. Benzerlik/Uzaklık Ölçüleri.....	30
1.6.3.2. Hiyerarşik Kümeleme .....	32
1.6.3.3. Hiyerarşik Olmayan Yöntemler.....	33
1.6.4. Birliktelik Kuralları.....	35
1.6.4.1. Apriori Algoritması.....	37
1.6.4.2. Pazar Sepet Analizi.....	39
1.6.5. Destek Vektör Makineleri .....	39
1.6.5.1 Lineer Destek Vektör Makineleri.....	43
1.6.5.2 Lineer Olmayan Destek Vektör Makineleri.....	48

## İKİNCİ BÖLÜM

### YAPAY SİNİR AĞLARI

<b>2.1. YAPAY SİNİR AĞLARININ YAPISI VE TEMEL BİLEŞENLERİ.....</b>	<b>53</b>
2.1.1. Girdiler .....	53
2.1.2. Ağırlıklar .....	54
2.1.3. Toplam Fonksiyonu .....	54
2.1.4. Aktivasyon Fonksiyonu.....	54
2.1.4.1. Doğrusal Fonksiyon .....	55
2.1.4.2. Basamak Fonksiyonu .....	56
2.1.4.3. Kutuplamalı Basamak Fonksiyonu .....	56
2.1.4.4. Parçalı Doğrusal Fonksiyon .....	57
2.1.4.5. Sigmoid Tipi Fonksiyon.....	57
2.1.4.6. Tanjant Hiperbolik Tipli Fonksiyon .....	58
2.1.4.7. Sinüs Tipli Fonksiyon .....	58
2.1.5. Hücrenin Çıktısı .....	59
<b>2.2. YAPAY SİNİR AĞLARININ YAPISI.....</b>	<b>59</b>
<b>2.3. YAPAY SİNİR AĞLARINDA ÖĞRENME .....</b>	<b>61</b>
2.3.1. Yapay Sinir Ağlarında Öğrenme Stratejileri .....	62
2.3.1.1. Danışmanlı Öğrenme .....	62
2.3.1.2. Danışmansız Öğrenme .....	63
2.3.1.3. Takviyeli Öğrenme (Reinforcement Learning) .....	64
2.3.2. Yapay Sinir Ağlarında Öğrenme Kuralları .....	66
2.3.2.1. Kohonen Öğrenme Kuralı .....	66
2.3.2.2. Hebb Öğrenme Kuralı .....	67
2.3.2.3. Hopfield Öğrenme Kuralı.....	68
2.3.2.4. Delta Kuralı (Windrow - Hoff Kuralı).....	69

2.3.2.5. Basit Algılayıcı (Perseptron) Öğrenme Kuralı .....	72
<b>2.4. YAPAY SİNİR AĞI MODELLERİ .....</b>	<b>73</b>
2.4.1. Tek katmanlı Algılayıcılar .....	73
2.4.1.1. Basit Algılayıcı (Perseptron) Modeli .....	76
2.4.1.2. ADALINE / MADALINE Modeli .....	77
2.4.2. Çok Katmanlı Algılayıcı .....	79
2.4.2.1. Geri Yayılma Ağları .....	80
2.4.2.2. İleri Besleme Aşaması .....	82
2.4.2.3. Geri Yayılma Aşaması .....	85
2.4.3. LVQ Modeli (Learning Vector Quantization) .....	86
2.4.4. ART (Adaptive Resonance Theory) Ağları .....	89
2.4.4.1. ART Ağlarının Çalışma Prensibi .....	90
2.4.4.2. ART Ağlarının Diğer Yapay Sinir Ağlarından Farkları .....	91
2.4.5. Hopfield Ağı .....	92
2.4.6. Jordan Ağı .....	93
2.4.7. Elman Ağı .....	93
2.4.8. Kohonen Ağı .....	95
<b>2.5. YAPAY SİNİR AĞ TASARIMI .....</b>	<b>96</b>
2.5.1. Öğrenme Algoritmasının Seçimi .....	97
2.5.2. Gizli Katman Sayısının Belirlenmesi .....	98
2.5.3. Gizli Katman İşlem Elemanı Sayısının Belirlenmesi .....	98
2.5.4. Yapay Sinir Ağı Parametrelerinin Belirlenmesi .....	99
2.5.4.1. Başlangıç Ağırlık Değerlerinin Atanması .....	100
2.5.4.2. Öğrenme Oranının Belirlenmesi .....	101
2.5.4.3. Momentum Oranının Belirlenmesi .....	101
2.5.4.4. Aktivasyon fonksiyonu seçimi .....	101



2.5.4.5. Verilerin Normalleştirilmesi.....	102
2.5.4.6. Eğitim ve Test Setlerinin Belirlenmesi .....	103
2.5.4.7. Yapay Sinir Ağının Eğitimi ve Testi .....	103
2.5.4.8. Eğitimin Sonlandırılması.....	105
2.5.4.9. Yapay Sinir Ağı Performans Ölçütünün Belirlenmesi.....	106
<b>2.6. YAPAY SİNİR AĞLARININ DİĞER UYGULAMALARDAN AYIRAN AVANTAJLARI .....</b>	<b>107</b>
<b>2.7. YAPAY SİNİR AĞLARININ DEZAVANTAJLARI .....</b>	<b>108</b>
<b>2.8. YAPAY SİNİR AĞLARI UYGULAMALARI .....</b>	<b>109</b>

## ÜÇÜNCÜ BÖLÜM

### UYGULAMA

<b>3.1. FİNANSAL BAŞARISIZLIK .....</b>	<b>113</b>
3.1.1. Finansal Başarısızlığın Nedenleri .....	117
3.1.1.1. İşletme Dışı (Dışsal) Faktörler .....	117
3.1.1.2. İşletme İçi Faktörler .....	119
3.1.2. Finansal Başarısızlık Tahminin Önemi .....	119
3.1.2.1. İşletme Yöneticileri Açısından Önemi.....	120
3.1.2.2. Kredi Kurumları Açısından Önemi.....	120
3.1.2.3. Yatırımcılar Açısından Önemi.....	121
3.1.2.4. Bağımsız Denetçi ve Analistler Açısından Önemi .....	122
<b>3.2. VERİLERİN DEĞERLENDİRİLMESİ VE ANALİZİ.....</b>	<b>123</b>
3.2.1. Uygulamada Kullanılan Yöntemler ve Yazılımlar .....	123
3.2.2. Uygulamanın Konusu ve Amacı.....	123
3.2.3. Analizde Kullanılan Değişkenlerin Seçimi .....	125
3.2.4. Finansal Başarısızlık Tahmini İçin Kurulan Modeller.....	127
3.2.4.1. Model 1 .....	127

3.2.4.2. Model 2 .....	127
3.2.4.3. Model 3 .....	131
3.2.4.4. Model 4 .....	132
3.2.5. Model 1: 28 Değişkenin Tümü İçin Analiz Sonuçları .....	134
3.2.5.1. Model 1 için Kullanılan YSA Mimarisinin Yapısı.....	134
3.2.5.2. Model 1 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı .	135
3.2.5.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	140
3.2.5.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	141
3.2.5.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	142
3.2.5.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	143
3.2.5.7. Model 1'in Sınıflandırma Sonuçlarının Karşılaştırılması .....	145
3.2.6. Model 2: 4 Değişken İçin Analiz Sonuçları .....	145
3.2.6.1. Model 2 için Kullanılan YSA Mimarisinin Yapısı.....	145
3.2.6.2. Model 2 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı .	146
3.2.6.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	149
3.2.6.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	150
3.2.6.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	151
3.2.6.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	152
3.2.6.7. Model 2'nin Sınıflandırma Sonuçlarının Karşılaştırılması .....	154
3.2.7. Model 3: 3 Değişken İçin Analiz Sonuçları .....	154

3.2.7.1. Model 3 için Kullanılan YSA Mimarisinin Yapısı.....	154
3.2.7.2. Model 3 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı .	155
3.2.7.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	160
3.2.7.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	161
3.2.7.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	162
3.2.7.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	163
3.2.7.7. Model 3'ün Sınıflandırma Sonuçlarının Karşılaştırılması .....	165
3.2.8. Model 4: 12 Değişken İçin Analiz Sonuçları .....	165
3.2.8.1. Model 4 için Kullanılan YSA Mimarisinin Yapısı.....	165
3.2.8.2. Model 4 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı .	166
3.2.8.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	170
3.2.8.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	172
3.2.8.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	173
3.2.8.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması .....	174
3.2.8.7. Model 4'ün Sınıflandırma Sonuçlarının Karşılaştırılması .....	175
3.2.9. Sınıflandırma Matrisinin Ayırıcı Gücünün Testi.....	176
3.2.10. 2010 Yılı Finansal Başarısızlık Tahmin Sonuçları .....	177
3.2.11. Tüm Yıllar İçin Tahmin Modellerinin Performansları.....	179
<b>SONUÇ.....</b>	<b>183</b>
<b>KAYNAKÇA.....</b>	<b>186</b>

<b>EKLER.....</b>	<b>196</b>
<b>Ek-1 Uygulama için Örnek Setlerini Oluşturan İşletmeler.....</b>	<b>197</b>
<b>Ek-2 Model 1-2-3-4 için Yapay Sinir Ağının Genel Mimari Yapısı .....</b>	<b>200</b>
<b>Ek-3 T-1 Dönemi için C5.0 Algoritması Veri Akışı .....</b>	<b>202</b>
<b>Ek-4 T-2 Dönemi için C5.0 Algoritması Veri Akışı .....</b>	<b>203</b>
<b>Ek-5 T-3 Dönemi için C5.0 Algoritması Veri Akışı .....</b>	<b>204</b>
<b>Ek-6 T-4 Dönemi için C5.0 Algoritması Veri Akışı .....</b>	<b>205</b>
<b>Ek-7 T-1 Dönemi için DVM Veri Akışı .....</b>	<b>206</b>
<b>Ek-8 T-2 Dönemi için DVM Veri Akışı .....</b>	<b>207</b>
<b>Ek-9 T-3 Dönemi için DVM Veri Akışı .....</b>	<b>208</b>
<b>Ek-10 T-4 Dönemi için DVM Veri Akışı .....</b>	<b>209</b>
<b>Ek-11 Başarısızlıktan 1, 2, 3 ve 4 Yıl Öncesi İçin YSA Analizinin Tahmin Sonuçları .....</b>	<b>210</b>
<b>Ek-12 Başarısızlıktan 1, 2, 3 ve 4 Yıl Öncesi İçin C5.0 Algoritmasının Tahmin Sonuçları .....</b>	<b>212</b>
<b>Ek-13 Başarısızlıktan 1, 2, 3 ve 4 Yıl Öncesi İçin Dvm Yönteminin Tahmin Sonuçları .....</b>	<b>214</b>
<b>ÖZGEÇMİŞ.....</b>	<b>216</b>

**ÖZET****DOKTORA TEZİ****VERİ MADENCİLİĞİ TEKNİKLERİNDEN C5.0 ALGORİTMASI VE DESTEK  
VEKTÖR MAKİNELERİ İLE YAPAY SİNİR AĞLARININ SINIFLANDIRMA  
BAŞARILARININ KARŞILAŞTIRILMASI:  
İMALAT SEKTÖRÜNDE BİR UYGULAMA****Emre YAKUT****Tez Danışmanı: Prof. Dr. Erkan OKTAY****2012, 217 sayfa****Jüri : Prof. Dr. Erkan OKTAY (Danışman)****Prof. Dr. M. Vedat PAZARLIOĞLU****Prof. Dr. Uğur YAVUZ****Doç. Dr. Suphi ÖZÇOMAK****Yrd. Doç. Dr. Emrah TALAŞ**

Günümüzde işletmeler çok büyük miktarlarda veri üretmekte ve bu veri içinde anlamlı ve yararlı bilgiler ortaya çıkarmakta sıkıntılar yaşamaktadırlar. Veriler tek başlarına bir anlam taşımadıkları için veri madenlerinde işlendiği zaman anlamlı hale gelebilmektedir. Bu nedenle, veri kaynaklarından gelen veriyi derlemek, yorumlamak ve işlerlik kazandırmak için veri madenciliğine ihtiyaç duyulmaktadır.

Bu çalışmanın amacı, İMKB’de işlem gören sanayi işletmelerinin 2002-2010 yılları arası verileri üzerinde, işletmelerin finansal başarısızlıklarını veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri ile birlikte yapay sinir ağları kullanılarak tahmin modellerini gerçekleştirmektir. Elde edilen modellerin dört yıl öncesine kadar ki performansları karşılaştırılarak ileriye yönelik 2010 yılına ait işletme başarısızlıklarının tahmini gerçekleştirilmiştir.

Analiz sonuçları ele alındığında, yapay sinir ağları yönteminin, hem C5.0 hem de destek vektör makineleri yöntemiyle karşılaştırıldığında, oluşturulan dört modelin tümü için daha iyi sonuçlar verdiği gözlenmiştir. Ayrıca modeller arası kıyaslama yapıldığında 3. Modelin 2. Modelden, 4. Modelin ise 1. Modelden daha yüksek sınıflandırma yüzdesiyle sonuçlar ürettiği görülmüştür.

**Anahtar Kelimeler:** Veri Madenciliği, C5.0 Algoritması, Destek Vektör Makineleri, Yapay Sinir Ağları

**ABSTRACT**  
**Ph. D. DISSERTATION**

**THE COMPARISON OF THE CLASSIFICATION SUCCESSES OF THE  
ARTIFICIAL NEURAL NETWORKS THROUGH DATA MINING TECHNIQUES OF  
C5.0 ALGORITHM AND SUPPORTING VECTOR MACHINES:  
AN APPLICATION IN MANUFACTURING SECTOR**

**Emre YAKUT**

**Advisor : Prof. Dr. Erkan OKTAY**

**2012 – PAGE: 217**

**Jury : Prof. Dr. Erkan OKTAY (Advisor)**

**Prof. Dr. M. Vedat PAZARLIOĞLU**

**Prof. Dr. Uğur YAVUZ**

**Assistant Professor Suphi ÖZÇOMAK**

**Assistant Professor. Emrah TALAŞ**

Today, businesses produce great amounts of data and have trouble in working out significant and useful information out of them. As the data don't have meaning alone, they become significant when they are processed in data mines. Therefore, data mining is needed to compile, interpret and make the data obtained from the data sources functional.

The goal of this is to develop prediction models for the financial failures of businesses by using artificial neural networks together with data mining techniques of C5.0 algorithm and supporting vector machines based on the data of the industrial businesses dealt at Istanbul Stock Exchange Market between the years of 2002 and 2010. By comparing the prior performances, as far as four years, of the obtained models a prediction of their upcoming business failures belonging to the year of 2010 was carried out.

When the results of the analysis were handled, it was observed that the method of artificial neural networks yielded better outcomes for the four formed models altogether when compared to C5.0 and the method of supporting vector machines. Besides, when a comparison was made between models, it was seen that the third model produced better results than the second one and the fourth model produced better results than the first one.

**Key Words:** Data mining, C5.0 Algorithm, Supporting Vektor Machines, Artificial Neural Networks

**KISALTMALAR DİZİNİ**

- IMKB : İstanbul Merkez Kıymetler Borsası  
VM : Veri Madenciliği  
DVM : Destek Vektör Makineleri  
YSA : Yapay Sinir Ağları  
LVQ : Learning Vector Quantization  
ART : Adaptive Resonance Theory

## ÇİZELGELER DİZİNİ

<b>Çizelge 1.1.</b> Kümeleme Analizinde Kullanılan Uzaklık Ölçüleri .....	31
<b>Çizelge 2.1.</b> Öğrenme algoritmalarının incelenmesi .....	65
<b>Çizelge 2.2.</b> Perceptron algoritmasının işleyişi .....	76
<b>Çizelge 2.3.</b> Ağ Türlerinin Sınıflandırılması.....	97
<b>Çizelge 2.4.</b> Öğrenme Algoritmaları ve Kullanım Alanları.....	98
<b>Çizelge 2.5.</b> Veri Normalizasyonunda Kullanılan Formüller .....	102
<b>Çizelge 2.6.</b> Yapay Sinir Ağları Performans Ölçüleri .....	106
<b>Çizelge 3.1.</b> Mali (Finansal) Başarısızlık/İflas Tanımı.....	116
<b>Çizelge 3.2.</b> Model Katsayıları İçin Omnibus Testi .....	128
<b>Çizelge 3.3.</b> Hosmer ve Lemeshow Uyum Testi Sonuçları .....	128
<b>Çizelge 3.4.</b> Modelin Özeti .....	129
<b>Çizelge 3.5.</b> Hosmer ve Lemeshow Kontenjans Çizelgesi .....	130
<b>Çizelge 3.6.</b> Lojistik Regresyon Analizi Sonucu Belirlenen Değişkenler .....	130
<b>Çizelge 3.7.</b> Özdeğerler.....	131
<b>Çizelge 3.8.</b> Wilk's Lambda Değeri .....	132
<b>Çizelge 3.9.</b> Standardize Kanonik Diskriminant Fonksiyon Katsayıları.....	132
<b>Çizelge 3.10.</b> Model 1 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri .....	135
<b>Çizelge 3.11.</b> Model 1 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	140
<b>Çizelge 3.12.</b> Model 1 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	141
<b>Çizelge 3.13.</b> Model 1 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	142
<b>Çizelge 3.14.</b> Model 1 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	143
<b>Çizelge 3.15.</b> Model 2 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri .....	146
<b>Çizelge 3.16.</b> Model 2 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	149
<b>Çizelge 3.17.</b> Model 2 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	150
<b>Çizelge 3.18.</b> Model 2 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	151
<b>Çizelge 3.19.</b> Model 2 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	152
<b>Çizelge 3.20.</b> Model 3 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri.....	155
<b>Çizelge 3.21.</b> Model 3 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	160



<b>Çizelge 3.22.</b> Model 3 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	161
<b>Çizelge 3.23.</b> Model 3 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	162
<b>Çizelge 3.24.</b> Model 3 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	163
<b>Çizelge 3.25.</b> Model 4 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri.....	166
<b>Çizelge 3.26.</b> Model 4 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	171
<b>Çizelge 3.27.</b> Model 4 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	172
<b>Çizelge 3.28.</b> Model 4 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	173
<b>Çizelge 3.29.</b> Model 4 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı .....	174
<b>Çizelge 3.30.</b> Modellerin Sınıflandırma Matrisinin Ayırma Gücünün Testi.....	177
<b>Çizelge 3.31.</b> 2010 yılı için YSA-C5.0-DVM Sınıflandırma Başarısı .....	178

**GRAFİKLER DİZİNİ**

<b>Grafik 3.1.</b> Yıllar itibari ile başarısız işletme sayısı .....	124
<b>Grafik 3.2.</b> Model 1'in Sınıflandırma Sonuçları .....	145
<b>Grafik 3.3.</b> Model 2'nin Sınıflandırma Sonuçları.....	154
<b>Grafik 3.4.</b> Model 3'ün Sınıflandırma Sonuçları.....	165
<b>Grafik 3.5.</b> Model 4'ün Sınıflandırma Sonuçları.....	175
<b>Grafik 3.6.</b> 2010 Yılı Sınıflandırma Sonuçlarının Karşılaştırılması.....	179
<b>Grafik 3.7.</b> Her Dört Modelin Tüm Dönemlere İlişkin Performans Ölçüleri .....	181

## ŞEKİLLER DİZİNİ

Şekil 1.1. Veri Madenciliği Süreci .....	7
Şekil 1.2. Veri Madenciliği ve İlişkide Olduğu Disiplinler .....	16
Şekil 1.3. X ve Y nitelikleri üzerine uygulanan testleri içeren basit bit karar ağacı .....	21
Şekil 1.4. Optimum ayırıcı düzlemler .....	41
Şekil 1.5. DVM Sınıflandırıcı .....	42
Şekil 1.6. Lineer Ayrılabilme Durumunda Optimum Ayırıcı Aşırı düzlem .....	44
Şekil 1.7. Birbirinden Doğrusal Olarak Ayrılamayan Veriler .....	46
Şekil 1.8. Lineer Ayrılamama Durumunda Optimum Ayırıcı Aşırıdüzlem .....	47
Şekil 1.9. Doğrusal Olmayan Sınıflandırıcılar .....	48
Şekil 1.10. İki Boyutlu Uzaydaki Verilerin Üç Boyutlu Uzaydaki Verilere .....	49
Şekil 1.11. Yüksek Dereceli Çekirdek Fonksiyonu ile Nitelik Uzayına Dönüşüm .....	50
Şekil 2.1. Yapay Sinir Hücresinin Yapısı .....	53
Şekil 2.2. Doğrusal veya Lineer Fonksiyon .....	55
Şekil 2.3. Basamak Fonksiyonları .....	56
Şekil 2.4. Tek Kutuplamalı Basamak Fonksiyonu .....	57
Şekil 2.5. Parçalı Doğrusal Fonksiyon .....	57
Şekil 2.6. Sigmoid Fonksiyonu .....	58
Şekil 2.7. Tanjant Hiperbolik Fonksiyonu .....	58
Şekil 2.8. Sinüs Tipli Fonksiyon .....	59
Şekil 2.9. Bir Yapay Sinir Ağının Genel Yapısı .....	60
Şekil 2.10. Danışmanlı Öğrenme Yapısı .....	63
Şekil 2.11. Danışmansız Öğrenme Yapısı .....	64
Şekil 2.12. Takviyeli Öğrenme Yapısı .....	65
Şekil 2.13. Tek Katmanlı Algılayıcı Yapısı .....	74
Şekil 2.14. Ağırlıkları ve sınıfları birbirinden ayıran doğrunun geometrik gösterimi ...	75
Şekil 2.15. Bir Perceptronun Yapay Sinir Modeli .....	76
Şekil 2.16. Bir ADALINE Yapısı .....	77
Şekil 2.17. Bir MADALINE Yapısı .....	78
Şekil 2.18. Örnek Geri Yayılma Ağı .....	82
Şekil 2.19. Bir Ağda Geri Yayılımlı Öğrenme Algoritmasının Uygulanması .....	86
Şekil 2.20. LVQ Ağ yapısı .....	88

<b>Şekil 2.21.</b> ART ağının genel yapısı.....	89
<b>Şekil 2.22.</b> Aşağıdan yukarıya bilgi işleme.....	90
<b>Şekil 2.23.</b> Yukarıdan Aşağıya Bilgi İşleme.....	91
<b>Şekil 2.24.</b> Hopfield Ağ Yapısı .....	92
<b>Şekil 2.25.</b> Jordan Ağı.....	93
<b>Şekil 2.26.</b> Elman Ağ Yapısı.....	94
<b>Şekil 2.27.</b> Elman ağının ayrıntılı gösterimi .....	95
<b>Şekil 2.28.</b> Kohonen SOM haritası.....	96
<b>Şekil 2.29.</b> Büyüyen Ağlar .....	99
<b>Şekil 2.30.</b> ÇKA Ağlarında Başlangıç Noktasının Etkisi .....	100
<b>Şekil 2.31.</b> Hata – Eğitim Devresi Sayısı.....	104
<b>Şekil 2.32.</b> Eğitimin Durdurma Bölgesi.....	105
<b>Şekil 3.1.</b> Finansal Başarısızlığın Oluşum Süreci .....	114
<b>Şekil 3.2.</b> Model 1'in Karar Ağacı Modeli .....	136
<b>Şekil 3.3.</b> Faizleri Karşılama Oranının Finansal Durum Sınıflandırılmasındaki Etkisi	137
<b>Şekil 3.4.</b> Cari Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	138
<b>Şekil 3.5.</b> Vergi Öncesi Kârdaki Büyüme Oranının Finansal Durum Sınıflandırılmasındaki Etkisi .....	138
<b>Şekil 3.6.</b> Finansal Giderlerin Net Satışlara Oranının Finansal Durum Sınıflandırılmasındaki Etkisi .....	139
<b>Şekil 3.7.</b> Model 2'nin Karar Ağacı Modeli .....	147
<b>Şekil 3.8.</b> Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	147
<b>Şekil 3.9.</b> Cari Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	148
<b>Şekil 3.10.</b> Model 3'ün Karar Ağacı Modeli.....	156
<b>Şekil 3.11.</b> Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	157
<b>Şekil 3.12.</b> Borç Oranının Finansal Durum Sınıflandırılmasındaki Etkisi .....	158
<b>Şekil 3.13.</b> Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	158
<b>Şekil 3.14.</b> Brüt Kârın Toplam Aktife Oranının Finansal Durum Sınıflandırılmasındaki Etkisi .....	159
<b>Şekil 3.15.</b> Model 4'ün Karar Ağacı Modeli.....	167
<b>Şekil 3.16.</b> Asit Test Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	168

<b>Şekil 3.17.</b> Faizleri Karşılama Oranının Finansal Durum Sınıflandırılmasındaki Etkisi .....	169
<b>Şekil 3.18.</b> Stok/Satışlar Oranının Finansal Durum Sınıflandırılmasındaki Etkisi .....	169
<b>Şekil 3.19.</b> Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi.....	170

## GİRİŞ

Günümüzde işletmeler çok büyük miktarlarda veri üretmekte ve bu veri içinde anlamlı ve yararlı bilgiyi ortaya çıkarmakta, verilerin depolanması ve kullanılması konusunda zorluklar yaşamaktadırlar. Geleneksel istatistiki yöntemlerle büyük hacimdeki verilerden anlamlı ve kullanışlı bilgiyi ortaya çıkarmak kolay değildir. Yığın halindeki verilerin anlamlı ve kullanışlı bilgiye dönüştürülmesi için farklı yöntemlere gereksinim duyulmuştur. Veri madenciliği teknikleri ve yapay sinir ağları bu gereksinimi karşılamak üzere ortaya çıkmıştır.

İşletme yöneticileri için sahip oldukları yığın halindeki verilerin anlamlı ve kullanışlı bilgiye dönüşmesi önemli bir rekabet avantajı sağlayacaktır. İşletmenin gelecekteki durumunun öngörülmesi geçmişteki verilerin doğru değerlendirilip analiz edilmesine bağlıdır. Dolayısıyla veriler arasındaki ilişkilerin ortaya çıkarılması için veri madenciliği teknikleri kullanılmaktadır. Bu ilişkilerin tespit edilmesinde, problemin karmaşıklığı ve analizin zorlaştığı durumlara göre farklı teknikler kullanılabilir.

Veri madenciliğinin son dönemlerde bilgi endüstrisinde önem kazanmasında en önemli etken, artan veri ve bu verinin yararlı bilgilere dönüştürülmesine duyulan ihtiyaçtır. Veri madenciliği ile ilgili literatürde çok sayıda tanım yapılmış olup ortak çıkan sonuç, veri ambarlarında bilginin keşfi, var olan verilerden yola çıkarak önceden bilinmeyen ilişkilerin ortaya konulmasıdır.

Teknolojinin hızla gelişmesiyle birlikte, insan beynini taklit eden sistemler kurmaya yönelik çalışma alanına sahip olan ve bu tez çalışmasında ele alınan diğer bir yöntem ise yapay sinir ağlarıdır. Yapay sinir ağları günümüzde pek çok problemleri çözmeye yaygın olarak kullanılmaktadır. Yapılan çalışmalardaki sonuçlar, istatistiki tekniklerin yetersiz kaldığı durumlarda, yapay sinir ağlarının daha iyi sonuçlar elde ettiği gözlenmiştir.

Bu çalışmanın amacı; İMKB'de işlem gören sanayi işletmelerinin finansal başarısızlıklarını tahmin etmede veri madenciliği tekniklerinden C5.0 algoritması ve Destek Vektör Makineleri ile Yapay Sinir Ağları karşılaştırılarak, en uygun yöntemi belirlemektir. Bu çalışmada denetimli öğrenme kapsamında yer alan geri yayımlı çok katmanlı yapay sinir ağı tercih edilmiştir.

İMKB’de işlem gören 60 başarısız ve 60 başarılı olmak üzere 120 sanayi işletmesi üzerinde bu üç yöntem kullanılarak tahmin modelleri geliştirilmiştir. Çalışmada, sanayi işletmelerinin 2002-2010 yılları arası verileri kullanılmıştır. Elde edilen modellerin dört yıl öncesine kadar olan performansları değerlendirilmiştir.

İşletmelerin finansal başarısızlıklarının tahmininde dört farklı model uygulanmıştır. Birinci modelde tüm değişkenlerin, ikinci modelde değişken seçimi için lojistik regresyon analizinden elde edilen dört değişkenin, üçüncü modelde diskriminant analizinden elde edilen üç değişkenin ve dördüncü modelde feature selection yöntemiyle elde edilen on iki değişkenin yer aldığı modellerin performans ölçütleri karşılaştırılmıştır.

Çalışma üç bölümden oluşmaktadır. Birinci bölümde veri madenciliğinin tanımı, süreci, veri madenciliğinde karşılaşılan sorunlar, veri madenciliğinin ilişkili olduğu diğer disiplinler ortaya konulduktan sonra veri madenciliği sınıflandırma yöntemlerinden Bayesyen Sınıflandırma, Karar Ağaçları ile Sınıflandırma tekniklerinden ID3 Algoritması, C4.5 Algoritması, C5.0 Algoritması ve CART, Kümeleme Analizi, Birliktelik Kuralı ve Destek Vektör Makineleri hakkında bilgiler verilmiştir.

İkinci bölümde ise yapay sinir ağları metodolojisi genel hatlarıyla tanıtılmış, yapay sinir ağlarının altyapısı, mimarisi, aktivasyon fonksiyonlarının tanıtımı, sinir ağlarında kullanılan temel öğrenme stratejileri ve öğrenme kuralları, yapay sinir ağları modelleri detaylı olarak incelendikten sonra yapay sinir ağının tasarımı, diğer uygulamalardan ayıran avantajları, son olarak yapay sinir ağlarının avantajları ve dezavantajlarından bahsedilmiştir.

Üçüncü bölüm iki kısımdan oluşmaktadır. İlkinde finansal başarısızlığın tanımı ve önemi ele alınmıştır. İkincisinde ise tezin ampirik çalışması yer almaktadır. İşletme başarısızlığının tahmini için bir sınıflandırma problemi oluşturulan uygulamanın veri seti üzerinde dört farklı model geliştirilmiş, başarısızlıktan bir, iki, üç ve dört sene öncesinden veri madenciliği tekniklerinden C5.0 Algoritması ve Destek Vektör Makineleri ile birlikte Yapay Sinir Ağlarıyla çalıştırılmıştır. Üç yöntemin sonuçları kıyaslanırken aynı zamanda oluşturulan modeller açısından da hangi modelin işletme başarısızlığını tahmin etmede daha anlamlı sonuçlar verdiği incelenmiş ve ileriye

yönelik 2010 yılına ait başarısızlık tahmini gerçekleştirilerek elde edilen sonuçlar grafiklerle gösterilmiştir. Son aşamada ise tüm yıllar için tahmin modellerinin performansları ele alınmıştır.

Tüm bu arařtırmaların bulgu ve yorumları, tezin sonuç kısmında ayrıntılı olarak açıklanmıştır.



## BİRİNCİ BÖLÜM

### VERİ MADENCİLİĞİ

#### 1.1. VERİ MADENCİLİĞİNE GİRİŞ

Veri madenciliği karar destek sistemlerinin kullandığı araçlardan biridir. Veri madenciliğinin tanımı farklı kaynaklardan araştırıldığında, yapılmış olan tanımların birbirine çok benzediği görülmektedir. Tüm kaynaklarda ortak olan nokta, veri madenciliğinin veri yığınları içerisinde, anlamlı ve faydalı ilişki ve modellerin keşfi süreci olduğudur. Bu işlemin “keşif” odaklı doğası sebebiyle kimi kaynaklar veri madenciliğini “veri tabanlarında bilgi keşfi süreci” olarak da adlandırmışlardır.<sup>1</sup>

Veri Madenciliği en basit tanımıyla bilgiyi büyük veri yığınları içerisinde çıkararak ya da “kazmaktır”.<sup>2</sup> Yirminci yüzyılın ortalarından itibaren teknolojinin gelişmesiyle birlikte ortaya çıkan dijitalleşme süreci ile veri kaynakları artmış, ancak bilgiye ulaşmak daha kolay hale gelmemiştir. Kamu alanında, bilimsel çalışmalarda ve iş hayatında devamlı büyüyen veri yığınları kaydedilmektedir. 1990 sonrası veritabanlarının hızlı gelişimi ile birlikte bilgileri keşfetme ve analiz etme, karışık bir işlem haline dönüşmüştür. Büyük veri yığınlarının çok küçük miktarı kullanılmaktadır. Çoğu durumda veriler kontrol edilebilirlik açısından çok büyük, analiz edilebilirlik açısından düzensiz ve anlamsız bir durumdadır.<sup>3</sup> Veri madenciliği bu çelişki üzerine kurulmuştur. Yöneticilerin elinde büyük miktarda veri yığınları bulunmasına rağmen yöneticiler için bir anlam ifade etmeyebilmektedir. Bu durum zamanla geliştirilen tekniklerle yeni bir uygulama alanı açmış ve veri madenciliği kavramını ortaya çıkarmıştır.

<sup>1</sup> Selwyn Piramuthu, “Evaluating feature selection methods for learning in data mining applications” [Veri Madenciliği Uygulamalarında Öğrenme Yöntemleri için Değerlendirme Özelliği Seçimi], *European Journal of Operational Research*, 156 (2), 2004, pp. 483, Erişim Tarihi: 14 Aralık 2011, ScienceDirect.

- Jing Luan, “Data Mining and Knowledge Management in Higher Education”, *Workshop and Presentation at 42nd Associate of Institutional Research International Conference*, Toronto, Canada 2002, pp.1.

- Dick Scott, Aleksandra Meeks, Mark Last, Horst Bunke ve Abraham Kandel, “Data mining in software metrics databases” [Yazılım Ölçümleri Veritabanında Veri Madenciliği], *Fuzzy Sets and Systems*, 145 (1), 2004, pp. 82, Erişim Tarihi: 14 Kasım 2011, ScienceDirect.

- F. Joseph, L. Hair, Ronald Tahtam, E. Rolph Anderson, and William Black, *Multivariate Data Analysis* [Elektronik Sürüm], Prentice Hall, New Jersey 1998, pp.674-675.

<sup>2</sup> J. Ham and M. Kamber, *Data Mining: Concepts and Techniques* [Elektronik Sürüm], Morgan Kaufmann Publisher, San Fransisco 2000, s.5.

<sup>3</sup> M. Kantardzic, *Data Mining Concepts, Models, Methods and Algorithms* [Elektronik Sürüm], IEEE Press, Piscataway- NJ 2001, s.7.

Veri madenciliği yaklaşımının başarısı üç farklı alandaki gelişme ile ilişkilidir. İlki, matematik alanında verimli ve hızlı yeni algoritmaların geliştirilmesidir. İkincisi, veritabanı teknolojilerindeki araştırma ve geliştirme çalışmalarının sonuçları. Son olarak da, bilgisayar teknolojisindeki yeni ve güçlü mimarilerin, çok büyük miktarlardaki veri kümelerini detaylı olarak incelemeye olanak vermesidir. Büyük miktardaki verilerin analiz edilip filtrelenmesi, bilgi ihtiyaçları her geçen gün artan işletme yöneticilerinin sadece kendileri ile ilgili bilgilere erişimini kolaylaştırmaktadır.<sup>4</sup>

Veri oluşturma ve toplama hızı her geçen gün daha da artmaktadır. Bilgi teknolojisindeki gelişmeler, büyük miktarlardaki verinin etkin ve kolay bir biçimde toplanıp depolanmasını olası kılmıştır. Birçok ticari ürün için kullanılmakta olan barkodlar, ticari işlemlerin ve resmi dairelerdeki işlemlerin neredeyse tamamen bilgisayarlara bağımlı hale gelmesi ve veri toplama araçlarındaki gelişmeler, bizi çok büyük miktarlarda veri ile karşı karşıya getirmiştir. İşletmecilikte, resmi kurumlarda, bilimsel ve mühendislik verilerinin yönetiminde milyonlarca veritabanı kullanılmaktadır. Kullanılan veritabanlarının sayısındaki hızlı artış, her geçen gün teknolojik olarak daha da geliştirilen güçlü ve ekonomik veritabanı sistemlerinin varlığı ile açıklanabilir. Veri ve veritabanı sistemlerindeki bu hızlı artış, işlenmiş veriyi kullanışlı bilgiye dönüştürebilecek yeni teknolojilere olan ihtiyacı arttırmıştır. Bu nedenle veri madenciliği her geçen gün önemi artan bir araştırma alanı haline gelmiştir.<sup>5</sup>

## 1.2. VERİ MADENCİLİĞİNİN TANIMI

Veri madenciliği kavramını anlayabilmek için işin en başında kelimelerin yalın anlamlarından yola çıkılabilir. Madencilik yeryüzünün gizli ve kıymetli kaynaklarının açığa çıkarılması süreci olup, bu kelimenin veri kelimesi ile ilişkilendirilmesi ise veri

<sup>4</sup> Jaesung Sim, *Critical Success Factors in Data Mining Projects*, (Yayımlanmamış Doktora Tezi), University of North Texas, Texas 2003, s.6., <http://www.umi.com>, UMI Number: 3106912

<sup>5</sup> Reijo Sund, "Utilization of Administrative Registers using Statistical Knowledge Discovery", International Workshop on "Mining Official Data", *National Research and Development Centre for Welfare and Health*, Helsinki- Finland 2002, s.1.

- Ming-Syan Chen, S. Jiawei, Philip Yu Han, "Data Mining: An Overview from Database Perspective", [Veri Madenciliği: Veritabanı Penceresinden Bir Bakış], *IEEE Transactions on Knowledge and Data Engineering*, 8 (6), 1996, pp. 866-869, Erişim Tarihi 08 Aralık 2011, [www.nyu.edu/classes/jcf](http://www.nyu.edu/classes/jcf)

yığınları içerisinde ilk bakışta fark edilemeyen kıymetli bilgilerin bulunması ve çıkartılması fikrini uyandırmaktadır.<sup>6</sup>

Veri madenciliği bir disiplin olarak ortaya çıktığında, veri tabanlarına depolanan verilerin, her 20 ayda bir, ikiye katladığı iddia edilmiştir.<sup>7</sup>

Literatürdeki veri madenciliği ile ilgili birçok tanım aşağıda verilmektedir:

- Veri madenciliği, geniş veritabanlarından bilgi çıkarımını hedeflemek için makine öğrenimi, örüntü tanıma, istatistik, veritabanı ve görselleştirme tekniklerini bir araya getiren disiplinler arası bir alandır.<sup>8</sup>
- Veri madenciliği, büyük ölçekli veriler arasından bilgiye ulaşma, bilgiyi madenleme işidir.<sup>9</sup>
- Veri madenciliği, veri sahibi için anlamlı ve kullanışlı olan beklenmeyen ilişkileri bulmak ve orijinal yollarla veriyi özetlemek için gözlenen veri kümelerinin (genellikle çok büyük) analizidir.<sup>10</sup>
- Veri madenciliği, otomatik veya yarı-otomatik biçimlerde verinin analiz edilerek gizli örüntülerin bulunmasıdır.<sup>11</sup>
- Veri madenciliği veride var olan örüntüleri keşfetme sürecidir. Süreç otomatik veya (daha çok) yarı otomatiktir. Keşfedilen örüntüler anlamlı olmalıdır ve genellikle ekonomik avantaj olmak üzere fayda sağlamalıdır.<sup>12</sup>
- Veri madenciliği, geniş veri tabanlarından bilginin çıkarımı konusunu adreslemek için, makine öğrenimi, örüntü tanıma, istatistik, veri tabanları ve görselleştirme alanlarından teknikleri birleştiren disiplinler arası bir alandır.<sup>13</sup>

<sup>6</sup> P. Giudici, *Applied Data Mining: Statistical Methods for Business and Industry* [Elektronik Sürüm], West Sussex, John Wiley& Sons, England 2003, s.1.

<sup>7</sup> Paul D. Scott, "How Size Matters: The Role of Sampling in Data Mining" [Veri Madenciliğinde Örnekleme Yöntemi: Örnek Hacmi Ne kadar Önemli], R.A. Sarker (Ed.), *Heuristic and Optimization for Knowledge Discovery*, Idea Group USA, 2002, pp. 122. Erişim Tarihi: 10 Aralık 2011, <http://www.idea-group.com>.

<sup>8</sup> P. Cabena, P. Hadjinian, R. Stadler, J. Verhees, and A. Zanasi, *Discovering Data Mining: From Concept to Implementation*, (1 edition), Prentice Hall PTR, NJ Upper Saddle River USA 1998, s.12.

<sup>9</sup> Han and Kamber, s.5.

<sup>10</sup> David Hand, Heikki Mannila and Padhraic Smyth, *Principles of Data Mining* [Elektronik Sürüm], The Mit Press, England 2001, s.1.

<sup>11</sup> Z. H. Tang, and J. MacLennan, *Data Mining with SQL Server 2005*, Wiley Publishing Inc, Indianapolis - USA 2005, s.2.

<sup>12</sup> I. H. Witten, and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques* [Elektronik Sürüm], Second Edition, Morgan Kaufmann Publishers, San Francisco 2005, s.5.

<sup>13</sup> Daniel T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining* [Elektronik Sürüm], Wiley and Sons publishing, USA 2005, s.2.

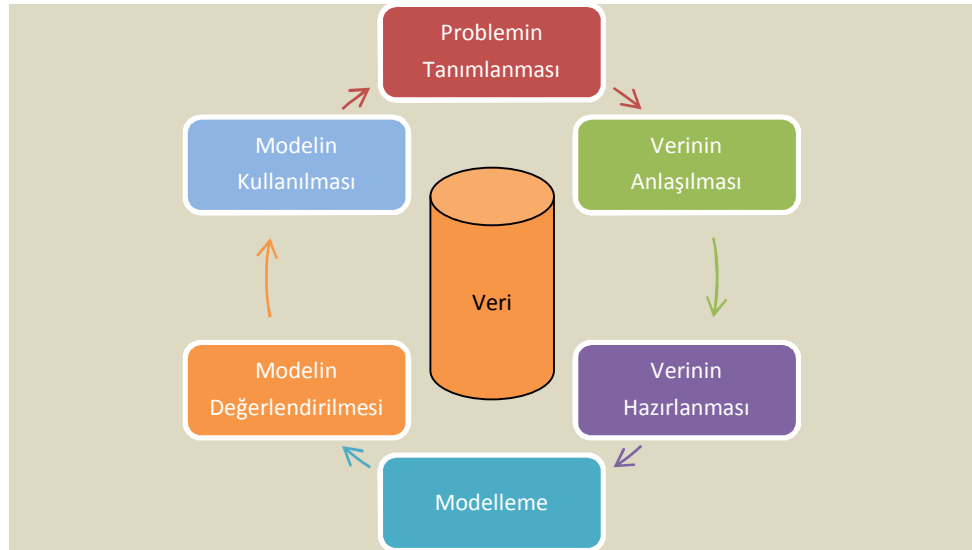
- Veri madenciliği, büyük veri depolarında faydalı bilgilerin otomatik olarak keşfedilmesi sürecidir.<sup>14</sup>

Literatürde yer alan tanımlardan ve veri madenciliği süreçlerinde yaşanan deneyimlerden yola çıkılarak şu tanım yapılabilir;

- Veri madenciliği, büyük veri setlerinde veritabanlarında bulunan veriler arasında var olan, bilinmeyen, klasik yöntemlerle görülmeyen ve sıradan olmayan ilişkileri, örüntüleri, belirli yapıları ve eğilimleri ortaya çıkarmak amacıyla istatistik, matematik, makine öğrenimi ve bilgisayar uygulamaları alanlarının birleşimi tekniklerini kullanarak analiz edilmesi ve sonuçların anlamlı bir şekilde özetlenmesi ve görselleştirilmesi işlevidir.

### 1.3. VERİ MADENCİLİĞİ SÜRECİ

Veri madenciliği uygulamalarına başlarken ilk adım problemi tanımlamaktır. Sonrasında verilerin hazırlanması gelir. Verileri hazırladıktan sonra ise model kurulup değerlendirilir, izlenir ve son olarak da model doğrulanır.<sup>15</sup>



**Şekil 1.1.** Veri Madenciliği Süreci

**Kaynak:** Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer ve Rüdiger Wirth, *CRISP-DM 1.0 Step-bystep data mining guide*, 2000, pp.13, Erşim Tarihi: 17 Ocak 2012, SPSS: [www.crisp-dm.org/](http://www.crisp-dm.org/).

<sup>14</sup> P. N. Tan, M. Steinbach and V. Kumar, *Introduction to Data Mining*, Pearson, Addison- Wesley, Boston -MA -USA 2006, s.2.

<sup>15</sup> Pete Chapman, Clinton Julian, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer and Rüdiger Wirth, *CRISP-DM 1.0 Step-bystep data mining guide*, 2000, s.13-15, Erşim Tarihi: 17 Ocak 2012, SPSS: [www.crisp-dm.org/](http://www.crisp-dm.org/).

- Sund, s.3-9.

### 1.3.1. Problemin Tanımlanması

Problemin tanımlanıp çözülmesi için, öncelikle problemin ilgili bilgi ve veri bölgesi ile eşleştirilmesi gerekir. Genellikle bir istatistikçi ya da veri analisti, problemin ilgili olduğu alanda uzman olmadığı için, ilgili alandaki uzmanlarla işbirliği içinde yapılacak bir çalışma bilgi keşfi işleminin önemli bir bölümüdür. Problemin karakteristiğinin istatistikçi olmayanlar için de kolay anlaşılır bir şekilde biçimlendirilmesi, farklı alanlardaki uzmanlar için ortak bir dil oluşturulmasını sağlamaktadır.<sup>16</sup>

Bir veri tabanından ne keşfedebileceği tam olarak tahmin edilemediği için, yüksek seviyeli bir veri madenciliği sorgusu bir inceleme aracı olarak kullanılıp daha detaylı araştırmalar için yol gösterici olabilecek bazı ilginç izleri ortaya çıkarabilir.

Kullanıcı, veri madenciliği isteklerini etkileşimli (interaktif) olarak filtrelemesini, verilerin odak noktasını dinamik olarak değiştirebilmeyi, veri madenciliği işlemini süreç içinde daha detaylı hale getirmeyi ve veri madenciliği sonuçlarını birden fazla özet seviyesinde ve değişik açılardan daha esnek olarak görüntüleyebilmeyi sağlayan, interaktif keşfe özendirilmelidir.<sup>17</sup>

Veri madenciliğinin keşfe yönelik doğası, problem tanımlama adımını oldukça basitleştirmektedir. Veri madenciliği ile problemi çok fazla özelleştirmeden araştırmaya başlanabilir. Diğer veri analiz yöntemlerinde, problem ne kadar net olursa çözüme de o kadar hızlı ulaşılabilir. Veri madenciliğinde ise böyle bir durum söz konusu değildir.<sup>18</sup>

### 1.3.2. Verilerin Hazırlanması

Verilerin hazırlanması, diğer bir deyişle veri ön işleme, veri madenciliği süreçlerinin en uzun zaman gerektirenidir. Büyük çaplı veri yığınları, genellikle tutarsızlıklar, problemler ya da ilgili oldukları konuya özel değerler içerdikleri için kullanımı mümkün olmayan ham veriler içerir. Bu safhada, operasyonel işlemler sonrasında elde edilen veriler, veri madenciliği uygulamalarına uygun hale gelmeleri amacıyla bir takım işlemlere tabi tutulurlar. Bu işlemler verinin kalitesini arttırırken,

---

<sup>16</sup> Sund, ss.3-8.

-Chapman vd., ss.13-18.

<sup>17</sup> Chen, Han and Yu, ss.2-7.

<sup>18</sup> Hair vd., s.678.

veri madenciliği uygulamasının da veri üzerinde daha rahat işlem yapabilmesini sağlar. Veri önışleme, genellikle veri yığınının bütününe erişimi gerektirdiği için, önışleme zamanı yığının fiziksel büyüklüğü ile doğru orantılıdır.<sup>19</sup>

Veri madenciliği uygulamaları ile veri ambarları bu süreçte kesişirler. Veri hazırlama işleminin hedefi, farklı sistemlerde depolanan düzensiz veri yığınlarını bir veri ambarı uygulaması ile düzenleyip bir araya getirmek olabileceği gibi, var olan veri ambarında gerekli düzenlemeleri yaparak veri madenciliği uygulaması için daha işlevsel hale getirmek de olabilmektedir.<sup>20</sup>

Sonuç olarak, veri önışlemenin hedefi, veri ambarı kullanarak, ya da var olan veri tabanlarını veri madenciliğine elverişli hale getirerek, veri madenciliği uygulamalarının performansını arttırmaktır. Bu işlem şu başlıklar altında incelenebilir;<sup>21</sup>

- Veri temizleme
- Veri normalleştirme
- Veri indirgeme
- Veri dönüşümü ve entegrasyonu

Yukarıdaki başlıklardan bir veya birkaçını veri hazırlama işlemi sırasında kullanılabilir.

### **1.3.2.1. Veri Temizleme**

Veri temizleme aşamasında seçilen analiz teknikleri için gerekli verinin kalitesi artırılmaya çalışılmaktadır. Bu aşama, verinin temiz altkümelerinin seçilmesi, eksik olan verilerin yerine uygun olan varsayılan değerlerin yerleştirilmesi, modelleme yöntemleriyle kayıp verilerin tahmini gibi faaliyetleri içermektedir.<sup>22</sup> Kirli veri olarak isimlendirilen kayıp ve gürültülü verilerle birlikte yanlış veya aşırı uçlarda bulunan verilerin de analiz dışında tutulması bu aşamada sağlanır.<sup>23</sup>

---

<sup>19</sup> Sund, s.8.

<sup>20</sup> Chapman vd., ss.20-22.

<sup>21</sup> Alok Khemka, *A Collaborative Predictive Data Mining Model*, (Yayımlanmamış Yüksek Lisans Tezi), Missouri: Faculty of University of Missouri-Kansas City 2003 ss.11-14.

<sup>22</sup> Chapman, vd., s.24.

<sup>23</sup> Pieter Adriaans ve Dolf Zantinge, *Data Mining* [Elektronik Sürüm], Addison-Wesley, England Boston 1997, s.84.

### 1.3.2.2. Veri Normalleştirme

Değişkenler kendi aralarında çok değişken aralıklara sahip olmaya eğilimlidir. Böylece, veri madencileri, her değişkenin etki aralığını standardize etmek için, nümerik değişkenleri normalize etmelidir. Normalizasyon için birçok teknik vardır. En yaygın olanları, min-maks normalizasyonu ve z-skor standardizasyonudur.<sup>24</sup>

**Min – Maks Normalizasyon:** Özellik değerinin en küçük değeri min (X), en büyük değer (X) olmak üzere, min-max normalizasyonu, Formül (1.1) ile hesaplanmaktadır. X, orijinal alan değerini, X\* normalize alan değerini göstermektedir.

$$X^* = (X - \min(X)) / (\max(X) - \min(X)) \quad (1.1)$$

**Z – Skor Standardizasyonu:** İstatistiksel analizlerde yaygın olarak kullanılır. Formül (1.2)'de verildiği gibi, z-skor değeri, alan değeri ile ortalamasının farkının, alan değerinin standart sapmasına bölümüne eşittir.

$$X^* = (X - \text{Ortalama}(X)) / \text{SD}(X) \quad (1.2)$$

**Kartiller Arası Açıklık:** Z-skor standardizasyonu formülündeki ortalama ve standart sapma, aykırı değerlerin varlığını tespit eder. Bir aykırı değer bir veri kümesine eklenirse, ortalama ve standart sapma bu yeni veri değeri tarafından boş yere etkilenecektir. Böylece aykırı değerleri değerlendirmek için aykırı değerlerin varlıklarına duyarlı ölçümler kullanmak uygun olmayabilir. Bu nedenle veri analistleri, kartiller arası açıklık gibi, aykırı değerlerin varlıklarına daha az hassas ve bu değerlerin tespiti için daha sıhhatli yöntemler geliştirmektedir. Bir veri kümesinin çeyrekleri, veri kümesini her biri verinin %25'ini içeren 4 parçaya bölmektedir. İlk kartil (Q<sub>1</sub>), 25. yüzdedir. İkinci kartil (Q<sub>2</sub>), 50. yüzdedir ve medyandır. Üçüncü kartil (Q<sub>3</sub>), 75. yüzdedir. Formül (1.3) ile hesaplanan kartiller arası açıklık değeri (IQR), verinin % 50'sinin yayılımını temsil etmeyi sağlamaktadır.

$$\text{IQR} = Q_3 - Q_1 \quad (1.3)$$

---

<sup>24</sup> Larose, ss.35-39.

### 1.3.2.3. Veri İndirgeme

Veri indirgeme teknikleri orijinal verinin bütünlüğüne yakın, hacim olarak daha küçük bir veri kümesinin azaltılmış bir temsilini elde etmek için uygulanabilir. İndirgenmiş veri kümesi üzerinde aynı veya kısmen aynı analitik sonuçları üreten madencilik daha verimlidir.

Veri indirgeme stratejileri aşağıdaki durumları içermektedir :<sup>25</sup>

- Veri küpü toplama: Bir veri küpündeki verilere toplama işlemleri uygulanır,
- Boyut indirgeme: Konu dışı, zayıf, alakalı, gereksiz özellikler veya boyutlar seçilebilir ve kaldırılabilir,
- Veri basıncı: Veri kümesinin büyüklüğünü azaltmak için şifreleme mekanizmaları kullanılır,
- Sayısal indirgeme: Verinin gerçek veri yerine alternatif olarak, sadece model parametrelerini saklamaya ihtiyaç duyan veya gruplama, örnekleme, histogram kullanımı gibi parametrik olmayan yöntemlerle daha küçük veri temsillerinin tahminlenmesi veya veri ile yer değiştirmesidir,
- Kesiklendirme ve kavram hiyerarşisi genelleme: Özellikler için satır veri değerleri aralıklar veya daha üst kavramsal seviyelerle yer değiştirir. Kavram hiyerarşileri, özetlemenin çoklu seviyelerinde verinin madenciliklerine izin verir. Bu veri madenciliği için güçlü bir araçtır.

### 1.3.2.4. Veri Dönüştürme ve Entegrasyonu

Verilerin birden fazla farklı kaynaktan elde edildiği durumlarda, benzer ya da aynı veriler, farklı şekilde ifade edilebilir. Veri entegrasyonunun temel amacı bu farklılıkları ortadan kaldırmaktır.<sup>26</sup> Veri grubundaki bazı özellikler alfabetik değerler taşıırken, kullanılan veri madenciliği modeli sayısal değerler bekliyor olabilir. Bu durumda alfabetik değerlerin sayısal değerlere çevrilmeleri gerekir. Bunun en basit yöntemi, ilgili özelliğin içerdiği değerlerin herbirine bir tamsayı atamak, ve daha sonra bunları istenilen ölçüğe göre normalleştirmektir<sup>27</sup>. Veri dönüştürme, veri ambarlama sırasında da kullanılan bir metoddur.

<sup>25</sup> Han and Kamber, s.10.

<sup>26</sup> Sund, s.8.

<sup>27</sup> Khemka, s.13.



Veri entegrasyonu, aslında veri ambarlama olarak da adlandırılabilir. Fakat veri ambarı kurmadan da veri entegrasyonu yapılabilir. Yani veri ambarı veri entegrasyonunu kapsar denilebilir. Veri ambarı kalıcı bir çözüm iken, veri entegrasyonu anlık olarak da yapılabilir. Yani veri madenciliği uygulamasını başlatmadan önce veriler aynı ortamda birleştirilip birbirleri ile uyumlu hale getirilebilir. Daha sonra bu veriler birleştirildikleri ortamdan kaldırılabilirler. Kalıcı bir veri ambarı kurmak ya da tek başına veri entegrasyonu yapmak arasındaki seçimi, kullanılacak veri madenciliği uygulaması, bu uygulamanın çalışma sıklığı, anlık veri transferinin hızı ve maliyeti, transfer edilecek verinin büyüklüğü gibi faktörler belirler.

### 1.3.3. Modelleme

Bu aşama VM'nin en çarpıcı bölümünü oluşturmaktadır. Bu aşamada analiz yöntemleri veriden bilginin elde edilmesi için kullanılmaktadır.<sup>28</sup> Modelleme aşamasında, çeşitli modelleme teknikleri seçilmekte, uygulanmakta ve optimumu değerlere ulaşabilmek için parametreleri ayarlanmaktadır. Bu aşamada, seçilen teknikler veri setleri üzerinde çalıştırılmakta ve çıkan matematiksel denklemler yorumlanmaktadır. Süreç tekrarlandıkça performans iyileşmekte, sonuçlar daha güvenilir hale gelmektedir. Bu aşamada seçilen VM aracının özellikli olarak hangi algoritmaları, teknikleri kullanacağına ve hangilerinin modele en uygun olduğuna karar verilmeye çalışılmaktadır. Genellikle, aynı VM problem tipi için birden fazla teknik bulunduğu ve bu tekniklerden bazıları verinin özel bir formunu gerektirdiğinden bu aşamadan veri hazırlama aşamasına geri dönüşler çoğunlukla gerekli olmaktadır.

#### 1.3.3.1. Modelleme Tekniğinin Seçimi

Modellemenin ilk aşaması olan modelleme tekniğinin seçiminde, kullanılacak olan teknik belirlenmektedir. İşin kavranması kısmında kullanılacak VM aracı seçilmesine rağmen bu aşama özellikli olarak hangi tekniğin kullanılacağına karar verildiği andır. Veri üzerine belirli varsayımlar yaparak çalışan tekniklerin varlığını dikkate alarak seçim yapılmaktadır. CHAID algoritması ile Karar Ağacı oluşturulması

---

<sup>28</sup> Ayşe Oğuzlar, "Veri Madenciliğinde Birliktelik Kuralları", *Marmara Üniversitesi Sosyal Bilimler Enstitüsü Hakemli Dergisi Öneri*, 6 (22), 2004, 316-317, Erişim Tarihi: 27 Aralık 2012, <http://tk.kutuphaneci.org.tr>.

veya ileriye doğru yayılımla YSA oluşturulması bu seçimlerden bazılarıdır.<sup>29</sup> Birden fazla teknik uygulanacaksa bu aşamada aynı işlem her teknik için ayrıca yapılmalıdır.

### **1.3.3.2. Test Tasarımının Oluşturulması**

Bir modelin tam olarak kurulmadan önce, geçerliliğinin ve kalitesinin test edilebileceği prosedürlerin oluşturulması gerekmektedir. Modelin geçerliliğinin ve kalitesinin test edilebilmesi için hazırlanan veri setinin bir bölümü kullanılmaktadır. Veri setinde model kurulan bölüm “model kurma örneği” iken, modelin geçerliliğinin test edildiği bölüm ise “sınama örneği”dir. Model kurma örneği üzerinde model çalıştırılıp eğitilirken sınama örneği üzerinde ise modelin kalitesi tahmin edilmektedir.<sup>30</sup>

### **1.3.3.3. Modelin Kurulması**

Modelin kurulması aşaması tekrarlanan, yinelenen bir süreçtir. İşletme problemlerinin çözümünde en kullanışlı modelin bulunması amacıyla alternatiflerin araştırılması gerekmektedir.<sup>31</sup>

### **1.3.4. Modelin Değerlendirilmesi**

Daha önceki aşamalar modelle ilgili doğruluk ve genellik gibi faktörlerle uğraşmaktadır. Değerlendirme aşaması ise modelin ne dereceye kadar işletme amaçlarına uygun olduğunu değerlendirirken aynı zamanda bu modelin yeterli kalmasına yol açan bir iş nedeni olup olmadığını belirlemeyi amaçlamaktadır. Bir başka değerlendirme seçeneği ise eğer zaman, bütçe ve teknoloji kısıtları elverirse modelin test uygulamalarının gerçek iş uygulamaları ile yapılmasıdır.

#### **1.3.4.1. Sonuçların Değerlendirilmesi**

Sonuçların değerlendirilmesi aşamasında, VM projesinin diğer sonuçlarının değerlendirmesi yapılır. Bu aşama sonunda projenin birincil işletme amaçlarını karşılayıp karşılamadığını gösteren sonuç ifadesi hazırlanır. Model değerlemesi

<sup>29</sup> Chapman, s.27.

<sup>30</sup> Şule Özmen, *Ağ-Ekonomisinde Yeni Ticaret Yolu: e-Ticaret*, İstanbul Bilgi Üniversitesi Yayınları, İstanbul 2003, s.198.

<sup>31</sup> B. Kiremitçi, *Veri Ambarlarında Veri Madenciliği ve Ulaştırma-Lojistik Sektöründe Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2005, s.67.

sonucunda iş amaçları bakımından başarılı bulunan modellerin onaylanması da yine bu aşamada gerçekleştirilmektedir.<sup>32</sup>

#### **1.3.4.2. Veri Madenciliği Sürecinin Gözden Geçirilmesi**

Bu aşamada, son modelin işletme amaçlarını tatmin edecek bir şekil aldığı umut edilmektedir. Bu adım kalite güvence konularını, modelin doğru kurulup kurulmadığını, kullanılma izni olan değişkenlerin, niteliklerin kullanıp kullanılmadığını ve gelecekte yapılacak analizlerde de bunların kullanıp kullanamayacağını kapsamaktadır. Bu aşama sonunda da gözden geçirme süreci ve gözden kaçırılan veya tekrar edilmesi gereken faaliyetler özetlenmektedir.<sup>33</sup>

#### **1.3.4.3. Sonraki Adımların Belirlenmesi**

Sürecin gözden geçirilmesi ve değerlendirme sonuçlarına göre projenin nasıl ilerleyeceğine bu aşamada karar verilmektedir. Projenin bu aşamada bitirilip uygulamaya, yayılım aşamasına geçilip geçilmeyeceği, projenin daha başka adımlarının, tekrarlamaların yapılıp yapılmayacağı veya yeni VM projelerin oluşturulup oluşturulmayacağı bu aşamada belirlenmektedir. Bundan sonra gerçekleştirilecek her adım için yapılacak olası faaliyetlerin nedenleriyle bu aşamada gösterilmektedir.

#### **1.3.5. Modelin Kullanılması**

Bir veri madenciliği modeli oluşturulduktan ve geçerliliği kabul edildikten sonra uygulama aşamasına geçilir. Veri madenciliği sonuçları modelin özelliğine göre iki şekilde uygulanabilir. Bunlardan ilki modelin sonuçlarına göre faaliyetlerin önerilmesidir. Örneğin madencilik modelinin oluşturduğu kümelere veya modeli tanımlayan kurallara bakılarak faaliyet planları oluşturulabilir. Ayrıca kaldıraç ve ROC grafikleri kullanılarak faaliyetlerin sağlayacağı faydalar vurgulanabilir. Veri madenciliği sonuçların modeli farklı veri kümelerine uygulamakta kullanılabilir. Model, verinin sınıflanmasına dayalı olarak bazı nesnelere ön plana çıkarabilir. Bu verilere ilişkin OLAP sistemi aracılığıyla daha ayrıntılı analizler yapılabilir.

Model uygulandıktan sonra sistemin ne kadar iyi çalıştığının ölçülmesi gerekir. Model ne kadar iyi çalışıyor olsa da modelin performansının sürekli olarak izlenmesi

<sup>32</sup> Colin Shearer, "The CRISP-DM Model: The New Blueprint for Data Mining" [CRISP-Veri Madenciliği Modeli: Veri Madenciliği için Yeni Tasarı], *Journal of Data Warehousing*, 5(4), 2000, pp. 16, Erişim Tarihi: 18 Ocak 2012, <http://mineracaodedados.files>.

<sup>33</sup> Chapman, s.31.

gerekir. Zaman içerisinde tüm sistemler değişime uğrar. Örneğin enflasyon oranı gibi dış etmenlerin değişmesi insanların davranış şekillerini değiştirebilir. Zamanla oluşturulan modelin değişen koşullara uyum sağlaması için test edilmesi, tekrar eğitilmesi ve gerekiyorsa yeniden oluşturulması gerekebilir. Tahmin edilen değerlerle gözlenen değerler arasındaki farkların grafikleri model sonuçlarının izlenmesinde mükemmel bir yoldur. Hesaplamanın yoğun olmadığı bu grafikleri kullanmak, anlamak kolaydır ve modeli uygulayan yazılımın içine yerleştirilmesi de mümkündür. Böylece sistem kendini izleyebilecektir.<sup>34</sup>

#### 1.4. VERİ MADENCİLİĞİNDE KARŞILAŞILAN SORUNLAR

Küçük veri kümelerinde ve ayıklanmış veri üzerinde hızlı ve doğru bir biçimde çalışan bir sistem çok büyük veritabanlarına uygulandığında sorun çıkabilir. Bir VM sistemi ayıklanmış veri üzerinde iyi çalışırken, aynı veriye gürültü eklendiğinde net olmayan sonuçlar oluşabilir.

##### 1.4.1. Veri Tabanı Boyutu

Veri tabanlarının boyutları sebebiyle, veri madenciliği yöntemlerinden herhangi birinin ham veriyle başarılı olma olasılığı yoktur. Veri madenciliği yöntemleri, bu şekilde elde edilen sonuçların tüm veri tabanını temsil edebileceğini umarak, veri tabanından bir örneğin çıkarılmasını gerektirebilir.

Bir veri tabanının boyutlarının küçültülmesi 2 yolla olabilir:<sup>35</sup>

- **Veri alanında örnekleme:** Genellikle rasgele bazı kayıtlar seçilir ve veri madenciliğinin sonraki aşamalarında kullanılır.
- **Özellik alanında örnekleme:** Her veri kaydının bazı özellikleri seçilir. Yine birçok özellik varsa, seçim rasgele yapılır.

##### 1.4.2. Gürültülü Veri

Nitelik değerlerindeki veya sınıf bilgilerindeki hatalar “gürültü” olarak adlandırılır. Bazı kayıtlarda aşırı uç değerler veya yanlış girilmiş değerler olabilir. Bu tür bilgilere de “gürültü” adı verilmektedir Herhangi bir veri toplama tekniğinin,

<sup>34</sup> Two Crows Corp., *Introduction to Data Mining and Knowledge Discovery*, USA 2005, pp.55, Erişim Tarihi: 11 Aralık 2011, [http:// www.twocrows.com](http://www.twocrows.com).

<sup>35</sup> U.T. Şimşek, *Veri Madenciliği ve Müşteri İlişkileri Yönetiminde (CRM) Bir Uygulama*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2006, s.12.

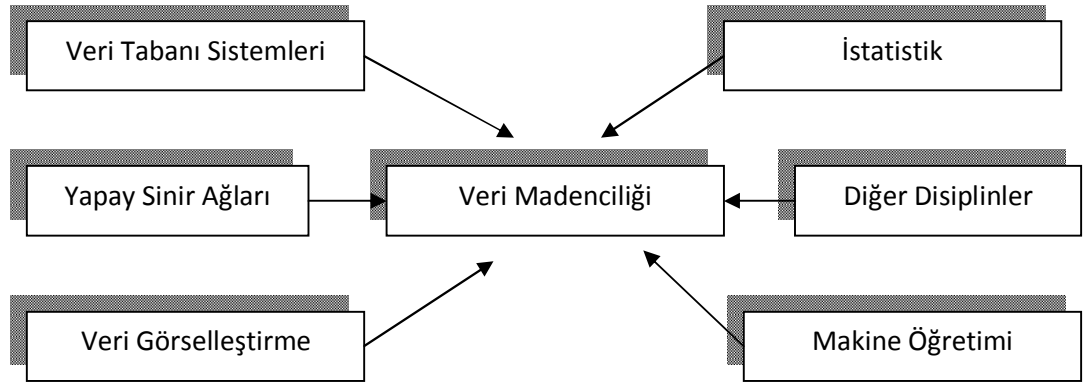
gürültüden tümüyle arınmış olması çok zordur. Fakat üretilmiş kuralların toplam doğruluğu için, sınıflanmış bilgidен gürültünün elenmesi istenir.<sup>36</sup>

### 1.4.3. Eksik ve Kesin Olmayan Değerler

Her nesnenin ayrıntılı bir biçimde tanımlandığı ve bu nesnelerin alabileceği değerler kümesinin belirtili olduğu durumlarda her bir nesnenin tanımı kesin ve yeterli olsaydı, sınıflama işlemi basitçe nesnelerin alt kümelerinden faydalanılarak yapılabilirdi. Bununla birlikte veriler kurum ihtiyaçları göz önünde bulundurularak düzenlenip toplandığında mevcut veri, gerçek hayatı yeterince yansıtmayabilir. Bu gibi koşullarda bilgi keşfi modeli belirli bir güvenlik derecesinde tahmini kararlar alabilmelidir.<sup>37</sup>

## 1.5. VERİ MADENCİLİĞİ VE İLİŞKİDE OLDUĞU DİĞER DİSİPLİNLER

Veri madenciliği, veri tabanı teknolojisi, istatistik, bilgisayar bilimleri, makine öğrenimi, örüntü tanıma ve görselleştirme gibi pek çok teknik alan arasında köprü görevi gören çok disiplinli bir alandır.<sup>38</sup> Veri madenciliği istatistiksel yöntemlerin bir uzantısıdır. Böylece, makine öğrenimi, veri tabanı yönetim sistemi gibi diğer alanlara ait teknikleri de kullanmaktadır. Ayrıca, veri madenciliği, yapay zeka alanını kullanabilmektedir.



**Şekil 1.2.** Veri Madenciliği ve İlişkide Olduğu Disiplinler

**Kaynak:** Al-Hudaİry, Hazem H.M. A.A., Data Mining and Decision Making Support in the Governmental Sector, (Yayımlanmamış Yüksek Lisans Tezi), University of Louisville Department of Computer Engineering and Computer Science, USA 2004, pp.4.

<sup>36</sup> Han Jiawei and Kamber Micheline, *Data Mining Concepts and Techniques* [Elektronik Sürüm], Morgan Kaufman Publishers, USA 2001, s.106.

<sup>37</sup> Ham and Kamber, s.5.

<sup>38</sup> Serhat Özkes, *Veri Madenciliği Uygulaması*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2002, s.3.

## **Veri Madenciliği ve Makine Öğrenimi**

Veri madenciliği ile ilişkili bir disiplin olan makine öğrenimi, bilgisayar bilimi ve yapay zeka ile bağlantılıdır. Ayrıca, genel doğruluklara çevrilebilen verideki ilişkileri ve düzenleri bulmakla ilgilidir. Makine öğreniminin amacı, analistlerin gözlenen veriden yeni ve henüz gözlenmemiş durumları genelleştirmesine imkan sağlaması ve veri genelleştirme sürecinin yeniden oluşturulmasıdır.<sup>39</sup>

Makine öğreniminin dezavantajı, yalnızca sembolik ve kategorik verileri kullanmasıdır. Gürültülü, eksik ve boş verileri işleyemez. Ayrıca, makine öğrenimi, veri madenciliğine kıyasla, düşük boyutlu deneysel verileri kullanmaktadır.

## **Veri Madenciliği ve İstatistik**

Veri madenciliği ve istatistiksel teknikler veriyi analiz etmek amacıyla kullanılmaktadır. İstatistik, veri analizi için daima var olan bir yöntemdir. İstatistiksel yöntemler ve makine öğrenimi arasındaki temel fark, istatistiksel yöntemlerin genelde analiz edilen veriyle ilişki içinde olması ve kavramsal referans örneğine göre geliştirilmesidir. Bu durum, istatistiksel yöntemleri uyumlu ve etkin yapmasına rağmen, istatistiksel yöntemlerin yeni bilgi teknolojisi ve yeni makine öğrenimi uygulamalarından gelişen yöntemlere hızlı bir şekilde adapte olma yeteneklerini sınırlamaktadır. Bu nedenle istatistikçiler, son yıllarda veri madenciliğine ilgi göstermektedir. Bu durum, veri madenciliğinin gelişimine yardımcı olmaktadır.<sup>40</sup>

## **Veri Görselleştirme**

Veri madenciliğinde girdi ve çıktı, kolay anlaşılır ve kullanıma uygun olmalıdır. Görselleştirme teknikleri verideki dağılımları, örüntüleri, kümeleri ve sınır dışı öğelerin gücünü daha çekici ve etkin bir hale getirebilmek amacıyla kullanılır. Bunun için bilgisayar grafikleri, veri dağılım haritaları, eğriler, üç boyutlu şekiller ve yüzeyler kullanılır.

Analiz için seçilecek olan görselleştirme metodu, eldeki veriye ve modele göre değişebilmektedir. Görsel görüntüler kullanıcılara, veri tabanındaki veri karakteristikleri için açık bir izlenim vermek ve genel bakış kazandırmak açısından da yardımcı olmaktadır.<sup>41</sup>

---

<sup>39</sup> Giudici, s.2.

<sup>40</sup> Giudici, s.5.

<sup>41</sup> Şimşek, s.16.

## **Diğer Disiplinler**

Veri madenciliği yukarıda sayılan disiplinler haricinde Veri Ambarları, Yapay Sinir Ağları ve Genetik Algoritmalar ile de ilişki içerisindedir.

### **1.6. VERİ MADENCİLİĞİ YÖNTEMLERİ**

Veri madenciliği konusunda çok sayıda yöntem ve algoritmalar geliştirilmiştir. Bu yöntemlerin birçoğu istatistiksel tabanlıdır. Söz konusu bu bölümde veri madenciliği yöntemlerinden;

- Sınıflandırma,
- Kümeleme,
- Birliktelik Kuralı
- Destek Vektör Makineleri

ele alınacaktır.

#### **1.6.1. Bayesyen Sınıflandırma**

Sınıflama modeli, sınıfı tanımlanmış mevcut verilerden yararlanarak sınıfı belli olmayan verilerin sınıfını tahmin etmek için kullanılan VM modelidir. Sınıflama iki adım içeren bir işlemdir. Birinci adımda tahmin için kullanılacak bir model oluşturulmaktadır. İkinci adımda, oluşturulan bu model sınıfı belli olmayan veriler üzerinde uygulanarak sınıflar tahmin edilmektedir.<sup>42</sup>

Bayes sınıflandırıcılar istatistiksel sınıflandırıcılardır ve belirli bir değişkenler grubunun belirli bir sınıfa ait sınıf üyeliği olasılıklarını tahmin ederler. Bu sınıflandırma Thomas Bayes'in teoremine dayanmaktadır.<sup>43</sup> Bayes teoremi iki rassal olayın koşullu ve marjinal olasılıklarını ilişkilendirir. Bayes sınıflandırıcılar temel olarak her özneliğin verilen sınıf etiketine göre sınıfsal koşullu olasılıklarını öğrenir.<sup>44</sup>

Bayesyen sınıflandırmanın amacı nesnelere ilgili istatistikî bilgiye göre hatalı sınıflandırmayı minimize etmektir. Yeni nesnelere ilişkili sınıflandırma kapasitesi daha önce popülasyon içerisinden rastgele gözlemlenen nesnelere bağlıdır. Yeni nesnelere

---

<sup>42</sup> Han and Kamber, s.35.

<sup>43</sup> Han and Kamber, s.31.

<sup>44</sup> C.C. Hsu, Y.P. Huang, K.W. Chang, "Extended Naive Bayes Classifier for Mixed Data" [Karışık Veri için Genişletilmiş Naive Bayes Sınıflandırıcıları], *Expert Systems with Applications*, 35(3), 2008, pp. 1081, Erişim Tarihi: 24 Ocak 2012, ScienceDirect.

ilgili tahmin tutarlılık seviyesi istatistiksel ölçümlerin miktarına ve önceki deneylerden elde edilen bilgilere bağlıdır.<sup>45</sup>

Navie Bayes sınıflandırıcı aşağıdaki şekilde çalışmaktadır :<sup>46</sup>

- D değişken grupları ve ilgili sınıf etiketlerinden oluşan bir eğitim seti olsun. Her değişkenler grubu bir n – boyutlu öznitelik vektörü  $X = (x_1, x_2, \dots, x_n)$  ile temsil edilmektedir ve  $A_1, A_2, \dots, A_n$  n adet öznitelikten oluşan bir değişkenler grubunda yapılan n adet ölçümü belirtmektedir.
- $C_1, C_2, \dots, C_m$  dan oluşan m tane sınıf olduğunu varsayalım. Bir bilinmeyen veri örnekleme olan X’de, sınıflandırıcı X’in X’e koşullu olmak üzere en yüksek ardıl olasılığa sahip olduğunu öngörür. Bu durum, Navie bayesyen sınıflandırıcının bilinmeyen örneklem X’i sadece aşağıda belirtilen durumda sınıf  $C_j$ ’ye atamaktadır.

$$P(C_m \setminus X) > P(C_j \setminus X) \quad 1 \leq j \leq m, j \neq a \quad (1.4)$$

Böylece  $P(C_m \setminus X)$  maksimize edilmiş olur.  $P(C_m \setminus X)$ ’nin maksimize edildiği sınıf  $C_m$  maksimum ardıl hipotez olarak adlandırılır. Bayes teoremine göre ise aşağıdaki şekilde formül ize edilir;

$$P(C_m / X) = \frac{P(X/C_m) P(C_m)}{P(X)} \quad (1.5)$$

- Tüm sınıflar için  $P(X)$  sabit olduğu için,  $P(X / C_m) P(C_m)$  maksimize edilmesi gerekmektedir. Eğer önceki sınıf olasılıkları i bilinmiyorsa, yaygın olarak kabul gören şekilde bütün sınıflar eşit olasılığa sahiptir. Bu durumda  $P(C_1) = P(C_2) = \dots = P(C_m)$  olur ve  $P(X/C_m)$  maksimize edilmiş olur. Aksi durumda  $P(X/C_m) P(C_m)$  maksimize edilmiş olacaktır.
- Çok fazla özniteliği n bulunduğu veri setlerinde,  $P(X/C_m)$  hesaplamak oldukça büyük bir kaynak kullanımını gerektirmektedir. Bu kullanımı azaltabilmek için Naive sınıf koşullu bağımsız değişkeni kullanılır. Böylece öznitelik değerleri

<sup>45</sup> J. Krysztof and W. P Horis, *Data Mining A Knowledge Discovery Approach* [Elektronik Sürüm], Springer Sciences Business Media, New York 2007, pp.476.

<sup>46</sup> Han and Kamber, s.311.



birbirinden koşullu olarak bağımsız sayılmaktadır. Böylece öznitelikler arasında herhangi bir ilişki bağılılığı yok sayılır.

$$P(X/C_i) = \prod_{k=1}^n P(x_k/C_i)$$

$$= P(x_1 / C_i) * P(x_2 / C_i) * \dots * P(x_n / C_i) \quad (1.6)$$

olasılıkları değişken grubundan hesaplanabilmektedir. Bunlar:

- Eğer  $A_k$  koşulsuz ise, bu durumda  $P(X_k / C_i)$ ,  $C_i$  sınıfının  $D$  içindeki  $A_k$ 'nın  $x_k$  değerini aldığı değişken grubu sayısının,  $C_i$  sınıfının  $D$  içindeki toplam değişken grubu sayısına oranıdır.
- Eğer  $A_k$  sürekli bir değişkense, özniteliğin Gaussian dağılımı gösterdiği varsayılır.

$$P(X/C_m) = g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1.7)$$

Gaussian yoğunluk fonksiyonu olan  $g(x, \mu, \sigma)$  ile Öznitelik  $A_k$  için ortalama ve standart sapmanın bulunmasını sağlamaktadır.

- Bilinmeyen örneklem  $X$ 'i sınıflandırmak için her sınıf  $C_i$  için  $P(X / C_m)$   $P(C_m)$  hesaplanır. Örneklem  $X$  Sınıf  $C_i$ 'ye aşağıdaki şartlar içerisinde atanabilir.

$$P(X / C_m) P(C_m) > P(X / C_j) P(C_j) \quad 1 \leq j \leq m, j \neq a \quad (1.8)$$

Başka bir deyişle,  $P(X / C_m) P(C_m)$  maksimum değer aldığı anda sınıf  $C_i$  atanabilmektedir.

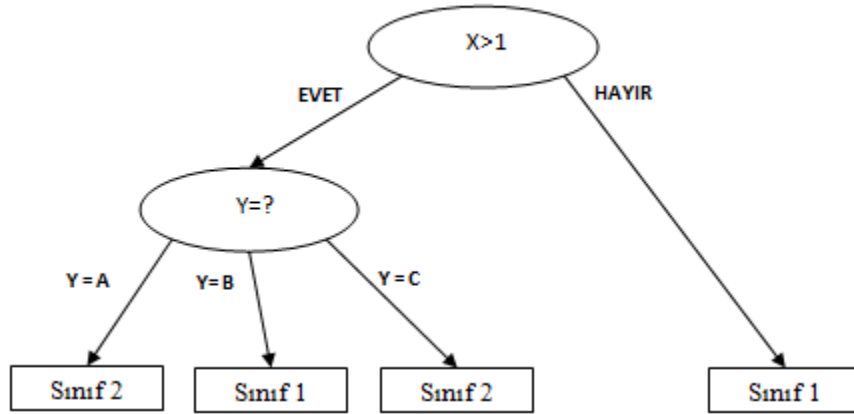
### 1.6.2. Karar Ağaçları İle Sınıflandırma

Karar ağaçları, sınıflandırma amacıyla veri madenciliğinde yaygın olarak kullanılmaktadır. Heterojen veri koleksiyonu ve veriyi açıklayan özelliklerin bir kümesi verildiğinde, karar ağaçları, bu özellik değerlerini, veri kümesini daha homojen küçük alt kümelere bölmede kullanmaktadır. Karar ağaçlarının gerçekleştirilmesi ağaçların

inşa ediliş şekli ile farklılık göstermektedir. Ağacı modelleme için birçok teknik bulunmaktadır.<sup>47</sup>

Karar ağaçları ve karar kuralları birçok gerçek yaşam uygulamasında sınıflandırma problemlerine güçlü çözümler üreten veri madenciliği yöntemleridir. Veriden sınıflandırıcılar üretmek için en etkin yöntemlerden biri karar ağaçlarıdır ve karar ağacı gösterimleri yaygın olarak kullanılan mantıksal yöntemlerdir. Bunlardan bazıları makine öğrenimi ve uygulamalı istatistik alanları olmak üzere literatürde tanımlanmış çok sayıda karar ağacı endüksiyon algoritması vardır. Bu algoritmalar bir girdi-çıkı örneklem setinden karar ağaçları yapılandırılan denetimli öğrenme metotlarıdır. Tipik bir karar ağacı öğrenme sistemi, arama uzayının bir bölümünde bir çözüm arayan yukarıdan aşağıya bir strateji kullanır. Karar ağacı, özneliklerin test edildiği düğümlerden (node) oluşur. Bir düğümün altında kalan dallar, düğümde test edilen tüm olası sonuçlara karşılık gelir.<sup>48</sup>

Karar ağaçları akış şemalarına benzeyen yapılardır. Herbir nitelik bir düğüm tarafından temsil edilir. Dallar ve yapraklar ağaç yapısının elemanlarıdır. En son yapı “yaprak”, en üst yapı “kök” ve bunların arasında kalan yapılar ise “dal” olarak isimlendirilir. Şekil 1.3 üzerinde bir karar ağacı yer almaktadır<sup>49</sup>.



**Şekil 1.3.** X ve Y nitelikleri üzerine uygulanan testleri içeren basit bit karar ağacı

<sup>47</sup> Fariial Shahnaz, *Decision Tree Based Algorithms* [Elektronik Sürüm], Michael W. Berry (Ed.), Lecture Notes in Data Mining, World Scientific Publisher, USA 2006, pp.79.

<sup>48</sup> Kantardzic, s.139-140.

<sup>49</sup> Yalçın Özkan, *Veri Madenciliği Yöntemleri*, Papatya Yayıncılık, İstanbul 2008, s.52.

X ve Y'den oluşan iki giriş niteliğine sahip bir örnek sınıfının basit bir karar ağacı yukarıdaki şekilde görülmektedir.  $X > 1$  ve  $Y = B$  değerlerini taşıyan örnekler sınıf 1'de;  $Y = A$  ve  $Y = C$  koşullarını uygun olanlar sınıf 2 sınıfında yer almaktadır. Ancak Y'nin değerini göz önüne almadan  $X \leq 1$  koşulu uygun örnekler sınıf 1'de yer almaktadır.

Karar ağacı temelli analizlerin yaygın olarak kullanıldığı alanlar şunlardır:<sup>50</sup>

- Belirli bir sınıfın olası üyesi olacak elemanların belirlenmesi,
- Çeşitli vakaların yüksek, orta, düşük risk grupları gibi çeşitli kategorilere ayrılması,
- Parametrik modellerin kurulmasında kullanılmak üzere çok sayıdaki değişkenden en önemlilerinin seçilmesi,
- Gelecekteki olayların tahmin edilebilmesi için kurallar oluşturulması,
- Sadece belirli alt gruplara özgü olan ilişkilerin tanımlanması,
- Kategorilerin birleştirilmesi ve sürekli değişkenlerin kesikli değişkenlere dönüştürülmesidir.

Çoğu karar ağacı algoritmalarının karşılaştığı konular şunlardır:<sup>51</sup>

- **Ayırma özelliklerinin seçilmesi:** Ayırma özellikleri için hangi özelliklerin kullanılacağı inşa edilmiş karar ağacının uygulanmasındaki performansını etkiler. Bazı özellikler diğerlerinden daha iyidir. Örneğin, bir veritabanında yer alan kaydın isim özelliği kesinlikle kullanılmamalıdır. Özelliğin seçimi, eğitim setindeki verinin incelenmesinin yanında alan uzmanlarının bilgilendirilmiş girdisini de gerektirir.
- **Ayırma özelliklerinin sıralanması:** Özelliklerin seçildiği sırada önemlidir. Bazı durumlarda birinci olarak seçilen bir özellik, gereksiz karşılaştırmaları içerebileceğinden, ikinci bir sefer incelenmeyi gerektirebilir.
- **Ayırma:** Özelliklerin sıralanması ile ilişkilendirilenler alınacak ayırmaların sayısıdır. Bazı özelliklerle birlikte, tanım kümesi küçüktür, dolayısıyla

<sup>50</sup> Haldun Akpınar, "Veri Tabanlarında Bilgi Kesfi ve Veri Madenciliği", *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 29(1), 2000, s.17.

<sup>51</sup> Levent Terlemez, *Eş İşlem Stratejisi Yöntemiyle İMKB'de Portföy Oluşturmada Veri Madenciliği Uygulaması*, (Yayımlanmamış Doktora Tezi), Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir 2008, s.49.

ayırmaların sayısı açıkça tanım kümesine dayanır. Buna rağmen, eğer tanım kümesi sürekli ise veya çok sayıda değere sahipse, kullanılacak ayırmaların sayısı kolaylıkla belirlenemez.

- **Ağaç yapısı:** Sınıflandırma için ağacın uygulanma performansını iyileştirmek için, en az seviyeye sahip dengelenmiş bir ağaç arzu edilir. Buna rağmen, bazı durumlarda çok yollu dallanmalı daha karmaşık karşılaştırmalara ihtiyaç duyulabilir. Bazı algoritmalar sadece ikili ağaçlar inşa eder.
- **Durma ölçütü:** Ağacın oluşturulması, veri seti mükemmel bir şekilde sınıflandırıldığında kesinlikle durur. Daha büyük ağaçların yaratılmasını önlemek için daha önce durulmasının arzu edilebileceği durumlar olabilir. Bu sınıflandırmanın doğruluğu ve performansı arasındaki bir alış verıştır. Buna ilave olarak daha erken durmak, gereğinden fazla uyumu bertaraf etmek içinde uygulanabilir.
- **Eğitim verisi:** Oluşturulan karar ağacının yapısı eğitim verisine bağlıdır. Eğer eğitim verisi çok küçük ise, o zaman oluşturulmuş ağaç daha genel veriyle düzgün çalışacak kadar belirli olmayabilir. Eğer eğitim verisi çok büyük ise, o zaman oluşturulan ağaç gereğinden fazla uyum gösterebilir.
- **Budama:** Ağaç oluşturulduktan sonra, sınıflandırma aşaması esnasında ağacın performansını iyileştirmek için ağaçta bazı değişikliklere ihtiyaç duyulabilir. Budama aşaması daha iyi performansa ulaşabilmek için, gereksiz karşılaştırmaları veya alt ağaçları kaldırabilir.

Bazı karar ağacı teknikleri ID3, C4.5, C5.0 ve CHART teknikleri sıralanabilir.

#### 1.6.2.1. ID3 Algoritması

Karar ağaçları yardımıyla sınıflandırma işlemlerini yerine getirmek üzere Quinlan tarafından gerçekleştirilmiştir. Bu yaklaşıma dayanan sistemler, her özelliğin ayırt edilebilen gücünün değerini belirlemek için entropi gibi bir bilgi ölçümü kullanmaktadır. Karmaşık bir karar verme sürecini basit kararların bir koleksiyonuna parçalaması ve kolay yorumlanabilir çözüm sağlaması, karar ağaçlarının en önemli özellikleridir. ID'ler, sembolik verinin sınıflandırması için popüler ve etkili karar verme yöntemleridir. Nümerik değerlerin olduğu durumlara genellikle uygun değildirler. Çoğu

gerçek hayat problemleri sembolik olmayan (sürekli) değerlerle çalıştığı için, özellik seçiminden önce kesiklendirme uygulanmalıdır.<sup>52</sup>

Bilgiyi nitelemek için kullanılan kavram entropidir. Entropi bir veri setindeki düzensizliğin veya rassallığın miktarını ölçmek için kullanılır.<sup>53</sup>

ID3, düğümleri ayırmak için entropiyi kullanmaktadır. Verilen t düğümü için entropi aşağıdaki formülde verilmiştir. Burada  $p_i$ , t düğümü içinde i sınıfındaki olasılıktır. Hedef niteliğini ifade eden T, hedef niteliği olmayan (yani sınıf niteliği olmayan) bir X niteliğinin değerine bağlı olarak T1, T2,.....Tn alt kümelere ayrılırsa, T'nin bir elemanının sınıfını belirlemek için gerekli bilgi, Ti'nin bir elemanının sınıfının belirlenmesinde gerekli olan bilginin ağırlıklı ortalaması kabul edilir. Bu ağırlıklı ortalama  $H(X,T)$  ile gösterilir. Kazanç(X,T) ölçütü hesaplanır. Kazanç(X,T) değerini ençoklama (maksimize) amaçlanır. En yüksek bilgi kazancını sağlayan karakteristik seçilir.<sup>54</sup>

$$Entropi = H(S) = - \sum_{i=1}^n p_i \log_2(p_i) \quad (1.9)$$

$$H(X,T) = \sum_{i=1}^n \frac{|T_i|}{|T|} H(T_i) \quad (1.10)$$

$$Kazanç(X,T) = H(T) - H(X,T) \quad (1.11)$$

### 1.6.2.2. C4.5 Algoritması

Quinlan tarafından ID3 algoritmasının geliştirilmesi amacıyla bulunmuştur. Bu algoritma sayısal değerlere sahip niteliklerin de karar ağaçlarını oluşturma olanağı sağlamıştır. Ayrıca bilinmeyen nitelik değerlerine sahip örnek kümeleri için karar ağacının nasıl oluşturulabileceği konusunda bir yol sunmaktadır.<sup>55</sup>

<sup>52</sup> E.K. Özyirmidokuz, *Veri Madenciliği Tekniklerini Kullanarak İmalat Verilerinin Modellenmesi ve Analizi*, (Yayımlanmamış Doktora Tezi), Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, Kayseri 2009, s.56.

<sup>53</sup> Terlemez, s.51.

<sup>54</sup> Özkan, s.58.

<sup>55</sup> Özkan, s.77.

C4.5, ID3'ten farklı olarak şu özelliklere sahiptir.<sup>56</sup>

- Sürekli değerli tahmin ediciler kullanılabilir,
- Budama kullanılır,
- Kural çıkarımı yapılır.

Karar ağacı algoritması C4.5 aşağıdaki yollarla ID3 algoritmasını iyileştirir:<sup>57</sup>

**Kayıp Veri:** Karar ağacı kurulduğunda, kayıp veri basitçe göz ardı edilir. Öyle ki, kazanım oranı sadece ilgilenilen özellik için değere sahip diğer kayıtlara bakılarak hesaplanır. Kayıp özellik değerine sahip bir kaydı sınıflamak için, o öge için değer diğer kayıtların özellik değerleri hakkında ne bilindiğine dayanarak kestirilebilir.

**Sürekli Veri:** Temel fikir veriyi, eğitim örneğinde ilgili öge için bulunan özellik değerlerine dayanarak aralıklara bölmektir.

**Budama:** C4.5'te önerilen iki birincil budama stratejisi vardır:

- 1-**Alt ağaçla yer değiştirme** ile eğer yer değiştirme orijinal ağaçtakine yakın bir hata oranıyla sonuçlanıyorsa yaprak düğümü bir alt ağaç ile yeri değiştirilir.
- 2-**Alt ağaç yükseltme** olarak adlandırılan başka bir budama stratejisi de, bir alt ağacı bu ağacın en çok kullanılan alt ağacı ile değiştirir. Burada bir alt ağaç o anki konumundan ağaçtaki daha yüksek bir konuma yükseltilir. Yine, bu yerleştirme için hata oranındaki artış belirlenmelidir.

**Kurallar:** C4.5 karar ağaçları ile veya karar ağaçlarından oluşturulan kurallarla sınıflandırmaya izin verir. İlave olarak, karmaşık kuralları basitleştirmek için bazı teknikler sunar. Bir yaklaşım, eğer eğitim setindeki tüm kayıtlar özdeş olarak işleniyorsa, kuralın sol tarafını daha basit bir sürümü ile değiştirmektir. Bir "diğer durum" tipindeki kural eğer diğer başka kurallar uygulanmayacaksa ne yapılması gerektiğini göstermekte kullanılabilir.

<sup>56</sup> A. S. Koyuncugil, *Bulanık Veri Madenciliği ve Sermaye Piyasalarına Uygulaması*, (Yayımlanmamış Doktora Tezi), Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara 2006, s.73.

<sup>57</sup> Terlemez, s.53.

### 1.6.2.3. C.5 Algoritması

C4.5 ve C5.0 (Quinlan 1993) ise ID3 algoritmasının geliştirilmiş versiyonları olup, ilk versiyonları kategorik veri türleri ile sınırlı olmakla birlikte, sürekli veri türleri üzerinde de uygulama olanağı sağlamaktadır.

Büyük veritabanlarına karşı kullanımı hedeflenmiştir. Karar ağacı oluşturulması C4.5'tekine yakındır, fakat kural oluşturma farklıdır. C4.5'ten farklı olarak, C5.0 için kullanılan kesin algoritmalar ifşa edilmemiştir. C5.0 kural oluşturma için iyileştirmeler içerir. Sonuçlar C5.0'in hafıza kullanımını %90 iyileştirdiğini, C4.5'ten 5,7 kez ve 240 kez arasında daha hızlı çalıştığını ve daha kesin kurallar ürettiğini göstermektedir.<sup>58</sup>

C5.0 karar ağacı türetme algoritması, tek bir düğümlle başlamakta ve en uygun sınıflayıcının tespiti için bilgi kazanımı adı verilen buluşsal ve entropi tabanlı bir ölçü kullanmaktadır. Bu ölçüye göre belirlenen değişkenin her bir değeri dallara dönüşmekte ve devam eden süreçte kalan değişkenlerin dikkate alınması ile aynı işlem sürdürülmektedir. Bu süreç, belirli bir düğümdeki tüm örneklerin aynı sınıfa ait olması, bölümlenecek yeni bir değişkenin kalmaması ve veri kümesinde sınıma değişkeninin söz konusu değerine sahip kayıt bulunmaması durumlarında son bulmaktadır.<sup>59</sup>

C5.0 algoritması kök düğümden yaprak düğüme uzanan akışla karar kuralları şeklinde de ifade edilebilmektedir. Özellikle çok büyük karar ağaçları yerine karar kurallarını kullanmak insan algısına daha uygun düşmekte ancak veri sayısı arttıkça işlem süresi uzamaktadır. Bir karar ağacında, ağacın alt seviyelerindeki değişken üst seviyelerdeki değişkene oranla daha az kullanılırken, karar kurallarında kuralın sıralamasının önem taşımamaktadır. Ayrıca, karar kuralları öngörü kesinliklerine göre sıralanabilmekte ve öngörü kesinliğini arttırmak amacıyla budanabilmektedir.

---

<sup>58</sup> Terlemeç, s.54.

<sup>59</sup> Özgür Çakır, *Veri Madenciliğinde Sınıflandırma Yöntemlerinin Karşılaştırılması: "Bankacılık Müşteri Veri Tabanı Üzerinde Bir Uygulama*, (Yayımlanmamış Doktora Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2008, s.61.

#### 1.6.2.4. Sınıflandırma ve Regresyon Ağaçları (CART)

CART algoritması Morgan ve Sonquist'in AID (Automatic Interaction Detection) adlı karar ağacı algoritmasının devamı niteliğine Breiman ve diğerleri tarafından 1984 yılında önerilmiştir. Hem sayısal hem de nominal veri türlerini, girdi ve kestirimsel değişken olarak kabul edebilen CART, sınıflandırma ve regresyon problemlerinde bir çözüm olarak kullanılabilir bir algoritmadır.<sup>60</sup>

CART analizi, cevap değişkeninin kategorik olması durumunda, sınıflandırma ağacı, (Classification Tree, CT) sürekli olması durumunda ise regresyon ağacı (Regression Tree, RT) olarak adlandırılır (Chang ve ark. 2006). Ancak, uygulamada çoğunlukla CART veya C&RT (Classification and Regression Tree) olarak kullanılmaktadır.

CART algoritmasında bir düğümde belirli bir kriter uygulanarak bölünme işlemi gerçekleştirilir. Bunun için önce tüm niteliklerin var olduğu değerler göz önüne alınır ve tüm eşleşmelerden sonra iki bölünme elde edilir. Bu bölünmeler üzerinde seçme işlemi uygulanır. Burada iki algoritma ele alınacaktır:<sup>61</sup>

1-Twoing Algoritması

2-Gini Algoritması

##### Twoing Algoritması

Aşağıdaki adımlar uygulanır:

**Adım 1.** Niteliklerin içerdiği değerler göz önüne alınarak eğitim kümesi iki parçaya ayrılır: ( $t_{sol}$ ,  $t_{sağ}$ )

**Adım 2.** Her bir bölüm için ayrı ayrı olasılıkları hesaplanır.

$$P_{sol} = \frac{t_{sol} \text{ daki herbir nitelik değerinin ilgili sütunundaki } i \text{ tekrar sayısı}}{\text{Eğitim setindeki örneklerin sayısı}} \quad (1.12)$$

<sup>60</sup> E.A. Sezer, A.S. Bozkır, S. Yağız, ve C. Gökçeoğlu, "Karar Ağacı Derinliğinin CART Algoritmasında Kestirim Kapasitesine Etkisi: Bir Tünel Açma Makinasının İlerleme Hızı Üzerinde Uygulama", *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu*, Kayseri: ASYU 21- 4 Haziran 2010, s.2.

<sup>61</sup> Özkan, s.89.



$$P(j/t_{sol}) = \frac{t_{sol} \text{daki kayıtların } j \text{ sınıfları sayısı}}{t_{sol} \text{daki her bir nitelik değerlerinin ilgili nitelik sütunundaki tekrar sayısı}} \quad (1.13)$$

$$P_{sağ} = \frac{t_{sağ} \text{daki her bir nitelik değerinin ilgili nitelik sütunundaki } i \text{ tekrar sayısı}}{\text{Eğitim setindeki örneklerin sayısı}} \quad (1.14)$$

$$P(j/t_{sağ}) = \frac{t_{sağ} \text{daki kayıtların } j \text{ sınıfları sayısı}}{t_{sağ} \text{daki her bir nitelik değerlerinin ilgili nitelik sütunundaki tekrar sayısı}} \quad (1.15)$$

**Adım 3.** t düğümündeki s aday bölünmesinin **uygunluk ölçüsü** hesaplanır.

$$\Phi(s | t) = 2P_{sol} P_{sağ} \sum_{j=1}^n |P(j | t_{sol}) - P(j | t_{sağ})| \quad (1.16)$$

**Adım 4.** Adım 1'e dönülerek ağacın alt kümesine aynı işlemler uygulanır.

### Gini Algoritması

İkili bölünmelere dayalı bir sınıflandırma yöntemidir. Nitelik değerlerinin sol veya sağ olmak üzere iki bölüme ayırma esasına dayanmaktadır:

**Adım 1:** Her nitelik değerleri ikili ağaç biçiminde gruplanır. Bu şekilde sol ve sağ bölünmelere karşılık gelen sınıf değerleri gruplandırılır.

**Adım 2:** Her bir nitelik ile ilgili sol ve sağ taraftaki için  $Gini_{sol}$  ve  $Gini_{sağ}$  değerleri hesaplanır.

$$Gini_{sol} = 1 - \sum_{i=1}^k \left( \frac{L_i}{|T_{sol}|} \right)^2 \quad (1.17)$$

k: Sınıf sayısı

T : Bir düğümdeki örnekler

$|T_{sol}|$  : Sol taraftaki örneklerin sayısı

$L_i$  : Sol taraftaki i kategorisindeki örneklerin sayısı

$$Gini_{sag} = 1 - \sum_{i=1}^k \left( \frac{R_i}{|T_{sag}|} \right)^2 \quad (1.18)$$

$|T_{sağ}|$  : Sağ taraftaki örneklerin sayısı

$R_i$  : Sağ taraftaki i kategorisindeki örneklerin sayısı

**Adım 3:** her j niteliği için, n eğitim kümesindeki satır sayısı olmak üzere aşağıdaki bağıntının değeri hesaplanır:

$$Gini = \frac{1}{n} \left( |T_{sol}| Gini_{sol} + |T_{sag}| Gini_{sag} \right) \quad (1.19)$$

**Adım 4:** Her j niteliği için hesaplanan  $Gini_j$  değerleri arasından en küçük olanı seçilir ve bölünme bu nitelik üzerinden gerçekleştirilir.

**Adım 5:** En baştaki adıma dönülerek işlemlere devam edilir.

### 1.6.3. Kümeleme Analizi

Kümeleme analizi birincil amacı, “gözlem birimlerini benzer özelliklerine göre gruplamak” olan çok değişkenli istatistik tekniklerinden biridir. Kümeleme analizinde, elde edilen bir küme içindeki gözlem birimleri, önceden belirlenmiş bir özellik bakımından birbirine benzemektedir. Dolayısıyla elde edilen kümedeki gözlem birimleri homojendir.<sup>62</sup>

Kümeleme modellerinde amaç küme üyelerinin birbirine çok benzediği, ancak özellikleri birbirinden çok farklı olan kümelerin elde edilmesi ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir.<sup>63</sup>

Kümeleme analizi en yaygın ve bilinen, tanımlayıcı veri madenciliği yöntemlerinden biridir.  $n$  adet gözlem ve  $p$  adet değişkenden oluşan bir veri matrisinde, kümeleme analizinin amacı gözlemleri kendi içinde homojen ve aralarında heterojen olan gruplara ayırmaktır.<sup>64</sup>  $n$  adet kayıt içeren bir veri seti için, hiyerarşik ve hiyerarşik

<sup>62</sup> Hair, s.473.

<sup>63</sup> Akpınar, s.6.

<sup>64</sup> Giudici, ss.75-76.

olmayan (amaç fonksiyonu temelli) şeklinde, iki genel tip kümeleme algoritması tanımlanabilir:<sup>65</sup>

- 1- Hiyerarşik yöntemler yığılmacı (agglomerative) veya ayırmacı (divisive) olabilir. Yığılmacı yöntemler  $n$  adet küme ile başlarlar ve benzer kümeleri bir tek küme elde edilene kadar sırayla birleştirirler. Ayırmacı yöntemler ise, bütün gözlemleri içeren bir tek küme ile başlamak üzere, tam tersi bir şekilde çalışırlar. Hiyerarşik yöntemler özellikle amacın, kümeleri doğal hiyerarşisi içerisinde düzenlemek olduğu durumlarda kullanışlıdır.
- 2- Hiyerarşik olmayan, örneğin k-ortalamlar gibi yöntemler önceden belirlenmiş sayıda küme kullanarak gözlemleri her bir kümeye atarlar. Bu yöntemler genellikle bilgisayar kaynakları ve hesaplamalar açısından daha az yoğundur. Bu yüzden çok büyük veri setlerinde tercih edilirler.

Kümeleme analizinin diğer analiz teknikleriyle arasındaki farklılıkları dikkate alındığında; ayırma analizinde gruplar önceden belirlenirken, kümeleme de bu belirleme, analiz sonucunda elde edilmektedir. Diğer taraftan, kümeleme analizinde değişkenleri, bağımlı (kriter) ve bağımsız (tahmin edici) değişkenler şeklinde ikiye ayırmaz. Bu yönüyle Faktör Analizine benzer. Faktör analizine benzeyen bir diğer vasfı da sınıflandırma özelliğidir. Kümelemenin, sınıflandırmadan farkı grupların ayrıştırılırken önceden tanımlanmış özelliklere göre sınıflandırılmamalarıdır. Yani bir değişkene bağımlı kalarak sınıflandırma söz konusu değildir. Örneğin, yaşa veya gelire göre bir sınıflandırma söz konusu değildir. Eğer müşteriler çeşitli özelliklerine göre bazı gruplara kendiliğinden ayrışıyorsa, bu kümeler onların bazı özellikleri açısından birbirlerine yakın olduklarını gösterir. Kendi içlerinde homojen olan, ancak birbirlerinden farklı özellikler taşıyan gruplar oluşur.<sup>66</sup>

### 1.6.3.1. Benzerlik/Uzaklık Ölçüleri

Kümeleme analizinin amacı, gözlenen birey ya da nesnelere arasındaki benzerlikleri ya da uzaklık/yakınlıkları tespit etmektir. Benzerlik, aslında uzaklık kavramının tersi olup, büyük bir sayı olduğunda iki nesnenin birbirine yakın olduğunu,

<sup>65</sup> S. Irmak, *Veri Madenciliği Yöntemleri ile Sağlık Sektörü Veritabanlarında Bilgi Keşfi: Tanımlayıcı ve Kestirimci Model Uygulamaları*, (Yayımlanmamış Doktora Tezi), Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü, Antalya 2009, ss.42-43.

<sup>66</sup> Özmen, s.189.

küçük bir sayı çıktığında da iki nesnenin birbirinden uzak olduğunu gösterir.<sup>67</sup> Uzaklık ölçüleri aşağıdaki çizelgede verilmiştir.

**Çizelge 1.1.** Kümeleme Analizinde Kullanılan Uzaklık Ölçüleri

<b>Uzaklık Fonksiyonu</b>	<b>Formülü</b>
<i>Öklit Uzaklığı</i>	$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
<i>Manhattan Uzaklığı</i>	$d(x, y) = \sum_{i=1}^n  x_i - y_i $
<i>Tchebyshev Uzaklığı</i>	$d(x, y) = \max_{i=1,2,\dots,n}  x_i - y_i $
<i>Minowski Uzaklığı</i>	$d(x, y) = \sqrt[p]{\sum_{i=1}^n  x_i - y_i ^p}, \quad p > 0$
<i>Canberra Uzaklığı</i>	$d(x, y) = \sum_{i=1}^n \frac{ x_i - y_i }{x_i + y_i}, \quad x_i, y_i > 0$
<i>Korelasyon Uzaklığı</i>	$d(x, y) = 1 - r_{(x,y)}^2$
<i>Mahalanobis Uzaklığı</i>	$d(x, y) = \sqrt{ x_i - y_i  \sum  x_i - y_i }$
<i>Açısal Ayrım</i>	$d(x, y) = \sum_{i=1}^n x_i y_i / \left[ \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 \right]^{1/2}$

**Kaynak:** Levent Terlemez, Eş İşlem Stratejisi Yöntemiyle İMKB’de Portföy Oluşturmada Veri Madenciliği Uygulaması, (Yayımlanmamış Doktora Tezi), Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir 2008, s.43.

<sup>67</sup> Nur Karadeniz, *Müşteri İlişkileri Yönetimi Açısından Veri Madenciliği Yöntemi ve Hizmet Sektörü Üzerine Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2008, s.48.

Öklit uzaklığı en yaygın kullanılan uzaklık ölçüsüdür. Ancak öklit uzaklığının üç temel özelliği kullanılırken dikkate alınmalıdır. Birincisi, yüksek oranda ölçeğe bağımlı olmasıdır. Bir değişkenin biriminin değiştirilmesinin sonuçlar üzerinde büyük etkisi olabilir. Standardizasyon bu sorun için bir çözüm olabilir. Öklit uzaklığının dikkat edilmesi gereken ikinci özelliği de ölçümler arasındaki ilişkileri dikkate almamasıdır. Bu yüzden eğer ölçümler güçlü bir şekilde ilişkili (ör. yüksek korelasyon) ise Mahalanobis uzaklığı gibi başka uzaklık fonksiyonları kullanılmalıdır. Üçüncü olarak ise eğer veri aykırı değerler içeriyor ve bunların giderilmesi mümkün değilse, Manhattan uzaklığı gibi daha sağlam uzaklık fonksiyonları tercih edilebilir.<sup>68</sup>

Kümeleme teknikleri Hiyerarşik ve Hiyerarşik Olmayan kümeleme teknikleri olarak ikiye ayrılır.

### 1.6.3.2. Hiyerarşik Kümeleme

Hiyerarşik (Aşamalı) kümeleme yöntemleri, birimlerin benzerliklerini dikkate alarak belirli düzeylerde (küme uzaklık ölçüleri) birbirleri ile birleştirmeyi amaçlayan yöntemlerdir.

Hiyerarşik kümeleme yöntemleri, kümelerin bir ana küme olarak ele alınması ve sonra aşamalı olarak içerdiği alt kümelere ayrılması veya ayrı ayrı ele alınan kümelerin aşamalı olarak bir küme biçiminde birleştirilmesi esasına dayanır. Bu grupta birçok hiyerarşik yöntem bulunmaktadır. Söz konusu yöntemlerden en yakın komşu algoritması ve en uzak komşu algoritması incelenecektir:<sup>69</sup>

**En Yakın Komşu Algoritması:** Tek bağlantı yöntemi “En Yakın Komşu” yöntemi olarak da adlandırılmaktadır. Bu yöntemde ilk belirlenen kümedeki elemanların her biriyle diğer elemanlar arasındaki mesafelerden minimum olanı seçilir ve bu iki eleman ilk kümeyi oluşturur. Daha sonra ya oluşturulan ilk kümeyle en yakın mesafedeki eleman seçilerek kümeyle eklenir ya da daha yakın mesafede olan iki eleman belirlenerek başka bir küme oluşturmaları sağlanır. Bu işlem bütün kümeler birleşerek

<sup>68</sup> G. Shmueli, N. R. Patel and P. C. Bruce, *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner*, John Wiley & Sons, Hoboken, NJ, USA 2007, pp.223.

<sup>69</sup> Özkan, s.136.

tek bir kümeye dönüşüncüye kadar devam eder.<sup>70</sup> Gözlemler arasındaki uzaklıkların belirlenmesinde Öklid uzaklık bağlantısı kullanılabilir ve daha sonra  $\text{Min } d(i,j)$  seçilir.

**En Uzak Komşu Algoritması:** Tam bağlantı yöntemi “En Uzak Komşu Yöntemi” olarak da adlandırılmaktadır. En yakın komşu yönteminin tam tersidir. İlk aşamada birbirine en yakın iki gözlem bir küme oluşturmaktadır. İki kümedeki tüm gözlem çiftleri arasındaki uzaklıklar içinden maksimum olanı, iki küme arasındaki uzaklık olarak tanımlanmıştır.

### 1.6.3.3. Hiyerarşik Olmayan Yöntemler

Hiyerarşik Olmayan Yöntemler, birimlerin kendi içinde homojen ve kendi aralarında heterojen olan kümelere ayrılmasını hedefleyen ve bu kümeler aracılığı ile alt popülasyonların parametre tahminlerini yapmayı amaçlayan yöntemlerdir. Hiyerarşik yöntemlerde hem birimler hem de değişkenler birbirleriyle değişik benzerlik düzeylerinde kümeler oluştururken, hiyerarşik olmayan yöntemlerde birimlerin uygun oldukları kümelere toplanmaları ve  $n$  birimin  $k$  adet kümeye parçalanması hedeflenmektedir.<sup>71</sup> Söz konusu yöntemlerden  $k$  ortalamalar, metoid ve yığma kümeleme yönetemi ele alınacaktır.

**K- Ortalamalar Yöntemi:** K-ortalamalar yöntemi 1967 yılında Mac Queen tarafından sunulmuştur. Uzun yıllar boyunca pek çok uygulama alanında yoğun olarak kullanılan bir kümeleme algoritmasıdır. Bu algoritmada  $k$  sayıda küme oluşmakta ve her küme içerisinde bulunan verilerin ağırlıklı ortalamaları sonucu bir değer ortaya çıkmaktadır. Küme içerisinde bu değere en yakın olan nokta değeri küme merkezi (centroid) olarak kabul edilmektedir.<sup>72</sup>

K-ortalamalar algoritması gürültülü veriye ve aykırı değerlere aşırı duyarlıdır. Bunun nedeni, bu türdeki az sayıda verinin bile ortalama değeri (mean) büyük ölçüde etkileyebilir olmasıdır. K-ortalamalar yönteminin aksine K-ortancalar yöntemi küme temsilcisi olarak gözlemlerin ortalama değerini almak yerine, bir küme içerisinde en

<sup>70</sup> Hair, s.494.

<sup>71</sup> Kazım Özdamar, *Paket Programlar ile İstatistiksel Veri Analizi*, Kaan Kitabevi, Eskişehir 2004, s.324.

<sup>72</sup> Berkhin Pavel, *Survey of Clustering Data Mining Techniques* [Elektronik Sürüm], Accrue Software, California 2002, s.54.

merkezi olarak yerleşmiş nesneyi (medyan) kullandığı için, gürültülü veriye ve aykırı değerlere daha az duyarlıdır.<sup>73</sup>

Bu yöntemde, daha başlangıçta belli sayıdaki küme için toplam ortalama hatayı minimize etmek amaçlanır. N boyutlu uzayda N örnekli küme içerisinde bu uzay {C1, C2, C3,.....Ck} biçiminde K kümeye ayrılır. Ck kümesinin ortalama vektörü  $M_k$  şu şekilde hesaplanır:<sup>74</sup>

$$M_k = \frac{1}{n_k} \sum_{i=1}^{n_k} X_{ik} \quad (1.20)$$

Burada  $X_k$  değeri  $C_k$  kümesine ait olan i. örnektir.  $C_k$  kümesi için kare-hata, her bir  $C_k$  örneği ile onun merkezi arasındaki Öklid uzaklıkları toplamıdır. Bu hataya küme içi değişme adı da verilir. Küme içi değişmeler şu şekilde hesaplanır:

$$e_i^2 = \sum_{i=1}^{n_k} (x_{ik} - M_k)^2 \quad (1.21)$$

K kümesini içeren bütün kümeler uzayı için kare-hata, küme içindeki değişmelerin toplamıdır. O halde söz konusu kare-hata değeri şu şekilde hesaplanır.

$$E_k^2 = \sum_{k=1}^K e_k^2 \quad (1.22)$$

Küme-hata kümeleme yönteminin amacı, verilen K değeri için  $E_k^2$  değerini minimize eden K kümeleri bulmaktır. O halde k-ortalama algoritmasında  $E_k^2$  değerinin bir önceki iterasyona göre azalması beklenir.

**Metoid Yönetmi:** n birimin, küme içi gözlemlerin benzer, kümelerarası gözlemlerin farklı olacağı biçimde iki veya daha fazla kümeye ayrılmasını amaçlar. Bu parçalamada metoid adı verilen kümeleri tanıtıcı merkez noktalar veya gözlemler yardımı ile n birimi k kümeye ayırma sağlanır. Metoid kümelemede en önemli sorun

<sup>73</sup> Kantardzic, ss.131-132.

<sup>74</sup> Özkan, ss.149-150.

merkez noktaların belirlenmesidir. Metoidler belirlendikten sonra her birim aralarındaki benzerliklerin maksimum ve farklılıkların minimum olduğu en yakın metoide sahip olan kümeye atanır.<sup>75</sup>

**Yığma Kümeleme Yöntemi:** Bu yöntemde birimler, en yakın ortalamalı kümeye atanma yerine önceden belirlenen bir istatistiksel kritere göre bir kümeden diğerine hareket ederek en uygun durum sağlandığında belirli bir kümede yer alırlar. Belirlenen istatistik kritere göre veri setindeki tüm birimler hangi kümede yer alacakları kesin olmaksızın atama işlemlerine tabi tutulur. Bu istatistiksel kriter, küme içi kovaryans matrisi ve kümelerarası kovaryans matrisi ile ilgili olarak geliştirilmiştir. Toplam kovaryans matrisinin determinantının, küme içi kovaryans matrisinin determinantına oranının maksimum olduğu küme durumunun logaritması alınarak hesaplanan katsayı (c), kümeleri ayırmada bir kriter olarak alınmaktadır. n birim ardışık olarak parçalanıp k kümeye ayrılarak c katsayısının optimum olduğu değere ulaşıldığında elde edilen k adet kümenin, n birimin en uygun kümelenmesi olacağı ileri sürülmüştür.<sup>76</sup>

#### 1.6.4. Birliktelik Kuralları

Veritabanı içinde yer alan kayıtların birbirleriyle olan ilişkilerini inceleyerek, hangi olayların eş zamanlı olarak birlikte gerçekleşebileceklerini ortaya koymaya çalışan veri madenciliği yöntemleri bulunmaktadır. Bu ilişkilerin belirlenmesiyle “birliktelik kuralları” (association rules) elde edilir.<sup>77</sup>

Bir birliktelik kuralı, veritabanı içerisinde belirli nesnelerin birlikte görülme olasılığı ifadesidir.<sup>78</sup>

Bu yöntemin temeli veri ambarlarındaki değişkenler ve veriler arasındaki ilginç ilişkileri ortaya çıkarmaya dayanmaktadır. Yöntem Pazar Sepet Analiz uygulamalarında yaygın olarak kullanılmaktadır.

Birliktelik kuralları yöntemlerinin uygulanmasında dikkat edilmesi gereken iki konu vardır. Birincisi, büyük bir işlemsel veri setinden örüntülerin çıkartılmasının bilgisayar kaynakları ve hesaplamalar açısından oldukça maliyetli ve zaman alıcı olmasıdır. İkincisi ise bulunan bazı örüntülerin şans eseri ortaya çıkma ve sahte olma

<sup>75</sup> Özdamar, s.325.

<sup>76</sup> Şimşek, s.137.

<sup>77</sup> Özkan, s.48.

<sup>78</sup> Hand, s.158.



ihtimalidir. Bunun için keşfedilen örüntüler ayrıca değerlendirme süreçlerinden geçirilmelidir.<sup>79</sup>

Birliktelik kuralları pazarlamaya, reklamcılığa, yer düzenlemeleri ve stok kontrolüne yardımcı olmak için sıklıkla perakende satış mağazaları tarafından kullanılır. Perakendecilik sektöründe direk uygulanabilirliği olmasına rağmen, iletişim ağlarındaki kusurları kestirmek gibi farklı amaçlar içinde kullanılmıştır. Birliktelik kuralları veri öğeleri arasındaki ilişkileri göstermek için kullanılır. Ortaya çıkartılmamış bu ilişkiler, fonksiyonel bağımlılıklarda olduğu gibi, verinin özünde değildir ve her hangi bir nedensellik türünü veya ilişkiyi temsil etmez. Onun yerine, birliktelik kuralları öğelerin genel kullanımını belirler. İçinde birliktelik kuralı bulunabilecek bir veritabanı, her kümenin bir öğeler kümesi taşıdığı, kaydı oluşturan veri kümesi olarak görülür.<sup>80</sup>

Bir öğenin (veya öğeler setinin) desteği, o öğenin (veya öğelerin) ortaya çıktığı işlemlerin yüzdesidir.

**Tanım 1:**  $I = \{I_1, I_2, \dots, I_m\}$  öğeler seti ve  $t_i = \{I_{i1}, I_{i2}, \dots, I_{ik}\}$ ,  $I_{ij} \in I$  iken  $D = \{t_1, t_2, \dots, t_n\}$  işlemler veritabanı verilsin, birliktelik kuralı  $X, Y \subset I$  öğe kümeleri olarak adlandırılan öğe kümeleri ve  $X \cap Y = \emptyset$  iken,  $X \Rightarrow Y$  formunun gerektirmesidir.

**Tanım 2:** Bu birliktelik kuralı  $X \Rightarrow Y$  için destek (s)  $XUY$  içeren bir veritabanındaki işlemlerin yüzdesidir.

**Tanım 3:** Bir birliktelik kuralı  $X \Rightarrow Y$  için güven veya güç (a),  $XUY$  içeren işlemlerin sayısının  $X$  taşıyan işlemlerin sayısına oranıdır. Genellikle tüm gerekliliklerle (implication) ilgilenilmez. Burada önem, Tanım 2 ve Tanım 3'te tanımlanan sıklıkla destek ve güven olarak isimlendirilen iki özellik ile ölçülür.

Birliktelik kurallarının seçimi Tanım 4'te birliktelik kuralı probleminin tanımında açıklandığı gibi bu iki değere dayanır. Destek kuralın veritabanında hangi sıklıkla ortaya çıktığını ölçerken, güven kuralın gücünü ölçer. Tipik olarak, büyük güven değerleri ve küçük bir destek değeri kullanılır. Destek kuralın veritabanının tümünde kaç kez ortaya çıktığının yüzdesini gösterdiğinden destek için küçük değerlere izin verilebilir. Birliktelik kuralında önceki ortaya çıktığında bağlı olan her zaman ortaya çıkar.

<sup>79</sup> Tan, vd., s.328.

<sup>80</sup> Terlemez, s.57.

**Tanım 4:** =  $\{I_1, I_2, \dots, I_m\}$  öğeler seti ve  $t_i = \{I_{i1}, I_{i2}, \dots, I_{ik}\}$ ,  $I_{ij} \in I$  iken  $D = \{t_1, t_2, \dots, t_n\}$  işlemler veritabanı verilsin, birliktelik kuralı problemi tüm  $X \Rightarrow Y$  birliktelik kurallarının bir minimum destek ve güvenle tanımlamaktır. Bu değerler  $(s, a)$  probleme girdi olarak verilir.<sup>81</sup>

#### 1.6.4.1. Apriori Algoritması

Agrawal tarafından 1994 yılında geliştirilen Apriori algoritması, birliktelik kuralları çıkarımında en çok bilinen algoritma olmuştur. Algoritmanın ismi, yaygın nesnelerin önsel bilgilerini kullanmasından yani bilgileri bir önceki adımdan almasından “önceki (prior)” anlamında aprioridir.

Apriori algoritmasının aşamaları şunlardır.<sup>82</sup>

- a) Birliktelik çözümlemesinin yapılabilmesi için öncelikle destek ve güven ölçülerini karşılaştırmak üzere eşik değerler belirlenir. Elde edilen sonuçların bu eşik değerlerden eşit veya büyük olması istenir.
- b) Veritabanı taranarak çözülemeye dahil edilecek her bir ürün için tekrar sayıları, yani destek sayıları hesaplanır. Bu destek sayıları eşik destek sayısı ile karşılaştırılır. Eşik destek sayısından küçük değere sahip satırlar çözülemeden çıkarılır ve koşula uygun kayıtlar göz önüne alınır.
- c) Yukarıdaki adımda seçilen ürünler bu kez ikişerli gruplandırılarak, bu grupların tekrar sayıları, yani destek sayıları elde edilir. Bu sayılar eşik destek sayıları ile karşılaştırılır. Eşik değerden küçük değerlere sahip satırlar çözülemeden çıkarılır.
- d) Bu kez üçerli, dörderli vb. gruplandırmalar yapılarak bu grupların destek sayıları elde edilir ve eşik değer ile karşılaştırılır, eşik değerlere uygun olduğu sürece işlemlere devam edilir.
- e) Ürün grubu belirlendikten sonra kural destek ölçütüne bakılarak birliktelik kuralları türetilir ve bu kuralların her birisiyle ilgili olarak güven ölçütleri hesaplanır.

<sup>81</sup> Terlemez, ss.58-59.

<sup>82</sup> Özkan, s.159.

➤ **Destek ve Güven ölçülerinin hesaplanması:**

Satılan ürünler arasındaki ilişkileri ortaya koymak için destek ve güven gibi iki ölçütten yararlanılır. Bu ölçütlerin hesaplanmasında destek sayısı adı verilen bir değer kullanılır.<sup>83</sup>

**Kural destek ölçütü** : Bir ilişkinin tüm alışverişler içinde hangi oranda tekrarlandığını belirler.

**Kural güven ölçütü** : Bir ürün grubunu satın alan müşterilerin başka bir ürün grubunu da satın alma olasılığını gösterir.

A ürün grubunu alanların B ürün grubunu da satın alma durumu, yani birliktelik kuralı  $A \rightarrow B$  biçiminde gösterilir. Bu durumda kural destek ölçütü aşağıdaki gibi ifade edilebilir:

$$destek(A \rightarrow B) = \frac{sayı(A, B)}{N} \quad (1.23)$$

Burada  $sayı(A, B)$ , A ve B ürün gruplarını birlikte içeren alışveriş sayısını göstermektedir. N ise tüm alışverişlerin sayısını göstermektedir.

A ve B ürün gruplarının birlikte satın alınması olasılığını ifade eden kural güven ölçütü aşağıdaki gibi ifade edilebilir:

$$güven(A \rightarrow B) = \frac{sayı(A, B)}{sayı(A)} \quad (1.24)$$

Bu ölçütleri karşılaştırmak için eşik değere ihtiyaç vardır. Hesaplanan destek ve güven ölçütlerinin destek(eşik) ve güven(eşik) değerlerinden büyük olması beklenir. Hesaplanan destek ve güven ölçüleri ne kadar büyük ise birliktelik kurallarının da o derece güçlü olduğu anlaşılır.

---

<sup>83</sup> J. Tan Wang, *Encyclopedia of Data Warehousing and Mining* [Elektronik Sürüm], Idea Group, USA 2006, pp.795.

#### 1.6.4.2. Pazar Sepet Analizi

Pazar Sepet Analizi uygulaması yaygın olarak kullanılan diğer bir birliktelik kuralı tekniğidir.

Pazar sepeti analizi müşterilerin alışveriş alışkanlıklarının veritabanındaki bilgiler aracılığıyla ortaya çıkartılması işlemidir. Bu alışveriş alışkanlıklarının ortaya çıkartılması alış-veriş merkezindeki ürünlerin yerleştirilmesi, marketin alanın tasarımı ve markette sergilenecek ve satılacak olan ürünlerin belirlenmesinde yardımcı olur.<sup>84</sup>

Bu yöntemde kurallar sol ve sağ kısım adı altında iki kısımdan oluşur. Sol kısım öncül veya koşul olarak adlandırılırken, sağ kısımda bağlı koşul olarak adlandırılır. Bu kural ilişkisinde sol taraf doğru ise sağ taraf da doğru olmaktadır. Buradaki genel form IF-THEN ilişkisi ile ifade edilmektedir. Bu durum;

**İF** <önkoşul> **THEN** <bağlı koşul>

şeklinde gösterilebilir. Burada ön koşul, bir veya birden fazla girdi değişkeninin değer veya değer aralığını, bağlı koşul da, bir veya birden fazla çıktı değişkeninin değer veya değer aralığını gösterir.<sup>85</sup>

#### 1.6.5. Destek Vektör Makineleri

Veri madenciliği sınıflama problemlerinde kullanılan bir diğer yöntem Destek Vektör Makinesi (Support Vector Machines) adını taşımaktadır. Bu yöntem sınıflandırmayı bir doğrusal ya da doğrusal olmayan bir fonksiyon yardımıyla yerine getirir. Destek vektör makinesi yöntemi, veriyi birbirinden ayırmak için en uygun fonksiyon tahmin edilmesi esasına dayanır. Daha çok makine öğrenmesi yöntemleri arasında yer alan bu yöntem günümüzde veri madenciliği alanında da tercih edilmeye başlanmıştır.<sup>86</sup>

Bir destek vektör makinesi, verileri optimal olarak iki kategoriye ayıran n – boyutlu bir hiperdüzlem oluşturmaktadır. DVM modelleri yapay sinir ağlarıyla yakından ilişkili olup, sigmoid bir kernel fonksiyonu kullanan DVM; iki katmanlı, ileri beslemeli bir yapay sinir ağına sahiptir.<sup>87</sup>

<sup>84</sup> Gökhan Silahtaroglu, *Veri Madenciliği*, Papatya Yayıncılık Eğitim, İstanbul 2008, s.83.

<sup>85</sup> Aysan Şentürk, *Veri Madenciliği, Kavram Ve Teknikler*, Ekin Yayınevi, Bursa 2006, s.19.

<sup>86</sup> Özkan, s.185.

<sup>87</sup> Haykin, s.318.

DVM, doğrusal öğrenme makineleri, genelleme teorisinin sağladığı kavrayışları göz önünde bulundurarak ve optimizasyon teorisini uygulayarak, kernele bağlı öznelik uzayında verimli bir şekilde eğitmek için geliştirilen bir sistemdir. Bu sistemlerin en önemli özelliklerinden biri genelleme teorisi ile elde edilen eğilimleri uygularken, hipotezin seyrek gösterimlerini üretmesi sonucunda çok verimli algoritmanın oluşmasıdır.<sup>88</sup>

DVM, tahminleyici bir değişken bir simge hiperdüzlemi tanımlamak için kullanılan bir simge ise bir özellik olarak nitelendirilmektedir. DVM modellemenin amacı; hedef değişkenin tek kategorili durumları düzlemin bir tarafında ve diğer kategorili durumları düzlemin diğer tarafında olacak şekilde, vektör kümelerini ayıran optimal hiper düzlemi bulmaktır.<sup>89</sup>

DVM'ler sadece bir istatistiksel kriter yerine marj tabanlı geometrik bir kriter kullanan bir sınıflandırma tekniği uygulamaktadır yani DVM'ler sınıflandırma işlevini istatistiksel teniklere gereksinim duymadan sınıflandırma işlevini marj maksimizasyonu tekniğiyle gerçekleştirmektedir.<sup>90</sup>

DVM'nin ilgi çekici özelliği; veri seti üzerinde ortalama hata karesini minimize ederek türetilen ampirik risk minimizasyonu prensibinden ziyade, istatistiksel öğrenme teorisindeki yapısal risk minimizasyonu niteliğinde çalışmasıdır. DVM'nin temel varsayımlarından biri, eğitim kümesindeki tüm örneklerin bağımsız ve benzer olarak dağılmış olmasıdır.<sup>91</sup>

DVM, optimum doğru ile sınıfları birbirinden ayırmayı amaçlamaktadır.

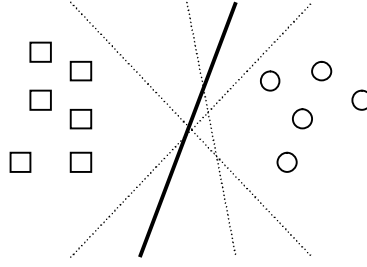
---

<sup>88</sup> Türker Tokan, *Destek Vektör Makinelerinin Mikrodalga Teori ve Tekniğindeki Uygulamaları*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2009, s.19.

<sup>89</sup> Witten and Frank, ss.41-49.

<sup>90</sup> Farid Melgani and Lorenze Bruzzone, "Classification of Hyperspectral Remote Sensing Images with Support Vector Machines" [Destek Vektör Makineleriyle Hiperspektral Uzaktan Algılama Resimlerinin Sınıflandırması], *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 2004, pp. 1778, Erişim Tarihi: 2 Mart 2012, IEEE.

<sup>91</sup> O. Song, W. Hu and W. Xie, "Robust Support Vector Machine with Bullet Hole Image Classification" [Kurşun Deliklerini Destek Vektör Makineleri ile Sınıflandırma], *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews*, 32(4), November 2002, pp. 440, Erişim Tarihi: 10 Şubat 2012, ScienceDirect.

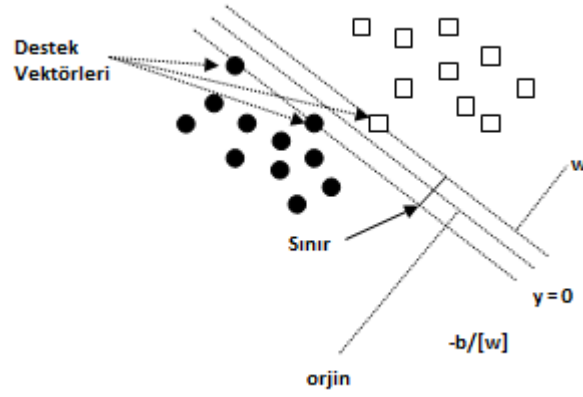


**Şekil 1.4.** Optimum ayırıcı düzlemler

Bu sınıflar arasında birçok düzlemin ihtimali vardır. Hata toleransını en düşük hale getirmek için çizilebilecek doğrular arasında her iki sınıfa en uzak olacak çizilen doğru bulunmalıdır.

DVM, sınıflandırma ve regresyon problemlerinde çalıştırılabilir. Sınıflandırma problemlerinin birkaç türü vardır. Bunlardan en temel olanı eğitim hatasının bulunmadığı doğrusal ayrılabilirlik durumudur. Bu durumda DVM giriş uzayı üzerinde doğrusal bir ayırıcı fonksiyon oluşturur. Bir başka durum, eğitim hatasının bulunmadığı doğrusal olmayan ayrılabilirlik durumudur. Bu durumda DVM sınıflandırıcısı giriş uzayı üzerinde doğrusal sınıflandırma işlemini gerçekleştiremez. Bu nedenle, ilk önce çekirdek fonksiyonlarının yardımıyla giriş uzayından nitelik uzayına dönüşüm yapılır. Daha sonra DVM sınıflandırıcısı nitelik uzayı üzerinde doğrusal ayırıcı fonksiyonunu oluşturur. Eğitim hatasının bulunduğu durumlarda da aynı işlemler yapılır. Fakat sisteme pozitif değerli bir esneklik parametresi eklenir. Hem sınıflandırma hem de regresyon işlemlerinde, öğrenme problemi ikinci dereceden amaç fonksiyonuna sahip bir optimizasyon problemi formunda temsil edilir. DVM regresyon metodundaki temel fikir, eldeki eğitim verilerinin karakterini mümkün olduğunca gerçeğe yakın bir şekilde yansıtan ve istatistiksel öğrenme teorisine uyan doğrusal ayırıcı fonksiyonun bulunmasıdır. Sınıflandırmaya benzer bir şekilde regresyonda da doğrusal olmayan durumların işlenebilmesi için çekirdek fonksiyonları kullanılır.<sup>92</sup>

<sup>92</sup> E. Çomak, *Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar*, (Yayımlanmamış Doktora Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2008, s.37.



**Şekil 1.5.** DVM Sınıflandırıcı

**Kaynak:** E. Çomak, Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar, (Yayımlanmamış Doktora Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2008, s.37.

Şekil 1.5’de basit bir DVM sınıflandırıcının yapısını göstermektedir. Amaç orjinin (optimum ayırıcı düzlem vektörünün) bulunmasıdır. Giriş verileri  $\{+1, -1\}$  değerini alan iki sınıfa ayrılırlar. Destek vektörler bulunduğu sınıfın orijine en yakın elemanıdır ve ayırıcı düzlemin optimum olması için destek vektörlerle ayırıcı düzlem arasındaki mesafenin maksimum olması gerekir.<sup>93</sup>

Destek Vektör Makinelerinde karşılaşılabilecek iki durum, sınıflandırma yapılırken verilerin lineer olarak ayrılacakları bir yapıda olması veya lineer olarak ayrılamayan yapıda olmasıdır. Ancak, gerçek yaşam problemlerinin büyük çoğunluğu birçok farklı bileşenden oluşan problemlerdir ve lineer olarak ayrılmış bir yapı halinde karşımıza çıkmazlar. Lineer olarak ayrılmış olan veriler arasında direkt olarak maksimum sınırın bulunması oldukça kolaydır, ancak lineer olarak ayrılamayan veriler öncelikle lineer olarak ayrılacakları farklı bir uzaya aktarılmalıdırlar. Bu bölüm altında bu işlemlerin nasıl gerçekleştirildiği açıklanacaktır.<sup>94</sup>

<sup>93</sup> E. Çomak, *Destek Vektör Makineleri Çoklu Sınıf Problemleri İçin Çözüm Önerileri*, (Yayımlanmamış Yüksek Lisans Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2004, s.25.

<sup>94</sup> D. A. Demirci, *Destek Vektör Makineleri ile Karakter Tanıma*, (Yayımlanmamış Yüksek Lisans Tezi), Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, Sakarya 2007, s.3.

### 1.6.5.1 Linear Destek Vektör Makineleri

Linear DVM'ler linear olarak ayrılabilen ve ardından iç içe geçmiş ve linear olarak ayrılamayan verilerle işlem yapan DVM'ler olarak ele alınacaktır.

**Linear Ayrılabilme Durumu:** Eğitim için kullanılacak N elemandan oluşan verinin  $\theta = \{x_i, y_i\}$ ,  $i = 1, 2, \dots, N$  olduğu varsayalım. Burada  $y_i \in \{-1, 1\}$  etiket değerleri ve  $x_i \in R^d$  özellikler vektörüdür. Linear olarak ayrılabilme durumunda, bu iki değerli veriler direkt olarak bir aşırı düzlem ile ayrılacaktır. Bu aşırı düzleme ayırıcı aşırı düzlem adı verilir. Destek Vektör Makinelerinin amacı bu aşırı düzlemin iki ayrı sınıfta bulunan örnek grubuna eş uzaklıkta olmasını sağlamaktır.

DVM ilk işlemi linear olarak ayrılamayan verileri yüksek boyutlu bir forma dönüştürür. Bu yüksek boyut verinin özelliklerinin boyutundan daha büyük derecede olacaktır. Bu işlemin yapılmasının amacı verilerin bir aşırı düzlem ile ayrılmasını sağlamaktır. Aşırı düzlem üzerindeki herhangi bir noktası, aşırı düzlemin normali ve  $|b|/\|w\|$  hiperuzayın orijine dik uzaklığı olmak üzere

$$w^T \cdot x + b = 0 \quad (1.25)$$

koşulunu sağlar. DVM yönteminde araştırılan, örnekler linear olarak ayrılabilirse, ve  $y_i = -1$  etiketli örneklere eşit uzaklıkta olan optimum ayırıcı aşırı düzlemin bulunmasıdır. Bunun için eğitim setinin aşağıdaki yapıları sağladığını varsayalım:

$$y_i = +1 \text{ için, } w^T \cdot x_i + b \geq +1 \quad (1.26)$$

$$y_i = -1 \text{ için, } w^T \cdot x_i + b \leq -1 \quad (1.27)$$

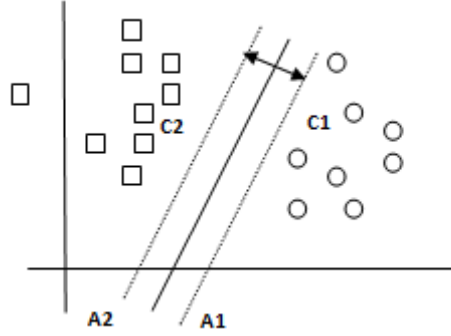
Bu eşitsizlikler bir arada ifade edilecek olursa,  $i = 1, 2, \dots, N$  için,

$$y_i(w^T \cdot x + b) \geq +1 \quad (1.28)$$

elde edilir. (1.28) koşulunu sağlayan aşırı düzlemin iki tarafındaki en yakın örneklere olan dik uzaklıkları toplamı sınır olarak adlandırılır. Optimum ayırıcı aşırı düzlem, sınırı



maksimum yapan aşırı düzlemdir. Optimum ayırıcı aşırı düzlemi bulmak için uygun  $w$  ve  $b$  değerleri hesaplanacaktır.<sup>95</sup>



**Şekil 1.6.** Lineer Ayrılabilme Durumunda Optimum Ayırıcı Aşırı düzlem

(1.26) eşitsizliğinin Şekil 2-6'daki C1 sınıfını ayıran A1 aşırı düzlemini oluşturan eşitsizlik, (1.27) eşitsizliğinin ise aynı şekildeki C2 sınıfını ayıran A2 aşırı düzlemini oluşturan eşitsizlik olduğunu varsayalım. Bu durumda A1 aşırı düzleminin orijine olan dik uzaklığı  $|-1 - b|/\|w\|$  olacaktır. Bu iki aşırı düzlemin optimal aşırı düzleme olan uzaklığı ise  $1/\|w\|$  kadardır, yani iki örnek kümesi arasındaki uzaklık A1 ve A2 aşırı düzlemlerinin birbirlerine paralel olmalarından dolayı  $2/\|w\|$  kadardır. A1 ve A2 aşırı düzlemleri arasında eğitim verilerine ait hiçbir örnek bulunmamaktadır. Bu iki aşırı düzlem arasındaki maksimum uzaklık ise  $\|w\|$  değerinin minimize edilmesiyle bulunabilir. DVM yöntemiyle yapılmaya çalışılan bu iki aşırı düzlemin arasındaki uzaklığın (sınırın) maksimum olmasını sağlamaktır. Aralarındaki sınır maksimum olan aşırı düzlemlerin üzerinde yer alan noktalara ise Destek Vektörleri (DV) adı verilmektedir.

Maksimum sınırın bulunması işlemi;

$$\min \frac{1}{2} \|w\|^2 \quad (1.29)$$

<sup>95</sup> A.U. Özkaya, *Intelligent Arrhythmia Classification Based On Support Vector Machines*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2003, s.19.

$$y_i(w^T x_i + b) \geq +1, \forall i \quad (1.30)$$

ile ifade edilir. Burada (1.29) çözülecek problem ve (1.30) problemin çözümü sırasında kullanılan koşuldur ve bu ifade ikinci dereceden optimizasyon problemidir. Problemin çözümü için problemin Lagrange formülasyonu yapılır. Bunun yapılmasının iki sebebi vardır; birincisi problemin Lagrange formülasyonu yapılarak Lagrange çarpanlarının hesaplanması daha kolaydır, ikincisi ise problemin lineer olmayan durumu içinde genelleştirilmesi için daha uygundur.<sup>96</sup>

Problemin Lagrange formülasyonu ise,

$$L_P = \frac{1}{2} \|W\|^2 - \sum_{i=1}^N a_i y_i (w^T x_i + b) + \sum_{i=1}^N a_i \quad (1.31)$$

şeklindedir. Bu formülasyonda  $a_i \geq 0$  değerleri pozitif Lagrange çarpanları olarak adlandırılır. (1.31)'deki formülasyonun çözülmesi için Karush-Kuhn-Tucker (KKT) koşulları kullanılarak dual problemine dönüştürülür. Bu problem için KKT koşulları:

$$\frac{\partial L_P}{\partial w} = 0 \rightarrow w = \sum_{i=1}^N a_i y_i x_i \quad (1.32)$$

$$\frac{\partial L_P}{\partial b} = 0 \rightarrow w = \sum_{i=1}^N a_i y_i = 0 \quad (1.33)$$

şeklindedir. Bu koşullar (1.31)'de yerine yazılacak olursa

$$\begin{aligned} L_P &= \frac{1}{2} (w^T w) - w^T \sum_{i=1}^N a_i y_i x_i - b \sum_{i=1}^N a_i y_i + \sum_{i=1}^N a_i \\ &= -\frac{1}{2} (w^T w) + \sum_{i=0}^N a_i \\ &= \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j x_i^T x_j \end{aligned} \quad (1.34)$$

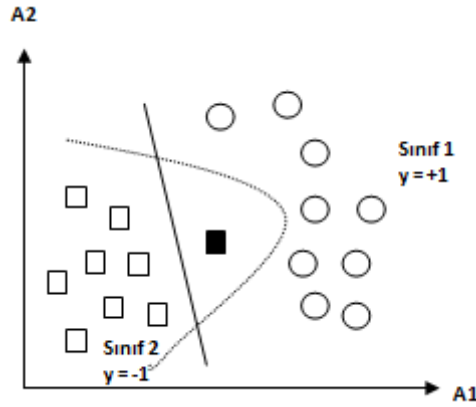
---

<sup>96</sup> C. J. C. Burges, *Tutorial on Support Vector Machines for Pattern Recognition*, Data Mining Knowledge Discovery, 2(2), 1998, pp.9.

$$a_i \geq 0, \forall i \quad (1.35)$$

ifadesi elde edilmiş olur. (1.34) probleminin (1.35) koşulları altında çözümü ikinci dereceden optimizasyon problemi ile gerçekleştirilir. Burada dikkat edilirse her eğitim örneği için bir tane Lagrange çarpanının olduğu görülür. Çözümde elde edilen Lagrange çarpanlarının büyük çoğunluğunun değeri sıfır olacaktır. Geriye kalan  $a_i > 0$  değerli  $x_i$  örnekleri DV'lerdir ve  $A_1$  ve  $A_2$  aşırıdüzlemlerinin arka taraflarında kalan örneklerdir.<sup>97</sup>

**Lineer Ayrılama Durumu:** Lineer ayrılabilme durumunda veriler iki sınıfa doğrusal bir düzlem ile ayrılabilirdi. Uygulamada bu durum her zaman geçerli olmayabilir. Yani doğrusal bir düzlem ile veriler birbirinden ayrılmayabilir.



**Şekil 1.7.** Birbirinden Doğrusal Olarak Ayrılamayan Veriler

**Kaynak:** Yalçın Özkan, Veri Madenciliği Yöntemleri, Papatya Yayıncılık, İstanbul 2008, s.195.

Eğer ki, örnekler lineer olarak tamamen ayrılabilir durumda değilse problemin çözümü için pozitif zayıflık değişkenleri,  $\xi_i$ ,  $i = 1, 2, \dots, N$  kullanılır. Aşağıdaki denklemler zayıflık değişkenleri ile yeniden tanımlanacak olursa, yeni ifadeler:

$$y_i = +1 \text{ için, } w^T \cdot x_i + b \geq +1 - \xi_i \quad (1.36)$$

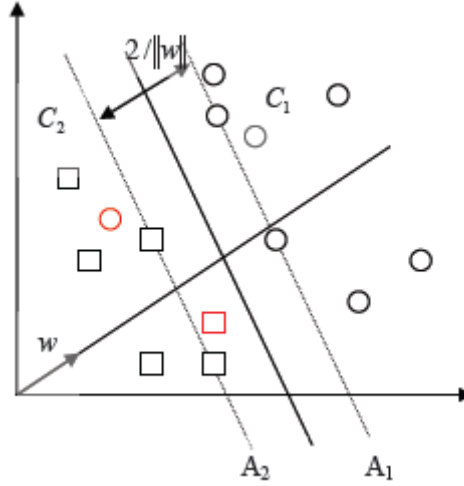
$$y_i = -1 \text{ için, } w^T \cdot x_i + b \leq -1 + \xi_i \quad (1.37)$$

$$\xi_i \geq 0, \forall i \quad (1.38)$$

<sup>97</sup> V. N. Vapnik, *The Nature of Statistical Learning Theory* [Elektronik Sürüm], Springer-Verlag, NewYork 1995, pp.135.

şeklinde olacaktır.  $\xi_i = 0$  olması durumunda  $x_i$  örneği doğru sınıflandırılmış,

$\xi_i < 0 < 1$  olması durumunda  $x$  örneği doğru sınıflandırılmış ancak  $A_1$  ve  $A_2$  aşırıdüzlemler arasında yer alıyor,  $\xi_i \geq 1$  ise yanlış sınıflandırılmış demektir.<sup>98</sup>



**Şekil 1.8.** Linear Ayrılama Durumunda Optimum Ayrıcı Aşırıdüzlem

Linear olarak ayrılama durumunda sistemin ezberlemesini, yani eğitim verisi içinde olası her durum için bir çözüm üretmemesi için sisteme bir  $C$  üst sınırı eklenir. Bu üst sınır Lagrange çarpanlarının alabilecekleri maksimum değeri göstermektedir. Bu şekilde Lagrange çarpanlarının  $0 \leq a_i \leq C$  aralığında kalması sağlamaktadır. Bu bilgilere göre Lagrange formülasyonu yeniden şu şekilde ifade edilecektir;

$$L_P = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N a_i \{y_i (w^T x_i + b) - 1 + \xi_i\} - \sum_{i=1}^N u_i \xi_i \quad (1.39)$$

Yukarıdaki formülasyonda  $u_i, \xi_i$ 'nin pozitif olmasını garanti etmek için kullanılmış olan Lagrange formülasyonunda çözülmesi zor olduğundan dolayı lineer ayrılabilir örneklerde olduğu gibi dual problemine dönüştürülmektedir. Bu problemde Karush-Kuhn-Tucher şartları uygulanırsa;

<sup>98</sup> E. Alpaydın, *Introduction to Machine Learning*, MIT, US 2004, s.221.

$$\frac{\partial L_p}{\partial w} = w - \sum_{i=1}^N a_i y_i x_i = 0 \quad (1.40)$$

$$\frac{\partial L_p}{\partial b} = - \sum_{i=1}^N a_i y_i = 0 \quad (1.41)$$

$$\frac{\partial L_p}{\partial \xi} = C - a_i - u_i = 0 \quad (1.42)$$

ifadeleri elde edilir. Bu ifadeler tekrar düzenirse:

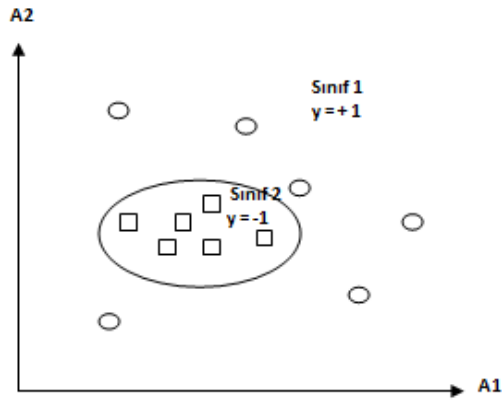
$$L_d = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j x_i^T x_j \quad (1.43)$$

$$0 \leq a \leq C, \forall_i \quad (1.44)$$

elde edilir. Bu problemin çözümünde  $0 \leq a \leq C$  aralığında yer alan Lagrange çarpanlarına karşılık gelen  $x_i$  değerleri destek vektörlerdir.<sup>99</sup>

### 1.6.5.2 Linear Olmayan Destek Vektör Makineleri

Veri kümelerinin doğrusal olarak ayıramadığı durumlar için gevşek değişkenler model koyularak çözüm arandı. Verilerin doğrusal olarak ayıramadığı durumlarda doğrusal sınıflandırıcı yerine doğrusal olmayan sınıflandırıcılar kullanılabilir.<sup>100</sup>



Şekil 1.9. Doğrusal Olmayan Sınıflandırıcılar

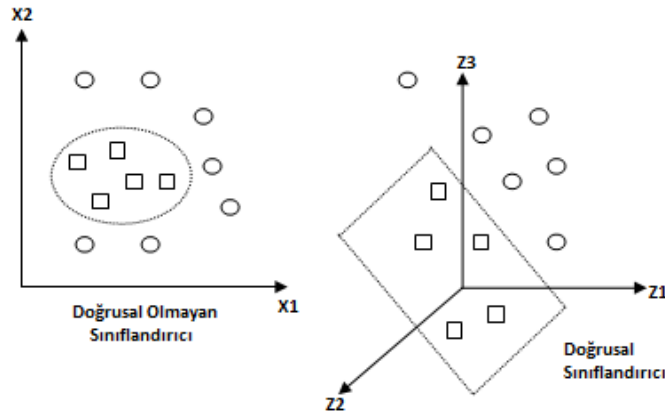
<sup>99</sup> Demirci, s.15.

<sup>100</sup> Özkan, s.199.

Yukarıdaki şekil üzerinde görüldüğü gibi iki farklı sınıfa ait verileri doğrusal olarak ayırtmak mümkün değildir. Bu durumda doğrusal olmayan sınıflandırıcıların kullanılması gerekmektedir.

**Doğrusal Olmayan Özellik Uzayı:**  $x \in R^n$  gözlem vektörünü daha yüksek dereceden bir uzayda  $z$  vektörüne dönüştürerek, bu yeni uzayda doğrusal sınıflandırıcıları elde etmek söz konusu olabilir. Bu  $z$  vektörünün yer aldığı özellik uzayı  $F$  ile gösterilsin. Bu durumda  $\emptyset$  ifadesi  $R^n \rightarrow R^F$  eşlemesini yapmak üzere  $z = \emptyset(x)$  biçiminde ifade edilebilir.

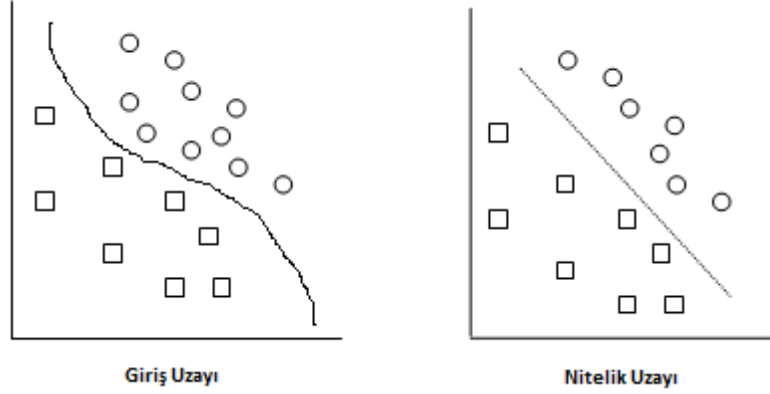
$$x \in R^n \rightarrow z(x) = [a_1, \emptyset_1(x), \dots \dots a_n, \emptyset_n(x)]^T \in R^F \quad (1.45)$$



**Şekil 1.10.** İki Boyutlu Uzaydaki Verilerin Üç Boyutlu Uzaydaki Verilere Dönüştürülmesi ve bir doğrusal Hiper Düzlem İle Sınıflandırılması

**Çekirdek Düzenlemesi ve Çekirdek Fonksiyonları:** Çekirdek fonksiyonlarının (kernel functions) dayandığı temel fikir, bir takım bileşenlerin belirli bir kurala göre dönüştürülmesidir. Bu dönüşüm aslında giriş uzayında da gerçekleştirilebilecek bazı işlemlerin potansiyel çok boyutlu bir uzayda gerçekleştirilmesi üzerine kurulmuştur. Böylece çekirdek fonksiyonlarını kullanan metotlar giriş uzayında karmaşık temsillerde sunulabilen birçok uygulama üzerinde de çalışabilmektedirler. Doğrusal olmayan ayrılabilirlik durumu düşünüldüğü zaman, eğitim örnekleri orijinal giriş uzayında doğrusal olarak ayrılamazlar. Bu tip durumlarda DVM, doğrusal olmayan haritalama fonksiyonu yardımıyla orijinal giriş uzayından doğrusal olarak kolayca sınıflandırma

yapabileceği yüksek boyutlu nitelik uzayına dönüşüm yapar. Bu yaklaşım Şekil 1.11’ de gösterilmektedir.<sup>101</sup>



**Şekil 1.11.** Yüksek Dereceli Çekirdek Fonksiyonu ile Nitelik Uzayına Dönüşüm

Çekirdek fonksiyonu kullanmanın temel avantajı, bütün değerlerin tekrar tekrar çarpım değerlerinin hesaplanarak bulunması yerine doğrudan çekirdek fonksiyonunda değerlerin yerine koyularak nitelik uzayındaki değerlerinin bulunmasıdır. Bu sayede, son derece yüksek boyutlu bir nitelik uzayı ile uğraşma olasılığı kalmaz. Seçilen çekirdek fonksiyonu ile DVM sonsuz boyutlu bir uzayda dahi oluşturulup çalıştırılabilir. Çekirdek fonksiyonlarının diğer bir avantajı da, eğitime aşamasında bir eğitim örneği için fonksiyon kurulup değerler bulunduktan sonra diğer örnekler için artık kalıp değerleri eğitim örneği dışında tamamen hazır olduğu için çok daha kolay hesaplanmasıdır.<sup>102</sup>

DVM yaygın olarak kullanılan dört çekirdek fonksiyonu vardır. Bu fonksiyonlar:<sup>103</sup>

- 1- Doğrusal Fonksiyon,
- 2- Polinomial Fonksiyon
- 3- Sigmoid Fonksiyon,
- 4- Radyal Tabanlı Fonksiyon

<sup>101</sup> Çomak, s.31.

<sup>102</sup> V. Kecman, *Learning and Soft Computing: Support Vector Machines. Neural Networks and Fuzzy Logic Models* [Elektronik Sürüm], MIT Press, Cambridge- Massachusetts 2001, s.169.

<sup>103</sup> Ü. Aydoğan, *Destek Vektör Makinelerinde Kullanılan Çekirdek Fonksiyonların Sınıflama Performanslarının Karşılaştırılması*, (Yayımlanmamış Yüksek Lisans Tezi), Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara 2010, s.40.

**Doğrusal Fonksiyon:** Girdi uzayında veriler doğrusal olarak ayrılabilir ise veriyi yüksek boyutta taşımaksızın doğrusal çekirdek fonksiyonu yardımıyla sınıflama işlemi yapılır. Bu fonksiyon herhangi bir boyut değeri ya da katsayı içermemektedir.

$$K(x_i, x_j) = x_i^T x_j \quad (1.46)$$

**Polinomial Fonksiyon:** Polinomial çekirdek fonksiyon,  $d$  gibi belirli bir derecede girdi vektörlerinin iç çarpımından oluşmaktadır. Fonksiyonun matematiksel gösterimi aşağıdaki şekildedir.

$$K(x_i, x_j) = (x_i x_j)^d \quad (1.47)$$

$d = 1$  olduğu durumlarda polinomial fonksiyon doğrusal fonksiyona dönüşmektedir.

**Sigmoid Fonksiyon:** Sigmoid fonksiyon  $k$  ve  $\delta$  gibi iki parametre içermektedir. Kaynaklar belirli parametreler için sigmoid radyal tabanlı fonksiyon çalıştığını göstermektedir.

$$K(x_i, x_j) = \tanh(kx_i - \delta) \quad (1.48)$$

**Radyal Tabanlı Fonksiyon:** Radyal tabanlı fonksiyon çekirdek fonksiyonlar arasında kullanımı en yaygın çekirdek fonksiyonudur. R programında sistem standart çıktılarını radyal tabanlı fonksiyona göre vermektedir.  $\gamma$  yarıçap kontrolünü sağlayan parametredir.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (1.49)$$



## İKİNCİ BÖLÜM

### YAPAY SİNİR AĞLARI

Yapay sinir ağlarının en genel tanımı: “İnsan beyni ve biyolojik sinir sisteminin işleyiş ve öğrenme mekanizmasını, elektronik ortamlar üzerinde taklit etmeye çalışan sistemlerdir” şeklinde yapılmaktadır.

Aşağıda, literatürde karşılan birkaç tanımlamaya yer verilmektedir.

Bir yapay sinir ağı, birbirlerine paralel olarak çalışan, birçok basit işlem elemanından oluşan ve fonksiyonu, ağın yapısı, bağlantı ağırlıkları ve elemanlarda gerçekleştirilen işlemler tarafından belirlenen bir sistemdir.<sup>104</sup>

Yapay sinir ağı, basit hesaplama elemanlarının yoğun bağlantılarından meydana gelmiş bir bilgi işleme elemanıdır.<sup>105</sup>

Yapay sinir ağları insan idrakinin ve biyolojik sinirlerinin matematiksel modelinin geliştirilmesi olarak tanımlanabilir.<sup>106</sup>

Bir sinir ağı, basit işlem birimlerinden oluşan, deneysel bilgileri biriktirmeye yönelik doğal bir eğilimi olan ve bunların kullanılmasını sağlayan yoğun bir şekilde paralel dağıtılmış bir işlemcidir.<sup>107</sup>

Bu işlemcinin insan beyni ile ilişkisi iki şekilde açıklanabilir;

- a) İşlemci tarafından öğrenilen bilgi, belirli bir öğrenme işlemi sayesinde edinilir.
- b) İşlemciyi oluşturan yapay nöronlar arasındaki bağlantı ağırlıkları, edinilen bilginin saklanması için kullanılır.

Yapay sinir ağları, insan beyninin özelliklerinden olan, öğrenme yolu ile yeni bilgiler türetebilme, yeni bilgiler oluşturabilme ve keşfedebilme gibi yetenekleri

<sup>104</sup> Bernard Widrow, *DARPA Neural Network Study* [Elektronik Sürüm], AFCEA International Press, USA 1988, s. 60, ISBN 0-916159-17-5.

<sup>105</sup> P.K. Simpson, *Neural Networks Theory, Technology and Applications* [Elektronik Sürüm], The Institute of Electrical and Electronics Engineers Press, New York 1990, s.1.

<sup>106</sup> L.V. Fausett, *Fundamentals of Neural Networks* [Elektronik Sürüm], Prentice Hall, Canada 1993, 3.

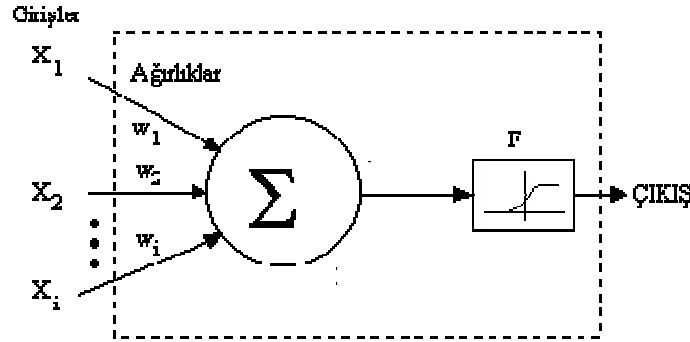
<sup>107</sup> S. Haykin, *Neural Networks: A Comprehensive Foundation* [Elektronik Sürüm], Prentice Hall Inc, New Jersey 1999, s.2.

herhangi bir yardım almadan otomatik olarak gerçekleştirmek amacıyla geliştirilen bilgisayar sistemleridir.<sup>108</sup>

YSA kavramı beynin çalışma ilkelerinin sayısal bilgisayarlar üzerinde taklit edilmesi fikri olarak ortaya çıkmış ve ilk çalışmalar beyni oluşturan biyolojik hücrelerin, ya da literatürdeki ismi ile nöronların matematiksel olarak modellenmesi üzerinde yoğunlaşır.<sup>109</sup>

Bu bilgiler ışığı altında; yapay sinir ağları, insan beyni ve biyolojik sinir sistemini bilgisayarlar üzerinde taklit etmeye çalışan; öğrenme, öğrendiklerini hafızasında saklama ve genelleme yapabilme gibi insani özellikleri taklit edebilen ve birçok temsili sinir hücresinin, birbirine bağlantı ağırlıkları ile bağlanması suretiyle oluşturulan bir sistem şeklinde ifade edilebilir.

## 2.1. YAPAY SİNİR AĞLARININ YAPISI VE TEMEL BİLEŞENLERİ



Şekil 2.1. Yapay Sinir Hücresinin Yapısı

Biyolojik sinir ağlarının hücreleri olduğu gibi yapay sinir ağlarının da hücreleri vardır. Yapay sinir ağı hücresi beş elemandan oluşur. Bunlar:

### 2.1.1. Girdiler

Yapay sinir hücresine (proses) elemanına dış dünyadan gelen bilgilerdir. Bilgiler dış dünyadan geldiği gibi başka hücrelerden olduğu gibi kendi kendisinden de gelebilir.

<sup>108</sup> E. Öztemel, *Yapay Sinir Ağları*, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.29.

<sup>109</sup> Z. Şen, *Yapay Sinir Ağları İlkeleri*, Su Vakfı Yayınları, İstanbul 2004, s.7.

### 2.1.2. Ağırlıklar

Ağırlıklar hücreye gelen bilginin hücre üzerindeki etki ve önemini gösterir. Ağırlığın küçük ve büyük olması önemli olup olmadığı anlamına gelmez.

### 2.1.3. Toplam Fonksiyonu

Bu fonksiyon hücreye gelen net girdiyi hesaplar. Toplam fonksiyon için farklı fonksiyonlar kullanılır. En yaygın olarak ağırlıklı toplam fonksiyonu kullanılır. Bu fonksiyon aşağıdaki şekilde formülize edilir.

$$\text{Net} = \sum G_i * A_i \quad (i= 1,2,\dots,n) \quad (2.1)$$

Burada G girdileri, A ağırlıkları, n ise girdi (proses) eleman sayısını göstermektedir. Bir proses için en uygun toplama fonksiyonunu belirleyecek bir formül yoktur.

Her proses elemanı bağımsız olarak farklı bir toplam fonksiyonuna sahip olabilecekleri gibi aynı toplama fonksiyonuna da sahip olabilirler.

➤ Çarpım :  $NET : \prod(G_i * A_i)$  (2.2)

➤ Maksimum :  $NET : \text{Max}(G_i * A_i)$  (2.3)

➤ Minimum :  $NET : \text{Min}(G_i * A_i)$  (2.4)

➤ Çoğunluk :  $NET = \sum \text{sign}(G_i * A_i)$  (2.5)

➤ Kümülatif Toplam :  $NET = NET_{\text{eski}} + \sum(G_i * A_i)$  (2.6)

### 2.1.4. Aktivasyon Fonksiyonu

Nöron davranışını belirleyen önemli faktörlerden biri nöronun aktivasyon fonksiyonudur. Toplama fonksiyonunun sonuçları aktivasyon fonksiyonu diye bilinen bir süreç yardımı ile çıktıya dönüştürülür. Literatürde, sıkıştırma veya eşik fonksiyonu olarak da adlandırılmaktadır.<sup>110</sup>

Bu fonksiyon, hücreye gelen net girdiyi işleyerek hücrenin bu girdiye karşılık üreteceği çıktıyı belirler. Toplama fonksiyonunda olduğu gibi aktivasyon fonksiyonu olarak da çıktıyı hesaplamak içinde değişik formüller kullanılmaktadır. Bazı modeller (mesela çok katmanlı algılayıcı) bu fonksiyonun türevinin alınabilir bir fonksiyon olmasını şart koşmaktadır. Toplama fonksiyonunda olduğu gibi aktivasyon

<sup>110</sup> G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting With Artificial Neural Networks: The State Of The Art" [Yapay Sinir Ağlarıyla Tahmin: ], *International Journal Of Forecasting*, 14(1) , 1998, pp. 35, Erişim Tarihi: 20 Aralık 2011, ScienceDirect.

fonksiyonunda da ađın proses elemanlarının hepsinin aynı fonksiyonu kullanması gerekmez. Bazı elemanlar aynı fonksiyonu diđerleri farklı fonksiyonları kullanabilirler. Bir problem için en uygun fonksiyonda yine tasarımcının denemeleri sonucunda belirleyebileceđi bir durumdur. Uygun fonksiyonu gösteren bir formül bulunmuş deđildir.<sup>111</sup>

Dođrusal olmayan fonksiyonların kullanılması yapay sinir ađlarının çok karmaşık ve farklı problemlere uygulanmasını sađlamıştır. En çok kullanılan aktivasyon fonksiyonları şunlardır.<sup>112</sup>

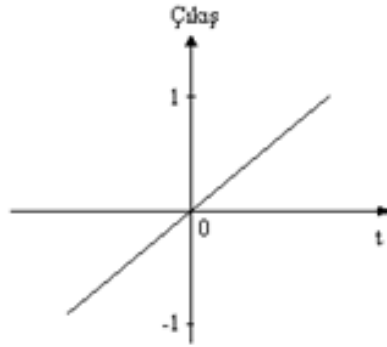
#### 2.1.4.1. Dođrusal Fonksiyon

Dođrusal fonksiyon, hücreye gelen girişleri olduđu gibi çıkışa iletir. Dođrusal fonksiyon Şekil 2.2.'de gösterilmiştir. Çođunlukla ADALINE olarak adlandırılan dođrusal işlemci eleman, klasik işaret işleme ve istatistiksel regresyon analizinde kullanılır. Formülü ise;

$$Net\ Toplam = v = \sum_i^n x_i w_i \text{ veya } Net\ Toplam = v = \sum_i^n x_i w_i + \theta \quad (2.7)$$

olmak üzere;  $F(NET) = y = F(v) = Av$

Formüldeki A sabit bir sayıdır.



Şekil 2.2. Dođrusal veya Lineer Fonksiyon

<sup>111</sup> Öztemel, s.50.

<sup>112</sup> Öztemel, ss.50-51.

- Ş. Sađırođlu, E. Beşdok, M. Erler, *Mühendislikte Yapay Zeka Uygulamaları I: Yapay Sinir Ađları*, Ufuk Kitap Kıratsiye-Yayıncılık, Kayseri 2003, ss.37-40.

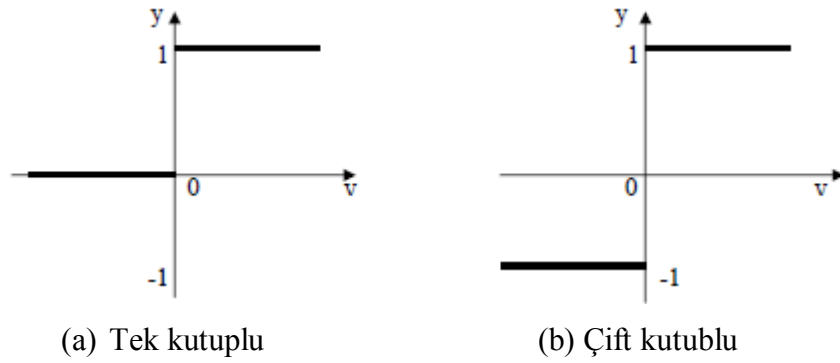
- Haykin, ss.10-12.

### 2.1.4.2. Basamak Fonksiyonu

Basamak fonksiyonu tek veya çift kutuplu fonksiyon olabilir. Bu tip fonksiyonların şekli Şekil 2.3'de, matematiksel ifadeleri gösterilmiştir. Perceptron (Basit Algılayıcı Model) olarak bilinen işlemci elemanları bu fonksiyon ile işlem görür.

$$F(NE\text{T}) = y = F(v) = \left\{ \begin{array}{l} 1 \quad v \geq 0 \\ 0 \quad v < 0 \end{array} \right\} \quad (2.8)$$

$$F(NE\text{T}) = y = F(v) = \left\{ \begin{array}{l} +1 \quad v \geq 0 \\ -1 \quad v < 0 \end{array} \right\} \quad (2.9)$$

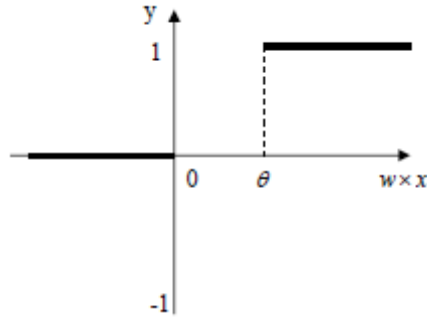


Şekil 2.3. Basamak Fonksiyonları

### 2.1.4.3. Kutuplamalı Basamak Fonksiyonu

Kutuplama değeri, tek kutuplu ve çift kutuplu basamak fonksiyonlarının her ikisinde de eklenmiş olabilir. Aktivasyon fonksiyonu eşik değeri olan  $\theta$ 'yı aştığı zaman nöron aktiftir denir. Tek kutuplamalı basamak fonksiyonunun matematiksel ifadesi aşağıda verilmiştir. Tek kutuplamalı basamak fonksiyonunun şekli ise Şekil 2.4.'de gösterilmiştir.

$$F(NE\text{T}) = y = F(v) = \left\{ \begin{array}{l} 1 \quad w \times x \geq \theta \\ 0 \quad w \times x < \theta \end{array} \right\} \quad (2.10)$$

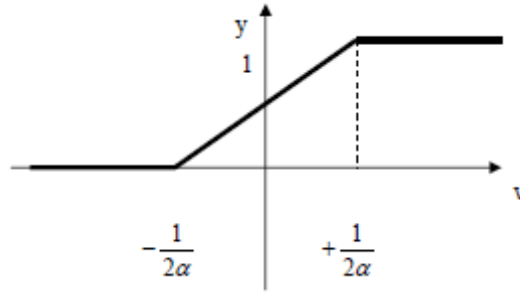


Şekil 2.4. Tek Kutuplamalı Basamak Fonksiyonu

#### 2.1.4.4. Parçalı Doğrusal Fonksiyon

Bu fonksiyon küçük aktivasyon potansiyeli için,  $\alpha$  kazancı olan bir doğrusal toplayıcı (ADALINE) olarak çalışır. Büyük aktivasyon potansiyeli için, nöron doyuma ulaşır ve çıkış işareti 1 olur. Büyük kazançlar için  $\alpha \rightarrow \infty$  iken, parçalı doğrusal fonksiyon basamak gibi davranır. Parçalı doğrusal fonksiyonun matematiksel ifadesi aşağıda verilmiş olup, Şekil 2.5.'de gösterilmiştir.

$$F(NE\text{T}) = y = F(v) = \left\{ \begin{array}{ll} 0 & v \leq -1/2\alpha \\ \alpha v + \frac{1}{2} & |v| < \frac{1}{2\alpha} \\ 1 & v \geq 1/2\alpha \end{array} \right\} \quad (2.11)$$

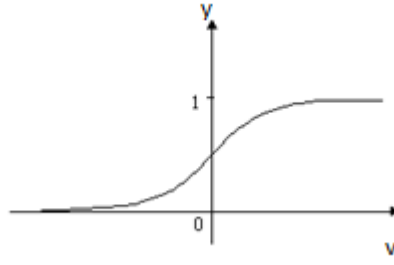


Şekil 2.5. Parçalı Doğrusal Fonksiyon

#### 2.1.4.5. Sigmoid Tipi Fonksiyon

Yapay sinir ağı uygulamalarında en çok kullanılan aktivasyon fonksiyonudur. Bu fonksiyonun en aktif bölgesinin 0,2 ile 0,8 arasında olduğu bilinmektedir. Literatürde tek kutuplu aktivasyon fonksiyonu olarak da bilinir. Sigmoid tipi fonksiyonun matematiksel ifadesi aşağıda verilmiş olup, Şekil 2.6.'da gösterilmiştir.

$$F(NE\text{T}) = y = F(v) = \frac{1}{1+e^{-v}} = \frac{1}{2} \left[ \tanh\left(\frac{v}{2}\right) + 1 \right] \quad (2.12)$$

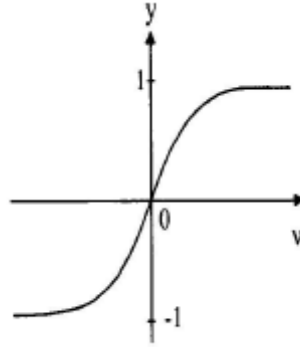


**Şekil 2.6.** Sigmoid Fonksiyonu

#### 2.1.4.6. Tanjant Hiperbolik Tipli Fonksiyon

Yapay sinir ağı uygulamalarında sık kullanılan fonksiyon tiplerinden biridir. Literatürde çift kutuplu fonksiyon olarak da bilinir. Tanjant hiperbolik tipli fonksiyonun matematiksel ifadesi aşağıda verilmiş olup, Şekil 2.7.'de gösterilmiştir.

$$F(NET) = y = F(v) = \frac{1-e^{-2v}}{1+e^{2v}} = \tanh(\beta v) \quad (2.13)$$

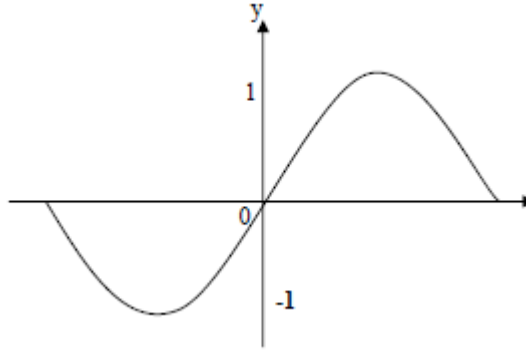


**Şekil 2.7.** Tanjant Hiperbolik Fonksiyonu

#### 2.1.4.7. Sinüs Tipli Fonksiyon

Öğrenilmesi düşünülen olaylar sinüs fonksiyonuna uygun dağılım gösteriyorsa bu gibi durumlarda sinüs fonksiyon kullanılır. Bu fonksiyonun matematiksel ifadesi ve şekli aşağıda verilmiştir.

$$F(NET) = y = F(v) = \text{Sin}(v) \quad (2.14)$$



**Şekil 2.8.** Sinüs Tipli Fonksiyon

### 2.1.5. Hücrenin Çıktısı

Aktivasyon fonksiyonu tarafından belirlenen çıktı değeridir. Üretilen çıktı dış dünyaya veya başka bir hücreye gönderilir. Hücre kendi çıktısını kendisine girdi olarak da gönderebilir. Bir proses elemanının birden fazla çıktısı olmasına rağmen sadece bir çıktısı olmaktadır. Ağ şeklinde gösterildiğinde bir proses elemanının birden fazla çıktısı varmış gibi görülmektedir. Bu sadece gösterim amacıyladır. Aslında bir proses elemanından çıkan tek bir çıktı değeri vardır. Aynı değer birden fazla proses elemanına girdi olarak gitmektedir.<sup>113</sup>

## 2.2. YAPAY SİNİR AĞLARININ YAPISI

Yapay sinir ağları, yapay sinir hücrelerinin katmanlar şeklinde bağlanmasıyla oluşturulan veri tabanlı sistemler olup insan beyninin öğrenme ve değişik koşullar altında çok hızlı karar verebilme gibi yeteneklerinin, basitleştirilmiş modeller yardımıyla karmaşık problemlerin çözülmesinde kullanılmasını amaçlamaktadır.<sup>114</sup>

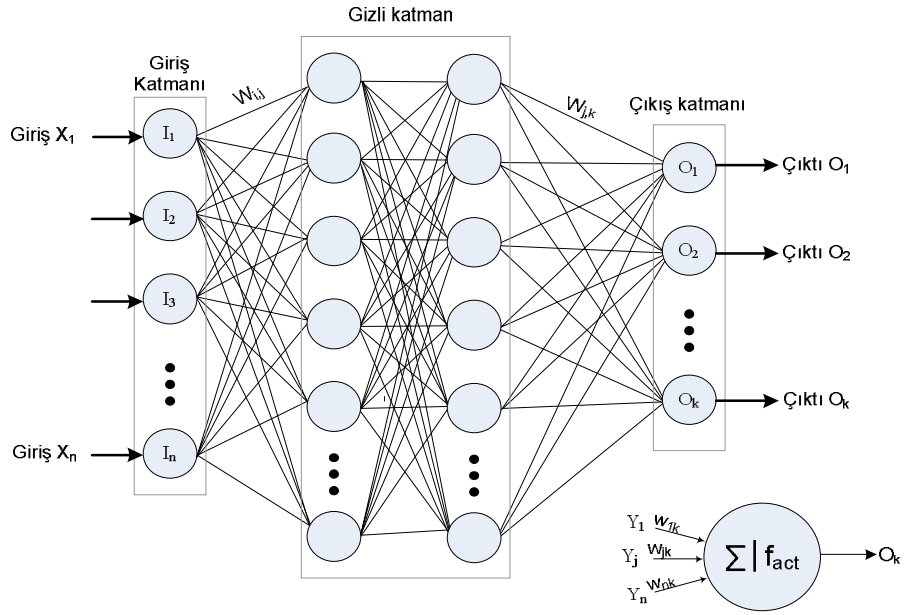
Yapay sinir ağlarında, yapay nöronlar basit bir şekilde kümelendirilmektedirler. Bu kümelendirme katmanlar halinde yapılmaktadır ve daha sonra bu katmanlar bir diğerine ilişkilendirilmektedir. Temel olarak, tüm yapay sinir ağları benzer bir yapıya sahiptirler. Böyle bir genel yapı Şekil 2.9.'da gösterilmiştir. Bu yapıda, bazı nöronlar girdileri almak için bazı nöronlar ise çıktıları iletmek için dış mekan ile bağlantılı

<sup>113</sup> Öztemel, s.51.

<sup>114</sup> M. L. Koç, C. E. Balas, A. Arslan, "Taş Dolgu Dalgakıranların Yapay Sinir Ağları İle Ön Tasarımı", *İMO Teknik Dergi*, 15(225), 2004, s. 3351, Erişim Tarihi: 12 Aralık 2011, [www.imo.org.tr/resimler/dosya\\_ekler](http://www.imo.org.tr/resimler/dosya_ekler)



haldedirler. Geri kalan tüm nöronlar ise gizli katmanlardadırlar, yani sadece ağ içinde bağlantıları vardır.<sup>115</sup>



**Şekil 2.9.** Bir Yapay Sinir Ağının Genel Yapısı

**Kaynak:** M. Bilgehan and P. Turgut, “The Use of Neural Networks In Concrete Compressive Strength Estimation”, [Beton Mukavemet Tahmininde Yapay Sinir Ağları Kullanımı], Computers and Concrete, Vol 7, No.3, 2010, pp. 275.

Bir yapay sinir ağının yapısı ve sinir hücrelerinin sayısı değişiklik göstermelerine rağmen, yapay sinir ağının oluşumu için kabul görmüş herhangi bir kural bulunmamaktadır. Gerekli gizli katman sayısından az gizli katmana sahip yapay sinir ağları komplike fonksiyonların çözümünde yetersiz kalırken, çok fazla gizli katmana sahip yapay sinir ağları ise istenmeyen kararsızlıklarla karşılaşmaktadır. Gizli katman sayısı belirlendikten sonra karşılaşılan problem ise her bir tabakada kaç tane nöronun yer alacağına karar vermede karşımıza çıkmaktadır. Girdi katmanı için bir sorun bulunmamaktadır; bu sayı sistem içerisindeki girdilerin sayısına eşittir. Aynı şekilde, çıktı katmanı da istenilen çıktı sayısı ile belirlenebilmektedir. Esas sorun, gizli katmanlarda nöron sayısını belirlemektir. Geleneksel matris algoritması, matris boyutlarının ya girdi sayısına ya da çıktı sayısına eşit olması gerektiğini söylemektedir. Ne yazık ki, gizli katmanda en verimli şekilde kaç tane nöronun bulunacağı konusunda

<sup>115</sup> D. Anderson and G. McNeill, *Artificial Neural Networks Technology*, Kaman Science Corporation, USA New York 1992, s.7.

herhangi bir matematiksel test bulunmamaktadır. Deneme ve yanılma yöntemi uygulanarak karar verilmelidir.<sup>116</sup>

### 2.3. YAPAY SİNİR AĞLARINDA ÖĞRENME

Yapay sinir ağları zaman içinde öğrenebilmektedirler. Bu bakımdan adaptif özellik taşırlar. Yani, sinir ağları geçmiş deneyimlerine dayanarak problem çözme yeteneklerini geliştirebilir. Yapay sinir ağlarındaki bu olay “öğrenme” olarak adlandırılır. Öğrenme işlemi istenilen çıktıları elde edebilmek amacıyla bağlantıların ağırlıklarının belirlenmesidir.

Öğrenmenin gerçekleşebilmesi için, yapay nöronlar arasındaki sinaptik ağırlıkların en uygun bir biçimde ayarlanması gerektiği ifade edilmiştir. Yapay sinir ağları klasik programlama gibi belirli bir algoritma çerçevesinde programlanmazlar.<sup>117</sup> Sinir ağları insanlar gibi örnekler ile eğitilirler. Yapay nöronlar, mevcut örnek kümesi üzerinde girdi ile çıktı arasındaki sinaptik ağırlıkların değiştirilmesiyle eğitilirler. Sunulan girdi kümesi için; aktivasyon tarafından sağlanan değerlere cevap olarak, bağlantı ağırlıklarının tamamının veya bir kısmının istenen çıktı ile ağ çıktısı arasındaki farkın ( $\epsilon_i$ ) belirli bir değere düşünceye kadar değiştirilmesidir.<sup>118</sup>

Bir sinir ağında öğrenmenin anlamı, ağına belirli bir probleme ait doğru çıktıları üretmesini sağlayacak optimum ağırlık değerlerinin bulunmasıdır. Bilgi, ağ boyunca bağlantılarda ağırlıklar şeklinde dağıtıldığı için tek bir bağlantı herhangi bir anlamlı bilgiyi ifade etmez. Daha doğrusu, anlamlı bir bilgi oluşturmak için işlem elemanları olan bir grup bağlantıyı tasarlamak gerekmektedir. Problemin çözümü için ağına, bağlantılarına ait doğru ağırlık değerlerine sahip olması gerekmektedir. Bu öğrenme veya eğitme olarak adlandırılan bir işlem vasıtasıyla gerçekleştirilir. Öğrenme, ağırlık değerlerinin nasıl değiştirilmesi gerektiğini ifade eden bir öğrenme kuralına dayanır. Geliştirilen birçok öğrenme kuralı vardır.<sup>119</sup>

<sup>116</sup> K. B. Detienne, D. H. Detienne and S.A. Joshi, “Neural Networks As Statistical Tools For Business Researchers”, [İşletmeciler Araştırmacıları için İstatistiksel Araç Olarak Yapay Sinir Ağları], *Organizational Research Methods*, 6(2), 2003, pp. 242-243, Erişim Tarihi: 18 Ocak 2011, ScienceDirect.

<sup>117</sup> Fu Li Min, *Neural Networks in Computer Intelligence* [Elektronik Sürüm], 2nd Edition, McGraw Hill, New York 1994, s.1114.

<sup>118</sup> Ömer Civalık ve Yusuf Calayır, “İnce Dikdörtgen Plakların Titreşim Frekanslarının Yapay Sinir Ağları ile Tahmini”, *İMO Teknik Dergi*, Yazı: 275, 2007, s.4164.

<sup>119</sup> P.D. Wasserman, *Advanced Methods in Neural Computing* [Elektronik Sürüm], Van Nostrand Reinhold, New York 1993, s.340.

McNeil yapay sinir ağlarının geribeslemeli ağlar, squasher fonksiyonları, radial basis fonksiyonu, ridgelet ağları, jump bağlantıları, çok katmanlı geri beslemeli ağlar ve recurrent ağlar şeklinde sınıflandırılabilceğini belirtmiştir.

### 2.3.1. Yapay Sinir Ağlarında Öğrenme Stratejileri

Yapay sinir ağlarında, işlemci elemanların hedeflenen çıktılara ulaşmak için bağlantıları ne şekilde değiştirecekleri öğrenme algoritması tarafından önceden tanımlanır. Tanımlanan öğrenme kuralına göre, system öngörü hatasını minimize edecek ağırlıkları deneyerek tespit eder. Bu bakış açısıyla, yapay sinir ağlarını kullanılan öğrenme kuralına göre sınıflandırmak mümkündür.

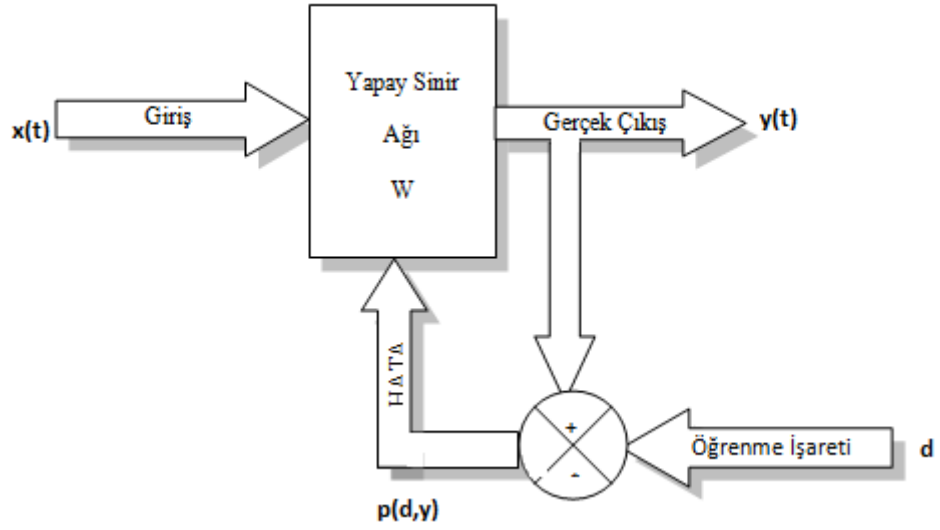
#### 2.3.1.1. Danışmanlı Öğrenme

Yapay sinir ağlarında yaygın olarak kullanılan bir öğrenme yöntemi olan danışmanlı öğrenmede, ağa örnek bir çıktı verilir ve bu çıktılarla ağın ürettiği çıktı değerlerinin karşılaştırılması amaçlanır. Bu öğrenmede, ağırlıklar rastgele olarak verilir ve ağ tarafından hata miktarı minimize edilerek döngüsel işlemler gerçekleştirilir.<sup>120</sup>

Bu öğrenme yönteminde, tahmin edilmek istenen durum gerçekleşene kadar döngü devam ettirilmeli, eğitim kümesi dışında oluşturulan bir test kümesi tarafından da ağın eğitiminin doğruluğu test edilerek kontrolü yapılmalıdır. Eğer ağın eğitilmesi başarı ile sonuçlanmamış ise test kümesinde de yanlış sonuçlar üretilmesi kuvvetle muhtemel olacaktır. Genelleştirilmiş delta kuralı, rastsal öğrenme kuralı ve geri beslemeli algoritmalar bu öğrenme metotlarına örnek olarak verilebilir.<sup>121</sup>

<sup>120</sup> Anderson and McNeill, s.10.

<sup>121</sup> H. Küçükönder, *Yapay Sinir Ağları ve Tarımda Bir Uygulama*, (Yayımlanmamış Doktora Tezi), Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş 2011, s.55.



**Şekil 2.10.** Danışmanlı Öğrenme Yapısı

**Kaynak:** Ş. Sağıroğlu, E. Beşdok ve M. Erler, Mühendislikte Yapay Zeka Uygulamaları I: Yapay Sinir Ağları, Ufuk Kitap Kıratsiye-Yayıncılık, Kayseri 2003, s.81.

Şekil 2.10.'da danışmanlı öğrenme yapısı gösterilmektedir. Bu öğrenme yönteminde öğrenmeye dışarıdan müdahale eden bir öğretmen, danışman vardır. Öğrenme danışmanın kontrolündedir. Danışman, eğitim kümesini ve hata değerini belirleyerek eğitimin ne kadar devam edeceğine karar verir. Bu yöntemin en önemli özelliği eğitim esnasında gerçek giriş, çıkış değerlerin kullanılmasıdır.<sup>122</sup>

### 2.3.1.2. Danışmansız Öğrenme

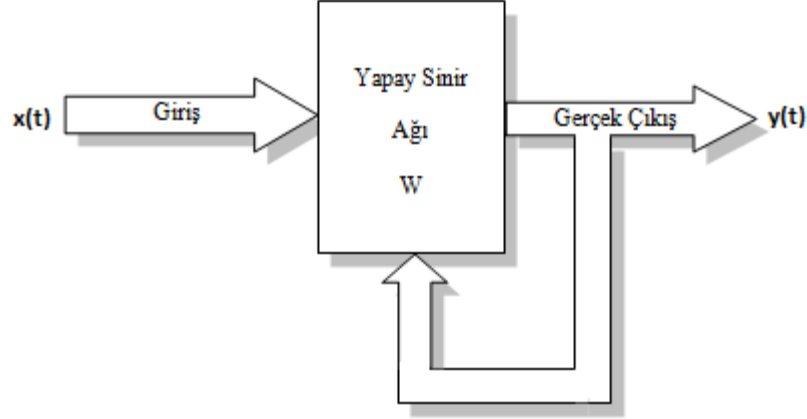
Bu öğrenme türünde sistemin öğrenmesine yardımcı olan herhangi bir danışman yoktur. Sisteme sadece girdi değerleri gösterilir. Örneklerdeki parametreler arasındaki ilişkileri sistemin kendi kendisine öğrenmesi beklenir. Bu, daha çok sınıflandırma problemleri için kullanılan bir öğrenme yöntemidir. Yalnız, sistemin öğrenmesi bittikten sonra çıktılarının ne anlama geldiğini gösteren etiketlendirmenin kullanıcı tarafından yapılması gerekmektedir.<sup>123</sup>

Danışmansız öğrenmede ağlar, performanslarını dışarıdan herhangi bir etkiye ihtiyaç duymadan içeriden gözlemledirler. Ağa bir verinin doğru veya yanlış olup olmadığı belirtilmeden, ağ kendisi bu durumu nasıl belirleyebileceği konusunda

<sup>122</sup> Şen, s.99.

<sup>123</sup> Öztemel, s.25.

birtakım bilgilere sahip olabilmektedir. Bu bilgi, ağ topolojisinin ve öğrenme kuralları içerisinde mevcuttur.<sup>124</sup> Şekil 2.11’de danışmansız öğrenme yapısı gösterilmiştir. ART (Adaptive Resonance Theory) ve SOM (Self Organizing Map) öğrenme kuralı danışmansız öğrenmeye örnek olarak verilebilir.



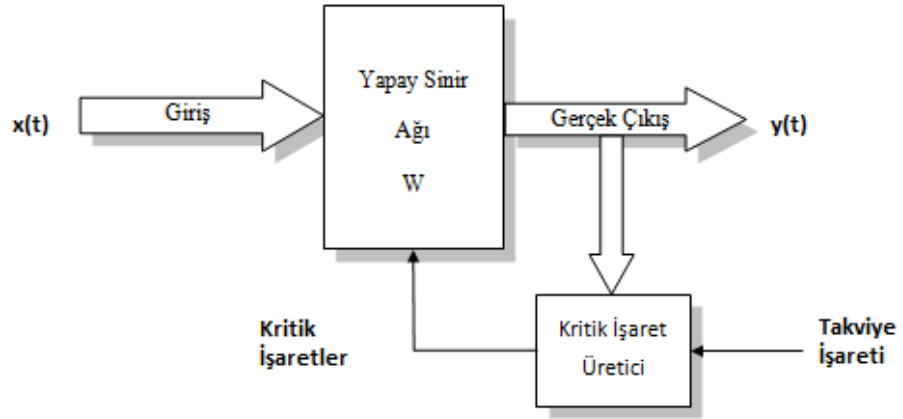
**Şekil 2.11.** Danışmansız Öğrenme Yapısı

**Kaynak:** Ş. Sağıroğlu, E. Beşdok ve M. Erler, Mühendislikte Yapay Zeka Uygulamaları I: Yapay Sinir Ağları, Ufuk Kitap Kıratsiye-Yayıncılık, Kayseri 2003, s.81.

### 2.3.1.3. Takviyeli Öğrenme (Reinforcement Learning)

Bu öğrenme kuralı danışmanlı öğrenmeye yakın bir metottur. Denetimsiz öğrenme algoritması, istenilen çıkışın bilinmesine gerek duymaz. Hedef çıktıyı vermek için bir “öğretmen” yerine, burada yapay sinir ağlarına bir çıkış verilmemekte fakat elde edilen çıkışın verilen girişe karşılık iyiliğini değerlendiren bir kriter kullanılmaktadır. Şekil 2.12.’de takviyeli öğrenme yapısı gösterilmiştir. Optimizasyon problemlerini çözmek için Hinton ve Sejnowski’nin geliştirdiği Boltzmann kuralı veya Genetik Algoritmalar takviyeli öğrenmeye örnek olarak verilebilir.

<sup>124</sup> Küçükönder, s.56.



**Şekil 2.12.** Takviyeli Öğrenme Yapısı

**Kaynak:** S. Burmaoğlu, Birleşmiş Milletler Kalkınma Programı Beşeri Kalkınma Endeksi Verilerini Kullanarak Diskriminant Analizi, Lojistik Regresyon Analizi ve Yapay Sinir Ağlarının Sınıflandırma Başarılarının Değerlendirilmesi, (Yayımlanmamış Doktora Tezi), Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Erzurum 2009, s.65.

Öğrenme algoritmalarını ağ yapısı, öğrenme algoritması ve kullanıldığı yer olmak üzere üç kategoriye göre sınıflandırılabilir.<sup>125</sup>

**Çizelge 2.1.** Öğrenme algoritmalarının incelenmesi

Öğrenme Yöntemi	Öğrenme Kuralı	Ağın Mimarisi	Öğrenme Algoritması	Kullanıldığı Yerler
Öğretmenli	Hata Düzeltme	Tek ya da çok Katmanlı	Perceptron Geriye Yayılım Adaline ve Madaline	Örüntü tanıma Tahmin Kontrol
	Boltzmann	Geri Dönüşlü	Boltzmann öğrenmesi	Örüntü sınıflandırma
	Hebbian	Çok Katmanlı İleri beslemeli	Lineer diskriminant	Veri analizi Sınıflandırma
Öğretmenli	Rekabetli	Rekabetli	Vektör kuantalama öğrenmesi	Veri sıkıştırma Sınıflandırma
		ART ağı	ART Haritalaması	Örüntü Sınıflandırma Sınıflandırma

<sup>125</sup> A.K. Jain and J. Mao, "Artificial Neural Networks: A Tutorial" [Yapay Sinir Ağları: Bir Eğitimi], *IEEE Computer Special Issue on Neural Computing*, IEEE Computer Society, 1996, pp. 38, Erişim Tarihi: 26 Kasım 2011, IEEEExplore.

Çizelge 2.1 Devamı

	Hata Düzeltme	Çok katmanlı İleri beslemeli	Sammon projeksiyonu	Veri Analizi
Öğretmensiz	Hebbian	İleri beslemeli Rekabetçi	Principal Component Analizi	Veri Analizi Veri Sıkıştırma
		Hopfield Ağı	Associative Memory Öğrenmesi	Associative Memory
	Rekabetli	Rekabetli	Vektör Kuantalması	Sınıflandırma Veri Sıkıştırma
		Kohonen SOM	Kohonen SOM	Sınıflandırma Veri Analizi
		ART ağı	ART1, ART 2	Sınıflandırma
Karma	Hata Düzeltme ve Rekabetli	Radyal Tabanlı	Radyal Tabanlı Öğrenme	Sınıflandırma Örüntü Tanıma Fonksiyon Yaklaşımı Tahmin Kontrol

**Kaynak:** A.K. Jain and J. Mao "Artificial Neural Networks: A Tutorial" [Yapay Sinir Ağları: Bir Eğitimi], IEEE Computer Special Issue on Neural Computing, IEEE Computer Society, 1996, pp. 38. Erişim Tarihi: 26 Kasım 2011, IEEEExplore.

### 2.3.2. Yapay Sinir Ağlarında Öğrenme Kuralları

Yapay sinir ağları gibi sistemlerde öğrenme bazı kurallara göre gerçekleşir. Bu kuralların bazıları çevrimiçi (on-line), bazıları çevrimdışı (off-line) çalışmaktadır.<sup>126</sup>

#### 2.3.2.1. Kohonen Öğrenme Kuralı

Kohonen kuralı, ağın kendi kendisini eğitebildiği Kohonen tabakasındaki süreç elemanlarının birbiri ile yarış halinde olması ilkesine dayanan bir kuraldır. Bu kural Kohonen (1982) tarafından geliştirilmiştir.

Bu kurala göre hücreler ağırlıklarını değiştirmek (öğrenmek) için birbirleri ile yarışır. Yarışma girdi vektörü ile ağırlık vektörleri (referans vektörleri) arasındaki öklid mesafesinin hesaplanmasına dayanır. Hangi proses elemanın referans vektörü girdi vektörüne en yakın ise yarışmayı o kazanır. Öğrenme sırasında sadece girdi katmanına bağlı ağırlık değiştirilir. Kazanan proses elemanı için iki durum söz konusudur:

<sup>126</sup> Öztemel, s.26.

Kazanan proses elemanı doğru sınıfın bir üyesidir. Bu durumda aynı proses elemanının tekrar kazanması için ilgili ağırlıklar girdi vektörüne biraz daha yaklaştırılırlar.  $\beta$  öğrenme katsayısı, A ağırlık vektörü, X girdi vektörü olmak üzere ağırlıkların değiştirilmesi formülü şu şekilde ifade edilir:

$$A_{\text{yeni}} = A_{\text{eski}} + \beta (X - A_{\text{eski}}) \quad (2.15)$$

Girdi vektörünün referans vektörüne çok yaklaşınca durması ve aksi yönde uzaklaşmaması için  $\beta$  öğrenme katsayısı sıfır değerine kadar zamanla azaltılır. Öğrenme katsayısı ne çok erken ne de çok geç sifira indirgenmemelidir. Öğrenmenin ne zaman durdurulacağına eğitim süreci ayrıntılı bir şekilde incelenerek karar verilir.

Kazanan proses elemanı yanlış sınıfın bir üyesidir. Bu durumda aynı proses elemanı tekrar kazanmaması için ilgili ağırlıklar girdi vektöründen biraz daha uzaklaştırılırlar.  $\beta$  öğrenme katsayısı, A ağırlık vektörü, X girdi vektörü olmak üzere ağırlıkların değiştirilmesi formülü şu şekilde ifade edilir:

$$A_{\text{yeni}} = A_{\text{eski}} + \beta (X - A_{\text{eski}}) \quad (2.16)$$

### 2.3.2.2. Hebb Öğrenme Kuralı

Bu kural, Hebb (1949) tarafından geliştirilen bir öğrenme kuralı olup, biyolojik öğrenmeyi esas almaktadır. Bu sinir ağında bağlantılar arası ağırlık değişimlerinin hesaplanmasında korelasyon kuralı kullanılmaktadır.<sup>127</sup>

Günümüzde kullanılan sinir ağının öncülerinden Donald Hebb, beynin öğrenme stratejilerini anlamaya yönelik yapmış olduğu çalışmalarında, sinir hücresinin yapısını ve işleyişini temel almıştır. İki sinir hücresinin birbirleriyle olan durumunu çözümlenmeye yönelik olarak da bu doğrultuda gerçekleştirilen Hebb kuralı, aynı zamanda iki sinir hücresinin aynı anda aktif olması durumunda bağlantıların güçleştiği bir kural olmuştur. Bu nedenle bu kuralda giriş yapıları oldukça önemli olup, başarılı

<sup>127</sup> N. Baş, *Yapay Sinir Ağları Yaklaşımı ve Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2006, s.35.



sonuç üretme bu yapılara bağlıdır. Giriş yapılarının dikey ve birbirleriyle ilişkisiz olduğu durumda başarılı sonuçlara ulaşabilmektedir.<sup>128</sup>

Hebb öğrenme kuralına göre öğrenme sinyali hücre çıkışına eşittir. Yani bir hücre aktif ise bağlı olduğu hücreyi aktif, pasif ise bağlı olduğu hücreyi pasif yapmaya çalışmaktadır.

$$\text{Öğrenme sinyali} = f(w_i^t x) \quad (2.17)$$

ile belirtilir.

Ağırlık vektöründeki artış ise;

$$\Delta w_i = c f(w_i^t x) x \quad (2.18)$$

ile veya

$$\Delta w_{ij} = c o_i x_j \quad (2.19)$$

ile belirtilir.

Bu öğrenme kuralında başlangıç olarak sıfıra yakın değerler ile başlanılmalıdır. Hebb öğrenme kuralına göre giriş ve çıkış değerlerinin çarpımı pozitif ise  $w_{ij}$  ağırlığında bir artış, negatif ise bir azalış gerçekleşir.<sup>129</sup>

### 2.3.2.3. Hopfield Öğrenme Kuralı

Bu kural Hebb Kuralına benzerdir. Yapay sinir ağlarının bağlantılarının ne kadar zayıflatılması veya ne kadar kuvvetlendirilmesi gerektiği belirlenir. Eğer istenilen çıkış ve girişin her ikisi aktif veya her ikisi de aktif değilse öğrenme oranı tarafından bağlantı ağırlığı arttırılır. Diğer durumlarda ise azaltılır. Ağırlıkların arttırılması veya azaltılması öğrenme katsayısı ile gerçekleştirilir. Genellikle bu katsayı 0 ile 1 arasında değerler almaktadır.<sup>130</sup>

<sup>128</sup> M. Özşahin, *Türkiye Otomotiv Sektörünün Kendini Örgütleyen Haritalar İle Finansal Analizi*, (Yayımlanmamış Yüksek Lisans Tezi), Çukurova Üniv., Fen Bilimleri Enstitüsü Endüstri Mühendisliği ABD, Adana 2009, s.32.

<sup>129</sup> F. Adıyaman, *Talep Tahmininde Yapay Sinir Ağlarının Kullanılması*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Teknik Üniversitesi Fen Bilimler Enstitüsü, İstanbul 2007, s.12.

<sup>130</sup> A.A. Hopgood, *Intelligent Systems for Engineers and Scientists* [Elektronik Sürüm], CRC Press, 2nd Edition, Florida 2001, s.207.

#### 2.3.2.4. Delta Kuralı (Windrow - Hoff Kuralı)

Hebb kuralının geliştirilmiş bir hali olan bu kural Widrow ve Hoff tarafından geliştirilmiştir.<sup>131</sup> Delta kuralı, hedeflenen çıktı ile yapay sinir ağından elde edilen çıktı arasındaki farkı (delta) azaltmak için ağırlıkların sürekli olarak değiştirilerek en iyi sonuca ulaşması varsayımına dayanır.

Bu kural hedef çıktı ile gerçekleşen çıktı arasındaki hata kareler ortalamasının minimize edilmesi olarak da bilinmektedir. Bu kurulda minimize edilme işlemi hatanın geriye dağıtılma işlemi ile gerçekleşir. Ağ'da hata bir önceki katmana göre çoğaltılarak, bu işlemin ilk katmana doğru ilerlemesi ile sonlanır. Delta Kuralı iki safhadan oluşur:

**a. İleri Doğru Hesaplama Safhası: Ağın çıktısının hesaplanması safhasıdır.**

**b. Geriye Doğru Hesaplama Safhası: Ağırlıkları değiştirme safhasıdır.**

##### a. İleri Doğru Hesaplama Safhası:

Bilgi işleme eğitim setindeki bir örneğin Girdi Katmanından (G1, G2....) ağa gösterilmesi ile başlar. Gelen girdiler olduğu gibi ara katmana gönderilir. Girdi katmanının k. Proses elemanın çıktısı  $C_k^i$ ,

$$C_k^i = G_k \quad (2.20)$$

k. girdi elemanın j. ara katmanına bağlayan bağlantının ağırlık değeri  $A_{kj}$ ,

Ara katmandaki proses elemanlarına net girdi  $NET_j^a$

$$NET_j^a = \sum A_{kj} C_k^i \quad (k=1, \dots, n) \quad (2.21)$$

j. eleman çıktısı türevi alınabilir bir fonksiyonla (genellikle sigmoid fonksiyon) hesaplanır.

$B_j$  ara katmanda bulunan j. elemana bağlanan eşik değer elemanın ağırlığı,

Sigmoid fonksiyonunun kullanılması halinde çıktı;

$$C_j^a = 1 / [ 1 + e^{-(NET_j^a + B_j)} ] \quad (2.22)$$

---

<sup>131</sup> Baş, s.36.

Eşik değer ünitesinin çıktısı sabit olup 1'e eşittir. Ağırlık fonksiyonu sigmoid fonksiyonunun oryantasyonunu belirler ve eğitim esnasında ağ bu değeri kendisi belirler.

Ara katmanın bütün proses elemanları ve çıktı katmanının proses elemanlarının çıktıları aynı şekilde kendilerine gelen NET girdinin hesaplanması ve sigmoid fonksiyonundan geçirilmesi sonucu belirlenirler. Çıktı katmanının çıktıları (Ç1, Ç2, ...) bulununca ağın ileri hesaplama işlemi tamamlanmış olur.

### **b. Geriye Doğru Hesaplama Safhası:**

Ağın çıktıları (Ç1, Ç2.....), beklenen çıktılarla (B1, B2.....) ile karşılaştırılır. Aradaki fark hata kabul edilir. Amaç bu hatayı düşürmektir. Geriye hesaplamada bu hata ağın ağırlık değerlerine dağıtılarak bir sonraki iterasyonda hatanın azaltılması sağlanır. Çıktı katmanındaki m. proses elemanı için oluşan hata (Em),

$$E_m = B_m - \text{Ç}_m \quad (2.23)$$

Bu bir proses elemanı için oluşan hatadır. Çıktı katmanı için oluşan toplam hatayı (TH) bulmak için bütün hataların toplanması gerekir. Bazı hata değerleri negatif olacağından toplamın sıfır olmasını önlemek amacı ile ağırlıkların kareleri hesaplanarak sonucun karekökü alınır. Ağın eğitilmesindeki amaç bu hatayı azaltmaktır.

$$TH = \frac{1}{2} \sum E_m^2 \quad (2.24)$$

Toplam hatayı minimum yapmak için hatanın kendisine neden olan proses elemanlarına dağıtılarak ağırlıklar değiştirilir. Bunun için;

Ara katman ile çıktı katmanı arasındaki ağırlıklar değiştirilir.

Ara katmanlar arası veya ara katman girdi katmanı arasındaki ağırlıklar değiştirilir.

Ara katman ile çıktı katmanı arasındaki ağırlıkların değiştirilmesi işleminde, Ara katmandaki j. Proses elemanını çıktı katmanındaki m. proses elemanına bağlayan bağlantının ağırlığındaki değişim miktarı  $\Delta A_{aj}$ 'nin, t. iterasyondaki ağırlığın değişim miktarı  $\Delta A_{ajm}(t)$ ;

m. çıktı ünitesinin hatası  $\delta_m$ ,

Öğrenme katsayısı (ağırlıkların değişim miktarı)  $\lambda$ ,

Momentum katsayısı  $\alpha$  olmak üzere,

$$\Delta A_{jma}(t) = \lambda \delta_m \zeta_{ja} + \alpha \Delta A_{jma}(t-1) \quad (2.25)$$

olur.

Momentum katsayısı ağın öğrenme sırasında yerel bir optimuma takılıp kalmaması için ağırlık değişim değerinin belirli bir oranda sonraki değişime eklenmesini sağlarlar.

Aktivasyon fonksiyonunun türevi  $f'(NET)$  olmak üzere ve Sigmoid fonksiyonunun kullanılması durumunda;

$$\delta_m = f'(NET) \cdot E_m, \quad (2.26)$$

$$\delta_m = \zeta_m(1-\zeta_m) \cdot E_m \quad (2.27)$$

olur.

t. iterasyondaki ağırlıkların yeni değerleri;

$$A_{jma}(t) = A_{jma}(t-1) + \Delta A_{ajm}(t) \quad (2.28)$$

olur.

Benzer şekilde eşik değer ünitesinin de ağırlıklarını değiştirmek gerekir. Çıktı katmanındaki proses elemanlarının eşik değer ağırlıkları  $\beta_\zeta$ ; bu ünitenin çıktısının sabit ve 1 olması nedeni ile değişim miktarı;

$$\Delta \beta_{m\zeta}(t) = \lambda \delta_m + \alpha \Delta \beta_{\zeta m}(t-1) \quad (2.29)$$

olur.

t. iterasyondaki ağırlıkların yeni değeri ise;

$$\beta_{m\zeta}(t) = \beta_{m\zeta}(t-1) + \Delta \beta_{m\zeta}(t-1) \quad (2.30)$$

olur.

Ara katmanların arası veya ara katman ile girdi katmanı arasındaki ağırlıklar değiştirilmesi:

Ağırlıklardaki değişim miktarı  $\Delta A_i$ ,

Hata terimi  $\delta_a$  olma üzere,

$$\Delta A_{kji}(t) = \lambda \delta_{ja} \zeta_{ki} + \alpha \Delta A_{kji}(t-1) \quad (2.31)$$

$$\delta_{ja} = f'(NET) \sum \delta_m A_{jm}^a \quad (2.32)$$

Aktivasyon fonksiyonu  $f(NET)$ , sigmoid fonksiyon düşünülürse;

$$\delta_{ja} = \zeta_{ja} (1 - \zeta_{ja}) \sum \delta_m A_{jm}^a \quad (2.33)$$

olur.

Ağırlıkların yeni değerleri ise;

$$A_{kji}(t) = A_{kji}(t-1) + \Delta A_{kji}(t-1) \quad (2.34)$$

olur.

Benzer şekilde eşik değer ünitesinin de ağırlıklarını değiştirmek gerekir. Ara katman eşik değer ağırlıkları  $\beta_a$ ,

$$\Delta \beta_{ja}(t) = \lambda \delta_m + \alpha \Delta \beta_{ja}(t-1) \quad (2.35)$$

olur.

t. iterasyondaki ağırlıkların yeni değeri ise,

$$\beta_{ja}(t) = \beta_{ja}(t-1) + \Delta \beta_{ja}(t-1) \quad (2.36)$$

olur.

Böylece bir iterasyonun ileri ve geri hesaplamaları tamamlanmış olur. İkinci bir örnek verilerek sonraki iterasyona başlanır ve aynı işlemler öğrenme tamamlanıncaya kadar yinelenir.<sup>132</sup>

### 2.3.2.5. Basit Algılayıcı (Perseptron) Öğrenme Kuralı

Basit algılayıcılar (perseptron) ağlarında kullanılır. Bu ağlarda sadece girdi ve çıktı katmanı vardır.

Ağa Girilen Değerler	X
Çıktı Değeri	Ç

<sup>132</sup> Öztemel, s.80.

Beklenen Değerler	B
Girdilerin Ağırlıkları	W
Eşik Değer	$\Phi$
Öğrenme Katsayısı	$\lambda$ ;

Ağa girdi seti ve beklenen çıktı gösterilir. Girdi değerleri birden fazla olabilir ( $x_1, x_2, \dots, x_n$ ). Çıktı değeri 1 veya 0 değerinden birisini alır.

Perseptron ünitesine gelen net girdi;

$$\text{NET} = \sum w_i x_i \quad (2.37)$$

olur.

Perseptron ünitesinin çıktısı ise;

$$\begin{aligned} \text{Ç} = & \quad 1 & \quad \text{Eğer } \text{NET} > \Phi \\ & \quad 0 & \quad \text{Eğer } \text{NET} < \Phi \end{aligned} \quad (2.38)$$

olur.

Çıktı ile beklenen çıktı aynı olursa herhangi bir değişiklik olmaz. Çıktı ile beklenen çıktı farklı ise ağırlık değerleri;

$$\begin{aligned} W_n = W_o - \lambda X & \quad \text{Eğer } B = 0 \text{ ve } \text{Ç} = 1 \\ & \quad \text{Eğer } B = 1 \text{ ve } \text{Ç} = 0 \end{aligned} \quad (2.39)$$

olur. Bu adımlar bütün girdi setindeki örnekler için sınıflandırma yapılmaya kadar devam eder.<sup>133</sup>

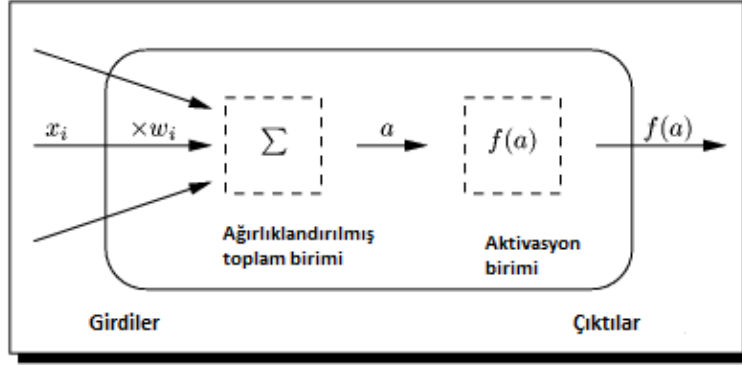
## 2.4. YAPAY SİNİR AĞI MODELLERİ

Bir yapay sinir ağında proses elemanlarının bağlanması sonucu oluşan topoloji, proses elemanlarının sahip oldukları toplama ve aktivasyon fonksiyonları, öğrenme stratejisi ve kullanılan öğrenme kurallarına göre çeşitli ağ modelleri vardır.

### 2.4.1. Tek katmanlı Algılayıcılar

Sadece girdi ve çıktı katmanından oluşan ağlardır. Her ağın bir veya daha fazla girdisi ve çıktısı vardır. Çıktı üniteleri (Ç) bütün girdi ünitelerine (X) bağlanmaktadır ve her bağlantının bir ağırlığı (W) vardır. Bu ağlarda proses elemanlarının değerinin sıfırdan farklı olmasını önleyen eşik değeri vardır ve daima 1'dir.

<sup>133</sup> Öztemel, s.63.



**Şekil 2.13.** Tek Katmanlı Algılayıcı Yapısı

**Kaynak:** R.M. Hristev, The ANN Book [Elektronik Sürüm], GNU Public License, Edition 1, Boston USA 1998, pp.4, "ftp://ftp.funet.fi/pub/sci/neural/books/".

Tek katmanlı yapay sinir ağlarında ağırlıklandırılmış giriş değerleri eşik değeri ( $\theta$ ) ile toplanarak aktivasyon fonksiyonundan geçerek çıktıya ulaşır. Bu durum matematiksel olarak,

$$Y = F\left(\sum_{i=1}^m W_i X_i + \theta\right) \quad (2.40)$$

şeklinde formülize edilir. Burada,

$W_i$  : i'ninci girdi değerinin ağırlığını,

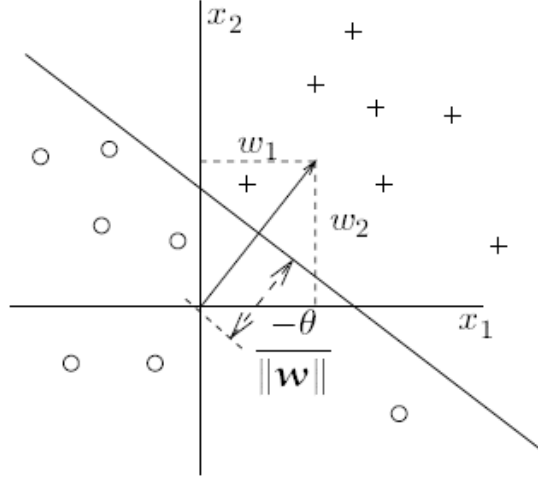
$\theta$  : Eşik değerini,

$Y$  : Çıktı değerini

göstermektedir. Bu ağ yapısında çıktı fonksiyonu doğrusal olup, ağa gösterilen örnekler iki sınıf arasında dağıtılarak, bu iki sınıfı birbirlerinden doğrusal olarak ayıran doğrunun bulunması hedeflenir. Bu nedenle de eşik değeri ( $\theta$ ) fonksiyonu kullanılır. Ağ çıktısı +1 ile -1 arasındadır.

$$F(s) = \begin{cases} 1, & s > 0 \\ -1, & \text{diğer durumlarda} \end{cases} \quad (2.41)$$

Ağ çıktısı + 1 değerini almış ise 1. sınıfta, -1 değerini almış ise de 2. sınıfta kabul edilerek gereken ayırım yapılır.<sup>134</sup> Ağırlıkların ve sınıfların birbirinden ayrımının geometrik gösterimi ise Şekil 2.14.'de gösterilmiştir.



**Şekil 2.14.** Ağırlıkları ve sınıfları birbirinden ayıran doğrunun geometrik gösterimi

**Kaynak:** B. Kröse and P. Van Der Smagt, An Introduction to Neural Networks [Elektronik Sürüm], The University of Amsterdam Press, Amsterdam 1996, s.24.

Bu doğrunun çizilebilmesi için kullanılan matematiksel denklem ise iki sınıfa ayırt etmek için,

$$W_1X_1 + W_2X_2 + \theta = 0 \quad (2.42)$$

şeklinde verilmektedir. Bu eşitlikten  $X_1$  ve  $X_2$  değerleri ise,

$$X_1 = \frac{W_2X_2 - \theta}{W_1} \quad (2.43)$$

$$X_2 = \frac{W_1X_1 - \theta}{W_2} \quad (2.44)$$

şeklinde hesaplanarak doğrunun çizimi yapılır.<sup>135</sup>

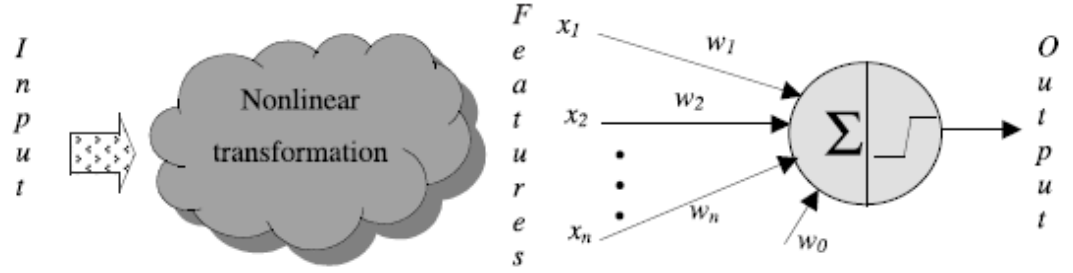
<sup>134</sup> Küçükönder, s.62.

<sup>135</sup> Küçükönder, s.63.



### 2.4.1.1. Basit Algılayıcı (Perseptron) Modeli

Rosenblat'ın (1958) yılında sınıflandırma amacı ile "Perceptron'u" geliştirmiştir. Perceptron, beyin işlevini modelliyebilmek amacıyla yapılan çalışmalar neticesinde ortaya çıkan, tek katmanlı, eğitilebilen ve tek çıkışa sahip olan bir yapay sinir ağıdır.<sup>136</sup>



**Şekil 2.15.** Bir Perceptronun Yapay Sinir Modeli

**Kaynak:** Y.H. Hu and J.N. Hwang, Handbook of Neural Network Signal Processing [Elektronik Sürüm], CRC Press, Florida 2002, pp.4.

Perceptron algoritması dört aşamada gerçekleşmekte olup, bu aşamalar sırasıyla Çizelge 2.2.'deki gibi özetlemek mümkündür.

**Çizelge 2.2.** Perceptron algoritmasının işleyişi

Aşamalar	Yapılacak olan işlemler
1.Aşama	Yapay sinir ağına girdi ve çıktı değerleri tanıtılır. Birden fazla girdi değerinin mevcut olduğu durumlarda buna karşılık gelen çıktı değeri 0 veya 1 değerinden birini alacaktır.
2.Aşama	NET girdi, $S = \sum_{i=1}^m W_i X_i$ Şeklinde hesaplanır. Burada, $W_i$ : i.ninci ağırlık değerini, $X_i$ : i.ninci girdi değerini, göstermektedir.

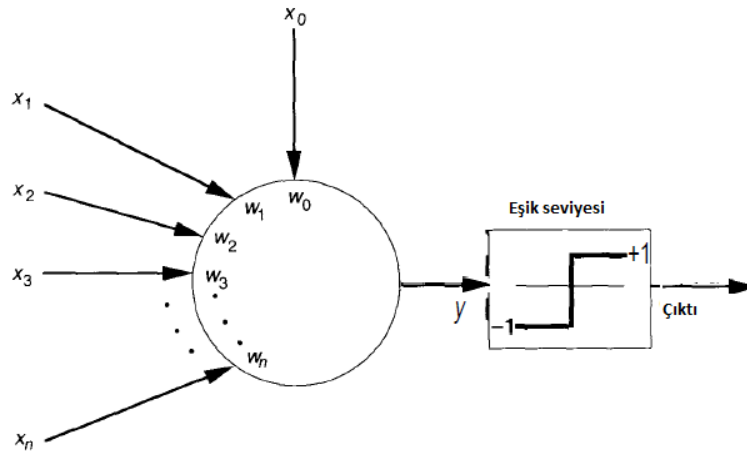
<sup>136</sup> Öztemel, s.27.

Çizelge 2.2. Devamı

<b>3.Aşama</b>	NET girdi değeri (S) eşik değeri ile, $Y = \begin{cases} 1, & S > \theta \\ 0, & S \leq \theta \end{cases}$ şeklinde mukayesesi yapılır. Gerçekleşen çıktı değeri ile beklenen çıktı değeri, aynı değeri alırsa ağırlıklarda değişiklik yapılmaz aksi durumlarda ise ağırlık değerleri duruma göre artırılır veya azaltılır.
<b>4.Aşama</b>	Bu aşamada, girdi setindeki örnekler için doğru sınıflandırma işlemi optimum oluncaya kadar iterasyon devam edilir. Doğru sınıflandırma yapılıncaya da işlem durdurulur.

### 2.4.1.2. ADALINE / MADALINE Modeli

ADALINE Modeli, Adaptif Doğrusal Eleman'ın (Adaptive Linear Element) kısaltılması şeklinde olup, bu model tek bir işlemci elemandan oluşan bir yapıya sahiptir. Eğitim algoritmasının bu modeli, Widrow ve Hoff (1959) tarafından en küçük kareler yöntemi (least mean square) olarak sunulmuş, ayrıca delta kuralı olarak da ifade edilmiştir.<sup>137</sup>



Şekil 2.16. Bir ADALINE Yapısı

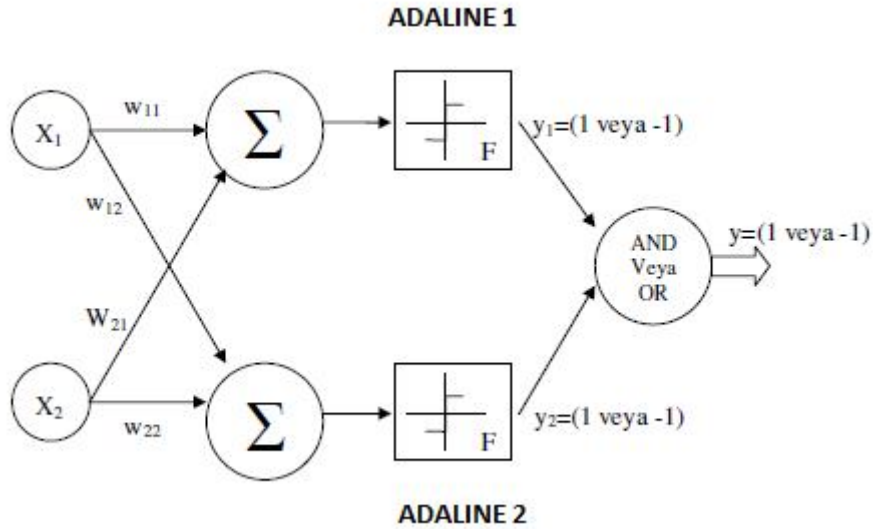
**Kaynak:** J.A. Freeman and D.M. Skapura, Computation and Neural Systems Series [Elektronik Sürüm], Addison-Wesley Publishing Company, USA 1991, pp.56.

<sup>137</sup> B. Kröse and P. Van Der Smagt, *An Introduction to Neural Networks* [Elektronik Sürüm], The University of Amsterdam Press, Amsterdam 1996, s.27.

En küçük kareler kuralı olarak da bilinen Delta kuralı gereğince, hata fonksiyonu minimize edilirken ağırlıklarda kademeli azaltma yöntemi kullanılarak, kısmi türevlerin negatif (-) değerlerinin kullanılmasıyla hatanın minimizasyon işlemi gerçekleştirilir.

ADALINE modeli, aktivasyon fonksiyonlarından Adım (Step) fonksiyonunun kullanıldığı ve çıktı fonksiyonunun değerinin 0'dan küçük olduğunda (-1) değerini, büyük olduğunda ise (+1) değerini aldığı bir modeldir.

MADALINE ağları birden fazla ADALINE ünitesinin bir araya gelerek oluşturdukları ağa verilen isimdir. MADALINE ağları genel olarak iki katmandan oluşmaktadır. Her katmanda değişik sayıda adaline ünitesi bulunmaktadır. Ağın çıktısı da yine 1 ve -1 değerleri ile gösterilmektedir.



**Şekil 2.17.** Bir MADALINE Yapısı

**Kaynak:** E Öztemel, Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.73.

Şekil 2.17.'deki MADALINE ağı, iki adet ADALINE hücresinin birleştirilmesiyle oluşturulmuş bir ağıdır. MADALINE, ADALINE 1 ve ADALINE 2'den gelen çıktıları, AND / OR mantıksal işlemcileri sayesinde birleştirerek, dış dünyaya nihai çıktısını vermektedir. Burada;

ADALINE 1'e gelen ağırlıklandırılmış (sinaptik ağırlıklar ile çarpılarak toplanmış) veri;

$$v_1 = \sum_{j=1}^2 w_{j,1}x_j + \theta_1 \quad (2.45)$$

ADALINE 2'ye gelen ağırlıklandırılmış veri;

$$v_2 = \sum_{j=1}^2 w_{j,2}x_j + \theta_2 \quad (2.46)$$

$\theta_1$  ve  $\theta_2$  ADALINE 1 VE ADALINE 2'nin eşik değerlerini ifade etmektedir.

ADALINE 1'in aktivasyon fonksiyonu

$$y_1 = f(v_1) \begin{cases} 1, & v_1 \geq 0 \text{ ise} \\ 0, & v_1 < 0 \text{ ise} \end{cases} \quad (2.47)$$

ADALINE 2'nin aktivasyon fonksiyonu

$$y_2 = f(v_2) \begin{cases} 1, & v_2 \geq 0 \text{ ise} \\ 0, & v_2 < 0 \text{ ise} \end{cases} \quad (2.48)$$

ADALINE1;  $y_1$  çıktısını, ADALINE 2;  $y_2$  çıktısını oluşturacak ve değerlendirilmesi için MADALINE yorumcusuna gönderecektir. MADALINE ise; kendi üzerine gelen  $y_1$  ve  $y_2$  çıktılarını 0 veya 1 çıktısı olarak iletacaktır.

#### 2.4.2. Çok Katmanlı Algılayıcı

Çok katmanlı algılayıcılar giriş ve çıkış katmanları arasında birden fazla katmanın kullanıldığı YSA sistemleridir. Bu katmanlar her iki girdi ve çıktı düğümleriyle direkt bağlantılı olmayan gizli düğümler bulundurabilirler.<sup>138</sup>

Yapay sinir ağlarında çok katmanlı perceptron, çözüme kavuşturulmak istenen problemlerde, girdi değerleri ile çıktı değerleri arasındaki ilişkinin doğrusal olmadığı durumlarda ADALINE modelinin yetersiz kalması nedeniyle geliştirilmiş bir modeldir.<sup>139</sup>

Bu model yapay sinir ağlarına olan ilgiyi bir anda artırmış ve sinir ağlarının tarihinde yeni bir döneme girilmesini sağlamıştır. Birçok öğrenme algoritmasının

<sup>138</sup> R.P. Lippmann, "An introduction to computing with neural nets" [Yapay Sinir Ağlarıyla Bir Hesaplama Giriş], *IEEE ASSP Magazine*, 4(2), 1987, pp. 15, Erişim Tarihi: 18 Aralık 2012, IEEE Journals & Magazines.

<sup>139</sup> Küçükönder, s.66.

kullanımına imkan sağlaması da bu modeli tercih edilebilir kılmıştır. Aynı zamanda, bir giriş katmanıya, bir veya birden fazla ara (gizli) katman ve bir tane de çıkış katmanından oluşan bir yapıya sahiptir. Delta kuralı olarak adlandırılan öğrenme kuralının yaygın olarak kullanılmasıyla problemleri çözüme kavuşturan bir modeldir. Delta kuralında, en küçük kareler yöntemi kullanılmakta olup amaç, ağıın beklenen çıktı değeri ile ürettiği çıktı değeri arasındaki hata miktarını azaltmaktır.

Minsky ve Papert doğrusal olmayan bir ilişki gösteren XOR probleminden yola çıkarak, günümüzde karşılaşılmaması muhtemel olan sorunların asıl çözümünde başarısız olduklarını araştırmacılara göstermiş olan öncü isimlerdendirler.<sup>140</sup>

#### 2.4.2.1. Geri Yayılma Ağları

Karmaşık verilerin sınıflandırılmasında kullanılan etkin YSA modellerinden birisi; ilk olarak Werbos tarafından düzenlenen daha sonra Parker, Rummelhart ve McClelland tarafından geliştirilen geri yayılım ağıdır (Back Propagation Network).<sup>141</sup> İlk uygulamaları yazılı metinden söz sentezi, robot kollarının kontrolüdür. Sınırları ise denetimli eğitim giriş ve çıkış örneklerinin çok sayıda olmasıdır.

“Geri yayılım” kelimesi gizli katmanlardaki hatanın, çıktı katmanlarında bulunan hatanın geriye doğru yayılmasıyla elde edilmesi sonucu ortaya çıkmıştır.<sup>142</sup> Geri yayılım öğrenme kuralı, yapay sinir ağları içerisinde hemen hemen her probleme rahatlıkla uygulanabilir bir yapıyı oluşturmaktadır. Geri beslemeli bir öğrenme mekanizması söz konusudur. Burada yapay sinir ağlarının yapısı ileri beslemeli olmasına karşın, hatanın geriye doğru yayılmasından kaynaklanan bir öğrenme kuralı olduğu için geri besleme söz konusu olmaktadır.<sup>143</sup>

Geri yayılma ağları (backpropagation networks); “katmanları arasında tam bağlantı (full connected) bulunan, ileri beslemeli (feedforward) ve denetimli (supervised) olarak eğitilen, çok katmanlı (multilayer) ağlardır.

<sup>140</sup> D. Michie, D.J. Spiegelhalter and C.C. Taylor, *Machine Learning, Neural and Statistical Classification* [Elektronik Sürüm], Ellis Horwood Publisher, London 1994, s.86.

<sup>141</sup> J.S.R. Jang, C.T. Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence* [Elektronik Sürüm], Printice Hall, Upper Saddle River NJ 1997, s.233.

<sup>142</sup> Fu, s.81.

<sup>143</sup> R.P. Lippmann, "An introduction to computing with neural nets" [Yapay Sinir Ağlarıyla Bir Hesaplama Giriş], *IEEE ASSP Magazine*, 4(2), 1987, pp. 5. Erişim Tarihi: 18 Aralık 2012, IEEE Journals & Magazines.

Geri yayılma ağlarında en sık olarak kullanılan öğrenme kuralı, genelleştirilmiş delta kuralıdır. Geri yayılma ağlarının işleyişi temel olarak iki kısma ayrılabilir.

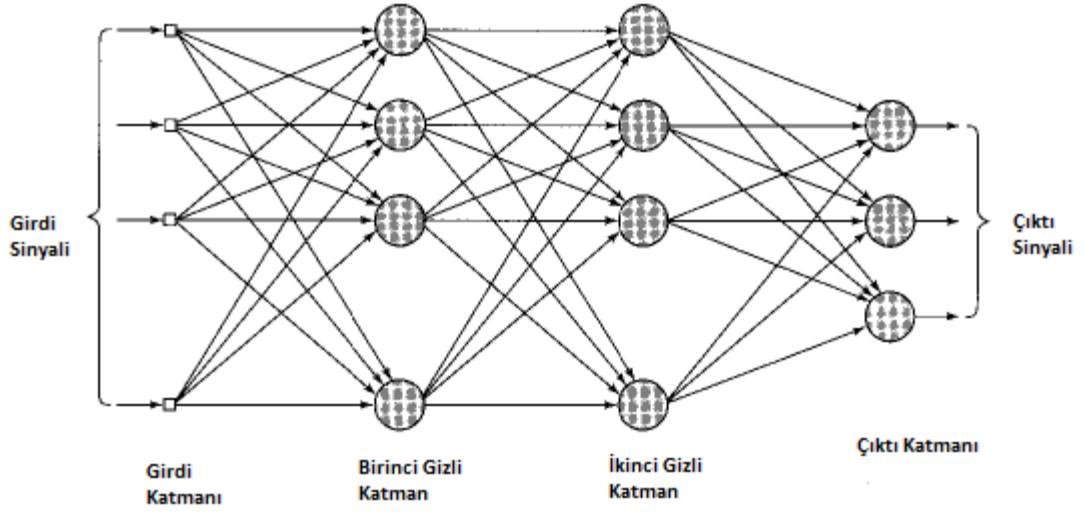
- İleri Besleme (feedforward)
- Geri Yayılma (Backpropagate)

İleri besleme aşamasında, ağın dış dünyadan aldığı girdiler, çıkış katmanına kadar ileri doğru transfer edilirken (forward), geri yayılma aşamasında ise; ağın çıktısı ( $y'$ ) ile istenen çıktı ( $y$ ) arasındaki fark ( $\epsilon_1$ ), çıkış katmanından başlayarak geriye doğru giriş katmanına kadar transfer edilir (backward). Bu hata; ağın sinaptik ağırlıklarının değiştirilerek hatanın minimize edilmesi için kullanılır.

Geri yayılma ağları hiyerarşik yapıda ağlardır. Bir katmandaki sinirler sadece kendilerinden bir sonra gelen katmana veri gönderebilirler. Bir katmanın atlanarak diğer katmana veri gönderilmesi bu tip ağlarda mümkün değildir. Giriş katmanındaki her bir sinir, kendinden sonra gelen gizli katmandaki her bir sinire, eğer birden çok gizli katman bulunuyorsa bir gizli katmandaki her bir sinir kendinden sonra gelen gizli katmandaki her bir sinire, çıkış katmanından önce gelen gizli katmandaki her bir sinir ise çıkış katmanındaki her bir sinire bağlıdır. Bir katmandaki hiçbir sinir kendi katmanındaki bir sinire bağlı olamaz. Her katmanın çıkış değeri bir sonra gelen katmanın giriş değeridir. Bu şekilde dış dünyadan alınan verinin, her katmanda aktive edilerek ağın içinde hiyerarşik bir yapıda çıkış katmanına kadar iletilmesine ileri besleme(feedforward) denir.<sup>144</sup>

---

<sup>144</sup> B. Çelik, *Yapay Sinir Ağları Metodolojisi ile Zaman Serisi: Teori ve Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul 2008, s.52.



**Şekil 2.18.** Örnek Geri Yayılma Ağı

**Kaynak:** S. Haykin Neural Networks: A Comprehensive Foundation [Elektronik Sürüm], Prentice Hall Inc, New Jersey 1999, pp.159.

#### 2.4.2.2. İleri Besleme Aşaması

Örnek bir geri yayılma ağı Şekil 2.18.'de gösterilmiştir. Şekil 2.18'de, giriş katmanındaki her bir nöron, gizli katman 1'deki her bir nörona  $w_{ij}$  sinaptik ağırlıklarıyla bağlıdır. Ayrıca; giriş katmanında bulunan hiçbir nöronun, giriş katmanından bir sonraki gizli katmana veya çıkış katmanına bağlantı bulunmamaktadır. Burada giriş katmanında bulunan nöronların çıktı değerleri, gizli katman 1'de bulunan sinirlerin girdi değerlerini oluşturmaktadır.

Bu ifadelerle göre; gizli katman 1'de bulunan  $N_{1,1}$  nöronuna ulaşan ağırlıklandırılmış net girdi,

$$v_{1,1} = \sum_{i=1}^3 x_{i,t} w_{i,1} = x_{1,t} w_{1,1} + x_{2,t} w_{2,1} + x_{3,t} w_{3,1} \quad (2.49)$$

Geri yayılma ağlarının sinaptik ağırlıklarının ( $w_{i,j}$ ,  $a_{i,j}$ ,  $b_{i,j}$ ) değerlerinin başlangıç değerleri, genellikle  $(-1,1)$  aralığında seçilen rassal değişkenler atanmaktadır.

Yine gizli katman 1'de bulunan  $N_{1,2}$  nöronuna gelen net girdi,

$$v_{1,2} = \sum_{i=1}^3 x_{i,1} w_{i,2} = x_{1,t} w_{1,2} + x_{2,t} w_{2,2} + x_{3,t} w_{3,2} \quad (2.50)$$

Bu aşamadan sonra  $N_{1,1}$  ve  $N_{1,2}$  nöronları kendilerine gelen girdileri aktive edeceklerdir. Bu örnek için seçilen aktivasyon fonksiyonları aşağıda verilmiştir.

$$y_{i,j} = f(v_{i,j}) = \frac{1}{1 + e^{-v_{i,j}}} \quad (2.51)$$

$y_{i,j}$  : i. katmadaki j. nöronun çıktısı

$v_{i,j}$  : i. katmandaki j. nöronun çıktısı, şeklindedir.

Bu açıklamaya göre, birinci katmanda  $N_{1,1}$  ve  $N_{1,2}$  nöronlarının çıktıları aşağıda gösterilmektedir.

$$f(v_{1,1}) = y_{1,1} = \frac{1}{1 + e^{-v_{1,1}}} \quad (2.52)$$

$$f(v_{1,2}) = y_{1,2} = \frac{1}{1 + e^{-v_{1,2}}} \quad (2.53)$$

Birinci katmanda bulunan nöronlar, ( $N_{1,1}$  ve  $N_{1,2}$ ) ikinci katmanda bulunan nöronlara ( $N_{2,1}$ ,  $N_{2,2}$ ,  $N_{2,3}$ )  $a_{ij}$  sinaptik ağırlıkları ile bağlıdır. Bu durumda ikinci katmanda bulunan ( $N_{2,1}$ ,  $N_{2,2}$ ,  $N_{2,3}$ ) nöronların girdileri aşağıda gösterilmektedir.

$$v_{2,1} = \sum_{i=1}^2 y_{1,i} a_{i,1} = y_{1,1} a_{1,1} + y_{1,2} a_{2,1} \quad (2.54)$$

$$v_{2,2} = \sum_{i=1}^2 y_{1,i} a_{i,2} = y_{1,1} a_{1,2} + y_{1,2} a_{2,2} \quad (2.55)$$

$$v_{2,3} = \sum_{i=1}^2 y_{1,i} a_{i,3} = y_{1,1} a_{1,3} + y_{1,2} a_{2,3} \quad (2.56)$$

(2.54), (2.55), (2.56)'da ifade edildiği gibi, "birinci katmanda bulunan nöronların çıktıları ( $y_{1,1}$ ,  $y_{1,2}$ ) ikinci katmanda bulunan nöronların girdilerini oluşturmakta ve girdiler ( $a_{ij}$ ) sinaptik ağırlıklarıyla çarpılarak ağırlıklandırılmış net girdiler ( $v_{2,1}$ ,  $v_{2,2}$ ,  $v_{2,3}$ ) hesaplanmaktadır.

Gizli katman 2'de bulunan nöronlar da kendi üzerlerine gelen net girdileri aktive edecek ve kendi çıktıları olan  $y_{2,1}$ ,  $y_{2,2}$  ve  $y_{2,3}$ 'ü oluşturacaklardır. Daha önceden de belirtildiği gibi, her bir nöronun aktivasyon fonksiyonu aynı olmak zorunda değildir ve hangi nöronda hangi aktivasyon fonksiyonunun kullanılacağı uygulayıcının kendi seçimidir. Aktivasyon fonksiyonları nöron bazlı olarak seçilebileceği gibi, katman bazlı



da seçilebilir (örneğin 1.katmanda sigmoid, 2.katmanda hiperbolik tanjant, çıktı katmanında doğrusal gibi). Bununla birlikte çoğu uygulamada her katmandaki (dolayısıyla tüm ağdaki) nöronların aktivasyon fonksiyonu olarak seçilmektedir. Gizli katman 2 için seçilen aktivasyon fonksiyonu aşağıda verilmiştir.

$$y_{i,j} = \frac{1}{1 + v_{i,j}} \quad (2.57)$$

Buna göre gizli katman 2'de bulunan nöronların çıktıları ise;

$N_{2,1}$ 'in çıktısı,

$$f(v_{2,1}) = y_{2,1} = \frac{1}{1 + v_{2,1}} \quad (2.58)$$

$N_{2,2}$ 'nin çıktısı,

$$f(v_{2,2}) = y_{2,2} = \frac{1}{1 + v_{2,2}} \quad (2.59)$$

$N_{2,3}$ 'ün çıktısı ise;

$$f(v_{2,3}) = y_{2,3} = \frac{1}{1 + v_{2,3}} \quad (2.60)$$

Gizli katman 2'de bulunan;  $N_{2,1}$ ,  $N_{2,2}$ , ve  $N_{2,3}$  nöronları ise,  $b_{i,j}$  sinaptik ağırlıkları ile çıkış katmanına bağlıdır. Dolayısıyla gizli katman 2'nin çıktı değerleri olan  $y_{2,1}$ ,  $y_{2,2}$  ve  $y_{2,3}$  çıktı katmanının girdilerini oluşturacaktır. Bu durumda çıkış katmanına gelen ağırlıklandırılmış net girdi aşağıda gösterilmiştir.

$$v_{3,1} = \sum_{i=1}^3 y_{2,i} b_{i,1} = y_{2,1} b_{1,1} + y_{2,2} b_{2,1} + y_{2,3} b_{3,1} \quad (2.61)$$

Son bir işlem olarak, çıktı katmanı, üzerine gelen ağırlıklandırılmış net girdiyi aktive edecek ve nihai sonucu çıkış katmanına gönderecektir. Çıkış katmanı için seçilen fonksiyonu;

$$y'_t = f(v_{3,1}) = \frac{e^{v_{3,1}} - e^{-v_{3,1}}}{e^{v_{3,1}} + e^{-v_{3,1}}} \quad (2.62)$$

Örnek yapay sinir ağının nihai çıktısı;  $y'_t$  olacaktır.

### 2.4.2.3. Geri Yayılma Aşaması

Geri yayılma aşında hatalar “aktivasyon fonksiyonun türevi tarafından, ileri besleme için yapılan bağlantılar üzerinden çıkış tabakasından giriş tabakasına doğru yayılmaktadır.” Geri yayılma aşında hata geriye yayma işleminin kullanılabilmesi için, kullanılan aktivasyon fonksiyonunun türevlenebilir olması gerek ve yeter şarttır. Geri yayılma ağları genelde öğrenme kuralı olarak geliştirilmiş delta öğrenme kuralını kullanırlar.

Öğrenme esnasında bağımsız değişkenler ( $x_{1,t}$ ,  $x_{2,t}$ ,  $x_{3,t}$  ..... $x_{n,t}$ ) ağa sırasıyla sunulur. Ağ, giriş tabakasından aldığı verileri gizli tabakaya iletir. Başlangıçta rastgele olarak atanan bağlantı ağırlıklarıyla çarpıp, ağırlıklandırılmış net girdiler, gizli tabakada aktivasyon fonksiyonuyla aktive edilirler. Daha sonra gizli tabaka da çıkış değeri olan  $y_{i,j}$ 'yi eğer önünde gizli tabaka varsa ona, yoksa çıkış tabakasına gönderir. Çıkış tabakası da  $y_{i,j}$ 'yi gizli tabaka ve çıktı tabakası arasındaki bağlantı ağırlıklarıyla çarpıp toplayarak nihai çıktı olan  $y_t$ 'yi üretir. Bağımsız değişkenlerden oluşan veri, ağ içinde ileri yönde hareket etmektedir.

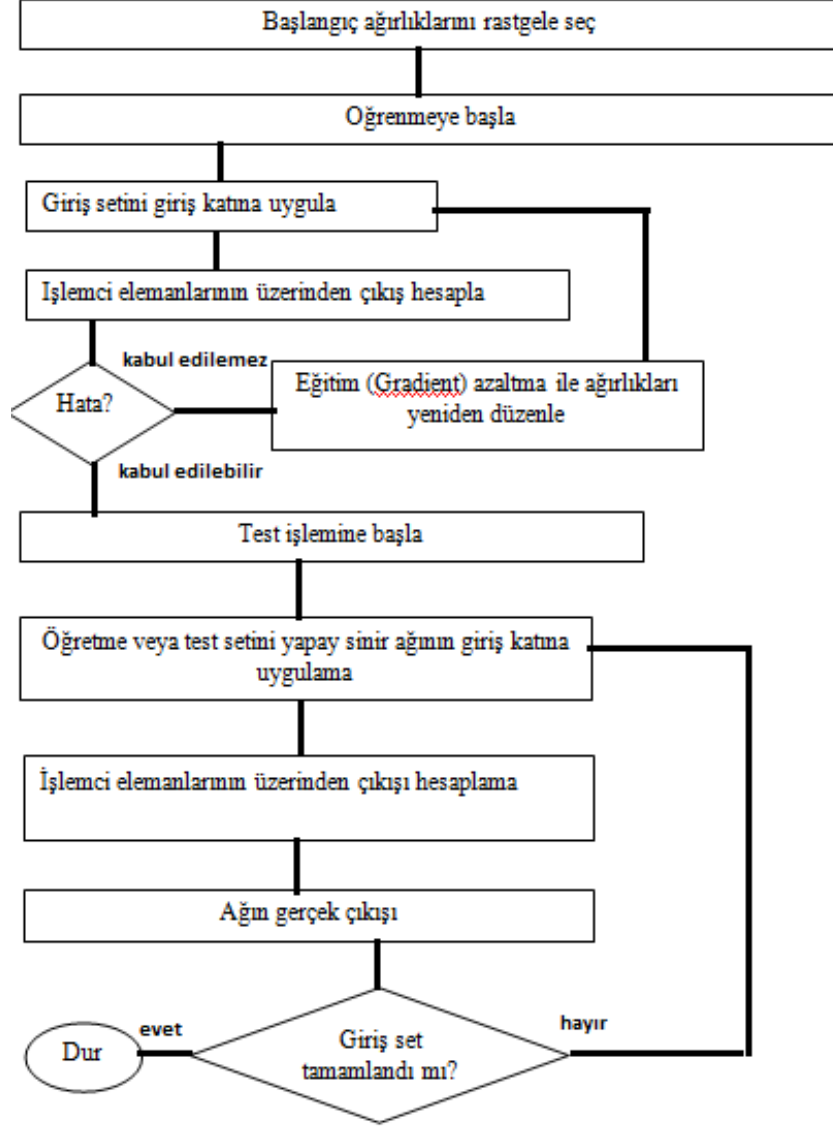
Hesaplanan çıktı ( $y'_t$ ) olması gereken çıktıyla ( $y_t$ ) karşılaştırılarak hata terimi ( $\varepsilon_t$ ) bulunur. Hata teriminin matematiksel ifadesi aşağıda verilmiştir.

$$\varepsilon_t = y_t - y'_t \quad (2.63)$$

Bulunan hata terimi, katmanlar arası bağlantı ağırlıklarının ayarlanmasında kullanılmak üzere, çıkış katmanından bir önce gelen gizli tabakaya iletilir. Burada ifade edilen geri besleme bağlantıları aslında katmanlar arası sinaptik bağlantılarla aynı bağlantılardır. Bu bağlantılar; veriyi iletmek için kullanıldığında katmanlar arası bağlantı, hatayı geri iletmek aracılığıyla kullanıldığında, geri besleme bağlantısı olarak adlandırılırlar.

Geri yayılma aşamasının daha iyi anlaşılması için, ağın öğrenme aşamasının akış şeması Şekil 2.19.'da gösterilmiştir. Ağın eğitilmesinin sağlıklı bir şekilde gerçekleştirilebilmesi için başlangıçta rastgele olarak atanan bağlantı ağırlıkları çok önemlidir. Rastgele olarak atanan bağlantı ağırlıkları eğitime hangi noktadan

başlayacağımızı belirlemektedir. Başlanan bu noktanın gerçek çözüme çok yakın veya uzak bir nokta olması tamamen başta belirleyeceğimiz değerlere bağlıdır.<sup>145</sup>



Şekil 2.19. Bir Ağda Geri Yayımlı Öğrenme Algoritmasının Uygulanması

**Kaynak:** Ş. Sağıroğlu, E. Beşdok, M. Erler, Mühendislikte Yapay Zeka Uygulamaları I: Yapay Sinir Ağları, Ufuk Kitap Kıratsiye-Yayıncılık, Kayseri 2003, s.80.

### 2.4.3. LVQ Modeli (Learning Vector Quantization)

Kohonen (1984) tarafından geliştirilen LVQ modeli destekleyici öğrenme modelini kullanan bir ağ yapısıdır. LVQ ağları genel olarak sınıflandırma problemlerinin çözümünde kullanılmaktadır.

<sup>145</sup> Sağıroğlu, s.80.

Öğrenme, girdi vektörünün hangi vektör seti (referans vektör) tarafından temsil edilmesi gerektiğinin bulunmasıdır. LVQ ağının görevi öğrenme yolu ile vektör setini yani girdi vektörlerinin üyesi olabilecekleri vektör sınıfını belirlemektir. Kohonen öğrenme kuralına göre öğrenir. Çıktılardan sadece birisi 1, diğerleri 0 değerlerini alır ve çıktı değerinin 1 olması girdinin ilgili çıktının temsil ettiği sınıfa ait olduğunu gösterir. LVQ ağı istatistiksel bir sınıflandırma ve ayırt etme metodu olarak kullanıldığı için amacı girdi verilerini sınıflandırmada kullanılır.<sup>146</sup>

LVQ ağında, bir giriş, bir çıkış, ve kohonen katmanı mevcut olup, giriş katmanındaki tüm nöronlar ara katmanlardaki tüm nöronlarla bağlantılı bir yapı sergilemektedir. Bu ağ da asıl amaç, n boyutlu bir vektörün, vektörler seti halinde haritalanmasıdır.<sup>147</sup>

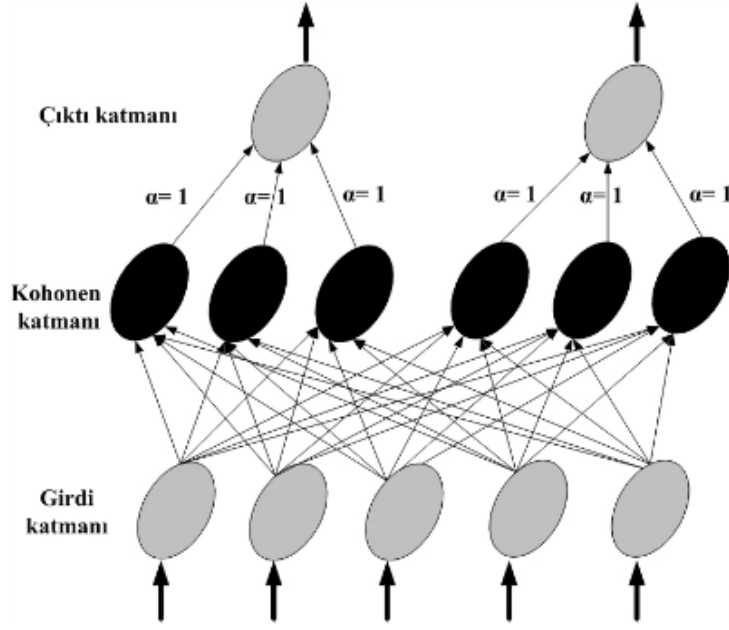
Eğitim sırasında girdi vektörünün sınıflara ayrılması en yakın komşu kuralına göre gerçekleştirilir. Girdi vektörü ile referans vektörleri arasındaki en kısa mesafe aranmakta ve girdi vektörünün en kısa mesafede bulunan vektör grubuna ait olduğu varsayılır ve ağın ağırlıkları değiştirilerek girdileri doğru sınıflara ayıracak referans vektörler belirlenir. Kullanılan öğrenme stratejisi destekleyici (reinforcement learning) öğrenmedir. Çıktı değerinin belirlenmesinde ise “kazanan her şey alır” stratejisi uygulanmaktadır. Ağ eğitilirken her iterasyonda ağın ürettiği çıktının değeri yerine sadece doğru olup olmadığı söylenir. Sadece girdi vektörüne en yakın olan vektörün (kazanan vektör) değerleri (ağın bu vektöre ait ağırlıkları) değiştirilir.<sup>148</sup>

---

<sup>146</sup> T. Kohonen, *Self-Organizing Maps* [Elektronik Sürüm], Springer Series in Information Sciences, 3rd Edition, New York 2001, s.245.

<sup>147</sup> Küçükönder, s.70.

<sup>148</sup> Adıyaman, s.37.



Şekil 2.20. LVQ Ağ yapısı

**Kaynak:** E. Öztemel, Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.116.

LVQ ağı üç katmandan oluşur:<sup>149</sup>

- **Girdi katmanı:** Bu katmanda bilgi işleme olmaz, gelen bilgiler girdi vektörünü oluşturur. Bu katmandaki her proses elemanı kohonen katmanındaki her proses elemanına bağlıdır. Girdi katmanı ve kohonen katmanındaki ağırlıklar değiştirilerek öğrenme sağlanır.
- **Kohonen katmanı:** Bu katmanda girdi setine en yakın olan ağırlık vektörü belirlenir. Bu katmandaki her eleman bir referans vektörünü gösterir ve girdi değerlerini Kohonen katmanındaki proses elemanlarına bağlayan bağlantıların ağırlık değerlerinden oluşur. Referans vektörünün eleman sayısı girdi katmanındaki eleman sayısı kadardır.
- **Çıktı katmanı:** Bu katmanda ise girdinin ait olduğu sınıf belirlenir. Kohonen katmanındaki proses elemanları çıktı katmanındaki bir tek proses elemanına bağlıdır. Kohonen katmanı ile çıktı katmanı arasındaki ağırlıklar ( $\alpha$ ) sabit olup 1'e eşittir.

<sup>149</sup> Adıyaman, s.38.

#### 2.4.4. ART (Adaptive Resonance Theory) Ağları

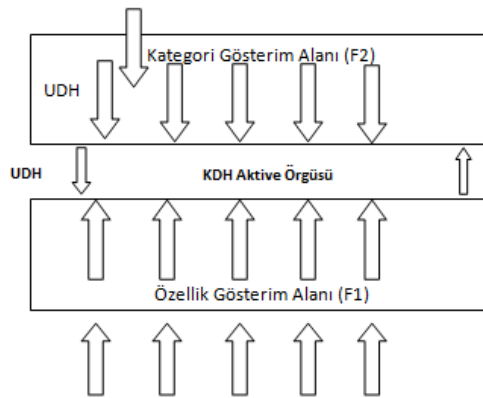
ART ağları Grosberg'in 1976 yılında biyolojik beynin fonksiyonlarına yönelik olarak yaptığı çalışmalar neticesinde ortaya çıkmıştır. Bu model, biyolojik sisteme ait 3 temel üzerine kurulmuştur.<sup>150</sup>

**1-Normalizasyon:** Çevredeki büyük değişikliklere karşı adaptif olma durumudur.

**2-Ayrıştırılabilme:** Çevredeki ayrıntılı seviyedeki değişiklikleri fark edebilme durumudur.

**3-Ayrıntıların saklandığı kısa dönemli hafıza:** Belirlenen farklılıklar önce hafızada saklanır ve sonra davranışa dönüşür. Anlık kararlarda ani değişimlerin tutulduğu kısa dönemli hafıza etkili olurken uzun karar vermede uzun dönemli hafıza etkili olmaktadır.

ART ağları genel olarak iki katmandan oluşur. Bu katmanlar girdilerin özelliklerinin gösterildiği F1 katmanı ve ayrıştırılmış sınıfların gösterildiği F2 katmanıdır. Girdi bilgileri F1 katmanından alınır ve sınıflandırma ise F2 katmanında yapılır. F1 katmanının aktivasyonu girdilerin özellikleri incelenerek belirlenir ve uzun dönemli hafızadaki bağlantı değerleri ile gelen bilgiler kategorilere ayrılarak F2 katmanına gönderilir. F2 katmanındaki sınıflandırma ile F1 katmanından gelen sınıflandırma birbirleri ile eşleştirilir. Örnek belirlenmiş bir sınıfa uyuyorsa o kategoride gösterilir uyumuyorsa yeni bir sınıf oluşturulur veya girdinin sınıflandırılması yapılmaz.



**Şekil 2.21.** ART ağının genel yapısı

**Kaynak:** B. Kröse and P. Van Der Smagt, An Introduction to Neural Networks [Elektronik Sürüm], The University of Amsterdam Press, Amsterdam 1996, s.70.

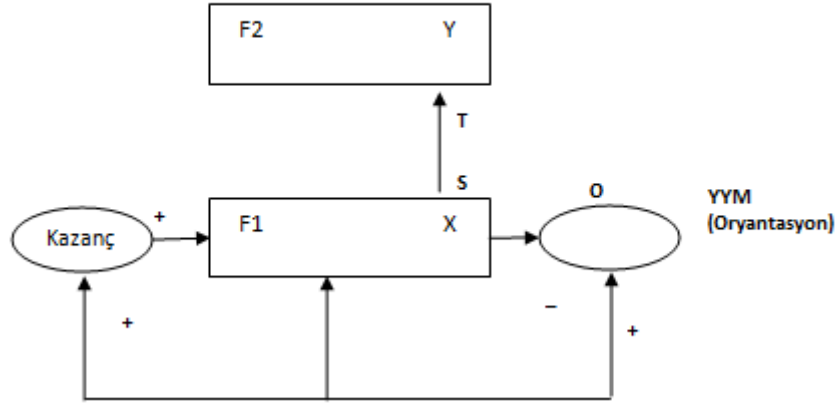
<sup>150</sup> Kröse and Van der Smagt, s.69.

#### 2.4.4.1. ART Ağlarının Çalışma Prensipleri

ART ağlarında F1 katmanından gelen bilgiler F2 katmanındaki kategorilerle eşleşir. Eşleşme sağlanmaz ise yeni bir kategori oluşturulur. ART ağlarının çalışması iki yönlü olmaktadır.

➤ **Aşağıdan yukarı (F1'den F2'ye) bilgi işleme:** Bir girdi örüntüsü ağa gösterilir.

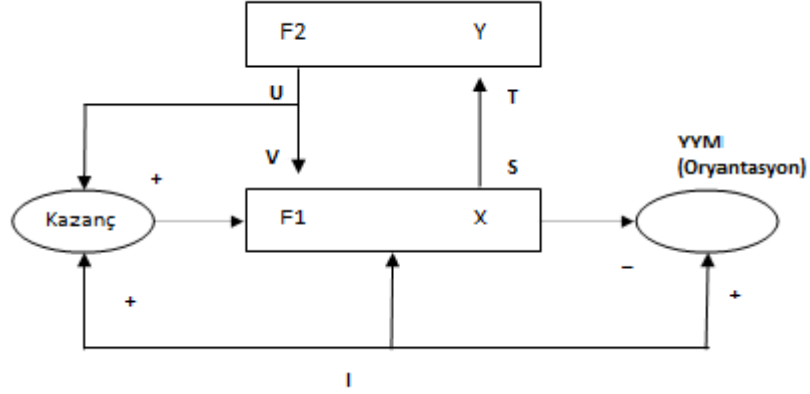
Bu örüntü F1 katmanında kısa dönemli hafızada aktivite örüntüsünü oluşturur ve yeniden yerleştirme modülünü aktif etmeye üzere bir işaret gönderir. Aktivite örüntüsü de yeniden yerleştirme modülüne men edici işaret gönderir ve F1 katmanının çıktısını oluşturur. Çıktı F2 katmanında bir girdi örüntüsüne dönüştürülür. Bu örüntüde F2 katmanının çıktı örüntüsünü oluşturur.



Şekil 2.22. Aşağıdan yukarıya bilgi işleme

**Kaynak:** E. Öztemel Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.142.

➤ **Yukarıdan aşağı (F2'den F1'e) bilgi işleme:** F2 katmanında oluşturulan çıktı örüntüsü yukarıdan aşağıya bir sinyal gönderir. Bu sinyal daha sonra beklenen şablon örüntüye dönüştürür. Aynı zamanda kontrol faktörü (kazanç) için men edici bir işaret oluşturur. Şablon örüntü ile girdi örüntüsünün eşleşip eşleşmeyeceğini denener. Eşleşme mümkün değilse F1 katmanında kısa dönemli hafıza örüntüsü oluşturulur. Bu örüntü de yeniden yerleştirme modülündeki men edici işaretin etkisini azaltır.



**Şekil 2.23.** Yukarıdan Aşağıya Bilgi İşleme

**Kaynak:** E. Öztemel Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.143.

#### 2.4.4.2. ART Ağlarının Diğer Yapay Sinir Ağlarından Farkları

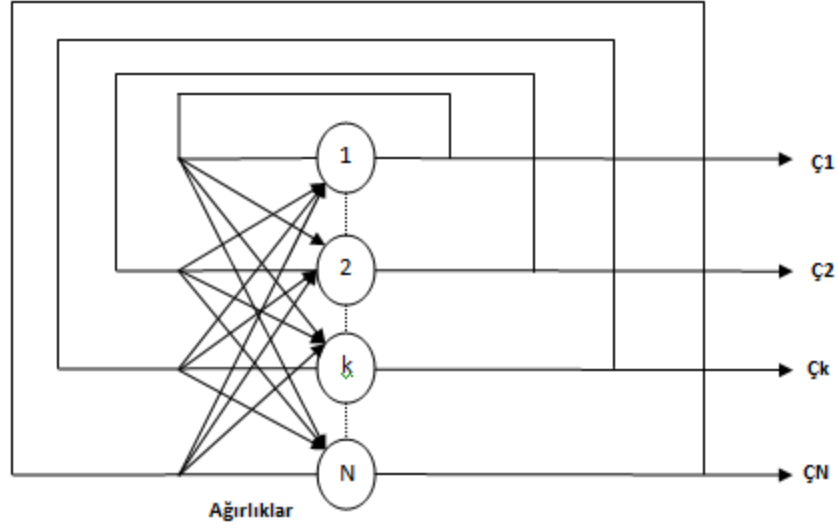
- ART ağları bir çok ağın tersine gerçek zamanlı, hızlı ve kararlı öğrenirler. Diğer ağlar ise çoğu zaman çevrimdışı öğrenirler ve ortama anlık olarak uyum sağlayamazlar.
- ART ağları öğretmensiz öğrenir.
- ART ağları ağa sunulan değişik durumlardaki örnekler karşısında kendi kendilerine kararlı bir yapı oluşturabilirler.
- ART ağlarında öğrenme süreklidir ve uzun dönemli hafızada bulunan ağırlıklar sürekli güncellenir.
- ART ağları girdi değerlerini otomatik olarak normalize ederler.
- ART ağlarında hem aşağıdan yukarı hem de yukarıdan aşağıya ağırlık değerleri vardır. Aşağıdan yukarı gelen bilgiler ile karşılaştırma kısa zamanlı hafızada (KDH) oluşur. Yukarıdan aşağı ağırlıklar KDH'da yapılan karşılaştırma ile aynı kategoride olmayan girdilerin o kategoriye girmesini önlerler.
- ART ağları yakın eşleme özelliği sayesinde hem hızlı hem de yavaş öğrenebilirler. Hızlı öğrenme uzun dönemli hafızada (UDH) bir denemede yeni bir dengenin oluşturulması ile gerçekleştirilir. Yavaş öğrenme ise bir dengenin oluşması için birden çok denemenin yapılmasıdır.<sup>151</sup>

<sup>151</sup> Adıyaman, s.42.



### 2.4.5. Hopfield Ağı

Hopfield ağı tek katmanlı ve geri dönüşümlü bir ağıdır. Proses elemanlarının tamamı hem girdi hem de çıktı elemanlarıdır. Proses elemanlarının hepsi birbirine bağlıdır. Ağ bağlantı değerleri bir enerji fonksiyonu olarak saklanır. Kesikli Hopfield Ağı ve Sürekli Hopfield Ağı olmak üzere iki tür modeli vardır. Zor optimizasyon problemlerinin çözümünde kullanılır.<sup>152</sup>



Şekil 2.24. Hopfield Ağ Yapısı

**Kaynak:** E. Öztemel Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.171.

Hopfield sinir ağında girişler ikili (0 veya 1) ve bipolar (-1 veya + 1) olarak kabul edilir. Bu ağda geri yayımlı ağın eğitiminden farklı olarak, örnek desen gruplarının seçimi ve seçilen bu desen gruplarının ağın başlangıç değerlerini belirlemek üzere kullanım durumları söz konusudur. Örnek desen grup seçimi ve bu desen grubuna göre başlangıç değeri belirlendikten sonra herhangi bir desen ağı sunularak, giriş desenine en çok benzeyen örnek desenlerden biriyle işlem sonlandırılır. Çıktı deseni birimlerinin durumuna göre ağdan okunmaları da mümkündür.<sup>153</sup>

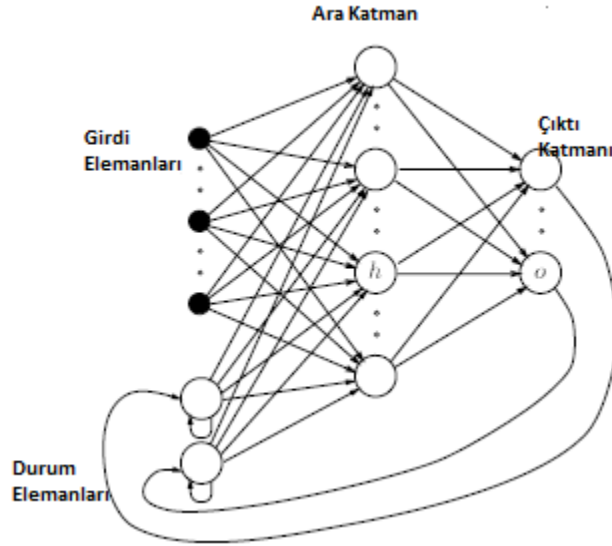
<sup>152</sup> J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities" [Acil Toplu Hesaplamalı Yeteneklerle Birlikte Fiziksel Sistemler ve Yapay Sinir Ağları], *Proceedings of National Academy of Science*, 79(8), 1982, pp. 2554, Erişim Tarihi: 06 Kasım 2011, ScienceDirect.

<sup>153</sup> Küçükönder, s.76.

### 2.4.6. Jordan Ağı

Jordan ağında yer alan durum elemanları, çıktı tabakasından aktivasyon değerlerini alarak bir sonraki iterasyonda bunu girdi olarak taşımakla görevli olan elemanlardır. Girdi, çıktı ve ara katmanlara ilaveten durum elemanları olarak adlandırılan özel işlemci elemanlarının bulunduğu, çok katmanlı geri beslemeli (yinelemeli) bir ağ yapısına sahiptir.<sup>154</sup>

Ara katman elemanlarının en önemli özelliği bu ağda hem doğrusal hem de doğrusal olmayan aktivasyon fonksiyonuna sahip olabilmeleridir. Öğrenme kuralı olarak çok katmanlı perceptron ağlarının kullandığı öğrenme kuralları bu ağda da kullanılabilir.<sup>155</sup>



Şekil 2.25. Jordan Ağı

**Kaynak:** B. Kröse and P. Van Der Smagt, An Introduction to Neural Networks [Elektronik Sürüm], The University of Amsterdam Press, Amsterdam 1996, s.48.

### 2.4.7. Elman Ağı

Çok tabakalı YSA yapısının tümüne sahip olan ilave olarak ara tabaka çıktılarını paralel bir girdi tabakası olarak yapısında bulduran YSA türüne Elman ağı denir.<sup>156</sup>

Elman ağı herhangi bir zamandaki girdi değerleri hem de ara katmanların önceki aktivite değerleri ağa girdi olarak verilirler. Ağın girdileri belirlendikten sonra ağ artık

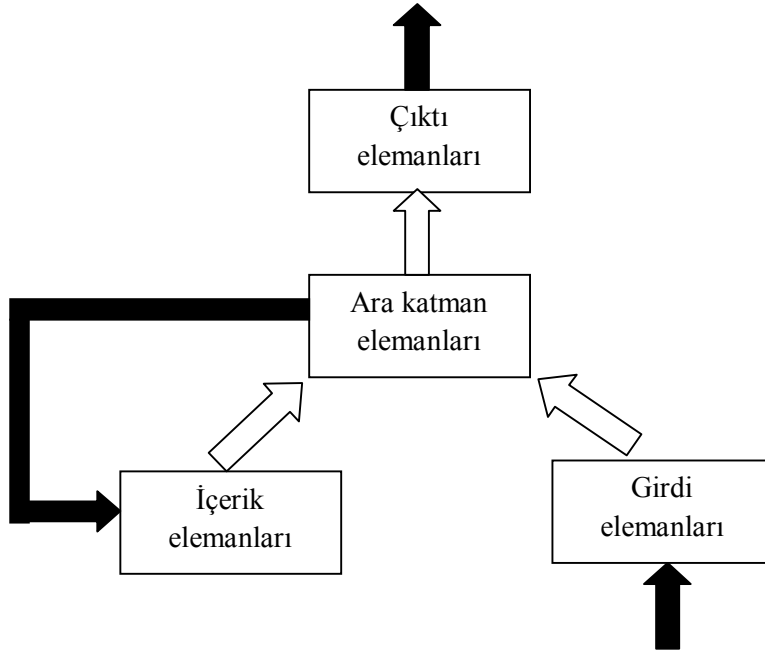
<sup>154</sup> Baş, s.79.

<sup>155</sup> Kröse and Van der Smagt, s.47.

<sup>156</sup> Şen, s.144.

ileri beslemeli birçok katmanlı algılayıcıya dönüşür. Bu girdiler kullanılarak ileri doğru ağın çıktıları belirlenir.<sup>157</sup>

Elman ağı, Jordan ağına oldukça benzer olmakla beraber aralarında iki önemli farklılıklar bulunmaktadır. Bu farklılıklardan ilki, geri besleme yaptıkları aktivasyon değerlerini çıktı katmanından almak yerine ara katmandan almaları, ikinci ise içerik elemanlarının kendilerine bağlantı durumlarının olmamasıdır.<sup>158</sup> Bu ağ şekil 2.26.'da gösterilmiştir.



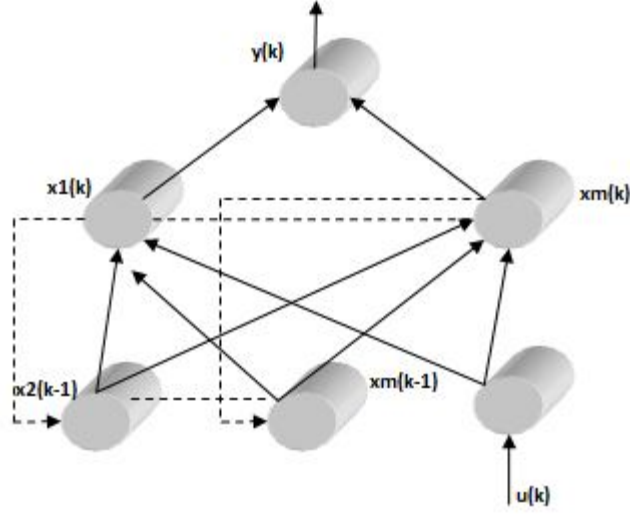
**Şekil 2.26.** Elman Ağ Yapısı

Elman ağında öğrenme, genelleştirilmiş delta öğrenme kuralına göre iki adımda gerçekleşmektedir. İşlem adımlarından ilki, ara katmanda yer alan işlemci elemanlara gelen net girdi değerinin, girdi katmanında yer alan eleman değerleriyle ağırlıklarının çarpılıp toplanmasıdır. İkinci adım ise içerik elemanlarından gelen bu bağlantı değerlerinin ara katmanlarda oluşan bir önceki aktivasyon değerleri ile çarpılıp

<sup>157</sup> J.L. Elman, "Finding structure in time", [Zamanında Yapı Keşfi], *Cognitive Science*, 14(2), 1990, pp, 182, Erişim Tarihi: 29 Ekim 2011, ScienceDirect.

<sup>158</sup> Küçükönder, s.78.

toplanmasıdır.<sup>159</sup> Elman ağının bu ayrıntılarının gösterdiği yapı ise şekil 2.27.'deki gibidir.



**Şekil 2.27.** Elman ağının ayrıntılı gösterimi

**Kaynak:** E. Öztemel, Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.167.

#### 2.4.8. Kohonen Ağı

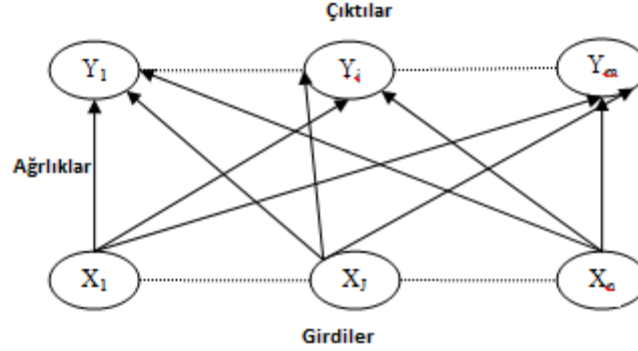
SOM (Self Organization Feature Map Network – Kendi Örgütlemeli Özellik Haritası) olarak da adlandırılan Kohonen ağı, beynin Neocortex tabakasında yaygın olan duyuşal haritalardan esinlenerek Kohonen (1972) tarafından geliştirilmiştir. Genel olarak sınıflandırma yapmak için kullanılan Kohonen ağının, girdi vektörlerini sınıflandırma ve girdi vektörlerinin dağılımını öğrenebilme yetenekleri oldukça yüksektir.

Kohonen ağlarının en temel özelliği olayları öğrenmek için bir öğretmen veya ağı üretmesi gereken çıktılarının ağa söylenme zorunluluğunun olmamasıdır. Yapısal olarak da bu ağla diğerlerinden farklıdır. Ağ girdi ve çıktı katmanından oluşmaktadır. Çıktı katmanı 2 boyutlu bir düzlemi göstermektedir.<sup>160</sup>

<sup>159</sup> Öztemel, s.168.

<sup>160</sup> Öztemel, s.181.

SOM ağırları yarışmayı kazanma ve kazanan elemanın 1 diğerlerinin 0 değerini alması ilkesine dayanmaktadır. Bir girdi verildiğinde çıktı uzayında yarışmayı kazanan ve onun etrafındaki komşuları eğitim sırasında ağırlıklarını değiştirmektedir.<sup>161</sup>



Şekil 2.28. Kohonen SOM haritası

**Kaynak:** L.V. Fausett, Fundamentals of Neural Networks [Elektronik Sürüm], Prentice Hall, Canada 1993, pp.170.

## 2.5. YAPAY SİNİR AĞ TASARIMI

Bir yapay sinir ağı modeli oluşturmak için uygulanması gereken temel aşamalar aşağıda listelenmiştir:<sup>162</sup>

- Öncelikle sistemin amacın diğer bir ifadeyle çıktısının ne olacağı belirlenmeli,
- Bu çıktıya etkisi olduğu düşünülen ve yapay sinir ağı sistemine girdi olarak alınacak değerler belirlenmeli,
- Bu girdiler ile ilgili bir veri seti oluşturulmalı, eğitim ve test seti olarak bölünmeli ve ağı uygun bir şekilde dönüştürülmeli,
- Ağ topolojisi belirlenmeli (kaç katman, katmanlarda kaç tane nöron, nöronların aktivasyon fonksiyonu),
- Öğrenme algoritması belirlenmeli,
- Ağın eğitilmesi ve test edilmesi,
- İstenilen seviyeye ulaşana kadar önceki adımların tekrarlanmasıdır.

<sup>161</sup> Fausett, s.169.

<sup>162</sup> Delvin.D. Hawley, J.D. Johnson and D. Raina, "Artificial Neural Systems: A new tool for Financial Decision Making", [Yapay Sinir Ağları Sistemleri: Yeni Bir Finansal Karar Verme Aracı], *Financial Analyst Journal*, 46(6), 1990, pp. 65, Erişim Tarihi: 12.02.2012, <http://www.cfapubs.org/doi/pdf/10.2469/faj.v46.n6.63>

Yapay sinir ağlarında çözüme ulaştırılmak istenen probleme göre ağ yapısı farklılık kazanmakta olup, kullanım amacına bağlı olarak istenilen problemin çözümlenmesindeki başarısı ağın doğru olarak belirlenmesine bağlıdır. Ağın türleri, kullanım alanları ve kullanım amacına göre tahmin, sınıflandırma ve veri ilişkilendirme olmak üzere üç farklı kategoride incelemek mümkündür.<sup>163</sup>

**Çizelge 2.3.** Ağ Türlerinin Sınıflandırılması

<b>Kullanım Amacı</b>	<b>Ağ türü</b>	<b>Kullanım alanı</b>
<b>Tahmin</b>	<ul style="list-style-type: none"> <li>•Çok katmanlı ağ</li> </ul>	Ağın girdilerinden bir çıktı değerinin tahmin edilmesi
<b>Sınıflandırma</b>	<ul style="list-style-type: none"> <li>•LVQ</li> <li>•ART</li> <li>•Counterpropagation</li> <li>•Olasılıklı sinir ağları</li> </ul>	Girdi değerlerinin ait oldukları sınıfları belirlemeye yönelik olarak sınıfların belirlenmesi
<b>Veri İlişkilendirme</b>	<ul style="list-style-type: none"> <li>•Hopfield</li> <li>•Boltzmann makinası</li> <li>•Bidirectional associative memory</li> </ul>	Girdi değerleri içinde bulunan eksik bilgilerin tespit edilerek tamamlanması ve girdi değeri içinde hatalı bilgilerin tespit edilmesi

**Kaynak:** T. Saraç, *Yapay Sinir Ağları*, Basılmamış Seminer projesi Gazi Üniversitesi, Ankara 2004, s.32.

### 2.5.1. Öğrenme Algoritmasının Seçimi

Yapay sinir ağlarında ağ mimarisinin seçimi açısından öğrenme algoritmasının doğru seçimi ağın başarılı sonuçlar üretmesini etkileyen önemli bir faktörlerden biridir.

Uygulama alanlarına göre algoritmaları, öngörü tanıma, sınıflandırma, veri ilişkilendirme ve veri kavramlaştırması olmak üzere dört farklı kategoriye ayırmak mümkündür. Bu kullanım alanlarına göre kullanılan algoritmalar aşağıdaki çizelgede gösterilmiştir.<sup>164</sup>

<sup>163</sup> T. Saraç, *Yapay Sinir Ağları*, Basılmamış Seminer projesi Gazi Üniversitesi, Ankara 2004, s.32.

<sup>164</sup> Saraç, s.33.

**Çizelge 2.4. Öğrenme Algoritmaları ve Kullanım Alanları**

<b>Uygulama alanları</b>	<b>Yapay sinir ağları</b>
<b>Öngörü tanıma</b>	Geri yayılım Delta bar delta Genelleştirilmiş delta bar delta Yönlendirilmiş rastsal tarama
<b>Sınıflandırma</b>	LVQ Olasılıklı yapay sinir ağları Counterpropogation
<b>Veri ilişkilendirme</b>	Hopfield Boltzmann
<b>Veri kavramlaştırma</b>	ART SOM

### 2.5.2. Gizli Katman Sayısının Belirlenmesi

Yapay sinir ağları, genelleştirme yeteneğini gizli katmanlardan almaktadır. Çoğu başarılı sinir ağı uygulamalarında, gizli katman sayısı ve bu katmanlarda bulunan işlem elemanı sayısı önemli rol oynar. Verilerdeki kalıpların kavranması, girdi ve çıktı değişkenleri arasındaki doğrusal olmayan eşleşmelere izin veren mekanizma, gizli katmanlardaki işlem elemanlarıdır.<sup>165</sup> Gizli katman sayısının ne olması gerektiği konusunda varılan nokta; çoğu problem için bir gizli katman veya en çok iki gizli katmanın yeterli olacağıdır.<sup>166</sup> Pratik çalışmalarda mantıksal zorunluluk dışında tek gizli katman kullanılır.<sup>167</sup>

### 2.5.3. Gizli Katman İşlem Elemanı Sayısının Belirlenmesi

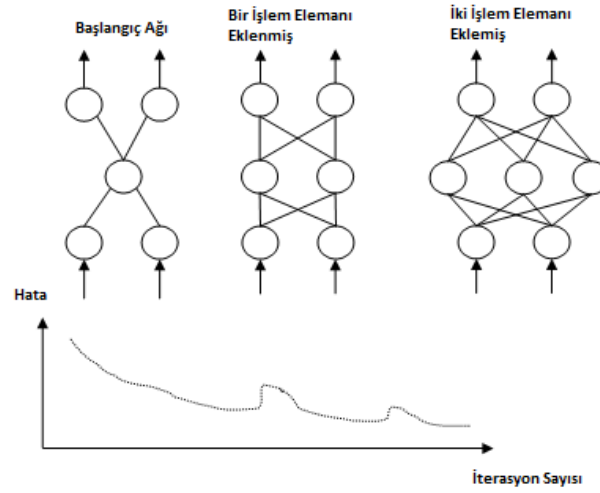
Gizli katmanda kullanılacak işlem elemanı sayısı, genellikle deneme-yanılma yöntemi kullanılarak, her deneme sonucunda ağın ürettiği hata grafiği dikkate alınarak, en uygun işlem elemanı sayısı belirlenmeye çalışılır. Uygun işlem elemanı sayısına

<sup>165</sup> Zhang, vd., s.42-44.

<sup>166</sup> Iebeling Kaastra, and S. Boyd Milton, "Designing a Neural Network for Forecasting Financial and Economic Time Series" [Finansal ve Ekonomik Zaman Serileri Tahmini için Bir Yapay Sinir Ağı Dizayını], *Neurocomputing*, 10(3), 1996, pp. 224, Erişim Tarihi: 24 Ocak 2012, ScienceDirect.

<sup>167</sup> Şen, s.87.

ulaşmak için ağın büyütülmesi veya budanması olarak ifade edilen iki değişik yöntem takip edilmektedir. Bu yöntemlerden ilki Şekil 2.29.'da gösterildiği gibi küçük bir ağ yapısından başlayıp, hata oranını dikkate alarak işlem elemanı sayısını artırmaktır. Ağın hata oranı yükselmeye başladığı noktada, gizli katman için uygun sayıda işlem elemanına ulaşıldığında karar verilir. İkinci yöntemde ise, tersi bir durum söz konusudur. Bu yöntemde, geniş bir ağ ile eğitime başlanır ve ağı genelleştirme yeteneğine bir düşme meydana gelinceye kadar, ağ küçülür. Böylece, hata oranı dikkate alınarak gizli katman için uygun sayıda işlem elemanı belirlenir.<sup>168</sup>



**Şekil 2.29.** Büyüyen Ağlar

**Kaynak:** E. Öztemel Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.105.

#### 2.5.4. Yapay Sinir Ağı Parametrelerinin Belirlenmesi

Yapay sinir ağı modeli geliştirilirken, önemli bir diğer konu da ağ parametrelerinin belirlenmesidir. Bu parametreler, başlangıç ağırlıklarının atanması, momentum ve öğrenme katsayılarının belirlenmesi, toplama ve aktivasyon fonksiyonlarının seçilmesidir. Uygun ağ parametrelerinin belirlenmesi, ağ başarısını

<sup>168</sup> Öztemel, s.105.

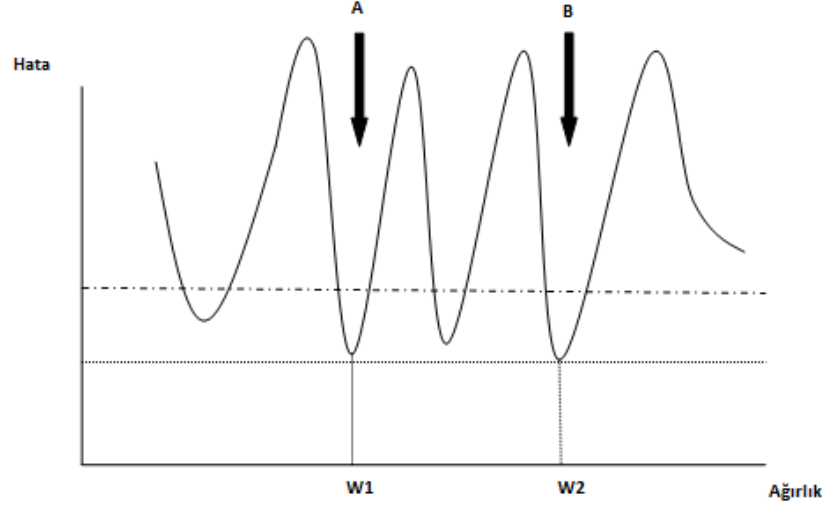
-A. Yalama, Y. Birol, ve Ünlü Ulaş, "İlk Halka Arzlarında Uzun Dönem Getirilerinin Tahmini: Yapay Sinir Ağları ile İMKB İçin Ampirik Bir Uygulama", *İstanbul Üni. İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, Sayı 10, 2009, ss. 40, Erişim Tarihi: 28 Kasım 2011, <http://eidergisi.istanbul.edu.tr/sayi10/iueis10m3.pdf>



önemli ölçüde etkilemektedir. Ağ mimarisinde olduğu gibi, bu parametrelerin belirlenmesinde de kesin formüller bulunmamaktadır. Deneme – yanılma yöntemi ile uygun parametre değerleri belirlenmeye çalışılmaktadır. Bu parametreler belirlenmeye çalışılırken dikkat edilmesi gereken noktalar aşağıda özetlenmiştir.<sup>169</sup>

#### 2.5.4.1. Başlangıç Ağırlık Değerlerinin Atanması

Yapay sinir ağları, eğitime başladığı ilk anda bağlantı ağırlıklarına ihtiyaç duymaktadır. Başlangıç ağırlıkları, ağın öğrenmeye başladığı noktayı ifade etmektedir. Bu değerlerin geniş bir aralıkta seçilmesi, ağın yerel noktalar arasında salınım göstermesine, dar bir aralığın seçilmesi durumunda ise ağın öğrenme hızının düşmesinde neden olmaktadır. Başlangıç ağırlık değerleri genelde -1 ile +1 arasında kalacak şekilde rastgele belirlenmektedir.



**Şekil 2.30.** ÇKA Ağlarında Başlangıç Noktasının Etkisi

**Kaynak:** E. Öztemel, Yapay Sinir Ağları, 1. Baskı, Papatya Yayıncılık, İstanbul 2003, s.99

Yapay sinir ağına başlangıç değerlerinin atanması ağ performansını etkilemektedir. Şekil 2.30'da görüldüğü gibi bir ÇKA ağı öğrenmeye A noktasından başlar ise yerel bir çözüme (W1) takılabilme olasılığı varken, B noktasından başladığı takdirde, en iyi çözümü (W2) bulması daha kolaylaşacaktır.<sup>170</sup>

<sup>169</sup> Öztemel, s.81.

- Zhang, pp.38.

- Kaastra and Boynd, pp.224.

<sup>170</sup> Öztemel, s.99.

#### **2.5.4.2. Öğrenme Oranının Belirlenmesi**

Öğrenme oranı, yapay sinir ağının hangi hızda öğrendiğini belirlemektedir. Bu oran, ağırlıkların ayarlanması için, ağırlık değerlerinde meydana gelen net değişimin hangi oranda yenileme hesabına katılacağını belirlemektedir. Öğrenme oranının küçük seçilmesi öğrenme sürecini yavaşlatırken, büyük öğrenme oranları ağın yerle çözümlere takılmasına neden olabilmektedir. Öğrenme oranı 0 ile 1 arasında bir değer almaktadır.

#### **2.5.4.3. Momentum Oranının Belirlenmesi**

Öğrenme oranı gibi, ağın eğitime başlamadan önce belirlenmesi gereken bir diğer önemli parametre de momentum katsayısıdır. Momentum katsayısı; bir önceki iterasyon sonucunda ağırlık değerlerinde oluşan net değişim miktarı, hangi oranda yeni değişim miktarına ekleneceğini belirlemektedir. Momentum terimi olmadan, ağın hata karelerin oluşturduğu çanağın dibine inilebilmesi için, çok düşük öğrenme oranına ihtiyaç duyulabilmektedir. Bu durum, eğitim sürecinin uzamasına neden olmaktadır.

Momentum teriminin kullanılmasının nedenlerinden biri de, ağların yerel çözümlerden kurtulmasına yardımcı olma işlevi görmesidir. Bu nedenle momentum katsayısının doğru belirlenmesi, ağın hızlı ve başarılı bir öğrenme süreci gerçekleştirmesine yardımcı olmaktadır. Momentum teriminin küçük seçilmesi, ağın yerel çözümlerden kurtulmasını zorlaştırırken, büyük değerler ise ağın tek bir çözüme ulaşmasında sorunlar yaratabilmektedir. Tecrübeler bu oranın 0.6 - 0.8 arasında seçilmesinin uygun olacağını göstermektedir. Fakat problemin niteliğine göre değişik oranlar seçilebilir.

#### **2.5.4.4. Aktivasyon fonksiyonu seçimi**

İşlem elemanına gelen net girdiyi hesaplamak için toplam fonksiyonu, hesaplanan bu girdiyi işleyip, bu girdiye karşılık, hücrenin çıktısını aktivasyon fonksiyonu üretmektedir.

Aktivasyon fonksiyonu seçiminde, bazı bulgucayıcı kurallar vardır. Klimasauskas (1991) problem yapısı içersindeki ortalama davranışın öğrenilmesi amaçlanan sınıflandırma problemlerinde sigmoid, eğer problem öngörü problemlerinde olduğu gibi, ortalamadaki sapmaların tespitini amaçlıyorsa, hiperbolik tanjant aktivasyon fonksiyonlarının kullanılmasını önermektedir. Bunun yanında farklı aktivasyon fonksiyonlarının ağ performansı üzerinde belirgin bir etkisinin olup olmadığı kesin

değildir. Literatürde en fazla sigmoid aktivasyon fonksiyonunun tercih edildiği görülmektedir.<sup>171</sup>

#### 2.5.4.5. Verilerin Normalleştirilmesi

Veriler, ağı sunulmadan önce ölçeklendirme işlemine tabi tutulması gerekir. Çünkü, veri setinde çok büyük veya çok küçük birkaç uç değer bulunması durumunda, ağı optimizasyon işlemini gerçekleştirmeyebilir. Veri normalleştirilmenin diğer bir nedeni de, aktivasyon fonksiyonlarının belirli aralıkların dışında ([0,1] , [-1,1] gibi) verilere hassas olmamasıdır. Bu nedenle, eğer veriler uygun bir aralık da ölçeklendirilmezse işlem elemanları büyük verileri aktivasyon fonksiyonunun en büyük değeri 1 olarak kabul edecek, küçük verileri de 0 (sigmoid fonksiyonu için) olacak kabul edeceğinden çok önemli miktarda veri kaybolacaktır.<sup>172</sup>

**Çizelge 2.5.** Veri Normalizasyonunda Kullanılan Formüller

Normalizasyon	Kullanılan matematiksel denklem
[ 0 1 ] aralığında dönüşüm	$x_n = (x_0 - x_{\min}) / (x_{\max} - x_{\min})$
[ a b ] aralığında dönüşüm	$x_n = (b-a) (x_0 - x_{\min}) / (x_{\max} - x_{\min})$
Basit normalizasyon	$x_n = (x_0) / (x_{\max})$
İstatistiksel normalizasyon	$x_n = (x_0 - \bar{x}) / S$

**Kaynak:** Y. Akdağ, Vadeli İşlem ve Opsiyon Borsası'nda TL/Dolar Vadeli İşlem Sözleşmelerinin Gün Sonu Uzlaşma Fiyatının Yapay Sinir Ağları ile Tahmini, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi, Bankacılık ve Sigortacılık Enstitüsü, İstanbul 2010, s.83.

Bu formüllerde yer alan,

$X_n$  : Normalleştirilmiş veri,

$X_0$  : Orijinal veri,

$X_{\min}$  : Veri matrisinde satır ve sütun boyunca minimum,

$X_{\max}$  : Veri matrisinde satır ve sütun boyunca maksimum,

$\bar{X}$  : Veri matrisinde satır ve sütun boyunca ortalama,

S : Satır ve sütun boyunca standart sapma

<sup>171</sup> Zhang, s.47.

<sup>172</sup> M. Kaplan, ve R. Tekeli, *Ekonomide Bekleyişler ve Tahmin: Yapay Sinir Ağları Uygulamaları*, Tablet Yayınları, Konya 2008, s.53.

değerlerini göstermektedir.

#### **2.5.4.6. Eğitim ve Test Setlerinin Belirlenmesi**

Yapay sinir ağı uygulamalarında genellikle veriler eğitim ve test seti olarak ikiye ayrıştırılır. Eğitim seti, ağı geliştirilmesi için, test seti ise modelin performansının değerlendirilmesi için kullanılır. Bazı çalışmalarda, aşırı uyum probleminden kaçınmak ve eğitim sırasında durma noktasını belirlemek için, onaylama veya geçerlilik seti olarak adlandırılan üçüncü bir seti de kullanılmaktadır.<sup>173</sup>

Literatürde bu ayrımı belirlemeye yönelik araştırmacılara, yol göstermesi açısından yapılan birtakım tavsiyeler bulunmaktadır. Bunlar, veri sayısının % 90'ının eğitim veri seti, % 10'unun da test veri seti olarak kullanılması tavsiye edilirken, bu oranlar % 80 eğitim verisi, % 20'si test verisi ve % 70'i eğitim verisi % 30'unun da test verisi olarak da ayrışımın yapılabileceği yönünde literatürde mevcut olan tavsiyelerdir.<sup>174</sup>

#### **2.5.4.7. Yapay Sinir Ağına Eğitimi ve Testi**

Yapay sinir ağlarında öğrenme, temel olarak sinir hücreleri arasındaki ağırlık değerlerinin belirlenmesi olarak ifade edilebilir. Ağa verilen girdilere karşılık, arzu edilen çıktılar elde edilmesi uygun ağırlık değerlerine bağlıdır. Ağı, uygun ağırlık değerlerine kavuşması için örneklerle tekrarlı bir şekilde eğitilmesi işlemine “ağına eğitimi” denir. Eğer yapay sinir ağları, örnekler arasındaki ilişkiyi bulup, genelleştirme yeteneğine kavuşturulmuşsa, ağ öğrenme işlemini gerçekleştirmiş olur.

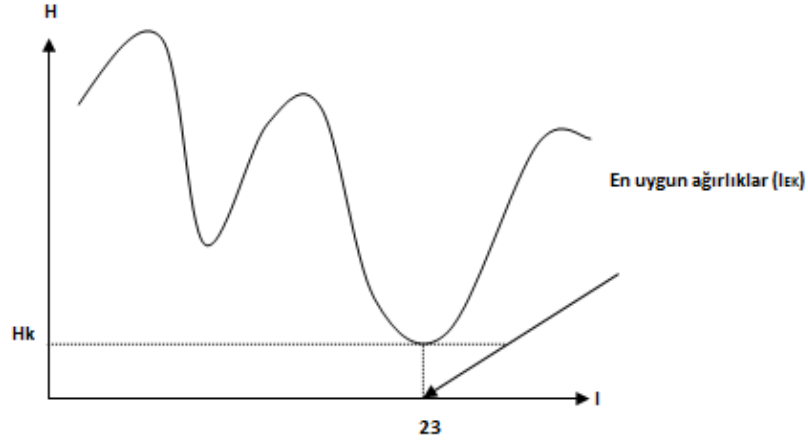
Yapay sinir ağlarında öğrenme, eğitim ve test olmak üzere iki aşamada gerçekleştirilir. Ağa sunulacak örnekler iki gruba ayrılır. Bu gruplardan biri, eğitim aşamasında ağa sunularak uygun ağırlık katsayılarının belirlenmesi sağlanır. Ağına verdiği çıktı değerleri, beklenen değere ulaştığında ağırlıklar sabitlenerek test aşamasına geçilir. Test aşamasında, ağına daha önce hiç görmediği örnekler karşısında ne tür bir davranış sergileyeceği izlenir. Eğer test aşamasında, ağına ürettiği sonuçların hatası ile eğitim aşamasındaki hatalar arasında fark kabul edilebilir bir noktada ise, ağ kullanıma

---

<sup>173</sup> W. Huang, K.K. Lai, Y. Nakamori, and S. Wang, “Forecasting Foreign Exchange Rates With Artificial Neural Network: A Review” [Yapay Sinir Ağlarıyla Döviz Kurlarının Tahmini: Bir Uygulama], *International Journal of Information Technology & Decision Making*, 3(1), 2004, pp. 151. Erişim Tarihi: 17 Aralık 2011, ScienceDirect.

<sup>174</sup> Zhang, s.50.

hazır demektir. Farkın çok yüksek olması durumunda, ağırlıkların tekrar gözden geçirilmesi gerekmektedir.



**Şekil 2.31.** Hata – Eğitim Devresi Sayısı

**Kaynak:** Z. Şen, Yapay Sinir Ağları İlkeleri, Su Vakfı Yayınları, İstanbul 2004, s.93.

Şekil 2.31.'de yapay sinir ağının eğitim sürecindeki çıktı değeri hatasının, eğitim devresi boyunca gösterdiği değişim gösterilmektedir. Geri yayılım algoritması ile eğitim aşamasında eğitim setindeki girdiler kullanılarak çıktılar üretilir. Üretilen çıktılar ile hedef çıktılar arasındaki hata hesaplanarak, bulunan bu hata değeri ağın bağlantı ağırlık değerlerinin değiştirilmesi için kullanılır. Başlangıç aşamasında bağlantı ağırlıklarının değeri rastgele seçildiği için hata kareleri toplamının aritmetik ortalaması H büyük olmaktadır. Eğitim süreci boyunca ileri besleme (iterasyon olarak da ifade edilmektedir) sayısı olan 1 ile ağın hatasının azaltılması beklenmektedir. Yukarıdaki şekilde uygun katsayılar 23.üncü iterasyonda elde edilmiştir.<sup>175</sup>

Yapay sinir ağlarının eğitiminde dikkat edilmesi gereken noktalardan biri de, elde edilen ağırlık katsayılarının en uygun ağırlık katsayılarına ulaşp ulaşmadığıdır. Yapay sinir ağlarının eğitim aşamasında, ağın global minimumu bulması kesin değildir. Çünkü; hata yüzeyi içerisinde, algoritmanın “sıkışabileceği” pek çok yerel minimum noktası bulunabilir.<sup>176</sup>

Yukarıda da ifade edildiği gibi, ağın uygun ağırlık katsayılarına ulaşması için başlangıç ağırlıklarının seçimi, öğrenme oranı ve momentum katsayısını belirleme

<sup>175</sup> Şen, s.93.

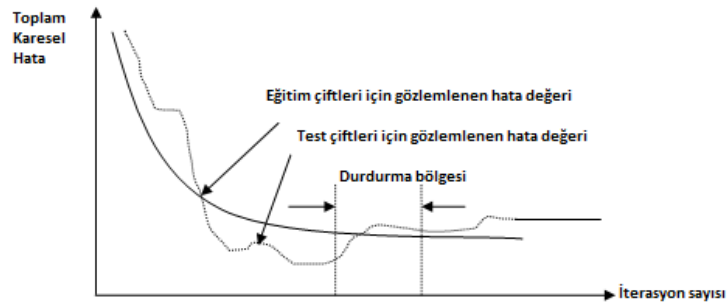
<sup>176</sup> Kaastra and Boyd, s.229.

sürecinde dikkatli karar vermek, ağın performansını etkilemektedir. Eğer eğitim aşamasında, bu oranlar doğru tespit edilirse, ağın doğru sonuçlar üretmesi, yani uygun ağırlık değerlerinin belirlenmesi kolaylaşacaktır.

#### 2.5.4.8. Eğitimin Sonlandırılması

Yapay sinir ağlarında, eğitim aşamasının sonlandırılacağı noktanın belirlenmesi, ağın genelleştirme yeteneğini etkileyebildiği gibi, ağın aşırı eğitilmesi durumunda da ezberleme problemine neden olabilmektedir. Bu nedenle eğitim işleminin durdurulacağı noktanın belirlenmesi, ağ başarısını etkilemektedir. Yapay sinir ağı modellerinde, eğitimin sonlandırılması için üç yöntem vardır. Birinci yöntemde, ağ önceden belirlenen iterasyon sayısına ulaşılması ile eğitim aşaması sonlandırılır. İkinci yöntemde ise, araştırmacı kabul edilebilir bir hata seviyesi belirleyerek eğitime işlemine başlar. Ağ önceden belirlenen bu eğitim seviyesine ulaştığı zaman, eğitim durdurulur.<sup>177</sup>

Son yöntem ise, çapraz geçerlilik olarak da bilinen onaylama seti için, hata oranının yükselmeye başladığı anda, eğitimin durdurulmasıdır. Bu yöntemde, ağın eğitimi için kullanılacak olan örnek seti üç kümeye ayrılır. Ağı eğitimi aşamasında, ağırlık değerleri belirlenirken, geçerlilik setinin hata oranı dikkate alınır. Şekil 2.32.'de görüldüğü gibi, geçerlilik setinin hata oranı, eğitim setinin hata oranını geçtiği noktada eğitim durdurulur. Bu noktada ağırlık değerleri, en uygun ağırlık değerleri olarak alınır.



**Şekil 2.32.** Eğitimin Durdurma Bölgesi

**Kaynak:** M. Ö. Efe, ve K. Okyay Yapay Sinir Ağları ve Uygulamaları, Boğaziçi Üniversitesi Yayınevi, İstanbul 2000, s.22.

<sup>177</sup> Chrispin Ntongo ve Boyd Milton "Commodity Futures Trading performance Using Neural Network Models Versus Arima Models" [Arıma Modellerine Karşı Yapay Sinir Ağları Modellerini Kullanarak Emtia Malların Ticari Performansı], *Journal Of Futures Markets*, 18(8), 1998, pp. 981. Erişim Tarihi: 12 Aralık 2011, ScienceDirect.

### 2.5.4.9. Yapay Sinir Ağı Performans Ölçütünün Belirlenmesi

Yapay sinir ağı modelinin, tahmin sonuçlarının değerlendirilmesi için bir performans ölçütüne ihtiyaç vardır. Performans ölçütü, ağıın verileri arasında mevcut olan ilişkiyi ne kadar öğrendiğini, diğer bir ifade ile ağıın gerçek değere hangi oranda bir hata ile yakınsaklık gösterdiğini belirlemektedir. Seçilen performans ölçütü dikkate alınarak, ağıın katsayıları değiştirilmektedir.

**Çizelge 2.6.** Yapay Sinir Ağları Performans Ölçüleri

Performans Fonksiyonları	Matematiksel açılımı
<b>Hata Kareler Ortalaması (Mean Squared Error – MSE)</b>	$HKO = \frac{1}{n} \sum_{i=1}^n (t_i - g_i)^2 = \frac{1}{n} \sum_{i=1}^n e_i^2$ <p><math>t_i</math> = i.inci birimin tahmin değeri,  <math>g_i</math> = i.inci birimin gerçek değeri,  <math>e_i</math> = Tahmin hatası  göstermektedir.</p>
<b>Hata Kareler Toplamı (Sum Squared Error – SSE)</b>	$HKT = \sum_{i=1}^n (t_i - g_i)^2 = \sum_{i=1}^n e_i^2$ <p><math>t_i</math> = i.inci birimin tahmin değeri,  <math>g_i</math> = i.inci birimin gerçek değeri,  <math>e_i</math> = Tahmin hatası  göstermektedir.</p>
<b>Ortalama mutlak hata yüzdesi (Mean Absolute Percentage Error – MAPE)</b>	$OMHY = 100 \times \left( \frac{\sum_{i=1}^n (t_i - g_i)^2 / t_i}{n} \right)$ <p><math>t_i</math> = i.inci birimin tahmin değeri,  <math>g_i</math> = i.inci birimin gerçek değeri,  <math>e_i</math> = Tahmin hatası  göstermektedir.</p>

## 2.6. YAPAY SİNİR AĞLARININ DİĞER UYGULAMALARDAN AYIRAN AVANTAJLARI

Yapay zekanın alt bir dalı olarak görülen yapay sinir ağları, teknolojik gelişimiyle birlikte, geniş bir uygulama alanı bulmuştur. Özellikle son yıllarda uygulama sonuçlarında sağlanan başarılar, yapay sinir ağlarına olan ilginin artmasına neden olmuştur. Bu ilginin artmasında, yapay sinir ağlarının uygulayıcılara sağladığı üstünlüklerin payı oldukça fazladır. Günümüzde birçok bilim alanında, aşağıda verilen özellikleri nedeniyle yapay sinir ağlarının kullanımında hızlı bir artış görülmektedir.<sup>178</sup>

**Genelleme yapabilmesi:** Genelleme; eğitim ya da öğrenme sürecinde karşılaşılmayan girişler için de ağına uygun tepkileri üretmesi olarak tanımlanır. Yapay sinir ağları, kendisine sunulan bilgileri öğrenmekte ve öğrendiği bilgiler sonucunda genelleştirme yapabilmektedir. Diğer bir ifadeyle, ağ görmediği bilgiler karşısında daha önce edindiği deneyimlerden yola çıkarak sonuçlar üretebilmektedir.

**Paralel olarak çalışma:** Yapay sinir ağını oluşturan bütün yapay sinir hücreleri eşanlı olarak çalışmaktadır. Bu özellik, yapay sinir ağlarını diğer geleneksel bilgisayarlardan daha hızlı hareket etme özelliği katmaktadır.

**Hata toleransı:** Yapay sinir ağları, çok sayıda hücrenin katmanlar şeklinde birbirine bağlanmasından oluştuğu için, paralel dağılmış bir hafızaya sahiptir. Ağına sahip olduğu bilgi, bütün bağlantılar üzerinde dağıtılmış durumdadır. Bu nedenle yapay sinir ağlarının bazı bağlantıları veya hücrelerinin etkisiz hale gelmesi, ağına doğru sonuçlar üretmesini önemli ölçüde etkilememektedir. Bu özelliği ile yapay sinir ağları, geleneksel yöntemlerle kıyaslandığında, hatayı tolere etme yetenekleri son derece yüksektir.

**Doğrusal olmama:** Yapay sinir ağlarının temel işlem elemanı olan yapay sinir hücreleri doğrusal değildir. Dolayısıyla yapay sinir hücrelerinin bağlantılar aracılığıyla oluşturduğu yapay sinir ağı da doğrusal değildir. Bu özelliği ile yapay sinir ağları, doğrusal olmayan karmaşık problemlerin çözümünde alternatif bir yöntemdir.

<sup>178</sup> E. Öztemel, s. 207.

- H. Yurtoğlu, *Yapay Sinir Ağları Metodolojisi ile Öngörü Modellemesi: Bazı Makroekonomik Değişkenler İçin Türkiye Örneği*, DPT yayın 2683, Ekonomik Modeller ve Stratejik Araştırmalar Genel Müdürlüğü, Ankara 2005, s.26.

- V.V. Nabyev, *Yapay Zeka : Problemler-Yöntemler - Algoritma*, Seçkin Yayıncılık, Ankara 2010, s.579.



**Herhangi bir varsayıma gerek duymama:** Yapay sinir ağları ile ilgili teoride herhangi bir varsayım bulunmamaktadır. Hem veri seçiminde ve hem de ağ mimarisinin seçiminde bir kısıtlama getirmemekte ve kısıtlama olmadan sınırsız sayıda veri ile çalışılabilmektedir. Dolayısıyla gerçek dünyanın karmaşık ilişkilerinin modellenmesinde başarılı sonuçlar üretebilmektedirler.

**Dağıtık hafıza:** Yapay sinir ağlarında öğrenme, ağın ağırlıklarının beklenen değeri vermesine kadar devam etmektedir. Ağ istenilen sonuca yakın değerler ürettiği anda, öğrenme durdurulmakta ve bağlantılardaki ağırlıklar sabitlenmektedir. Bu ağırlıklar, ağın doğru sonuçlar üretmesine yardımcı olan bilgi birikimi olarak kabul edilir. Yani ağ, bilgileri sinir hücrelerini birbirine bağlayan bağlantılarda saklamaktadır. Yapay sinir ağlarında çok sayıda bağlantının bulunması ve bu bağlantıların her birinin bir ağırlık değerinin bulunması, ağa dağıtık hafıza özelliği katmaktadır.

**Uyarlanabilirlik:** Yapay sinir ağlarında her bağlantı değerinin bir ağırlığı vardır. Bu ağırlıklar ağın hafızasını oluşturur. Bir problemi çözmek için eğitilmiş yapay sinir ağı, problemde meydana gelen değişimlere göre ağın bağlantı değerlerini bu yeni ortama göre ayarlayabilmektedir. Yapay sinir ağı için bu özellik, gerçek hayatta devamlı değişim halinde olan problemlerin çözümünde önemli bir avantaj getirmektedir. Bu özelliği ile yapay sinir ağları, uyarlamalı örnek tanıma, sinyal işleme, sistem tanılama ve denetim gibi alanlarda etkin olarak kullanılmaktadır.

## 2.7. YAPAY SİNİR AĞLARININ DEZAVANTAJLARI

Yapay sinir ağlarının yukarıda bahsedilen üstünlüklerinin yanında, önemli dezavantajları da bulunmaktadır. Yapay sinir ağlarının dezavantajlarından önemli bir kısmı uygulama ile ilgili problemlerin çözümüne yardımcı olacak ve doğru sonuçların bulunmasını sağlayan kurallar setinin olmamasından kaynaklanmaktadır. Aşağıda bu dezavantajlar kısaca açıklanmıştır.<sup>179</sup>

**Optimum çözümün garanti edilememesi:** Yapay sinir ağları, çözülecek problemlerle ilgili, bulunan çözümün en uygun olduğunu garanti etmez. Çünkü problemin çözümünde kullanılacak ağın yapısı ve ağın bağlantı ağırlıkları deneme-yanılma yöntemi ile belirlenmektedir. Denemeler sonucunda ağın ürettiği sonuçlar,

---

<sup>179</sup> Öztemel, s.208.  
Kaastra and Boynd, pp.233.

gerçek değerlerle karşılaştırılmaktadır. Kabul edilebilir hata seviyesinde, ağıın yapısı belirlenmekte ve eğitimi durdurulmaktadır. Bu yöntem hem uzun zaman almakta hem de optimum çözümün bulunduğunu garanti etmemektedir.

**Uygun eğitim seviyesinin belirlenme sorunu:** Yapay sinir ağları beklenen sonuçları üretmek için eğitim verileri ile tekrarlı bir şekilde eğitilmektedir. Ağıın hatası belirlenen bir seviyenin altına indiğı zaman eğitim durdurulmaktadır. Bu seviyenin neresi olduğı ile ilgili bir kuralın bulunmaması, ağıın en uygun öğrenmeyi gerçekleştirip gerçekleştirmediğı sorusunu akla getirmektedir.

**Açıklama eksikliği:** Yapay sinir ağlarının önemli bir eksiliğı de giriş ve çıkış değışkenlerinin ikişerli ilişkilerinin bilinmesinin imkansızlığıdır. Bunun nedeni; bütün yapay sinir ağlarının bir kara kutu oluşudur.<sup>180</sup> Herhangi bir problemle ilgili üretilen çözümün nasıl ve neden üretildiğı konusunda bilgi vermemektedir. Bu, ağıın ürettiğı sonuca olan güveni azaltmaktadır.<sup>181</sup>

## 2.8. YAPAY SİNİR AĞLARI UYGULAMALARI

Yapay sinir ağları üzerinde binlerce uygulama yapılmış ve başarılı sonuçlar üretmiştir. Geniş bir uygulama alanına sahiptir. Temel olarak yapay sinir ağları uygulamaları ile aşağıdaki işlemler yapılmaktadır:<sup>182</sup>

**Tahmin:** Ağa sunulan giriş değerlerine karşılık çıkış değerlerini tahmin edebilirler.

**Sınıflandırma:** Kendilerine sunulan bilgileri kategorize ederler.

**Veri ilişkilendirme:** Ağlar, ağıın sunulan verilerin hatalı ya da eksik olup olmadığını belirler. Eksik bilgi varsa tamamlar.

**Veri filtreleme:** Ağlar birçok veri arasından uygun verileri bulurlar.

**Tanım ve eşleştirme:** Değışik şekil ve örüntülerin tanınması, eksik, belirsiz bilgilerin eşlenmesi işlemini yaparlar.

**Teşhis:** Ağlar sistemlerin olumsuzluklarını ortaya çıkarır ve problemlerin teşhis edilmesini sağlar.

---

<sup>180</sup> Şen, s.18.

<sup>181</sup> Öztemel, s.35.

<sup>182</sup> Elmas, s.161.

**Yorumlama:** Bir olay hakkında toplanan verilerle ađ eđitildikten sonra bu bilgileri kullanarak yeni olayları yorumlarlar.

Yukarıdaki verilen fonksiyonları gerekleřtiren yapay sinir ađları ok farklı alanlarda kullanılmaktadır. Alanlara gre yapay sinir ađları uygulamaları:<sup>183</sup>

### **Endüstriyel Uygulamalar**

- Kalite denetimi,
- Kalite geliştirme,
- Hidroelektrik santralinde bulunan 4 büyük senkron kompanzatörün gü dağıtım şebekelerinde kullanımı,
- Ark fırın denetim sistemi ve bu sistemin tasarrufu,
- Güneş enerjisi pilleri üretiminde konum kalınlık hatalarının saptanması,
- Petrol endüstrisi için rafine işlemleri,
- Elektrik motorlarında arıza tespiti,
- Konuşmayı metne çevirme,
- Otomatik dil çevirisi,
- İşitme ve bedensel engelliler için protez geliştirilmesi,
- Sesle kumanda etme,
- Doğal diş işleme,
- Fırınların ürettiđi gaz miktarını tahmin etme,
- Ürün tasarımı,
- Makinelerin bakımı,
- Kimyasal proseslerin dinamik modellenmesi,
- Otomobillerde otomatik rehber sisteminin geliştirilmesi,
- Robotlarda görme sistemi ve kontrolü,
- Cep telefonlarında ses ile çalışabilme,
- Araba pistonlarının üretim şartlarının belirlenmesi
- Elektronik yonga hata analizleri,
- Optimizasyon çalışmaları,
- Müşteri tahmini, pazar verilerinin deđerlendirilmesi ve analiz edilmesi,
- Kömür gü istasyonları için çevrimii karbon akımı ölçülmesi,

---

<sup>183</sup> Öztemel, s.203.  
Elmas, s.

- İşlerin makinelere atanması ve çizelgeleme,
- Gezgın satıcı problemlerin çözümü

#### **Mühendislik Uygulamaları**

- Kontrol akış örüntülerinin sınıflandırılması,
- Günlük sıcaklık tahmini,
- Endüstriyel robot sensör modellemede sinerji,
- Sistem kimliklendirme,
- Robotik uygulamalar,
- Mikroşerit antenler,
- Dairesel mikroşerit anten tasarımı,
- Dikdörtgen mikroşerit anten tasarımı,
- Raster verilerde geometrik dönüşüm,
- Katı modelleme,
- Sayısal görüntülerin kodlanması,
- Uzaktan algılama,
- İşaret işleme

#### **Finansal Uygulamalar**

- Makro ekonomik tahminler,
- Borsa benzetim çalışmaları endekslerinin tahmini,
- Kredi kart hilelerinin tespiti,
- Kredi kartı kurumlarında iflas tahminleri,
- Banka kredilerinin değerlendirilmesi,
- Emlak kredilerinin değerlendirilmesi,
- Döviz kuru tahminleri,
- Risk analizleri,
- El yazısı ve imza tanıma

#### **Askeri Uygulamalar**

- Hedef tanıma ve takip sistemleri,
- Yeni sensörlerin performans analizleri,
- Radar ve görüntü sinyalleri işleme,
- Sensör fizyonu,

- Askeri uçakların uçuş yörüngelerinin belirlenmesi (optimizasyonu)
- Mayın dedektörleri

### **Sağlık Uygulamaları**

- Solunum hastalıklarının teşhisi,
- EEG, ECG analizleri,
- Transplant zamanlarının optimizasyonu,
- Hastalıkların teşhisi ve resimlerden tanınması,
- Kardiyovasküler sistemlerin modellenmesi ve teşhisi,
- CTG izleme,
- Hamile kadınların karınlarındaki çocukların kalp atışlarının izlenmesi,
- Yumurtalık kanserinin izlenmesi,
- Üroloji uygulamaları,
- RNA ve DNA'daki aminosait incelemesi,
- Hastaların ilaçlara karşı tepkilerinin incelenmesi,
- Kazalarda sakatlıklardan korunma,

### **İletişim Sanayi Uygulamaları**

- Faks ve telefon hatlarının güncellenmesi,
- Hatlarda olabilecek yankıların giderilmesi,
- Veri sıkıştırma

### **Ulaştırma ve Havacılık Sanayi Uygulamaları**

- Fren denetimi,
- Mühendislik hataları,
- Etkin süspansiyon,
- Uzay mekiğinde manevra denetimi,
- Uçaklarda titreşim seviyeleri ve sesin görüntülenerek motor sorunlarının erken uyarı sistemi,
- Kara taşıtın sürücüsüz seyri,
- Bir helikopterin denetimi,
- Hava alanlarında bomba tanıma ve ortaya çıkarma,
- Uçak parçalarının hata teşhislerinin yapılması,
- Rotolama sistemleri.

## ÜÇÜNCÜ BÖLÜM

### UYGULAMA

Üçüncü bölüm iki kısımdan oluşmaktadır. İlkinde finansal başarısızlığın tanımı ve önemi ele alınmıştır. İkincisinde ise tezin ampirik çalışması yer almaktadır.

#### 3.1. FİNANSAL BAŞARISIZLIK

İşletmeler temel misyonları olan “kâr elde etmek” ya da “firma sahiplerinin refahını maksimize etmek” hedeflerini kimi zaman kendi içlerinde yaşadığı krizler nedeniyle kimi zaman dış çevredeki faktörler veya kaynaklar nedeniyle kimi zaman da buldukları ülke ekonomisindeki krizler nedeniyle yerine getirememektedirler. Zaman içerisinde işletmeyi ana amacından uzaklaştıran bu faktörler artarak kurumun başarısızlığa sürüklenmesine neden olabilmektedir.

Sözlüklerde genel olarak başarısızlık terimi, “eksik olmak, yetersiz olmak”, finansal başarısızlık ise “iflas etme, acze düşme” olarak tanımlanmaktadır.<sup>184</sup>

Finansal başarısızlığı bütün yönleriyle tanımlamak bir hayli zordur. Ekonomilerde küçük ölçekli, büyük ölçekli bütün firmalar başarısızlık olgusuyla yüz yüze gelebilirler. Finansal başarısızlık; firmaların politikalarında, aldıkları finansal kararlarda ve firmanın diğer alanlarında olan başarısızlıkların sonucunda ortaya çıkan hedeflere ulaşamama durumudur. Firma içinde bir bölümün başarısız olması, firmanın alt bir kuruluşunun başarısız olması veya firmanın tüm bölümlerinin bütün olarak başarısız olması durumu iş dünyasında her zaman görülebilecek durumlardır. Firmaların hedeflerine ulaşamaması, borçlarını zamanında ödeyememesi, kredibilitesini düşürmesi ya da kaybetmesi yani sorumluluklarını kısmen ya da tamamen yerine getirememesi başarısızlık olarak nitelendirilebilir.<sup>185</sup>

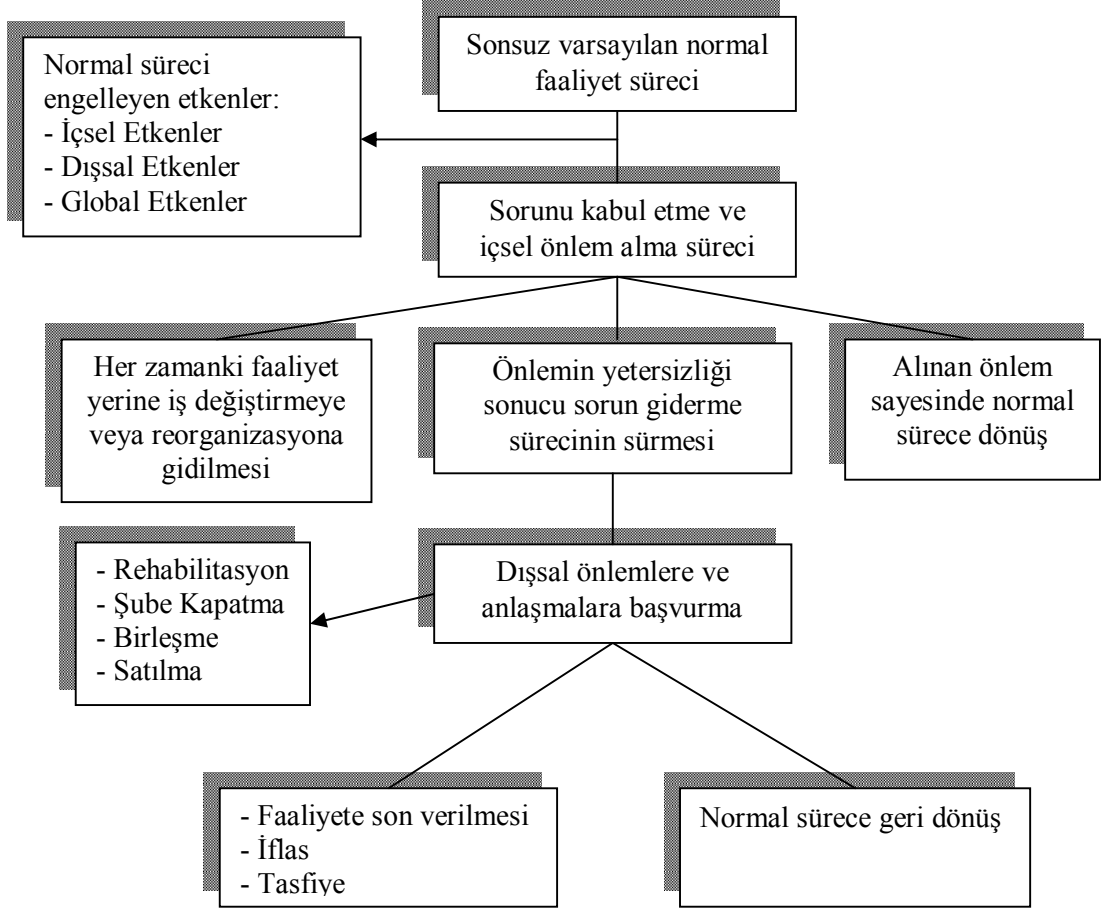
Finansal başarısızlık literatürde genel olarak iflastan önce gerçekleşen ve şirketin iflas sürecindeki yollardan biri olarak ele alınmıştır. Finansal başarısızlığa düşmüş bir işletme normal faaliyet sürecine dönmek için tüm yolları denemiş ve sonuç alamamışsa,

<sup>184</sup> Karl Gratzer, “Business Failure and The New Economy” [İşletme Başarısızlığı ve Yeni Ekonomi], *EBHA Conference 2001*, Business and Knowledge: July 2001, 20p, Oslo Norway 2001, pp.2.

<sup>185</sup> O. Okka, *Analitik Finansal Yönetim*, Nobel Yayın Dağıtım, Ankara 2009, s.928.

işletmenin varlığını sürdürmesi zorlaşır ve tasfiyeye gitmekten başka çaresi kalmayabilir.<sup>186</sup>

İşletmelerin finansal başarısızlıklarının oluşum süreci, karşılaştığı bu soruna karşı aldıkları önlemler ve ulaşılabilecek sonuçları şekil 3.1’de gösterilmiştir.



**Şekil 3.1.** Finansal Başarısızlığın Oluşum Süreci

**Kaynak:** Y. Kılıç, Finansal Başarısızlık Tahmininde Veri Madenciliğinin Kullanılması: İMKB’de Bir Uygulama, (Yayımlanmamış Yüksek Lisans Tezi), Gaziantep Üniversitesi Sosyal Bilimler Enstitüsü, Gaziantep 2011, s.5.

<sup>186</sup> Y. Kılıç, *Finansal Başarısızlık Tahmininde Veri Madenciliğinin Kullanılması: İMKB’de Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Gaziantep Üniversitesi Sosyal Bilimler Enstitüsü, Gaziantep 2011, s.5.

Buna göre işletmenin normal faaliyet süreci içsel, dışsal ve global etkenler dolayısıyla kesintiye uğramıştır. Normal faaliyet sürecine tekrar dönmek için işletmede bazı önlemler alınmıştır. Alınan önlemlerin başarılı ya da başarısız sonuçlar doğurmasına göre işletme normal sürece dönmüş, iş değiştirmeye gitmiş ya da sorun giderme sürecine devam etmiştir. Devam eden sorunları engellemek için işletme dışsal önlemlere başvurmuş ve bu durum sonucunda işletme ya tamamen normal sürece dönmüştür ya da faaliyetlerine son vermek zorunda kalmıştır.

Finans literatüründe yaygın olarak kullanılan finansal başarısızlık türleri ve bunların tanımları aşağıda yer almaktadır.<sup>187</sup>

**Ekonomik Başarısızlık:** Ekonomik açıdan başarısızlık; bir işletmenin gelirlerinin, sermaye maliyetini de içeren toplam maliyetlerini karşılayamamasını ifade etmektedir. Ekonomik açıdan başarısız olan işletmeler, ancak yatırımcıların ilave sermaye sağlamaya istekli olmaları durumunda veya işletme sahiplerinin piyasanın altında getiri oranlarını kabul etmeye razı olmaları halinde faaliyetlerini sürdürebilmektedir. Fakat eninde sonunda ilave sermaye sağlanamaz hale gelir ve işletme varlıkları tükenir. Bu durumdaki işletmeler ya kapanır ya da normal bir getirinin sağlanacağı daha küçük bir işletmeye dönüşürler.

**İşletme Başarısızlığı:** İşletme başarısızlığı terimi, dünyanın en güvenilir uluslararası derecelendirme kuruluşlarından biri olan Dun&Bradstreet tarafından “bir işletmenin kredi verenlerin zarar etmelerine neden olacak şekilde faaliyetlerini durdurması” olarak tanımlanmaktadır. Böylece bir işletme, yasal iflas sürecine girmese bile başarısız olarak değerlendirilebilir. Bu tanıma göre, borç verenleri zarara uğratmadıkça bir işletmenin faaliyetlerini durdurması veya kapanması işletme başarısızlığı olarak görülmemektedir.

**Teknik Acizlik:** Bir işletme, vadesi geldiğinde cari yükümlülüklerini karşılayamıyorsa teknik olarak borçlarını ödeyemez (aciz) sayılır. Teknik acizlik, belirli bir zamandaki geçici likidite noksanlığını gösterebilir. Dolayısıyla teknik olarak aciz olan bir işletme nakit sıkıntısı sorununu çözebilir, yükümlülüklerini tamamıyla yerine

<sup>187</sup> T. Torun, *Finansal Başarısızlık Tahmininde Geleneksel İstatistikî Yöntemlerle Yapay Sinir Ağlarının Karşılaştırılması ve Sanayi İşletmeleri Üzerinde Uygulama*, (Yayımlanmamış Doktora Tezi), Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, Kayseri 2007, s.5.



getirebilir ve faaliyetlerini sürdürebilir. Diğer taraftan, teknik acizlik ekonomik başarısızlığın belirtisi olarak ortaya çıktığında, işletme için ciddi sorunlara neden olabilmektedir.

**Negatif Net Değere Sahip Olma:** Toplam yükümlülüklerinin defter değeri, varlıklarının gerçek değerinden daha fazla olan işletme negatif net değere sahiptir. Bu, teknik acizlikten daha önemli bir durumdur. Çünkü, negatif net değere sahip olma genellikle ekonomik başarısızlığın bir işaretidir ve sıklıkla işletmenin tasfiyesiyle sonuçlanır. Ancak, negatif net değere sahip olan işletmelerin, kimi zaman yasal iflas sürecine girmeyeceğini burada belirtmek gerekmektedir.

**İflas:** İflas terimi, hukuki bir özellik taşımaktadır. Bazen, başarısız bir işletmeyi belirtmek için kullanılmasına rağmen, mahkeme tarafından ilan edilmedikçe hiçbir işletme hukuki olarak iflas etmiş sayılmaz.

**Çizelge 3.1. Mali (Finansal) Başarısızlık/İflas Tanımı**

Yazar	Kullanılan Kavram	Tanım
Altman	İflas	Yasal olarak iflas etmiş ve kayyum atanmış ya da ulusal iflas yasası hükümlerince reorganizasyon hakkı verilmiş işletmeler.
Beaver	Başarısızlık	Vadesi gelen finansal yükümlülüklerini ödeyememe. Araştırma kapsamına mali başarısızlık olarak şu olaylar alınmıştır: İflas, tahvil faizinin ödenmemesi, karşılıksız çek yazılması, imtiyazlı hisse senetlerine temettü dağıtılmaması
Blum	Başarısızlık	Vadesi gelen borçlarını ödeyememe, iflas sürecine girme, alacaklılarla borçların azaltılması konusunda açık anlaşma yapma.
Deakin	Başarısızlık	İflas etmiş, ya da alacaklıların isteği üzerine tasfiye edilmiş işletme.
Edmister	Başarısızlık	Hem Beaver hem de Blum'un tanımlarını kullanmıştır.
Elam	İflas	İflas yasası hükümlerine göre iflas etmiş sayılan işletmeler.
Taffler	İflas	Başarısızlık tasfiye, alacaklıların isteği üzerine tasfiye ve mahkeme kararıyla faaliyetine son vermek olarak tanımlanmıştır.

**Kaynak:** R. Aktaş, M. Doğanay ve B. Yıldız, "Mali Başarısızlığın Öngörülmesi İstatistiksel Yöntemler ve Yapay Sınır Ağı Karşılaştırması", Ankara Üniversitesi Siyasal Bilgiler Fakültesi Dergisi, 58(4), 2003, s.7.

### 3.1.1. Finansal Başarısızlığın Nedenleri

İşletmelerde finansal başarısızlık hem işletmeleri hem de işletmenin içinde bulunduğu ülkeyi yakından ilgilendirir. Ülke açısından yaklaşıldığında finansal başarısızlık, istihdam sorunu yaratmakta, tasarruf sahiplerini üzerinde olumsuz bir etki bırakmakta, elverişli kredi kullanılabilirliğini düşürmekte kısaca ülke ekonomisini olumsuz yönde etkilemektedir.<sup>188</sup>

Uluslararası derecelendirme kuruluşu Dun&Bradstreet tarafından, 1987 yılında yürütülen çalışmaya göre, işletme başarısızlıkları aşağıdaki beş temel faktörden kaynaklanmaktadır:<sup>189</sup>

- Ekonomik faktörler, Yönetimin tecrübesi, Satışların azalması, Giderlerin artması,
- Diğer muhtelif faktörler.

Söz konusu rapora göre, işletme başarısızlığının en önemli nedeni olarak ekonomik faktörler gösterilirken, ikinci sırada yönetim tecrübesi yer almaktadır.

Herhangi bir işletmenin başarısı ya da başarısızlığı, iki temel faktör kümesinin etkileşiminin bir sonucudur. İlk olarak, bir işletmenin performansı işletme dışı faktörler tarafından etkilenir ve bu faktörler işletme yöneticilerinin kontrolü dışındadır. Ekonominin büyüme oranı, enflasyon, döviz kuru, faiz oranı, tercihlerin, tutumların ve tüketici davranışlarının değişmesi, pazardaki faaliyet karakteristiklerinin değişmesi gibi çevresel koşullar açık bir şekilde işletmenin pazar gücünü ve kârlılığını etkiler.<sup>190</sup> Bir işletmenin performansını etkileyen diğer temel faktörler ise işletmenin kendi içerisinde ortaya çıkan ve kontrol edilmesi mümkün görülen faktörlerdir.

#### 3.1.1.1. İşletme Dışı (Dışsal) Faktörler

İşletmelerin normal faaliyet süreçlerini toplumsal faktörler, yasal ve politik çevre, ekonomik faktörler ve doğal çevre işletme dışından etkileyebilmektedir.

<sup>188</sup> R. Aktaş, M. Doğanay, ve B. Yıldız, “Mali Başarısızlığın Öngörülmesi İstatistiksel Yöntemler ve Yapay Sinir Ağı Karşılaştırması”, *Ankara Üniversitesi Siyasal Bilgiler Fakültesi Dergisi*, 58(4), 2003, s.3.

<sup>189</sup> Torun, s.8.

<sup>190</sup> Subhash Sharma and Mahajan Vijay, “Early Warning Indicators of Business Failure” [İşletme Başarısızlığının Erken Uyarıcı Göstergeleri], *Journal of Marketing*, 44, Fall 1980, pp.82, Erişim Tarihi: 24 Şubat 2012, Web of Science

- **Toplumsal Faktörler:** İşletmeler başarılı olabilmek için toplumun beklentilerini bilmek ve bu beklentilere uygun olarak faaliyetlerini devam ettirmek zorundadırlar. Tekelci uygulamalardan kaçınmak, kaliteli mal veya hizmet üretmek, tüketici haklarına saygılı olmak, çevreye karşı duyarlı olmak toplumsal çevrenin beklentilerinden bazılarıdır.<sup>191</sup>
- **Ekonomik Faktörler:** İşletmeler ekonomik sistemin bir parçasıdır ve faaliyette buldukları ülkenin ekonomik koşullarından etkilenirler. İşletmenin olumsuz durumlar ile karşılaşmasına neden olabilecek faktörler; faiz oranlarındaki ani yükseliş veya azalış, enflasyon oranındaki beklenmedik değişiklikler, döviz kurlarındaki ani dalgalanmalar, ithalat ve ihracat rejiminde yapılan değişiklikler ve uygulanan para politikaları şeklinde sıralanabilir.<sup>192</sup>
- **Yasal ve Politik Çevre:** Hukuki ve politik düzenlemelere uymayan işletmelere cezai yaptırımlar getirilmiştir. İşletmelerin cezai müeyyideler ile karşı karşıya kalmaları, ağır cezaların ödenmesine neden olabileceği gibi, toplumda işletmenin küçük düşmesine de neden olabilir. Ayrıca hukuki ve politik çevre faktörlerine işletmelerin uyum sorunu göstermeleri, söz konusu işletmelerin başarısız olmalarına neden olabilir.<sup>193</sup>
- **Doğal Çevre:** İşletmeler açısından doğal çevre, üretimde kullanılan doğal kaynakları ifade etmektedir. Doğal çevrenin sürekli olarak değişmesi ve gelişmesi, işletmeler için bazı fırsatlar sunabileceği gibi başarısız olmalarına da neden olabilmektedir. Doğal çevre, özellikle doğal kaynakların tükenmesi ve çevre kirliliği gibi nedenlerden dolayı işletme faaliyetleri üzerinde etkili olmaktadır. Ayrıca, deprem, su baskını gibi işletmenin başarısızlığına neden olabilecek doğal afetleri de bu faktörler arasında saymak gerekmektedir.<sup>194</sup>

<sup>191</sup> R. Metin Türko, *Finansal Yönetim*, Alfa Yayın, İstanbul 1999, s.599.

<sup>192</sup> S. Büker, R. Aşıkoğlu, ve G. Sevil, *Finansal Yönetim*, (2. Baskı), Anadolu Üniversitesi, Eskişehir 1997, s.526.

<sup>193</sup> Büker, vd., s.525.

<sup>194</sup> Türko, s.600.

### 3.1.1.2. İşletme İçi Faktörler

İşletmenin kontrolü altındaki işletme içi faktörlerin sayısı oldukça fazla olmakla birlikte, genel hatlarıyla aşağıdaki başlıklar altında toplanabilirler:<sup>195</sup>

- Kötü yönetim,
- Çevredeki değişime uyum gösterememe,
- Yetersiz iletişim,
- Sağlıksız büyüme,
- Ana projelerde başarısızlık.

Bir firmanın başarısızlığa uğramasına, finansal yapısının zayıflamasına yol açan etmenlerin başında yönetim hataları gelmektedir. Firma yöneticilerinin, bir yöneticide bulunması gerekli yetenek ve niteliklerden yoksun oluşları firmanın varlığını tehlikeye düşürebilmektedir.<sup>196</sup>

### 3.1.2. Finansal Başarısızlık Tahminin Önemi

Bir ülkede fazla sayıda başarısız işletme bulunması, o ülkede refah kayıplarına ve makroekonomik sorunlara neden olabilir. Ülkedeki işletmelerin finansal açıdan sorunlar yaşaması bankaların sermayesini aşındırarak bankacılık sistemini zayıflatabilir ve finansal krizlere yol açabilir. Başarısız işletmelerin sayısının artması kullanılabilir kredilerin azalmasına neden olacak ve kaynakların daha değerli yatırımlarda kullanılmasını engelleyerek verimliliğin düşmesine yol açacaktır. Ayrıca başarısızlıktaki bu artış ülkedeki işsizlik oranının artmasına da neden olacaktır.<sup>197</sup>

İşletme başarısızlığı, sadece işletmenin doğrudan çevresi için değil, genel ekonomi için de çeşitli maliyetler ortaya çıkarmaktadır. Geniş çevreye sahip bir işletmenin başarısız olmasının maliyetleri, “yayıma etkisiyle” tüm ekonomide negatif bir etki oluşturabilir. Böylece, işletme başarısızlığı, istihdam ve ekonomik refahla ilgili olumsuz sonuçlara neden olabilir. Sonuç olarak, işletme başarısızlığının tahmini, sadece bireysel bakış açısıyla belirli kişi ya da gruplar için değil, bir bütün olarak toplum için de önemlidir.<sup>198</sup>

<sup>195</sup> Yasemin Keskin, *İşletmelerde Finansal Başarısızlığın Tahmini, Çok Boyutlu Model Önerisi ve Uygulaması*, (Yayımlanmamış Doktora Tezi), Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara 2002, s.16.

<sup>196</sup> Öztin Akgüç, *Finansal Yönetim*, (Yenilenmiş 7. Baskı), Avcıol Basım-Yayın, İstanbul 1998, s.948.

<sup>197</sup> Kılıç, s.22.

<sup>198</sup> Torun, s.15.

Finansal başarısızlık tahmin modelleri yardımıyla yatırımların verimli ve uygun alanlara aktarılması ülke ekonomisi açısından önemli katkılar sağlayacaktır.

### 3.1.2.1. İşletme Yöneticileri Açısından Önemi

Başarılı bir şekilde büyümek isteyen her işletme, mevcut durumunu dönemsel olarak nesnel bir biçimde gözden geçirmek zorundadır. Bu gözden geçirme sonucunda, işletmenin önemli denebilecek kuvvetli ve zayıf yönleri saptanabilir. Bu değerlendirme, mevcut yönetim politikalarında değişiklik yapıp yapılmaması konusunda yöneticilere yardımcı olacaktır. İşletmenin yönetim politikasında değişiklikler yapılması, uzun vadede başarıya ulaşması açısından çok önemlidir. Doğru, tutarlı ve düzenli aralıklarla sürekli olarak finansal başarısızlık tahmininin kullanılması durumunda ileride ortaya çıkabilecek sorunların önceden bilinmesini ve gerekli önlemlerin alınmasını mümkün kılan bir “erken uyarı sistemi”dir.<sup>199</sup>

Finansal başarısızlığı öngörebilecek bir modelin yöneticiler tarafından kendi işletmeleri için kullanılması, zamanında fon bulma, yeni işbirlikleri kurma, işletmenin piyasa değerinde bir düşüş olmadan devredilme yollarının araştırılması gibi konularda, yöneticilere avantaj sağlayacaktır. İşletme yöneticileri finansal başarısızlık öngörü modelini ticari ilişki içerisinde buldukları kişi ya da kuruluşlara da uygulayabilirler ve finansal başarısızlık yaşayacağı öngörülen kuruluşlar ile ilişkilerini yeniden düzenleyebilirler. Aynı zamanda kredili satış yaptıkları müşterilerini de değerlendirebilirler. Diğer taraftan tedarikçi işletmelerin yaşayacağı bir finansal başarısızlık da işletmeyi de zor duruma düşürebilir. Finansal başarısızlık öngörü modeli tedarikçi işletmeler üzerinde de uygulanabilir. Zaman zaman finansal başarısızlık öngörü modeli sayesinde tedarikçi işletmeler kontrol edilebilir. Böylece işletme yöneticileri işletme faaliyetlerindeki olası aksamaları engelleyebilirler.<sup>200</sup>

### 3.1.2.2. Kredi Kurumları Açısından Önemi

Kredi kararlarının doğruluğu, kredi verenler açısından olduğu kadar, genel ekonomi açısından da son derece önemlidir. Yanlış bir karar, kredi verenin faiz gelirinden mahrum kalmasına ve anaparayı kaybetmesine neden olabileceği gibi, batık

<sup>199</sup> R. Aktaş, “Endüstri İşletmeleri İçin Mali Başarısızlık Tahmini (Çok Boyutlu Model Uygulaması)”, *Türkiye İş Bankası Kültür Yayınları*, Yayın No: 323, Ankara 1993, s.14.

<sup>200</sup> S. Akkoç, *Finansal Başarısızlığın Öngörülmesinde Sinirsel Bulanık Ağ Modelinin Kullanımı ve Ampirik Bir Çalışma*, (Yayınlanmamış Doktora Tezi), Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü, Kütahya 2007, s.26.

kredilerin sayısındaki artış genel ekonomi açısından da ciddi sorunlar doğurabilecektir.<sup>201</sup>

Kredi veren kurumlar, özellikle bankalar, yeni bir kredi verecekleri zaman ya da daha önceden vermiş olduğu bir kredisinin vadesini uzatmadan önce, müşterisinin krediyi geri ödemede temerrüde düşme ihtimalinin tahmin etmeye çalışırlar.<sup>202</sup> Kredi kurumları tarafından kurulabilecek finansal başarısızlık tahmin modeli, doğru işletmeleri seçmeleri bakımından kredi kurumlarına büyük fayda sağlar. Böyle bir modelle müşterilerine verecekleri kredilerin değerlendirmesinde daha başarılı ve daha hızlı sonuçlar üretebilirler.

Kredi kurumlarının başarısız olan bir işletmeye kredi kullandırmalarının fırsat maliyeti de vardır. Öncelikle söz konusu kredi verimsiz bir şekilde değerlendirilmiş olacak, diğer taraftan da bu kredinin başarılı bir işletmeye kullanılması halinde gerçekleşecek olan üretimden de vazgeçilmiş olacaktır. Kıt olan kaynakların etkin kullanımının gerçekleştirilememesi ekonomik açıdan bazı sorunları beraberinde getirecektir.<sup>203</sup>

### 3.1.2.3. Yatırımcılar Açısından Önemi

İşletmeler sermaye ihtiyaçlarını genel olarak kredi kuruluşlarından karşılamaktadırlar. Bunun yanında hisse senedi ve tahvil gibi menkul kıymetler aracılığıyla çok sayıda küçük yatırımcıdan da fon gereksinimini sağlayabilmektedirler. İşletme içerisindeki mevcut yatırımcılar ya da potansiyel hisse senedi alıcıları gibi çok sayıdaki yatırımcı, yatırım kararlarını vermeden önce işletme ile ilgili bilgilerden yararlanmaktadırlar. Mali başarısızlığın henüz başarısızlık ortaya çıkmadan, verimli ve doğru bir şekilde tahmin edilebilmesi bu yatırım kararlarının verilmesinde büyük önem taşımaktadır.<sup>204</sup>

Finansal başarısızlık öngörü modeli yatırımcılara başarılı işletmelere yatırım yapma kararı konusunda yardımcı olabileceği gibi başarısızlık yaşayan işletmelerin başarısızlıktan kurtulabileceğini tahmin etmelerini sağlayarak bu işletmelerin hisse senetlerini düşük fiyatlarla alıp, yüksek kâr elde etmelerini de sağlayabilecektir. Yatırım

<sup>201</sup> Torun, s.17.

<sup>202</sup> A. F. Atiya, "Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results", [Yapay Sinir Ağlarını Kullanarak Kredi Riski İçin Başarısızlık Tahmini, Bir Uygulama ve Sonuçlar], *IEEE Transaction on Neural Networks*, 12(4), 2001, pp.932.

<sup>203</sup> Akkoç, s.26.

<sup>204</sup> Kılıç, s.24.

konusundaki bu yararlarına ek olarak, zararlı ve verimsiz alanlara yatırım yapılmasını önlemek suretiyle ülkenin yatırılabılır fonlarının optimum kullanımına da yardımcı olabilir.<sup>205</sup>

İşletmelerin büyümek amacıyla gerçekleştirdikleri yatırımlar, fon ihtiyacını doğurur. Söz konusu fon gereksinimi kredi kuruluşlarından karşılanabileceği gibi sermaye piyasasından da karşılanabilir. İşletmeler sermaye piyasası aracılığıyla hisse senedi ihraç ederek yeni ortak sahibi olabileceği gibi tahvil ihraç ederek de uzun vadeli borçlanabilirler. Bu durumdaki bir işletmeye ortak olan ya da borç veren yatırımcılar için işletmenin ileride yasayabileceği finansal başarısızlık büyük önem arz eder.<sup>206</sup>

#### **3.1.2.4. Bağımsız Denetçi ve Analistler Açısından Önemi**

Genellikle finansal sıkıntı içinde bulunan işletmelere ait finansal çizelgelerin doğruluğu ve güvenilirliği konusunda önemli riskler bulunmaktadır. Bunun anlamı daha yüksek kontrol ve bulgu riski demektir. Finansal başarısızlığın öngörülmesi bağımsız denetçilere inceledikleri işletmenin finansal durumu ve dolayısıyla denetim riski hakkında daha fazla bilgi sağlayacaktır. Böylelikle denetçi müşteri işletmeyi seçme aşamasında veya görüşünde daha doğru sonuçlara ulaşmış olacaktır. Denetçiler denetim yaptıkları işletmenin finansal başarısızlık potansiyeli taşıyıp taşımadıklarını bir model aracılığı ile öngörebildiklerinde, bağımsız denetim için başvuran işletmelerin değerlendirilmesi ve denetimin planlanması hızlanacak ve kolaylaşacaktır.<sup>207</sup>

Finansal çizelgeler kullanılarak gerçekleştirilen finansal başarısızlık öngörü çalışmaları dış denetçiler için ayrı bir önem taşımaktadır. Bu kesim finansal başarısızlık öngörü çalışmalarının doğrudan kullanıcısı olmasa bile, genel bilgi düzeyindeki yatırımcıların bilgilendirilmesi işlevinde önemli bir görevi yerine getirmektedirler.<sup>208</sup>

---

<sup>205</sup> Aktaş, s.19.

<sup>206</sup> Akkoç, s.27.

<sup>207</sup> Yıldız, s.23.

<sup>208</sup> Akkoç, s.56.

### **3.2. VERİLERİN DEĞERLENDİRİLMESİ VE ANALİZİ**

Bu çalışmada, veri olarak hisse senetleri İstanbul Menkul Kıymetler Borsası'nda (İMKB) işlem gören sanayi işletmeleri ele alınmıştır. Bunun temel nedeni İMKB'de işlem gören sanayi işletmelerine olan ulaşım kolaylığıdır. İMKB'nin web sayfasında şirketlerin 1998 yılından itibaren 3,6,9 ve 12 aylık mali tablolarından ve dipnotları sürekli güncellenerek, kullanıcıların hizmetine sunulmaktadır. Veri seti olarak İMKB'de işlem gören şirketlerden mali kuruluşlar, ticaret ve hizmet işletmeleri farklı özellikler taşıdığı için örneklem dışı bırakılmış ve kapsama 120 sanayi işletmesinin 2002-2010 yılları arasındaki verileri ele alınmıştır.

#### **3.2.1. Uygulamada Kullanılan Yöntemler ve Yazılımlar**

Finansal başarısızlığı tahmin edebilecek modeller kurulurken; yapay sinir ağı, veri madenciliği tekniklerinden entropiye dayalı C5.0 algoritması ve destek vektör makineleri yöntemleri kullanılmıştır. Yapay sinir ağları modelinin kurulmasında STATISTICA 7 yazılımından, C5.0 algoritması ve destek vektör makineleri yöntemleri için CLEMENTINE 12.0 ve model kurulma aşamasında SPSS 18.0 for Windows paket programından yararlanılmıştır. Değişken sayılarını azaltmak için lojistik regresyon analizi ve diskriminant analizine ait modellerin kurulmasında stepwise yöntemi ele alınmıştır.

#### **3.2.2. Uygulamanın Konusu ve Amacı**

Bir işletmenin operasyonel durumu periyodik olarak açıklanan finansal bildirimlerle gösterilir. Bu finansal bildirimlerinden yola çıkarak işletmelerin finansal başarı durumları tespit edilir. Finansal başarısızlığa uğramış işletmelerin sayısındaki artış işletmelerin hem kendi kaynaklarının hem de ülke kaynaklarının iyi kullanılmadığının bir göstergesi durumundadır. Bu nedenle finansal başarısızlığın tahmin edilmesi önem arz etmektedir.

Çalışmada 2002-2010 yılları arasında, hisse senetleri İMKB'de işlem gören 120 sanayi işletmesi ele alınmıştır. Ele alınan işletmelerin bilanço ve gelir tabloları incelenerek belli kriterlere göre "finansal başarılı" ve "finansal başarısız" olarak sınıflandırma yapılmıştır. Finansal başarısız işletmeler aşağıdaki kriterlere göre tespit edilmiştir:



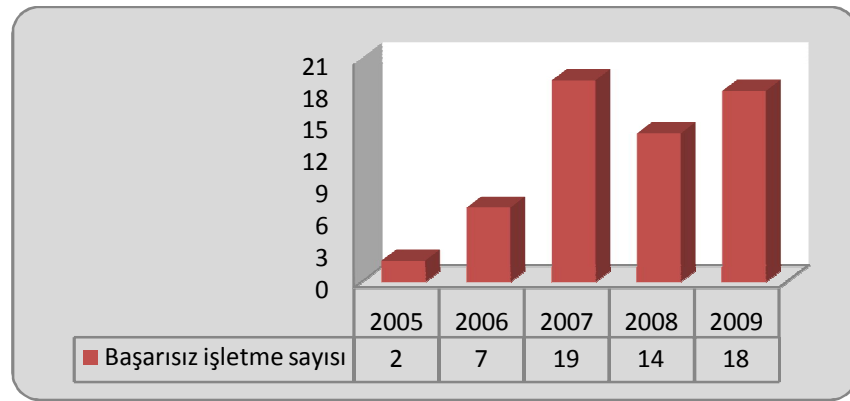
- İflas etmiş olmak,
- Borsada tahtası kapanmış olmak,
- Faaliyetlerini durdurmuş olmak,
- Aktif tutarının %10'unu kaybetmiş olmak
- Üst üste iki yıl veya daha fazla yıl zarar etmiş olmak.

Yukarıdaki kriterlere göre eğer bir işletme;

- İflas etmişse başarısız kabul edilmiştir.
- Aktif tutarının %10'unu geçtiği yıllarda başarısız kabul edilmiştir.
- Son olarak aralıksız iki yıl üst üste zarar etmiş ise başarısız kabul edilmiştir.

Başarısız olmayan işletmeler ise yukarıdaki kriterlere uymayan işletmelerdir. Başarısız olmayan işletmeler arasında bir yıl zarar etmiş ancak sonrasında faaliyetlerini normal olarak sürdürmüş işletmeler de yer almıştır.

Bu çalışma için belirlenen başarısız veya başarılı örnek kriterlerine uymayan ve bu kriterlere uydukları halde bazı yıllarda mali tablo kalemlerinden bazıları eksik olan işletmeler de örneklem dışı bırakılmıştır. Geriye kalan 175 işletmeden 60 başarısız, 115'i ise başarılı işletme özelliklerini taşımaktadır. Her iki grubun örnek içerisinde eşit sayıda temsil edilmesini sağlamak amacıyla, tesadüfi olarak seçilen 55 başarılı işletme kapsam dışı bırakılmıştır. Ele alınan uygulamada kullanılan örneğin yarısı başarılı ve yarısı başarısız olmak üzere toplam 120 sanayi işletmesinden oluşmaktadır.



**Grafik 3.1.** Yıllar itibari ile başarısız işletme sayısı

Yukarıdaki grafikte görüldüğü gibi 5 yıllık dönem içerisinde toplam 60 firma finansal açıdan başarısız kabul edilmiştir. En düşük finansal açıdan başarısızlık 2005 yılı ile 2, en yüksek finansal açıdan başarısızlık 2007 yılı ile 19 firmanın başarısız oldukları gözlenmiştir.

#### ➤ **Başarılı İşletmeler için Baz Yılı'nın Belirlenmesi**

İşletmelerin başarılı ve başarısız olarak sınıflandırılması sonrası başarısızlıktan bir, iki, üç ve dört öncesine ait modellerin kurulabilmeleri için başarılı işletmelerde bir temel (baz) yıl seçimi gerekmektedir. Başarısız işletmeler için başarısız oldukları yıllardan bir, iki, üç ve dört yıl öncesine ait oranları bulmada bir sorun olmamasına karşın başarılı işletmelerde 2005, 2006, 2007, 2008 ve 2009 olmak üzere beş farklı yılın baz yıl olarak seçilebileceği görülmektedir. Bu yıllardan birinin baz yıl olarak seçilmesiyle bu baz yıldan geriye bir, iki, üç ve dört yıl öncesine ait oranların bulunması mümkün olacaktır.

Yapılan diğer çalışmalarda en çok başarısızlık hangi yıl olmuşsa o yılın başarılı işletmeler için baz yıl olarak alınması yoluna gidilmektedir. Bu çalışmada en çok başarısızlık 19 firma ile 2007 yılı ve onu 18 firma ile 2009 yılı takip etmektedir. Dolayısıyla başarılı ve başarısızlık tahmini için başarılı işletmelerin baz yılı olan 2007 yılından önceki 2006, 2005, 2004 ve 2003 verileri kullanılarak analizler gerçekleştirilmiştir.

Bu çalışmada işletmeleri başarılı ve başarısız olarak sınıflandırırken dört farklı model kullanılmıştır. Bu modellerin kullanılmasında yapay sinir ağları ve veri madenciliği tekniklerini uygulamak için örnek grubunu oluşturan veriler eğitim ve test seti olmak üzere iki alt kümeye ayrılmıştır. Çalışmada, 120 işletmeden oluşan toplam örneğin %80'ni eğitim seti ve geri kalan %20'sini test seti oluşturmuştur. Dolayısıyla 96 işletme eğitim setini ve 24 işletme test setini oluşturmada ele alınmıştır.

#### **3.2.3. Analizde Kullanılan Değişkenlerin Seçimi**

Çalışmada bağımlı değişken olarak 120 işletmenin finansal başarısızlık durumları ele alınmıştır. İşletme finansal açıdan başarısız ise değişkenin değeri 0, finansal açıdan başarısız değilse 1 olarak alınmıştır. Bağımsız değişken olarak literatürde yaygın olarak kabul edilen finansal oranlar kullanılmıştır. Beş farklı kategoride 28 finansal oran

belirlenmiştir. Söz konusu bu oranlar, işletmelerin 12 aylık bilanço ve gelir tabloları yardımıyla formüller kullanılarak hesaplanmıştır.

#### **A) Likidite Oranları**

**X1:** Cari Oran: Dönen Varlıklar / KVYK

**X2:** Asit-test Oranı: (Hazır değerler+Menkul Değerler+Kısa Süreli Alacaklar)/KVYK

**X3:** Nakit oranı: (Hazır değerler + Menkul değerler)/ KVYK

**X4:** Stok / toplam aktif oranı

**X5:** Dönen varlıkların toplam aktiflere oranı

**X6:** Dönen varlıkların / toplam yabancı kaynaklara oranı

#### **B) Finansal Yapı Oranları**

**X7:** Özkaynak / borç oranı

**X8:** Borç / özkaynak oranı

**X9:** Borç oranı: Toplam Borcun/Toplam varlıklara

**X10:** Nakit akışı / toplam borç oranı

**X11:** Nakit akışı / uzun süreli borç oranı

#### **C) Faaliyet Oranları**

**X12:** Aktif devir hızı: Satışlar/ Aktif Toplamı

**X13:** Özkaynak devir hızı: Satışlar/Özsermaye

**X14:** Duran varlık devir hızı: Net Satış Tutarı/Maddi duran varlıklar

**X15:** Stok / satışlar oranı

**X16:** Duran varlıklar / toplam varlık oranı

**X17:** Net Çalışma (İşletme) Sermayesi Devir Hızı Oranı: Net Satışlar/(Dönen Varlıklar-KVYK)

**X18:** Stok devir hızı oranı: SMM/(Ortalama Stok)

**X19:** Alacak devir hızı oranı: Kredili Satışlar/Ortalama Ticari Alacak

**D) Kârlılık Oranları**

**X20:** Özsermaye kârlılığı:  $\text{Net kâr} / \text{Öz Sermaye Toplamı}$

**X21:** Aktif kârlılığı:  $\text{Net Kâr} / \text{Toplam Aktif}$

**X22:** Vergi öncesi kârdaki büyüme:  $(\text{VÖK}-\text{VÖK}-1)/\text{VÖK}-1$

**X23:** Faizleri kârşılama oranı:  $(\text{Vergiden önceki kâr}+\text{Faiz Giderleri})/ \text{Faiz Giderleri}$

**X24:** Finansman Giderleri / Net Satışlar

**X25:** Brüt kâr oranı marjı:  $(\text{Satışlar}-\text{SMM})/\text{Satışlar}$

**X26:** brüt kâr / toplam aktif oranı

**E) Büyüme Oranları**

**X27:** Brüt kâr büyüme oranı:  $(\text{Brüt kâr}-\text{Brüt kâr}-1)/\text{Brüt kâr}-1$

**X28:** Satış gelirleri büyüme oranı:  $(\text{Satışlar}-\text{Satışlar}-1)/\text{Satışlar}-1$

**3.2.4. Finansal Başarısızlık Tahmini İçin Kurulan Modeller**

Uygulama, analiz öncesi sürece göre birinci kısımda, tüm değişkenler analize sokulmuştur. İkinci kısımda ise değişken seçimi yapılmış ve bunun sonucunda modelde kalmasına karar verilen değişkenler analizde yer almıştır. Diğer değişkenler ise modelin dışında bırakılmışlardır.

**3.2.4.1. Model 1**

Bu analizde tüm değişkenler modelde yer almaktadır. Değişken seçimi yapılmamıştır. Bağımsız ve bağımlı değişkenler seçildikten sonra yapay sinir ağları ve veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri kullanılarak finansal açıdan başarılı ve başarısız firmaların başarısızlıktan bir, iki, üç ve dört yıl öncesinden sınıflandırma işlemlerinin tahmini gerçekleştirilmiştir.

**3.2.4.2. Model 2**

Bu analizde değişken seçimi için lojistik regresyon analizinden elde edilen 4 değişken belirlenmiştir. Diğer geri kalan 24 değişken modelin dışında bırakılmışlardır. Dört bağımsız değişken ve bağımlı değişken seçildikten sonra yapay sinir ağları ve veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri yöntemleri

kullanılarak finansal açıdan başarılı ve başarısız firmaların başarısızlıktan bir, iki, üç ve dört yıl öncesinden sınıflandırma işlemlerinin tahmini gerçekleştirilmiştir.

#### - Lojistik Regresyon Analizi Sonucu Belirlenen Değişkenler

Verilerin değerlendirilmesi SPSS 18.0 istatistiksel analiz programıyla yapılmıştır.

Kapsama alınan 120 sanayi işletmesinin verileri lojistik regresyon analizine girilmiştir. Lojistik regresyon analizinde bağımlı değişken olarak finansal açıdan başarılı ve başarısız sınıflandırması yapılmıştır. Araştırmada başarısız işletmelere 0 ve başarılı işletmelere 1 kodlanmıştır.

Lojistik regresyon analizi yapılırken finansal oranlar arasında çoklu bağlantı ihtimalinin yüksek olması nedeniyle aşamalı seçim (stepwise) yöntemi kullanılarak 4 değişken belirlenmiştir.

#### Model 2 için Lojistik Regresyon Analizi Sonuçları

Çizelge 3.2. Model Katsayıları İçin Omnibus Testi

		Ki-Kare	Sd	Anl.
Adım 6	Adım	-,040	1	,842
	Blok	119,251	4	,000
	Model	119,251	4	,000

Geleneksel Ki-Kare metodu kullanılarak modelin anlamlılığı bu aşamada test edilmiştir. Çizelge 3.2’de değerler incelendiğinde aşamalı seçim yapılması nedeniyle blok ve model değerlerinin aynı olduğu tespit edilmiştir. Yapılan analiz sonucunda önem derecelerinin 0,01’den küçük olduğu için ( $p<0,01$ ) modelin anlamlı olduğu söylenebilir.

Ayrıca modelin uygunluğunun test edilmesinde Hosmer ve Lemeshow Testi de kullanılmıştır.

Çizelge 3.3. Hosmer ve Lemeshow Uyum Testi Sonuçları

Adım	Ki-Kare	Sd	Sig.
6	6,700	8	,569

Çizelge 3.3.'de elde edilen değerler görülmektedir. Hosmer ve Lemeshow Testi'nde modelde tahmin edilen değerlerle gerçekte gözlenen değerler arasında fark yoktur sıfır hipotezi test edilmektedir.

**H<sub>0</sub>** : Teorik model verileri iyi temsil etmektedir.

**H<sub>a</sub>** : Teorik model verileri iyi temsil etmemektedir.

Elde edilen anlamlılık seviyesinin 0,569 olması sebebi ile ( $p>0,05$ ) modelin tahmin ettiği verilerin istenen anlamlılık düzeyinde kabul edilebilir olduğu söylenilebilir.

**Çizelge 3.4.** Modelin Özeti

Adım	-2 Log likelihood	Cox & Snell R Kare	Nagelkerke R Kare
6	47,104	,630	,840

Çizelge 3.4.'de görüldüğü gibi 6'ncı iterasyonda en uygun model bulunmuştur. -2 Log olabilirlik istatistiği modelin ne kadar güçlü ve zayıf kararlar vereceğini ifade etmektedir. Yüksek değerler alması durumunda modelin zayıf olduğu, çok küçük değerler aldığımda ise modelin iyi olduğu belirtilmektedir.

Cox ve Snell R kare istatistiği regresyondaki R kare değeri ile aynı olduğu ifade edilebilir. Bağımsız değişkenlerin bağımlı değişkeni %63'le açıklayabildiği söylenebilir. Nagelkerke R Kare istatistiği ise Cox ve Snell R Kare istatistiğinin düzeltilmiş halidir. Snell R Kare istatistiğinden yüksek bir değer almaktadır. Nagelkerke R Kare istatistiği bağımsız değişkenlerin bağımlı değişkeni %84'le açıklayabildiği söylenebilir.

**Çizelge 3.5.** Hosmer ve Lemeshow Kontenjans Çizelgesi

		Finansal Açıdan Başarısız		Finansal Açıdan Başarılı		Toplam
		Gözlenen	Beklenen	Gözlenen	Beklenen	
<b>Adım 6</b>	1	12	12,000	0	,000	12
	2	12	11,995	0	,005	12
	3	12	11,850	0	,150	12
	4	12	11,038	0	,962	12
	5	8	8,111	4	3,889	12
	6	1	3,437	11	8,563	12
	7	3	1,249	9	10,751	12
	8	0	,284	12	11,716	12
	9	0	,036	12	11,964	12
	10	0	,002	12	11,998	12

Çizelge 3.5.'de Hosmer ve Lemeshow kontenjans çizelgesinde altıncı adımda örneklemin 10 grup halinde oluşturularak yapılan tahminler görülmektedir.

**Çizelge 3.6.** Lojistik Regresyon Analizi Sonucu Belirlenen Değişkenler

		B	Standart Hata	Wald	Sd	Anl.	Exp(B)
<b>Adım 6e</b>	X1 : Cari oran	-,672	,292	5,303	1	,021	,511
	X9 : Borç oranı	-10,581	2,538	17,374	1	,000	,000
	X21: Aktif karlılık oranı	28,687	6,868	17,445	1	,000	
	X26 : Brüt kar / toplam aktif	8,126	3,388	5,752	1	,016	3380,136
	Sabit	5,326	1,730	9,480	1	,002	205,566

Çizelge 3.6.'da değişkenler incelendiğinde X1 (cari oran) ve X9 (borç oranı) negatif etki yaptığı; X21 (aktif kârlılık oranı) ve X26 (brüt kâr/toplam aktif oranı) pozitif etki yaptığı gözlenmektedir.

Elde edilen Lojistik regresyon denklemi aşağıdaki gibidir.

$$L_i = 5,326 - 0,672 X_1 - 10,581 X_9 + 28,687 X_{21} + 8,126 X_{26}$$

Denklemden

$L_i$  = Lojistik regresyon denklemi,

$X_1$  = Cari Oran (Dönen Varlıklar / KVK)

$X_9$  = Borç Oranı (Toplam Borcun / Toplam Varlıklara)

$X_{21}$  = Aktif Kârlılık Oranı (ROA) (Net kâr / Toplam Aktif)

$X_{26}$  = Brüt kâr / Toplam Aktif

### 3.2.4.3. Model 3

Bu analizde değişken seçimi için diskriminant analizinden elde edilen 3 değişken belirlenmiştir. Diğer geri kalan 25 değişken modelin dışında bırakılmışlardır. Üç bağımsız değişken ve bağımlı değişken seçildikten sonra yapay sinir ağları ve veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri yöntemleri kullanılarak finansal açıdan başarılı ve başarısız firmaların başarısızlıktan bir, iki, üç ve dört yıl öncesinden sınıflandırma işlemlerinin tahmini gerçekleştirilmiştir.

#### - Diskriminant Analizi Sonucu Belirlenen Değişkenler

Verilerin değerlendirilmesi SPSS 18.0 istatistiksel analiz programıyla yapılmıştır. 28 finansal oran bağımsız değişken olarak programa sunulmuş, aşamalı seçim (stepwise) yöntemi kullanılarak 3 değişken belirlenmiştir.

#### Model 2 için Diskriminant Analizi Sonuçları

Çizelge 3.7. Özdeğerler

Fonksiyon	Özdeğer	Varyans %'si	Birikimli %	Kanonik Korelasyon
1	,764	100,0	100,0	,658

Çizelge 3.7.'de görüldüğü gibi başlangıçta belirlenen iki grup (finansal açıdan başarılı ve başarısız) olduğu için 1 diskriminant fonksiyonu türetilmiştir. Özdeğerin (Eigenvalue) büyük olması bağımlı değişkendeki varyansın daha büyük bir kısmının elde edilen fonksiyon tarafından açıklanabildiğini göstermektedir. Kesin bir değer olmamakla birlikte 0,40'ın üzerindeki değerler iyi olarak kabul edilmektedir. Çizelge 3.7'de görülebileceği gibi modelde özdeğer 0,764 bulunmuş ve varyansın %100'ünü açıklamaktadır. Kanonik Korelasyon Katsayısı 0,658 olarak bulunmuştur. Katsayının karesi 0,43'dür. Bağımsız değişkenlerin bağımlı değişkeni %43 oranında açıkladığı söylenebilir.



**Çizelge 3.8.** Wilk's Lambda Değeri

Fonksiyon Testi	Wilks' Lambda	Ki-Kare	Sd	Anl.
1	,567	66,132	3	,000

Çizelge 3.8. Wilk's Lambda istatistiğinin diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanmayan kısmını (oranını) göstermektedir. Modelde 0,567 yani toplam varyansın %56,7'si gruplar tarafından açıklanamamaktadır.

**Çizelge 3.9.** Standardize Kanonik Diskriminant Fonksiyon Katsayıları

	Fonksiyon
	1
<b>X9 : Borç oranı</b>	-,732
<b>X21 : Aktif karlılık oranı</b>	2,979
<b>X26: Brüt kar / toplam aktif</b>	4,449
<b>Sabit</b>	-,253

Çizelge 3.9'da görüleceği üzere 28 bağımsız değişken içerisinde 3 tanesi istatistiksel açıdan anlamlı bulunarak fonksiyona dahil edilmiştir. Değişkenler incelendiğinde X9 (borç oranı) negatif etki yaptığı; X21 (aktif kârlılık oranı) ve X26 (brüt kâr/toplam aktif oranı) pozitif etki yaptığı gözlenmektedir.

Elde edilen diskriminant denklemi aşağıdaki gibidir.

$$Z_i = -0,253 - 0,732 X_9 + 2,979 X_{21} + 4,449 X_{26}$$

Denklemden

$Z_i$  = Diskriminant skoru,

$X_9$  = Borç Oranı (Toplam Borcun / Toplam Varlıklara)

$X_{21}$  = Aktif Kârlılık Oranı (ROA) (Net kâr / Toplam Aktif)

$X_{26}$  = Brüt kâr / Toplam Aktif

#### 3.2.4.4. Model 4

Bu analizde değişken seçimi için feature selection yönteminden elde edilen 12 değişken belirlenmiştir. Diğer geri kalan 16 değişken modelin dışında bırakılmışlardır.

On iki bağımsız değişken ve bağımlı değişken seçildikten sonra yapay sinir ağları ve veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri yöntemleri kullanılarak finansal açıdan başarılı ve başarısız firmaların başarısızlıktan bir, iki, üç ve dört yıl öncesinden sınıflandırma işlemleri gerçekleştirilmiştir.

#### - Feature Selection Analizi Sonucu Belirlenen Değişkenler

Verilerin değerlendirilmesi CLEMENTINE 12.0 programından yararlanılmıştır. 28 finansal oran bağımsız değişken olarak programa sunulmuş, feature selection yöntemi kullanılarak 12 değişken belirlenmiştir.

Bu değişkenler;

**X1:** Cari Oran: Dönen Varlıklar / KVYK

**X2:** Asit-test Oranı: (Hazır değerler+Menkul Değerler+Kısa Süreli Alacaklar)/KVYK

**X7:** Özkaynak / borç oranı

**X9:** Borç oranı: Toplam Borcun/Toplam varlıklara

**X10:** Nakit akışı / toplam borç oranı

**X11:** Nakit akışı / uzun süreli borç oranı

**X12:** Aktif devir hızı: Satışlar/ Aktif Toplamı

**X15:** Stok / satışlar oranı

**X21:** Aktif kârlılığı: Net Kâr / Toplam Aktif

**X23:** Faizleri karşılama oranı : (Vergiden önceki kâr+Faiz Giderleri)/ Faiz Giderleri

**X25:** Brüt kâr oranı marjı: Satışlar-SMM)/Satışlar

**X26:** brüt kâr / toplam aktif oranı

İşletmelerin finansal başarısızlık tahmini için oluşturulan modeller, 1 yıllık (t-1), 2 yıllık (t-2), 3 yıllık (t-3) ve 4 yıllık (t-4) verilerinden oluşmaktadır. (t-1), (t-2), (t-3) ve (t-4) başarısızlıktan önceki 1, 2, 3 ve 4 yılı temsil etmektedir. Dolayısıyla işletmelerin başarısız oldukları (t-1), (t-2), (t-3) ve (t-4) dönemleri yıllara bağlı olarak değişmektedir. Başarılı bankalar için 2007 yılı verileri t yılı verileri olarak ele alınmıştır.

Veri setindeki deęişkenlerin aldığı deęerler geniş bir aralıkta yer aldığı için veriler normalize edilmiştir. Normalizasyonun amacı, deęişkenlerin sahip olduğu çok büyük ve çok küçük deęerler de çözümlerin sağlıklı biçimde yapılmasını engellediğinden, dolayısıyla veriler belirli bir yöntemle normalleştirilerek bu durumdan kaçınılmıştır.

Uygulamadaki dört model için kullanılan DVM sınıflandırma modeli aşağıda verilmiştir:

$$\min \frac{1}{2} \|w\|^2$$

$$y_i(w^T \varphi x_i + b) \geq +1, \forall i$$

Bu ifade ikinci dereceden optimizasyon problemidir. Eğitim verileri  $\theta = \{x_i, y_i\}$ ,  $i = 1, 2, \dots, N$ , burada  $y_i \in \{-1, 1\}$  etiket deęerleri ve  $x_i \in R^d$  özellikler vektörüdür.  $\varphi$ , pozitif yarı tanımlı bir matristir.

$\varphi_{ij} = y_i y_j K(x_i, x_j)$  ve  $K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$  kerneldir.  $x_i$  eğitim vektörleri,  $\Phi$  fonksiyonu (kernel) tarafından belirlenmektedir. Karar fonksiyonu aşağıdaki gibi ifade edilmektedir.

$$\text{sig}\left(\sum_{i=1}^l y_i a_i K(x_i, x_j) + b\right)$$

Uygulanan modellerde radyal tabanlı fonksiyon denenmiştir.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$$

### 3.2.5. Model 1: 28 Deęişkenin Tümü İçin Analiz Sonuçları

#### 3.2.5.1. Model 1 için Kullanılan YSA Mimarisinin Yapısı

Uygun yapay sinir ağı yönteminin belirlenmesinde yaygın olarak deneme yanılma yöntemi kullanılmakta ve çok sayıda test edilmektedir. Bu kapsamda gizli katman sayısı, gizli katmanlardaki düğüm sayısı, momentum terimi, aktivasyon fonksiyonu, devir sayısı gibi parametrelerin çeşitli kombinasyonları denenerek gerek eğitim seti üzerinde gerekse test seti üzerinde daha iyi performans gösteren ağ saptanmıştır. Yapay sinir ağlarının öğrenmesini sağlamak amacıyla t yılına ait verilerden yola çıkarak ağların eğitimi gerçekleştirilmiştir.

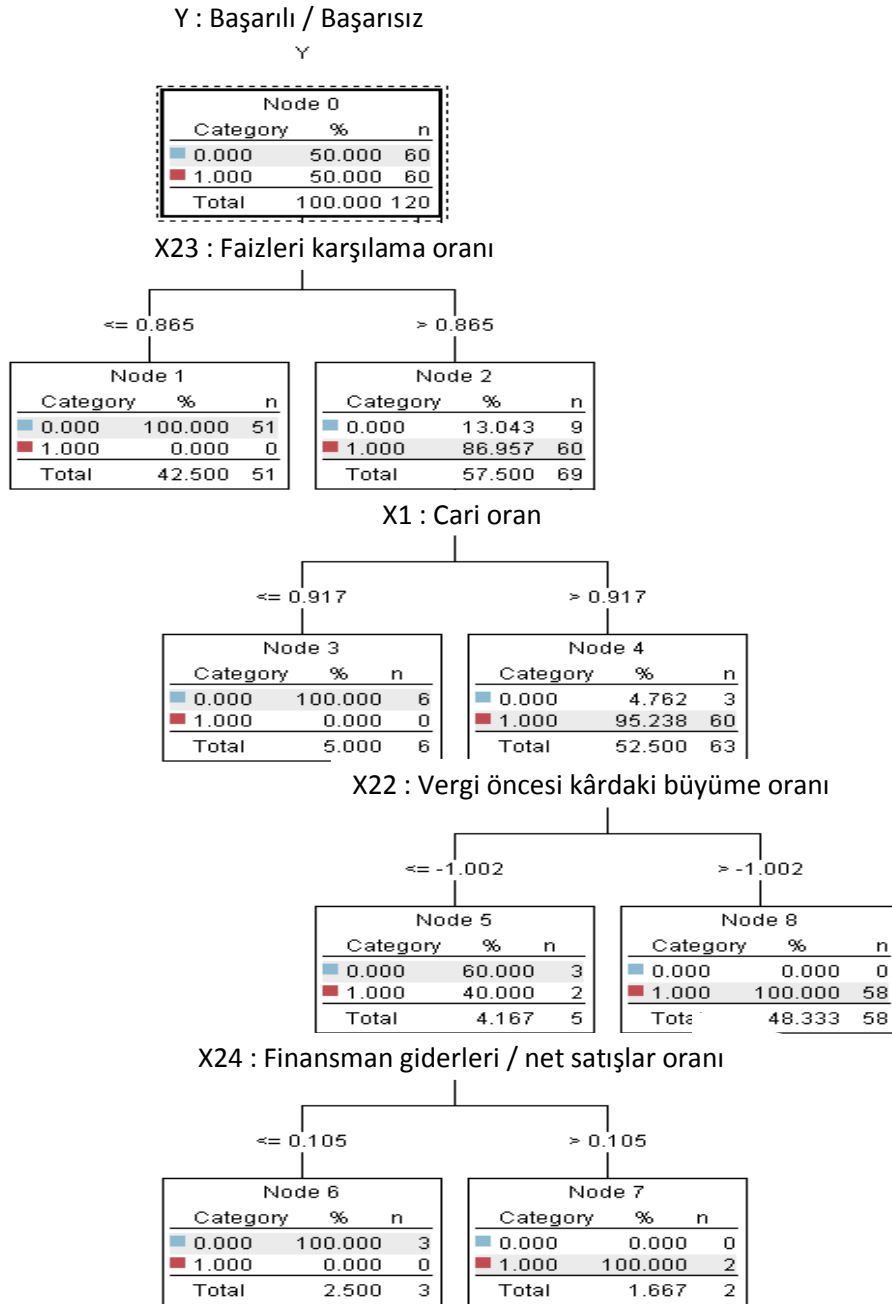
Deneme-yanılma yöntemi sonunda belirlenen en iyi performansa sahip yapay sinir ağının özellikleri aşağıdaki Çizelge 3.10.'da yer almaktadır.

**Çizelge 3.10.** Model 1 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri

Ağın türü	Çok Katmanlı Perseptron
Öğrenme Algoritması	Geri Yayılım
Öğrenme Kuralı	Momentum
Giriş Katmanındaki Düğüm Sayısı	28
Gizli Katman Sayısı	3
1.Gizli Katmandaki Düğüm Sayısı	12
2.Gizli Katmandaki Düğüm Sayısı	15
3.Gizli Katmandaki Düğüm Sayısı	17
Çıkış katmanındaki Düğüm Sayısı	1
Öğrenme Oranı	0,01
Devir Sayısı	500
Gizli Katmanlar için Transfer Fonksiyonu	Tansig
Çıktı Katman İçin Transfer Fonksiyonu	Purelin
Geri Yayılım Ağının Eğitim Fonksiyonu	Trainlm

### 3.2.5.2. Model 1 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı

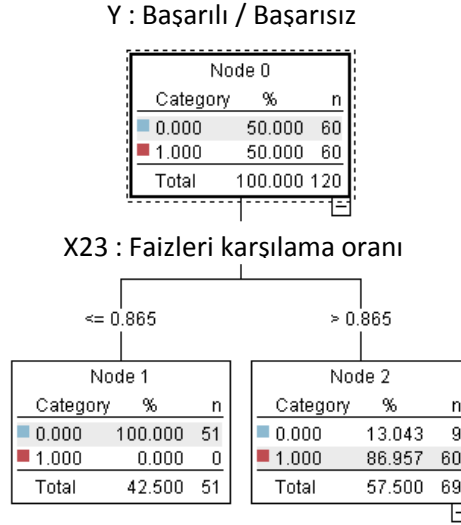
Model 1 için tüm değişkenlerin yer aldığı, C5.0 algoritmasının öğrenmesini sağlamak amacıyla t yılına ait verilerden elde edilen karar ağacı Şekil 3.2.'de gösterilmiştir.



Şekil 3.2. Model 1'in Karar Ağacı Modeli

### Faizleri Karşılama Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Karar ağacının kökünde bulunan en kuvvetli etkiye sahip olan değişken X23 değişkeni “faizleri karşılama oranıdır”. Bu oran, işletmeleri sınıflandırmada en yüksek dereceli etkiye sahip olduğu için karar ağacının kökünde yer almaktadır. Karar ağacının kökünde 60'ı başarısız ve 60 başarılı olmak üzere 120 firma mevcuttur.



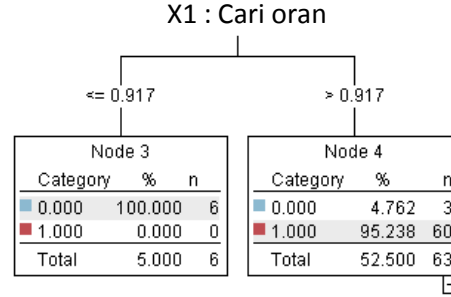
**Şekil 3.3.** Faizleri Karşılama Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.3'e göre, 120 finansal durumdan % 42,5'inin faizleri karşılama oranı 0,865 eşit ya da bu orandan küçükse o firma % 100 başarısızdır. Çalışmada bu kriterlere uyan 51 finansal durumdan 51'i de başarısızdır. % 100 başarısız sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Faizleri karşılama oranının 0,865'ten yüksek olduğu durumda ise toplam 69 işletme vardır. Bu finansal oranının 0,865'ten yüksek olduğu durumlarda % 86,957'si başarılıdır. Bu durumda firmaların 60'ı (%86,957) başarılı iken 9'u (%13,043) başarısızdır.

### **Cari Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Burada yapılacak yorumların ön koşulu faizleri karşılama oranının 0,865'ten büyük olması durumudur. Eğer bir işletmenin faizleri karşılama oranının 0,865'ten büyükse aşağıdaki yorumlar o firma için geçerlidir. Bu koşulu sağlayan 69 işletme mevcuttur.



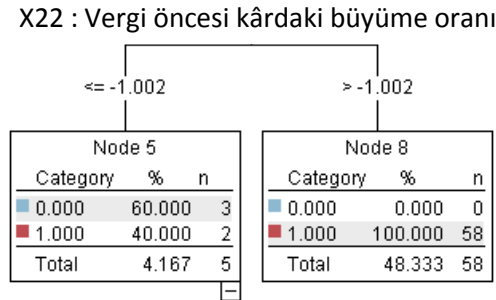
**Şekil 3.4.** Cari Oranın Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.4.'e göre bir firmanın X1 değişkeni olan cari oranının 0,917 eşit yada bu orandan küçükse, o firma % 100 başarısızdır. Çalışmada bu kriterlerde olan 6 finansal durumdan 6'sı da başarısızdır. % 100 başarısız sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Cari oranının 0,917'den yüksek olduğu durumda ise toplam 63 işletme vardır. Bu finansal oranın 0,917'den yüksekse, o işletme %95,238 oranla başarılıdır. Çalışmada finansal oranları bu şekilde olan toplam 63 finansal durumdan 60'ı (%95,238) başarılı, 3'ü (%4,762) ise başarısızdır.

### **Vergi Öncesi Kardaki Büyüme Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Faizleri karşılama oranının 0,865'ten yüksek ve cari oranının 0,917'den yüksek olmak ön koşulunu sağlayan 63 işletme mevcuttur.



**Şekil 3.5.** Vergi Öncesi Kârdaki Büyüme Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.5.'e göre X22 değişkeni olan veri öncesi kârdaki büyüme oranının -1,002'den büyükse, o firma % 100 başarılıdır. Çalışmada bu kriterlerde olan 58 finansal durumdan 58'i de başarılıdır. % 100 başarılı sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Vergi öncesi kârdaki büyüme oranının -1,002'ye eşit yada küçük olduğu durumlarda ise 5 finansal durum bulunmaktadır. Bu durumda olan firmaların 3'ü (%60) başarısız iken, 2'si (%40) başarılıdır.

### Finansal Giderlerin Net Satışlara Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Faizleri karşılama oranının 0,865'ten yüksek, cari oranının 0,917'den yüksek ve vergi öncesi kârdaki büyüme oranının -1,002'ye eşit ya da küçük olmak ön koşulunu sağlayan 5 işletme mevcuttur.

X24 : Finansman giderleri / net satışlar oranı

<= 0.105				> 0.105			
Node 6				Node 7			
Category	%	n		Category	%	n	
0.000	100.000	3		0.000	0.000	0	
1.000	0.000	0		1.000	100.000	2	
Total	2.500	3		Total	1.667	2	

Şekil 3.6. Finansal Giderlerin Net Satışlara Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.6.'ya göre X24 değişkeni olan finansal giderlerin net satışlara oranının 0,105'ten büyükse, o firma % 100 başarılıdır. Bu kriterler de olan 2 finansal durumdan 2'si de başarılıdır. % 100 başarılı sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Finansal giderlerin net satışlara oranının 0,105'e eşit ya da küçük olduğu durumlarda ise 3 finansal durum bulunmaktadır. Bu durumda olan 3 firmanın 3'ü de başarısızdır. % 100 başarısız sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.



### 3.2.5.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 1'in başarısızlıktan bir yıl öncesinden (t-1) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.11.** Model 1 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>53</b>	<b>7</b>	<b>60</b>	<b>88</b>
	<b>Başarılı</b>	<b>8</b>	<b>52</b>	<b>60</b>	<b>87</b>
	<b>Toplam</b>	<b>61</b>	<b>59</b>	<b>120</b>	<b>87,5</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>55</b>	<b>5</b>	<b>60</b>	<b>92</b>
	<b>Başarılı</b>	<b>11</b>	<b>49</b>	<b>60</b>	<b>82</b>
	<b>Toplam</b>	<b>66</b>	<b>54</b>	<b>120</b>	<b>86,6</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>49</b>	<b>11</b>	<b>60</b>	<b>82</b>
	<b>Başarılı</b>	<b>6</b>	<b>54</b>	<b>60</b>	<b>90</b>
	<b>Toplam</b>	<b>55</b>	<b>65</b>	<b>120</b>	<b>85,8</b>

Çizelge 3.11'de görüldüğü gibi 120 işletmenin yer aldığı model 1'in başarısızlıktan bir yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 53'ü ve 60 başarılı işletmeden 52'si doğru tahmin edilmiş; başarısız işletmeler için %88 ve başarılı işletmeler için % 87'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %87,5 olarak elde edilmiştir.

Model 1'in başarısızlıktan bir yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 55'i ve 60 başarılı işletmeden 49'u doğru tahmin edilmiş; başarısız işletmeler için %92 ve başarılı işletmeler için % 82'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %86,6 olarak elde edilmiştir.

Model 1'in başarısızlıktan bir yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 49'u ve 60 başarılı işletmeden 54'ü doğru tahmin edilmiş; başarısız işletmeler için %82 ve başarılı işletmeler için % 90'lık doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarısı %85,8 olarak elde edilmiştir.

### 3.2.5.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 1'in başarısızlıktan iki yıl öncesinden (t-2) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.12.** Model 1 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>46</b>	14	60	77
	<b>Başarılı</b>	9	<b>51</b>	60	85
	<b>Toplam</b>	55	65	120	<b>80,8</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>49</b>	11	60	82
	<b>Başarılı</b>	13	<b>47</b>	60	78
	<b>Toplam</b>	62	58	120	<b>80</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>45</b>	15	60	75
	<b>Başarılı</b>	12	<b>48</b>	60	80
	<b>Toplam</b>	57	63	120	<b>77,5</b>

Çizelge 3.12'de görüldüğü gibi 120 işletmenin yer aldığı model 1'in başarısızlıktan iki yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 46'sı ve 60 başarılı işletmeden 51'i doğru tahmin edilmiş; başarısız işletmeler için %77 ve başarılı işletmeler için % 85'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %80,8 olarak elde edilmiştir.

Model 1'in başarısızlıktan iki yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 49'u ve 60 başarılı işletmeden 47'si doğru tahmin edilmiş; başarısız işletmeler için %82 ve başarılı işletmeler için %78'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

Model 1'in başarısızlıktan iki yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 45'i ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %75 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %77,5 olarak elde edilmiştir.

### 3.2.5.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 1'in başarısızlıktan üç yıl öncesinden (t-3) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.13.** Model 1 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	47	13	60	78
	<b>Başarılı</b>	11	49	60	82
	<b>Toplam</b>	58	62	120	80
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	46	14	60	77
	<b>Başarılı</b>	10	50	60	83
	<b>Toplam</b>	56	64	120	80
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	38	22	60	63
	<b>Başarılı</b>	8	52	60	87
	<b>Toplam</b>	46	74	120	75

Çizelge 3.13.'de görüldüğü gibi 120 işletmenin yer aldığı model 1'in başarısızlıktan üç yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 47'si ve 60 başarılı işletmeden 49'u doğru tahmin edilmiş; başarısız işletmeler için %78 ve başarılı işletmeler için %82'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

Model 1'in başarısızlıktan üç yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 46'sı ve 60 başarılı işletmeden 50'si doğru tahmin edilmiş; başarısız işletmeler için %77 ve başarılı işletmeler için %83'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

Model 1'in başarısızlıktan üç yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 38'i ve 60 başarılı işletmeden 52'si doğru tahmin edilmiş; başarısız işletmeler için %63 ve başarılı işletmeler için %87'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %75 olarak elde edilmiştir.

### 3.2.5.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 1'in başarısızlıktan dört yıl öncesinden (t-4) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.14.** Model 1 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>40</b>	20	60	67
	<b>Başarılı</b>	10	<b>50</b>	60	83
	<b>Toplam</b>	50	70	120	<b>75</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>34</b>	26	60	57
	<b>Başarılı</b>	25	<b>35</b>	60	58
	<b>Toplam</b>	59	61	120	<b>57,5</b>

Çizelge 3.14. Devamı

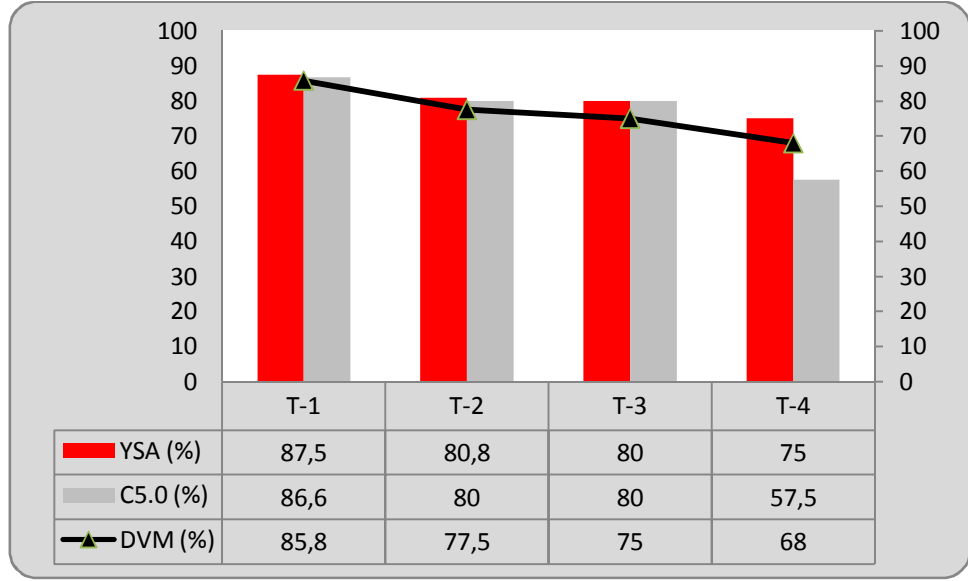
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	Başarısız	30	30	60	50
	Başarılı	8	52	60	87
	Toplam	38	82	120	68,3

Çizelge 3.14.'de görüldüğü gibi 120 işletmenin yer aldığı model 1'in başarısızlıktan dört yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 40'ı ve 60 başarılı işletmeden 50'si doğru tahmin edilmiş; başarısız işletmeler için %67 ve başarılı işletmeler için %83'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %75 olarak elde edilmiştir.

Model 1'in başarısızlıktan dört yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 34'ü ve 60 başarılı işletmeden 35'i doğru tahmin edilmiş; başarısız işletmeler için %57 ve başarılı işletmeler için %58'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %57,5 olarak elde edilmiştir.

Model 1'in başarısızlıktan dört yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 30'u ve 60 başarılı işletmeden 52'si doğru tahmin edilmiş; başarısız işletmeler için %50 ve başarılı işletmeler için %87'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %68,3 olarak elde edilmiştir.

### 3.2.5.7. Model 1'in Sınıflandırma Sonuçlarının Karşılaştırılması



**Grafik 3.2.** Model 1'in Sınıflandırma Sonuçları

Grafik 3.2'de Model 1'in yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri ile başarısızlıktan bir, iki, üç ve dört yıl öncesinden elde edilen sonuçların karşılaştırmalı grafiği verilmiştir. Başarısızlıktan bir, iki, üç ve dört yıl öncesinden ayrı ayrı analiz edilen Model 1'de, yapay sinir ağları diğer iki yönteme göre daha yüksek doğru sınıflandırma elde ettiği, bunu takiben sırasıyla C5.0 ve destek vektör makineleri olduğu görülmektedir. Dört ayrı veri seti içinden en iyi sonucu ise, %87,5'lik doğru sınıflandırma ile başarısızlıktan bir yıl öncesindeki verilerle yapılan sınıflandırma vermektedir. 60 başarısız işletmeden 53'ü ve 60 başarılı işletmeden 52'si; toplamda da 120 işletmeden 105'i yapay sinir ağları ile doğru sınıflandırılmıştır.

### 3.2.6. Model 2: 4 Değişken İçin Analiz Sonuçları

#### 3.2.6.1. Model 2 için Kullanılan YSA Mimarisinin Yapısı

Bu modelde gizli katman sayısı, gizli katmanlardaki düğüm sayısı, momentum terimi, aktivasyon fonksiyonu, devir sayısı gibi parametreler aynen kalarak, sadece lojistik regresyon analizi sonucu elde edilen 4 değişken giriş katmandaki düğüm sayısı olarak değiştirilmiştir. Yapay sinir ağlarının öğrenmesini sağlamak amacıyla t yılına ait verilerden yola çıkarak ağların eğitimi gerçekleştirilmiştir.

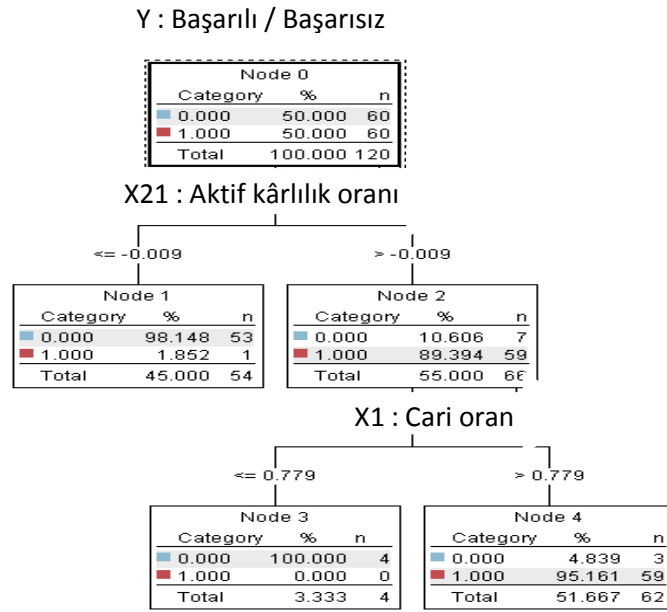
Deneme-yanılma yöntemi sonunda belirlenen en iyi performansa sahip yapay sinir ağının özellikleri aşağıdaki Çizelge 3.15.'de yer almaktadır.

**Çizelge 3.15.** Model 2 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri

Ağın türü	Çok Katmanlı Perseptron
Öğrenme Algoritması	Geri Yayılım
Öğrenme Kuralı	Momentum
Giriş Katmanındaki Düğüm Sayısı	4
Gizli Katman Sayısı	3
1.Gizli Katmandaki Düğüm Sayısı	12
2.Gizli Katmandaki Düğüm Sayısı	15
3.Gizli Katmandaki Düğüm Sayısı	17
Çıkış katmanındaki Düğüm Sayısı	1
Öğrenme Oranı	0,01
Devir Sayısı	500
Gizli Katmanlar için Transfer Fonksiyonu	Tansig
Çıktı Katman İçin Transfer Fonksiyonu	Purelin
Geri Yayılım Ağının Eğitim Fonksiyonu	Trainlm

### 3.2.6.2. Model 2 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı

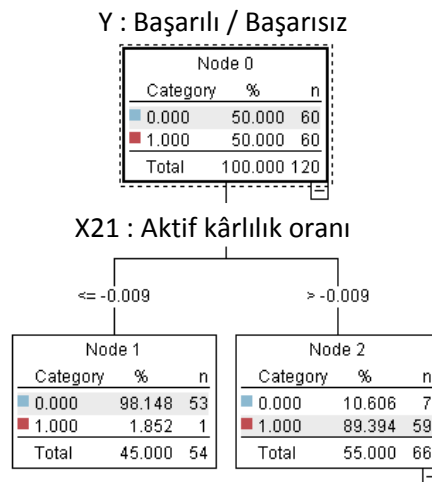
Model 2 için dört değişkenin yer aldığı, C5.0 algoritmasının öğrenmesini sağlamak amacıyla t yılına ait verilerden elde edilen karar ağacı Şekil 3.7.'de gösterilmiştir.



Şekil 3.7. Model 2'nin Karar Ağacı Modeli

### Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Karar ağacının kökünde bulunan en kuvvetli etkiye sahip olan değişken X21 değişkeni “ aktif kârlılık oranıdır”. Bu oran, işletmeleri sınıflandırmada en yüksek dereceli etkiye sahip olduğu için karar ağacının kökünde yer almaktadır. Karar ağacının kökünde 60'ı başarısız ve 60'ı başarılı olmak üzere 120 firma mevcuttur.



Şekil 3.8. Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

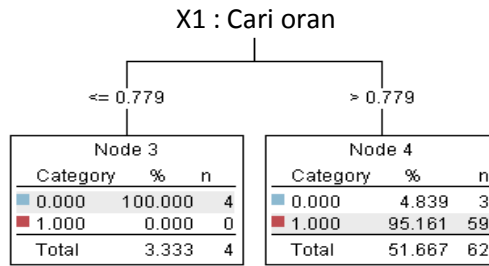


Şekil 3.8.'e göre 120 finansal durumdan %45'inin aktif kârlılık oranı -0,009 eşit ya da bu orandan küçük 54 işletme mevcuttur. Çalışmada bu kriterlere uyan 54 finansal durumdan 53 firma (%98,148) başarısız iken 1 firma (%1,852) başarılıdır.

Aktif kârlılık oranı -0,009'dan yüksek olduğu durumda ise toplam firmaların % 55'ini oluşturan 66 firma mevcuttur. Bu finansal oranının -0,009'dan yüksek olduğu durumlarda firmaların %89,394'ü başarılıdır. Bu durumda firmaların 59'u (%89,394) başarılı iken, 7'si (%10,606) başarısızdır.

### Cari Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Burada yapılacak yorumların ön koşulu aktif kârlılık oranının -0,009'dan büyük olmasıdır. Eğer bir işletme aktif kârlılık oranının -0,009'dan büyükse aşağıdaki yorumlar o firma için geçerlidir. Bu koşulu sağlayan 66 işletme mevcuttur.



**Şekil 3.9.** Cari Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.9'a göre, bir firmanın X1 değişkeni olan cari oranının 0,779 eşit ya da bu orandan küçükse, o firma % 100 başarısızdır. Çalışmada bu kriterlerde olan 4 finansal durumdan 4'ü de başarısızdır. % 100 başarısız sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Cari oranın 0,779'dan yüksek olduğu durumda ise toplam 62 firma mevcuttur. Bu finansal oranının 0,779'dan yüksekse, o firma % 95,161 oranla başarılıdır. Çalışmada finansal oranları bu şekilde olan toplam 62 finansal durumdan 59'u (%95,161) başarılı iken, 3'ü (%4,839) başarısızdır.

### 3.2.6.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 2'nin başarısızlıktan bir yıl öncesinden (t-1) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.16.** Model 2 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>60</b>	0	60	100
	<b>Başarılı</b>	5	<b>55</b>	60	92
	<b>Toplam</b>	65	55	120	<b>95,8</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>60</b>	0	60	100
	<b>Başarılı</b>	8	<b>52</b>	60	87
	<b>Toplam</b>	68	60	120	<b>93,3</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>51</b>	9	60	85
	<b>Başarılı</b>	9	<b>51</b>	60	85
	<b>Toplam</b>	60	60	120	<b>85</b>

Çizelge 3.16.'da görüldüğü gibi 120 işletmenin yer aldığı model 2'nin başarısızlıktan bir yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 60'ı ve 60 başarılı işletmeden 55'i doğru tahmin edilmiş; başarısız işletmeler için %100 ve başarılı işletmeler için %92'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %95,8 olarak elde edilmiştir.

Model 2'nin başarısızlıktan bir yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 60'ı ve 60 başarılı işletmeden 52'si doğru tahmin edilmiş; başarısız işletmeler için %100 ve başarılı işletmeler için %87'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %93,3 olarak elde edilmiştir.

Model 2'nin başarısızlıktan bir yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 51'i ve 60 başarılı işletmeden 51'i doğru tahmin edilmiş; başarısız işletmeler için %85 ve başarılı işletmeler için %85'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarısı %85 olarak elde edilmiştir.

#### 3.2.6.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 2'nin başarısızlıktan iki yıl öncesinden (t-2) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.17.** Model 2 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>55</b>	5	60	92
	<b>Başarılı</b>	5	<b>55</b>	60	92
	<b>Toplam</b>	60	60	120	<b>91,7</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>57</b>	3	60	95
	<b>Başarılı</b>	11	<b>49</b>	60	82
	<b>Toplam</b>	68	52	120	<b>88,3</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>46</b>	14	60	77
	<b>Başarılı</b>	12	<b>48</b>	60	80
	<b>Toplam</b>	58	62	120	<b>78,3</b>

Çizelge 3.17.'de görüldüğü gibi 120 işletmenin yer aldığı model 2'nin başarısızlıktan iki yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 55'i ve 60 başarılı işletmeden 55'i doğru tahmin edilmiş; başarısız işletmeler için %92 ve başarılı işletmeler için %92'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %91,7 olarak elde edilmiştir.

Model 2'nin başarısızlıktan iki yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 57'si ve 60 başarılı işletmeden 49'u doğru tahmin edilmiş; başarısız işletmeler için %95 ve başarılı işletmeler için %82'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %88,3 olarak elde edilmiştir.

Model 2'nin başarısızlıktan iki yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 46'sı ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %77 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %78,3 olarak elde edilmiştir.

### 3.2.6.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 2'nin başarısızlıktan üç yıl öncesinden (t-3) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.18.** Model 2 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	43	17	60	72
	<b>Başarılı</b>	4	56	60	93
	<b>Toplam</b>	47	73	120	<b>82,5</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	48	12	60	80
	<b>Başarılı</b>	12	48	60	80
	<b>Toplam</b>	60	60	120	<b>80</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	39	21	60	65
	<b>Başarılı</b>	7	53	60	88
	<b>Toplam</b>	46	74	120	<b>76,7</b>

Çizelge 3.18.'de görüldüğü gibi 120 işletmenin yer aldığı model 2'nin başarısızlıktan üç yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 43'ü ve 60 başarılı işletmeden 56'sı doğru tahmin edilmiş; başarısız işletmeler için %72 ve başarılı işletmeler için %93'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %82,5 olarak elde edilmiştir.

Model 2'nin başarısızlıktan üç yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 48'i ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %80 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

Model 2'nin başarısızlıktan üç yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 39'u ve 60 başarılı işletmeden 53'ü doğru tahmin edilmiş; başarısız işletmeler için %65 ve başarılı işletmeler için %88'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %76,7 olarak elde edilmiştir.

### 3.2.6.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 2'nin başarısızlıktan dört yıl öncesinden (t-4) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.19.** Model 2 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>41</b>	19	60	68
	<b>Başarılı</b>	15	<b>45</b>	60	75
	<b>Toplam</b>	56	64	120	<b>71,7</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>38</b>	22	60	63
	<b>Başarılı</b>	17	<b>43</b>	60	72
	<b>Toplam</b>	55	65	120	<b>67,5</b>

Çizelge 3.19. Devamı

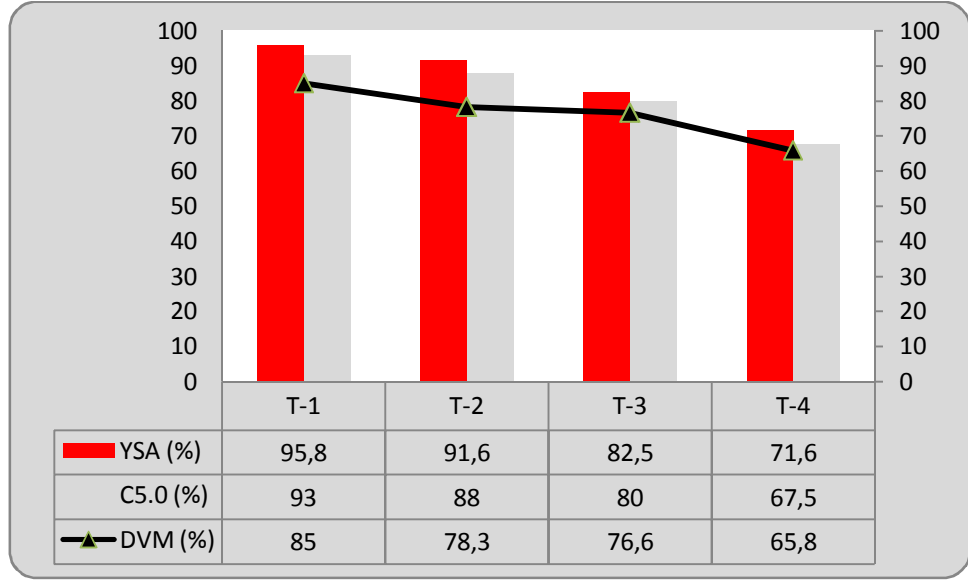
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>33</b>	27	60	55
	<b>Başarılı</b>	14	<b>46</b>	60	77
	<b>Toplam</b>	47	73	120	<b>65,8</b>

Çizelge 3.19’da görüldüğü gibi 120 işletmenin yer aldığı model 2’nin başarısızlıktan dört yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 41’i ve 60 başarılı işletmeden 45’i doğru tahmin edilmiş; başarısız işletmeler için %68 ve başarılı işletmeler için %75’lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %71,7 olarak elde edilmiştir.

Model 2’nin başarısızlıktan dört yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 38’i ve 60 başarılı işletmeden 43’ü doğru tahmin edilmiş; başarısız işletmeler için %63 ve başarılı işletmeler için %72’lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %67,5 olarak elde edilmiştir.

Model 2’nin başarısızlıktan dört yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 33’ü ve 60 başarılı işletmeden 46’sı doğru tahmin edilmiş; başarısız işletmeler için %55 ve başarılı işletmeler için %77’lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %65,8 olarak elde edilmiştir.

### 3.2.6.7. Model 2'nin Sınıflandırma Sonuçlarının Karşılaştırılması



**Grafik 3.3.** Model 2'nin Sınıflandırma Sonuçları

Grafik 3.3.'de Model 2'nin yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri ile başarısızlıktan bir, iki, üç ve dört yıl öncesinden elde edilen sonuçların karşılaştırmalı grafiği verilmiştir. Başarısızlıktan bir, iki, üç ve dört yıl öncesinden ayrı ayrı analiz edilen Model 2'de, yapay sinir ağları diğer iki yöntemle göre daha yüksek doğru sınıflandırma elde ettiği, bunu takiben sırasıyla C5.0 ve destek vektör makineleri olduğu görülmektedir. Dört ayrı veri seti içinden en iyi sonucu ise, %95,8'lik doğru sınıflandırma ile başarısızlıktan bir yıl öncesindeki verilerle yapılan sınıflandırma vermektedir. 60 başarısız işletmeden 60'ı ve 60 başarılı işletmeden 55'i; toplamda da 120 işletmeden 115'i yapay sinir ağları ile doğru sınıflandırılmıştır.

### 3.2.7. Model 3: 3 Değişken İçin Analiz Sonuçları

#### 3.2.7.1. Model 3 için Kullanılan YSA Mimarisinin Yapısı

Bu modelde gizli katman sayısı, gizli katmanlardaki düğüm sayısı, momentum terimi, aktivasyon fonksiyonu, devir sayısı gibi parametreler aynen kalarak sadece diskriminant analizi sonucu elde edilen 3 değişken giriş katmandaki düğüm sayısı olarak değiştirilmiştir. Yapay sinir ağlarının öğrenmesini sağlamak amacıyla t yılına ait verilerden yola çıkarak ağların eğitimi gerçekleştirilmiştir.

Deneme-yanılma yöntemi sonunda belirlenen en iyi performansa sahip yapay sinir ağının özellikleri aşağıdaki Çizelge 3.20.'de yer almaktadır.

**Çizelge 3.20.** Model 3 için Deneme-Yanılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri

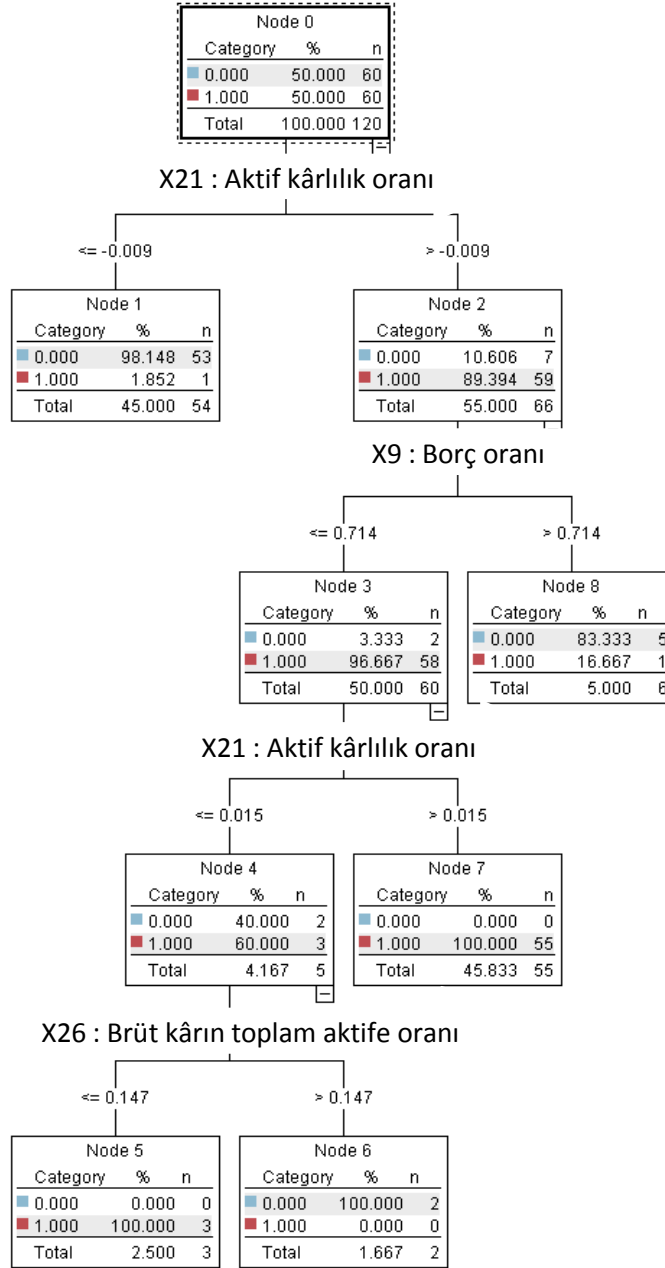
Ağın türü	Çok Katmanlı Perseptron
Öğrenme Algoritması	Geri Yayılım
Öğrenme Kuralı	Momentum
Giriş Katmanındaki Düğüm Sayısı	3
Gizli Katman Sayısı	3
1.Gizli Katmandaki Düğüm Sayısı	12
2.Gizli Katmandaki Düğüm Sayısı	15
3.Gizli Katmandaki Düğüm Sayısı	17
Çıkış katmanındaki Düğüm Sayısı	1
Öğrenme Oranı	0,01
Devir Sayısı	500
Gizli Katmanlar için Transfer Fonksiyonu	Tansig
Çıktı Katman İçin Transfer Fonksiyonu	Purelin
Geri Yayılım Ağının Eğitim Fonksiyonu	Trainlm

### 3.2.7.2. Model 3 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı

Model 3 için üç değişkenin yer aldığı, C5.0 algoritmasının öğrenmesini sağlamak amacıyla t yılına ait verilerden elde edilen karar ağacı Şekil 3.10.'da gösterilmiştir.



Y : Başarılı / Başarısız

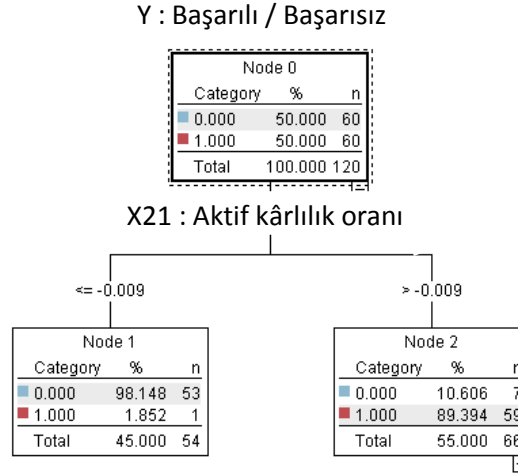


Şekil 3.10. Model 3'ün Karar Ağacı Modeli

### Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Karar ağacının kökünde bulunan en kuvvetli etkiye sahip olan değişken X21 değişkeni “aktif kârlılık oranıdır”. Bu oran, işletmeleri sınıflandırmada en yüksek

dereceli etkiye sahip olduğu için karar ağacının kökünde yer almaktadır. Karar ağacının kökünde 60 başarısız ve 60 başarılı olmak üzere 120 firma mevcuttur.



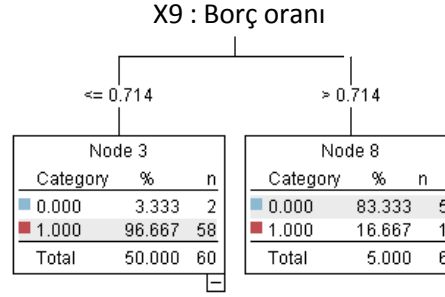
**Şekil 3.11.** Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.11'e göre, 120 finansal durumdan %45'inin aktif kârlılık oranı -0,009'a eşit ya da bu orandan küçük toplam 54 işletme mevcuttur. Çalışmada bu kriterlere uyan 54 finansal durumdan 53 firma başarısızdır. Bu durumda firmaların 53'ü (%98,148) başarısız iken 1'i (%1,852) başarılıdır.

Aktif kârlılık oranının -0,009'dan büyük olduğu durumda ise 66 işletme mevcuttur. Bu finansal oranının -0,009'dan yüksek olduğu durumlarda % 89,394'ü başarılıdır. Bu durumda firmaların 59'u (%89,394) başarılı iken 7'si (%10,606) başarısızdır.

### **Borç Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Burada yapılacak yorumların ön koşulu aktif kârlılık oranının -0,009'dan büyük olması durumudur. Eğer bir işletmenin aktif kârlılık oranı -0,009'dan büyükse aşağıdaki yorumlar o firma için geçerlidir. Bu koşulu sağlayan 66 firma mevcuttur.



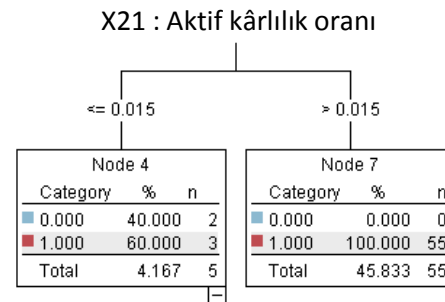
**Şekil 3.12.** Borç Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.12'ye göre bir firmanın X9 değişkeni olan borç oranının 0,714'ten büyük olduğu durumda 6 işletme mevcuttur. Çalışmada bu kriterlere uyan 6 finansal durumdan 5 firma başarısızdır. Bu durumda firmaların 5'i (%83,333) başarısız iken, 1'i (16,667) başarılıdır.

Borç oranının 0,714'e eşit ya da bu orandan küçükse toplam 60 firma mevcuttur. Bu finansal oradan 0,714'e eşit ya da bu orandan küçükse, o işletme % 96,667 oranla başarılıdır. Çalışmada finansal oranları bu şekilde olan toplam 60 finansal durumdan 58'i (96,667) başarılı iken, 2'si (%3,333) başarısızdır.

### **Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Aktif kârlılık oranının -0,009'dan büyük ve borç oranının 0,714'e eşit ya da bu orandan küçük olmak ön koşulu olması durumdur. Borç oranının 0,714'e eşit ya da bu orandan küçükse aşağıdaki yorumlar o firma için geçerlidir. Bu koşulu sağlayan 60 firma mevcuttur.



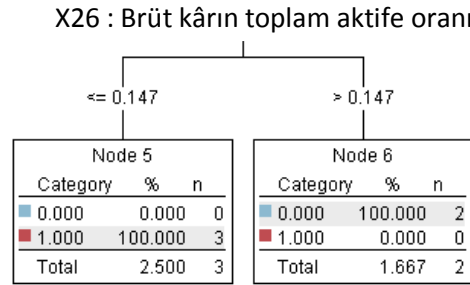
**Şekil 3.13.** Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.13.'e göre bir firmanın X21 değişkeni olan aktif kârlılık oranının 0,015'ten büyükse o firma, % 100 başarılıdır. Çalışmada bu kriterler de olan 55 finansal durumdan 55'i de başarılıdır. %100 başarılı sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Aktif kârlılık oranının 0,015 eşit ya da bu orandan küçük olduğu durumda ise toplam 5 firma mevcuttur. Bu oranının 0,015 eşit ya da bu orandan küçükse, o işletme %60 oranla başarılıdır. Çalışmada finansal oranları bu şekilde olan toplam 5 finansal durumdan 3'ü (%60) başarılı iken 2'si (%40) başarısızdır.

### **Brüt Kârın Toplam Aktife Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Aktif kârlılık oranının -0,009'dan büyük ve 0,015 küçük olmak ayrıca borç oranının 0,714'e eşit ya da bu orandan küçük olmak ön koşulu olması durumdur. Aktif kârlılık oranının 0,015 eşit ya da bu orandan küçükse aşağıdaki yorumlar o firma için geçerlidir. Bu koşulu sağlayan 5 firma mevcuttur.



**Şekil 3.14.** Brüt Kârın Toplam Aktife Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.14'e göre bir firmanın X26 değişkeni brüt kârın toplam aktife oranının 0,147 eşit ya da bu orandan küçükse, o firma %100 başarılıdır. Çalışmada bu kriterler de olan 3 finansal durumdan 3'ü de başarılıdır.

Brüt kârın toplam aktife oranının 0,147'den yüksek ise toplam 2 firma vardır. 2 firmanın 2'si de başarısızdır.

### 3.2.7.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 3'ün başarısızlıktan bir yıl öncesinden (t-1) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.21.** Model 3 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>59</b>	1	60	98
	<b>Başarılı</b>	4	<b>56</b>	60	93
	<b>Toplam</b>	63	57	120	<b>95,8</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>60</b>	0	60	100
	<b>Başarılı</b>	8	<b>52</b>	60	87
	<b>Toplam</b>	68	52	120	<b>93,3</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>53</b>	7	60	88
	<b>Başarılı</b>	9	<b>51</b>	60	85
	<b>Toplam</b>	62	58	120	<b>86,7</b>

Çizelge 3.21.'de görüldüğü gibi 120 işletmenin yer aldığı model 3'ün başarısızlıktan bir yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 59'u ve 60 başarılı işletmeden 56'sı doğru tahmin edilmiş; başarısız işletmeler için %98 ve başarılı işletmeler için %93'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %95,8 olarak elde edilmiştir.

Model 3'ün başarısızlıktan bir yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 60'ı ve 60 başarılı işletmeden 52'si doğru tahmin edilmiş; başarısız işletmeler için %100 ve başarılı işletmeler için %87'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %93,3 olarak elde edilmiştir.

Model 3'ün başarısızlıktan bir yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 53'ü ve 60 başarılı işletmeden 51'i doğru tahmin edilmiş; başarısız işletmeler için %88 ve başarılı işletmeler için %85'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarısı %86,7 olarak elde edilmiştir.

### 3.2.7.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 3'ün başarısızlıktan iki yıl öncesinden (t-2) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.22.** Model 3 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>55</b>	5	60	92
	<b>Başarılı</b>	7	<b>53</b>	60	88
	<b>Toplam</b>	62	58	120	<b>90</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>57</b>	3	60	95
	<b>Başarılı</b>	11	<b>49</b>	60	82
	<b>Toplam</b>	68	52	120	<b>88,3</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>44</b>	16	60	73
	<b>Başarılı</b>	12	<b>48</b>	60	80
	<b>Toplam</b>	56	74	120	<b>76,7</b>

Çizelge 3.22.'de görüldüğü gibi 120 işletmenin yer aldığı model 3'ün başarısızlıktan iki yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 55'i ve 60 başarılı işletmeden 53'ü doğru tahmin edilmiş; başarısız işletmeler için %92 ve başarılı işletmeler için %88'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %90 olarak elde edilmiştir.

Model 3'ün başarısızlıktan iki yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 57'si ve 60 başarılı işletmeden 49'u doğru tahmin edilmiş; başarısız işletmeler için %95 ve başarılı işletmeler için %82'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %88,3 olarak elde edilmiştir.

Model 3'ün başarısızlıktan iki yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 44'ü ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %73 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %76,7 olarak elde edilmiştir.

### 3.2.7.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 3'ün başarısızlıktan üç yıl öncesinden (t-3) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.23.** Model 3 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>49</b>	11	60	82
	<b>Başarılı</b>	12	<b>48</b>	60	80
	<b>Toplam</b>	61	59	120	<b>80,8</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>50</b>	10	60	83
	<b>Başarılı</b>	13	<b>47</b>	60	78
	<b>Toplam</b>	63	57	120	<b>80,8</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>42</b>	18	60	70
	<b>Başarılı</b>	6	<b>54</b>	60	90
	<b>Toplam</b>	48	72	120	<b>80</b>

Çizelge 3.23.'de görüldüğü gibi 120 işletmenin yer aldığı model 3'ün başarısızlıktan üç yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 49'u ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %82 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80,8 olarak elde edilmiştir.

Model 3'ün başarısızlıktan üç yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 50'si ve 60 başarılı işletmeden 47'si doğru tahmin edilmiş; başarısız işletmeler için %83 ve başarılı işletmeler için %78'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80,8 olarak elde edilmiştir.

Model 3'ün başarısızlıktan üç yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 42'si ve 60 başarılı işletmeden 54'ü doğru tahmin edilmiş; başarısız işletmeler için %70 ve başarılı işletmeler için %90'lık doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

### 3.2.7.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 3'ün başarısızlıktan dört yıl öncesinden (t-4) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.24.** Model 3 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>38</b>	22	60	63
	<b>Başarılı</b>	12	<b>48</b>	60	80
	<b>Toplam</b>	50	70	120	<b>71,7</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>41</b>	19	60	68
	<b>Başarılı</b>	17	<b>43</b>	60	72
	<b>Toplam</b>	58	62	120	<b>70</b>



Çizelge 3.24. Devamı

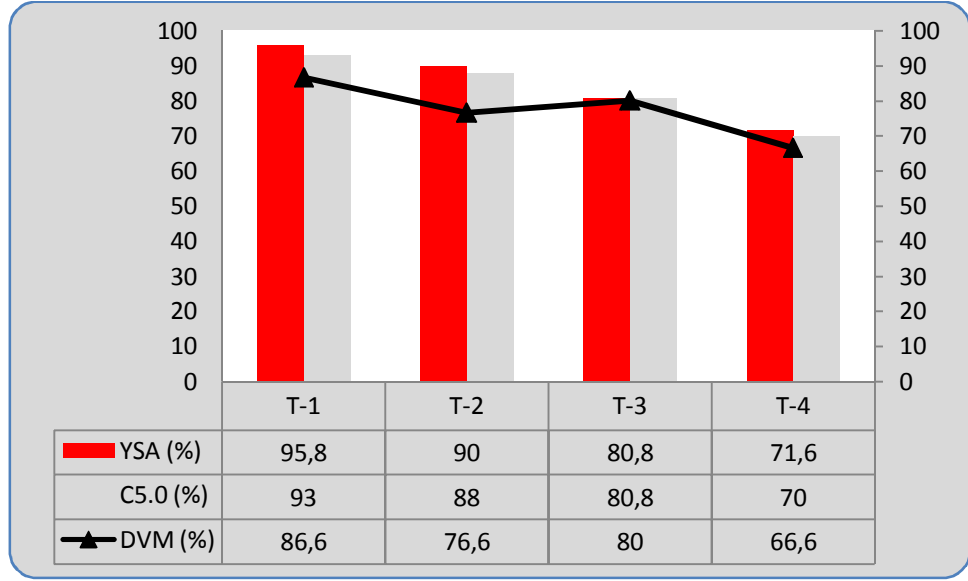
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	33	27	60	55
	<b>Başarılı</b>	13	47	60	78
	<b>Toplam</b>	46	74	120	<b>66,7</b>

Çizelge 3.24.'de görüldüğü gibi 120 işletmenin yer aldığı model 3'ün başarısızlıktan dört yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 38'i ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %63 ve başarılı işletmeler için % 80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %71,7 olarak elde edilmiştir.

Model 3'ün başarısızlıktan dört yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 41'i ve 60 başarılı işletmeden 43'ü doğru tahmin edilmiş; başarısız işletmeler için %68 ve başarılı işletmeler için %72'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %70 olarak elde edilmiştir.

Model 3'ün başarısızlıktan dört yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 33'ü ve 60 başarılı işletmeden 47'si doğru tahmin edilmiş; başarısız işletmeler için %55 ve başarılı işletmeler için %78'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %66,7 olarak elde edilmiştir.

### 3.2.7.7. Model 3'ün Sınıflandırma Sonuçlarının Karşılaştırılması



**Grafik 3.4.** Model 3'ün Sınıflandırma Sonuçları

Grafik 3.4.'de Model 3'ün yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri ile başarısızlıktan bir, iki, üç ve dört yıl öncesinden elde edilen sonuçların karşılaştırmalı grafiği verilmiştir. Başarısızlıktan bir, iki, üç ve dört yıl öncesinden ayrı ayrı analiz edilen Model 3'de, yapay sinir ağları diğer iki yöntemle göre daha yüksek doğru sınıflandırma elde ettiği, bunu takiben sırasıyla C5.0 ve destek vektör makineleri olduğu görülmektedir. Dört ayrı veri seti içinden en iyi sonucu ise, %95,8'lik doğru sınıflandırma ile başarısızlıktan bir yıl öncesindeki verilerle yapılan sınıflandırma vermektedir. 60 başarısız işletmeden 59'u ve 60 başarılı işletmeden 56'sı; toplamda da 120 işletmeden 115'i yapay sinir ağları ile doğru sınıflandırılmıştır.

### 3.2.8. Model 4: 12 Değişken İçin Analiz Sonuçları

#### 3.2.8.1. Model 4 için Kullanılan YSA Mimarisinin Yapısı

Bu modelde gizli katman sayısı, gizli katmanlardaki düğüm sayısı, momentum terimi, aktivasyon fonksiyonu, devir sayısı gibi parametreler aynen kalarak, sadece feature selection analizi sonucu elde edilen 12 değişken giriş katmandaki düğüm sayısı olarak değiştirilmiştir. Yapay sinir ağlarının öğrenmesini sağlamak amacıyla t yılına ait verilerden yola çıkarak ağların eğitimi gerçekleştirilmiştir.

Deneme-yanılma yöntemi sonunda belirlenen en iyi performansa sahip yapay sinir ağının özellikleri aşağıdaki Çizelge 3.25.'de yer almaktadır.

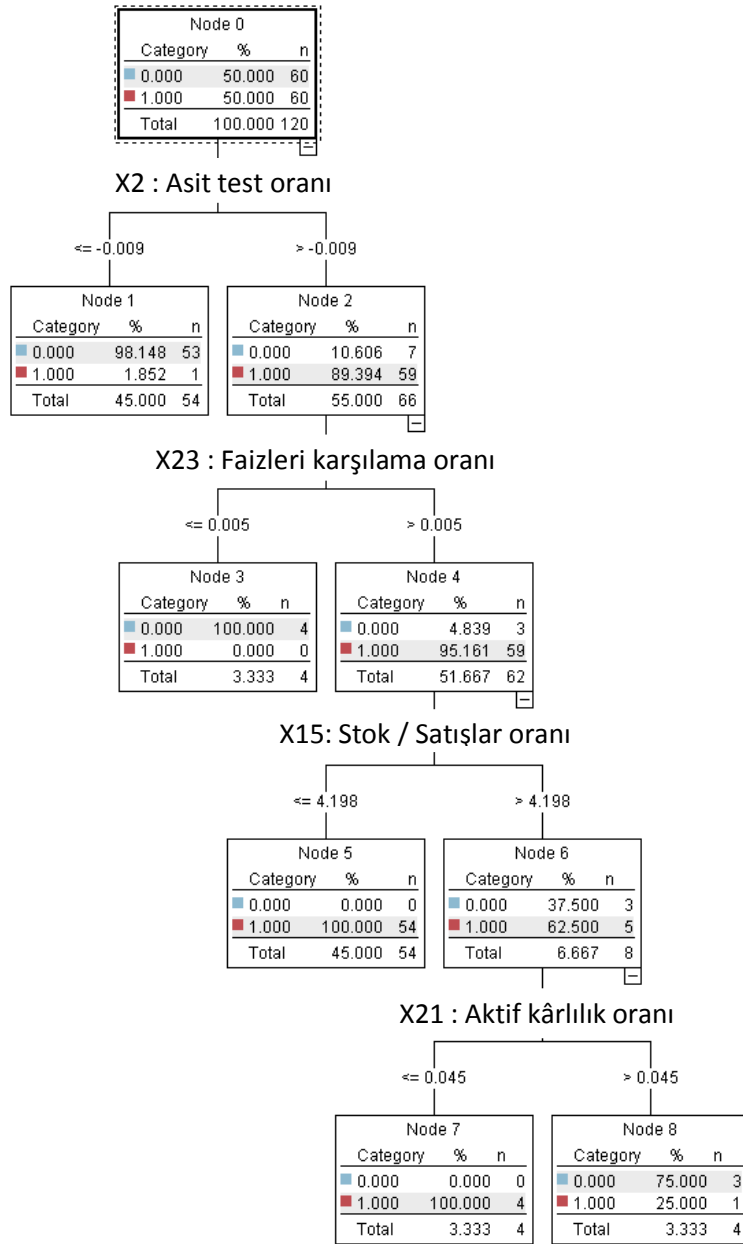
**Çizelge 3.25.** Model 4 için Deneme-Yanıılma Yöntemi Sonunda Belirlenen En İyi Performansa Sahip Yapay Sinir Ağının Özellikleri

Ağın türü	Çok Katmanlı Perseptron
Öğrenme Algoritması	Geri Yayılım
Öğrenme Kuralı	Momentum
Giriş Katmanındaki Düğüm Sayısı	12
Gizli Katman Sayısı	3
1.Gizli Katmandaki Düğüm Sayısı	12
2.Gizli Katmandaki Düğüm Sayısı	15
3.Gizli Katmandaki Düğüm Sayısı	17
Çıkış katmanındaki Düğüm Sayısı	1
Öğrenme Oranı	0,01
Devir Sayısı	500
Gizli Katmanlar için Transfer Fonksiyonu	Tansig
Çıktı Katman İçin Transfer Fonksiyonu	Purelin

### 3.2.8.2. Model 4 için C5.0 Algoritması Tarafından Elde Edilen Karar Ağacı

Model 4 için on iki değişkenin yer aldığı, C5.0 algoritmasının öğrenmesini sağlamak amacıyla t yılına ait verilerden elde edilen karar ağacı Şekil 3.15.'de gösterilmiştir.

Y : Başarılı / Başarısız

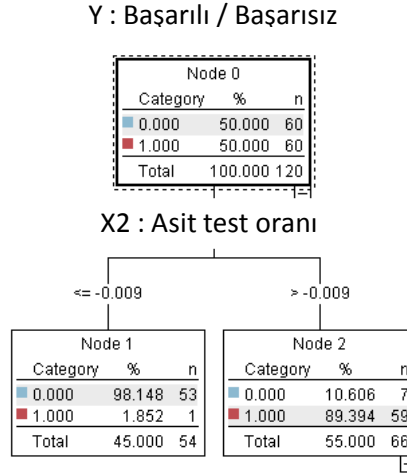


Şekil 3.15. Model 4'ün Karar Ağacı Modeli

### Asit Test Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Karar ağacının kökünde bulunan en kuvvetli etkiye sahip olan değişken X2 değişkeni “asit test oranıdır”. Bu oran, işletmeleri sınıflandırmada en yüksek dereceli

etkiye sahip olduğu için karar ağacının kökünde yer almaktadır. Karar ağacının kökünde 60'ı başarısız ve 60'ı başarılı olmak üzere 120 firma mevcuttur.



**Şekil 3.16.** Asit Test Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.16.'ya göre, 120 finansal durumdan %45'inin asit test oranı -0,009'a eşit ya da küçüktür. Çalışmada bu kriterlere uyan 54 finansal durumdan 53'ü (%98,148) başarısız iken 1'i (%1,852) başarılıdır.

Asit test oranının -0,009'dan yüksek olduğu durumda ise toplam 66 işletme vardır. Bu finansal oranının -0,009'dan yüksek olduğu durumlarda % 89,394'ü başarılıdır. Bu durumda firmaların 59'u (%89,394'ü) başarılı iken 7'si (%10,658) başarısızdır.

### **Faizleri Karşılama Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Burada yapılacak yorumların ön koşulu asit test oranının -0,009'dan yüksek olması durumudur. Eğer bir işletmenin asit test oranının -0,009'dan büyük ise aşağıdaki yorumlar o firma için geçerlidir. Bu koşulu sağlayan 66 işletme mevcuttur.

**X23 : Faizleri karşılama oranı**

$\leq 0.005$	$> 0.005$																								
Node 3	Node 4																								
<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Category</th> <th style="text-align: right;">%</th> <th style="text-align: right;">n</th> </tr> </thead> <tbody> <tr> <td style="background-color: #e6f2ff;">0.000</td> <td style="text-align: right;">100.000</td> <td style="text-align: right;">4</td> </tr> <tr> <td style="background-color: #ffe6e6;">1.000</td> <td style="text-align: right;">0.000</td> <td style="text-align: right;">0</td> </tr> <tr> <td style="border-top: 1px solid black;">Total</td> <td style="text-align: right; border-top: 1px solid black;">3.333</td> <td style="text-align: right; border-top: 1px solid black;">4</td> </tr> </tbody> </table>	Category	%	n	0.000	100.000	4	1.000	0.000	0	Total	3.333	4	<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Category</th> <th style="text-align: right;">%</th> <th style="text-align: right;">n</th> </tr> </thead> <tbody> <tr> <td style="background-color: #e6f2ff;">0.000</td> <td style="text-align: right;">4.839</td> <td style="text-align: right;">3</td> </tr> <tr> <td style="background-color: #ffe6e6;">1.000</td> <td style="text-align: right;">95.161</td> <td style="text-align: right;">59</td> </tr> <tr> <td style="border-top: 1px solid black;">Total</td> <td style="text-align: right; border-top: 1px solid black;">51.667</td> <td style="text-align: right; border-top: 1px solid black;">62</td> </tr> </tbody> </table>	Category	%	n	0.000	4.839	3	1.000	95.161	59	Total	51.667	62
Category	%	n																							
0.000	100.000	4																							
1.000	0.000	0																							
Total	3.333	4																							
Category	%	n																							
0.000	4.839	3																							
1.000	95.161	59																							
Total	51.667	62																							

**Şekil 3.17.** Faizleri Karşılama Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.17'ye göre, bir firmanın X23 değişkeni olan faizleri karşılama oranının 0,005'ten küçük olması durumunda 4 firma mevcuttur. Çalışmada bu kriterlere uyan 4 finansal durumdan 4'ü de başarısızdır. %100 başarısız sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Faizleri karşılama oranının 0,005'ten büyük olduğu durumda ise toplam 62 firma vardır. Bu finansal oranının 0,005'ten büyükse, o işletme %95,161 oranla başarılıdır. Çalışmada finansal oranları bu şekilde olan toplam 62 finansal durumdan 59'u (%95,161) başarılı, 3'ü (%4,839) başarısızdır.

### **Stok/Satışlar Oranının Finansal Durum Sınıflandırılmasındaki Etkisi**

Asit test oranının -0,009'dan büyük ve faizleri karşılama oranının 0,005'ten büyük olmak ön koşulunu sağlayan 62 firma mevcuttur.

**X15 : Stok / satışlar oranı**

$\leq 4.198$	$> 4.198$																								
Node 5	Node 6																								
<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Category</th> <th style="text-align: right;">%</th> <th style="text-align: right;">n</th> </tr> </thead> <tbody> <tr> <td style="background-color: #e6f2ff;">0.000</td> <td style="text-align: right;">0.000</td> <td style="text-align: right;">0</td> </tr> <tr> <td style="background-color: #ffe6e6;">1.000</td> <td style="text-align: right;">100.000</td> <td style="text-align: right;">54</td> </tr> <tr> <td style="border-top: 1px solid black;">Total</td> <td style="text-align: right; border-top: 1px solid black;">45.000</td> <td style="text-align: right; border-top: 1px solid black;">54</td> </tr> </tbody> </table>	Category	%	n	0.000	0.000	0	1.000	100.000	54	Total	45.000	54	<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Category</th> <th style="text-align: right;">%</th> <th style="text-align: right;">n</th> </tr> </thead> <tbody> <tr> <td style="background-color: #e6f2ff;">0.000</td> <td style="text-align: right;">37.500</td> <td style="text-align: right;">3</td> </tr> <tr> <td style="background-color: #ffe6e6;">1.000</td> <td style="text-align: right;">62.500</td> <td style="text-align: right;">5</td> </tr> <tr> <td style="border-top: 1px solid black;">Total</td> <td style="text-align: right; border-top: 1px solid black;">6.667</td> <td style="text-align: right; border-top: 1px solid black;">8</td> </tr> </tbody> </table>	Category	%	n	0.000	37.500	3	1.000	62.500	5	Total	6.667	8
Category	%	n																							
0.000	0.000	0																							
1.000	100.000	54																							
Total	45.000	54																							
Category	%	n																							
0.000	37.500	3																							
1.000	62.500	5																							
Total	6.667	8																							

**Şekil 3.18.** Stok/Satışlar Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.18.'e göre X15 değişkeni olan stok/satışlar oranının 4,198'den küçük ya da eşit olması durumunda, o firma %100 başarılıdır. Çalışmada bu kriterlerde olan 54

finansal durumdan 54'ü de başarılıdır. % 100 başarılı sonuca ulaşıldığı için karar ağacında bu dal devam etmemektedir.

Stok/satışlar oranının 4,198'den büyük olması durumunda 8 finansal durum bulunmaktadır. Bu durumda olan firmaların 5'i (%62,5) başarılı iken 3'ü (%37,5) başarısızdır.

### Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Asit test oranının -0,009'dan büyük, faizleri karşılama oranının 0,005'ten büyük ve stok/satışlar oranının 4,198'den büyük olmak ön koşulunu sağlayan 8 işletme mevcuttur.

X21 : Aktif kârlılık oranı

<= 0.045

Node 7		
Category	%	n
0.000	0.000	0
1.000	100.000	4
<b>Total</b>	<b>3.333</b>	<b>4</b>

> 0.045

Node 8		
Category	%	n
0.000	75.000	3
1.000	25.000	1
<b>Total</b>	<b>3.333</b>	<b>4</b>

**Şekil 3.19.** Aktif Kârlılık Oranının Finansal Durum Sınıflandırılmasındaki Etkisi

Şekil 3.19'a göre X21 değişkeni olan aktif kârlılık oranının 0,045'ten küçükse, o firma %100 başarılıdır. Çalışmada bu kriterler de olan 4 finansal durumdan 4'ü de başarılıdır.

Aktif kârlılık oranının 0,045'ten büyük olması durumlarda ise 4 finansal durum söz konusudur. Bu durumda olan firmaların 3'ü (%75) başarısızken, 1'i (25) başarılıdır.

### 3.2.8.3. Başarısızlıktan Bir Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 4'ün başarısızlıktan bir yıl öncesinden (t-1) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir:

**Çizelge 3.26.** Model 4 T-1 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>60</b>	0	60	100
	<b>Başarılı</b>	8	<b>52</b>	60	87
	<b>Toplam</b>	68	52	120	<b>93,3</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>55</b>	5	60	92
	<b>Başarılı</b>	10	<b>50</b>	60	83
	<b>Toplam</b>	65	55	120	<b>87,5</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>51</b>	9	60	85
	<b>Başarılı</b>	6	<b>54</b>	60	90
	<b>Toplam</b>	57	63	120	<b>87,5</b>

Çizelge 3.26.'da görüldüğü gibi 120 işletmenin yer aldığı model 4'ün başarısızlıktan bir yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 60'ı ve 60 başarılı işletmeden 52'si doğru tahmin edilmiş; başarısız işletmeler için %100 ve başarılı işletmeler için %87'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarısı %93,3 olarak elde edilmiştir.

Model 4'ün başarısızlıktan bir yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 55'i ve 60 başarılı işletmeden 50'si doğru tahmin edilmiş; başarısız işletmeler için %92 ve başarılı işletmeler için %83'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarısı %87,5 olarak elde edilmiştir.

Model 4'ün başarısızlıktan bir yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 51'i ve 60 başarılı işletmeden 54'ü doğru tahmin edilmiş; başarısız işletmeler için %85 ve başarılı işletmeler için %90'lık doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarısı %87,5 olarak elde edilmiştir.



### 3.2.8.4. Başarısızlıktan İki Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 4'ün başarısızlıktan iki yıl öncesinden (t-2) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.27.** Model 4 T-2 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>50</b>	10	60	83
	<b>Başarılı</b>	6	<b>54</b>	60	90
	<b>Toplam</b>	56	64	120	<b>86,7</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>49</b>	11	60	82
	<b>Başarılı</b>	13	<b>47</b>	60	78
	<b>Toplam</b>	62	58	120	<b>80</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>47</b>	13	60	78
	<b>Başarılı</b>	13	<b>47</b>	60	78
	<b>Toplam</b>	60	60	120	<b>78,3</b>

Çizelge 3.27.'de görüldüğü gibi 120 işletmenin yer aldığı model 4'ün başarısızlıktan iki yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 50'si ve 60 başarılı işletmeden 54'ü doğru tahmin edilmiş; başarısız işletmeler için %83 ve başarılı işletmeler için %90'lık doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarıları %86,7 olarak elde edilmiştir.

Model 4'ün başarısızlıktan iki yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 49'u ve 60 başarılı işletmeden 47'si doğru tahmin edilmiş; başarısız işletmeler için %82 ve başarılı işletmeler için %78'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarıları %80 olarak elde edilmiştir.

Model 4'ün başarısızlıktan iki yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 47'si ve 60 başarılı işletmeden 47'si doğru tahmin edilmiş; başarısız işletmeler için %78,3 ve başarılı işletmeler için %78,3'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %78 olarak elde edilmiştir.

### 3.2.8.5. Başarısızlıktan Üç Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 4'ün başarısızlıktan üç yıl öncesi (t-3) için yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir:

**Çizelge 3.28.** Model 4 T-3 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>50</b>	10	60	83
	<b>Başarılı</b>	14	<b>46</b>	60	77
	<b>Toplam</b>	64	56	120	<b>80</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>46</b>	14	60	77
	<b>Başarılı</b>	10	<b>50</b>	60	83
	<b>Toplam</b>	56	64	120	<b>80</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>37</b>	23	60	62
	<b>Başarılı</b>	12	<b>48</b>	60	80
	<b>Toplam</b>	49	71	120	<b>70,8</b>

Çizelge 3.28.'de görüldüğü gibi 120 işletmenin yer aldığı model 4'ün başarısızlıktan üç yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 50'si ve 60 başarılı işletmeden 46'sı doğru tahmin edilmiş; başarısız işletmeler için %83 ve başarılı işletmeler için %77'lik doğru sınıflandırma

gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

Model 4'ün başarısızlıktan üç yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 46'sı ve 60 başarılı işletmeden 50'si doğru tahmin edilmiş; başarısız işletmeler için %77 ve başarılı işletmeler için %83'lük doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %80 olarak elde edilmiştir.

Model 4'ün başarısızlıktan üç yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 37'si ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %62 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan üç yıl öncesi için toplam sınıflandırma başarısı %70,8 olarak elde edilmiştir.

### 3.2.8.6. Başarısızlıktan Dört Yıl Öncesi İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 4'ün başarısızlıktan dört yıl öncesinden (t-4) yapay sinir ağları, veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri için sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.29.** Model 4 T-4 Dönemi için YSA-C5.0-DVM Sınıflandırma Başarısı

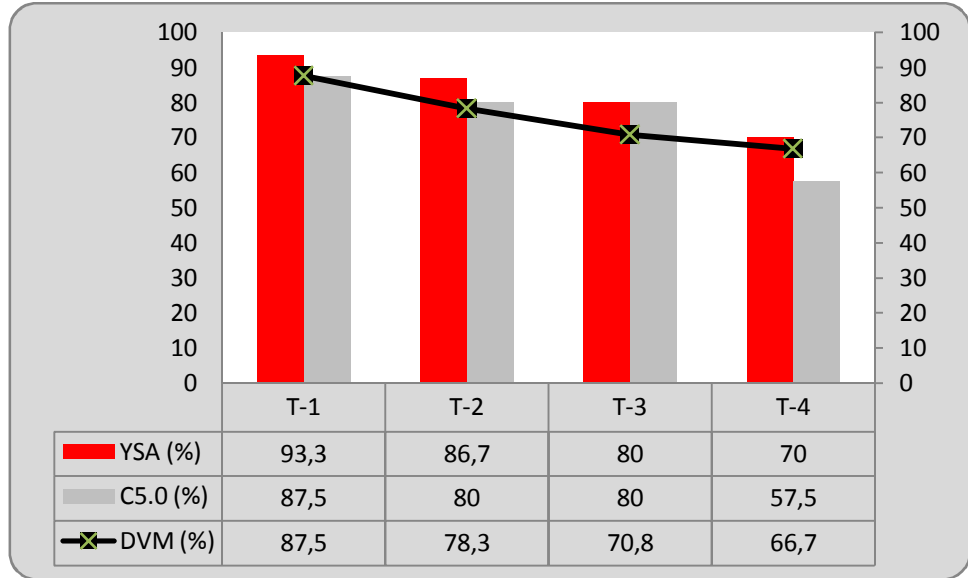
YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	36	24	60	60
	<b>Başarılı</b>	12	48	60	80
	<b>Toplam</b>	48	72	120	70
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	34	26	60	57
	<b>Başarılı</b>	25	35	60	58
	<b>Toplam</b>	59	61	120	57,5
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	25	35	60	42
	<b>Başarılı</b>	5	55	60	92
	<b>Toplam</b>	30	90	120	66,7

Çizelge 3.29.'da görüldüğü gibi 120 işletmenin yer aldığı model 4'ün başarısızlıktan dört yıl öncesi için yapay sinir ağları yöntemi kullanılarak 60 başarısız işletmeden 36'sı ve 60 başarılı işletmeden 48'i doğru tahmin edilmiş; başarısız işletmeler için %60 ve başarılı işletmeler için %80'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan iki yıl öncesi için toplam sınıflandırma başarısı %70 olarak elde edilmiştir.

Model 4'ün başarısızlıktan dört yıl öncesi için C5.0 algoritması yöntemi kullanılarak 60 başarısız işletmeden 34'ü ve 60 başarılı işletmeden 35'i doğru tahmin edilmiş; başarısız işletmeler için %57 ve başarılı işletmeler için %58'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %57,5 olarak elde edilmiştir.

Model 4'ün başarısızlıktan dört yıl öncesi için destek vektör makineleri yöntemi kullanılarak 60 başarısız işletmeden 25'i ve 60 başarılı işletmeden 55'i doğru tahmin edilmiş; başarısız işletmeler için %42 ve başarılı işletmeler için %92'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan dört yıl öncesi için toplam sınıflandırma başarısı %66,7 olarak elde edilmiştir.

### 3.2.8.7. Model 4'ün Sınıflandırma Sonuçlarının Karşılaştırılması



**Grafik 3.5.** Model 4'ün Sınıflandırma Sonuçları

Grafik 3.5.'de Model 4'ün yapay sinir ağıları, C5.0 algoritması ve destek vektör makineleri ile başarısızlıktan bir, iki, üç ve dört yıl öncesinden elde edilen sonuçların karşılaştırmalı grafiği verilmiştir. Başarısızlıktan bir, iki, üç ve dört yıl öncesinden ayrı ayrı analiz edilen Model 4'de, yapay sinir ağıları diğer iki yönteme göre daha yüksek doğru sınıflandırma elde ettiği, bunu takiben sırasıyla C5.0 ve destek vektör makineleri olduğu görülmektedir. Dört ayrı veri seti içinden en iyi sonucu ise, %93,3'lük doğru sınıflandırma ile başarısızlıktan bir yıl öncesindeki verilerle yapılan sınıflandırma vermektedir. 60 başarısız işletmeden 60'ı ve 60 başarılı işletmeden 52'si; toplamda da 120 işletmeden 112'si yapay sinir ağıları ile doğru sınıflandırılmıştır.

### 3.2.9. Sınıflandırma Matrisinin Ayırıcı Gücünün Testi

Sınıflandırma matrisinin ayırıcı gücünü araştırmada kullanılan istatistik testlerinden biri aşağıda formülü verilen Press's Q istatistiğidir. Hesaplanan değer bir kritik değerle (arzulanan anlamlılık seviyesinde 1 serbestlik dereceli Ki kare değeri) karşılaştırılır. Eğer hesaplanan değer kritik değeri aşarsa, o zaman sınıflandırmanın şansa göre istatistik olarak anlamlı olduğu kabul edilir.<sup>209</sup>

$$Press'sQ = \frac{[N - n(K)]^2}{N(K - 1)}$$

Burada;

N = Toplam örnek büyüklüğü,

n = Doğru Sınıflandırılan birimlerin sayısı,

K = Grup sayısı = 2 olmak üzere

<sup>209</sup> Bülent Öz, Türkiye'de Ticari Bankaların Başarısızlığında Etkili Olan Faktörlerin Çok Değişkenli İstatistik Yöntemlerle İncelenmesi, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2005, s.178.

**Çizelge 3.30. Modellerin Sınıflandırma Matrisinin Ayırma Gücünün Testi**

Model 1	YSA	C5.0	DVM	Ki kare	Serbestlik	Anlamlılık
	Press's Q				Derecesi	Seviyesi
T-1	67,500	64,533	61,633	6,635	1	0,01
T-2	45,633	43,200	36,300	6,635	1	0,01
T-3	43,200	43,200	30,000	6,635	1	0,01
T-4	30,000	2,700	16,133	6,635	1	0,01
Model 2	YSA	C5.0	DVM	Ki kare	Serbestlik	Anlamlılık
	Press's Q				Derecesi	Seviyesi
T-1	100,833	90,133	58,800	6,635	1	0,01
T-2	83,333	70,533	38,533	6,635	1	0,01
T-3	50,700	43,200	34,133	6,635	1	0,01
T-4	22,533	14,700	12,033	6,635	1	0,01
Model 3	YSA	C5.0	DVM	Ki kare	Serbestlik	Anlamlılık
	Press's Q				Derecesi	Seviyesi
T-1	100,833	90,133	64,533	6,635	1	0,01
T-2	76,800	70,533	34,133	6,635	1	0,01
T-3	45,633	45,633	43,200	6,635	1	0,01
T-4	22,533	19,200	13,333	6,635	1	0,01
Model 4	YSA	C5.0	DVM	Ki kare	Serbestlik	Anlamlılık
	Press's Q				Derecesi	Seviyesi
T-1	90,133	67,500	67,500	6,635	1	0,01
T-2	64,533	43,200	38,533	6,635	1	0,01
T-3	43,200	43,200	20,833	6,635	1	0,01
T-4	19,200	2,700	13,333	6,635	1	0,01

Test sonuçları incelendiğinde başarısızlıktan bir, iki, üç ve dört yıl öncesi dört farklı modeller için Model 1 ve Model 4'e ait dört yıl öncesi C5.0 algoritmasının hesaplanan test istatistik değerleri hariç diğer tüm test istatistik değerlerinin kritik ki-kare değerinden büyük olması nedeniyle kurulan modellerin şans modeline göre anlamlı olduğu anlaşılmaktadır.

### 3.2.10. 2010 Yılı Finansal Başarısızlık Tahmin Sonuçları

Analizde kullanılan dört modelde karşılaştırılmış ve Model 3'ün diğer üç modele göre daha yüksek doğru yüzdelerle tahmin yaptığı belirlenmiştir. Buradan hareketle 2010 yılına ait işletmelerin finansal başarısızlık tahminlerini belirlemede diğer üç modele göre kıyasla daha başarılı olan Model 3'ün bir yıl öncesine ait verileri kullanılarak tahmin gerçekleştirilmiştir.

2010 yılı finansal başarısızlık tahminleri gerçekleştirilirken 2009-2010 yılında faaliyetlerini devam ettirmeyen RAKSE, RKSEV, UZEL ve UKIM kodlu işletmeler

analiz kapsamından çıkartılarak 60 olan başarısız işletme sayısı 56'ya, aynı şekilde her iki grubun örnek içerisinde eşit sayıda temsil edilmesini sağlamak amacıyla, tesadüfi olarak seçilen 4 başarılı işletme CMBTN, VESTEL, YATAŞ ve PETKIM kapsam dışı bırakılarak başarılı işletme sayısı 56'ya çekilmesiyle toplamda 112 işletme analiz kapsamına dahil edilmiştir.

### 2010 Yılı İçin Tahmin Yöntemlerinin Karşılaştırılması

Model 3'ün ileriye yönelik tahmin amacıyla, başarısızlıktan bir yıl öncesi 2009 yılı verileri kullanılarak yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri ile elde edilen sınıflandırma başarıları aşağıda verilmiştir.

**Çizelge 3.31.** 2010 yılı için YSA-C5.0-DVM Sınıflandırma Başarısı

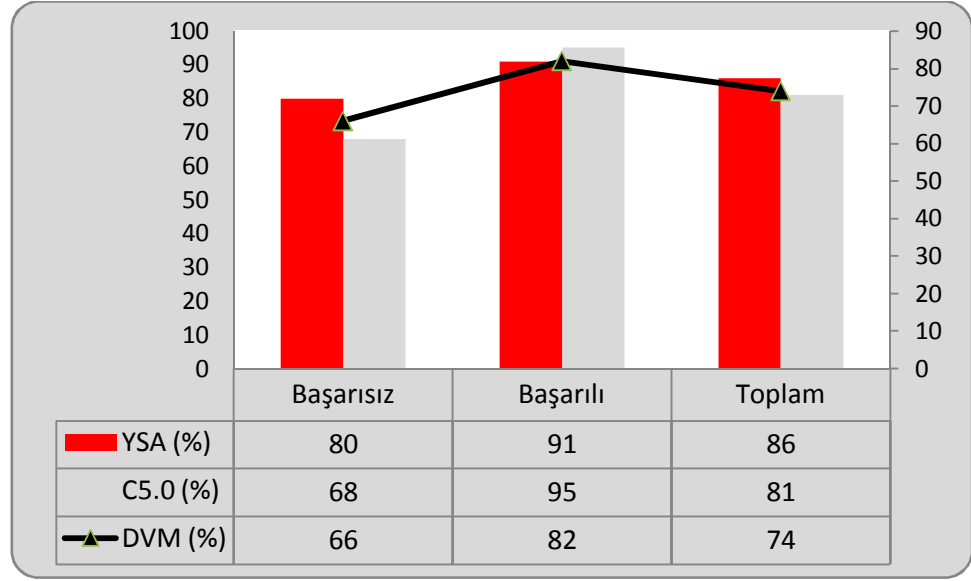
YSA Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>45</b>	11	56	80
	<b>Başarılı</b>	5	<b>51</b>	56	91
	<b>Toplam</b>	50	62	112	<b>86</b>
C5.0 Algoritmasının Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>38</b>	18	56	68
	<b>Başarılı</b>	3	<b>53</b>	56	95
	<b>Toplam</b>	41	71	112	<b>81</b>
Destek Vektör Makineleri Analizi		Tahmin Edilen Grup			
		Başarısız	Başarılı	Toplam	Doğruluk Yüzdesi
Gözlenen Grup	<b>Başarısız</b>	<b>37</b>	19	56	66
	<b>Başarılı</b>	10	<b>46</b>	56	82
	<b>Toplam</b>	47	65	112	<b>74</b>

Çizelge 3.31.'de görüldüğü gibi 112 işletmenin yer aldığı model 3 ait yapay sinir ağları yöntemi kullanılarak 56 başarısız işletmeden 45'i ve 56 başarılı işletmeden 51'i doğru tahmin edilmiş; başarısız işletmeler için %80 ve başarılı işletmeler için %91'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %86 olarak elde edilmiştir.

Model 3 ait C5.0 algoritması yöntemi kullanılarak 56 başarısız işletmeden 38'i ve 56 başarılı işletmeden 53'ü doğru tahmin edilmiş; başarısız işletmeler için %68 ve

başarılı işletmeler için %95'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %81 olarak elde edilmiştir.

Model 3 ait destek vektör makineleri yöntemi kullanılarak 56 başarısız işletmeden 37'si ve 56 başarılı işletmeden 46'sı doğru tahmin edilmiş; başarısız işletmeler için %66 ve başarılı işletmeler için %82'lik doğru sınıflandırma gerçekleştirilmiştir. Başarısızlıktan bir yıl öncesi için toplam sınıflandırma başarıları %74 olarak elde edilmiştir.



**Grafik 3.6.** 2010 Yılı Sınıflandırma Sonuçlarının Karşılaştırılması

Grafik 3.6.'da 2010 yılı yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri ile başarısızlıktan bir yıl öncesi 2009 yılı verileri kullanılarak elde edilen sonuçların karşılaştırmalı grafiği verilmiştir. Başarısızlıktan bir yıl öncesinden yapılan analiz kapsamında, yapay sinir ağlarının diğer iki yönteme göre daha yüksek doğru sınıflandırma elde ettiği gözlenmiştir. 112 işletmenin yer aldığı veri setinde YSA toplamda %86'lık bir doğru sınıflandırma yüzdesi ile ileriye yönelik bir sınıflandırma tahmininde bulunabileceğini göstermiştir.

### 3.2.11. Tüm Yıllar İçin Tahmin Modellerinin Performansları

İşletme finansal başarısızlıklarının tahmin edilmesine yönelik olarak yapılan uygulamada dört ayrı model oluşturulmuştur. Oluşturulan modellerin başarısızlıktan bir, iki, üç ve dört yıl öncesinden tahmin güçleri araştırılmış ve yapay sinir ağları, C5.0

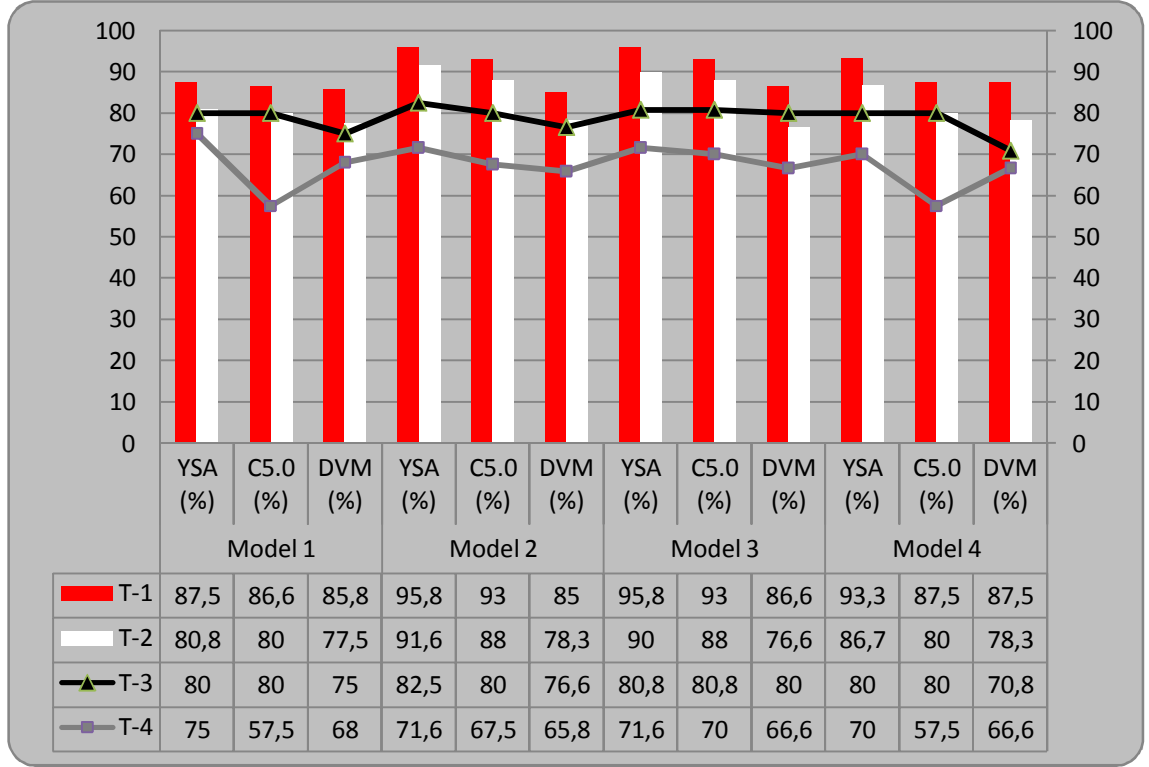


algoritması ve destek vektör makineleri yöntemiyle elde edilen sonuçlar karşılaştırılmıştır. Analiz sonuçları ele alındığında yapay sinir ağlarının, hem C5.0 algoritması hem de destek vektör makineleriyle karşılaştırıldığında, tüm modeller için genel anlamda daha iyi sonuçlar verdiği söylenebilir. Doğru sınıflandırma yüzdeleri sırasıyla yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri şeklinde sıralamak mümkündür.

Analizde incelenen bir, iki, üç ve dört yıl öncesi yapay sinir ağları, C5.0 algoritması ve destek vektör makineleri ile elde edilen sonuçların değiştiği gözlenmiştir. Her üç yönteme göre başarısızlığın gerçekleştiği yıldan uzaklaştıkça elde edilen tahmin sonuçlarının nispeten azaldığı benzer şekilde başarısızlığın gerçekleştiği yıla yaklaştıkça elde edilen tahmin sonuçlarının daha yüksek doğruluk yüzdelerine ulaştığını ifade etmek mümkündür. Her üç yönteme göre başarısızlıktan bir yıl öncesine ait tahmin sonuçlarının, sırasıyla iki, üç ve dört yıl öncesine ait sonuçlara göre daha yüksek sınıflandırma yüzdelerini elde ettiğini söyleyebiliriz.

Analiz sonuçları incelendiğinde, değişkenlerin modelden çıkarılmasının analizde kullanılan yöntemlerin tahmin gücünü artırdığı gözlenmektedir. Finansal başarısızlıktan bir, iki, üç ve dört yıl öncesinden verilerle yapılan tahminlerde, 3. Modelin 2. Modelden, 4. Modelin ise 1. Modelden daha yüksek doğru sınıflandırma yüzdesiyle tahmin başarıları gösterdiği görülmektedir.

Yukarıda yapılan açıklamalar ile birlikte, değişken seçimi için diskriminant analizinden elde edilen 3 değişkenli Model 3'ün, doğru sınıflandırma yüzdesi en iyi tahmin modeli olarak ortaya çıkmaktadır. Başarısızlıktan bir yıl öncesi verilerle yapılan analizde %95,8'lik doğru sınıflandırma yüzdesi elde edilmiştir. Bunu takiben yapay sinir ağları yöntemini kullanarak başarısızlıktan iki, üç ve dört yıl öncesi verilerle yapılan analizde sırasıyla % 90, %80,8 ve %71,6'lık doğru sınıflandırma yüzdesi elde edilmiştir.



**Grafik 3.7.** Her Dört Modelin Tüm Dönemlere İlişkin Performans Ölçüleri

Grafik 3.7.'de her dönem itibari ile dört modelin de toplam sınıflandırma doğruluğuna sahip değerler gösterilmiştir. Başarısızlıktan önceki birinci yılda en iyi performansı Model 3 göstermiştir. Model 3'ün başarısızlıktan bir yıl öncesindeki yapay sinir ağlarıyla %95,8, C5.0 algoritmasıyla %93 ve DVM yönteminin kullanılmasıyla %86,6'lık doğru sınıflandırma yüzdesi elde edilmiştir. Başarısızlıktan önceki ikinci yılda en iyi performansı Model 2 göstermiştir. Model 2'nin başarısızlıktan iki yıl öncesindeki yapay sinir ağlarıyla %91,6, C5.0 algoritmasıyla %88 ve DVM yönteminin kullanılmasıyla %78,3'lük doğru sınıflandırma yüzdesi elde edilmiştir. Başarısızlıktan önceki üçüncü yılda en iyi performansı Model 3 göstermiştir. Model 3'ün başarısızlıktan üç yıl öncesindeki yapay sinir ağlarıyla %80,8, C5.0 algoritmasıyla %80,8 ve DVM yönteminin kullanılmasıyla %80'lik doğru sınıflandırma yüzdesi elde edilmiştir. Başarısızlıktan önceki dördüncü yılda en iyi performansı Model 3 göstermiştir. Model 3'ün başarısızlıktan dört yıl öncesindeki yapay sinir ağlarıyla %71,6, C5.0 algoritmasıyla %70 ve DVM yönteminin kullanılmasıyla %66,6'lık doğru sınıflandırma yüzdesi elde edilmiştir. Her dört modelde yapay sinir ağları yönteminin diğer iki yönteme göre performansı biraz daha iyidir.

İşletme ile ilgili çeşitli kararları vermede, gerekli önlemleri alabilmede ve revizyon yapabilmeye iki yıl gibi bir sürenin oldukça uzun olduğu dikkate alınırsa yüksek doğruluk performansına sahip yapay sinir ağlarının faydalı bir yöntem olduğu ifade edilebilir. Bununla birlikte C5.0 algoritmasının performans sonuçları da gözden kaçırılmamalıdır. Başarısızlığın gerçekleştiği yıldan uzaklaştıkça C5.0 algoritmasının performansının, yapay sinir ağları performansına yakın sonuçlar gösterdiğini dikkate almak gerekir.

Uygulanan dört model kendi içinde tutarlı sonuçlar vermektedir. Buradan hareketle 2010 yılına ait Model 3'ün bir yıl öncesi 2009 verileri kullanılarak ileriye yönelik tahmin yapıldığında yapay sinir ağlarıyla %86, C5.0 algoritmasıyla %81 ve DVM yönteminin kullanılmasıyla %74'lük doğru sınıflandırma yüzdesi elde edilmiştir.

## SONUÇ

Bir işletmenin operasyonel durumu periyodik olarak açıklanan finansal bilgilerle gösterilir ve bu bilgilerden yola çıkılarak işletmelerin finansal başarı durumları tespit edilir. Finansal başarısızlığa uğramış işletmelerin sayısındaki artış işletmenin hem kendi kaynaklarının hem de ülke kaynaklarının etkin ve verimli kullanılmadığının bir göstergesidir. Buradan hareketle bu çalışmada işletme başarısızlıklarının önceden tahmini için veri madenciliği tekniklerinden C5.0 Algoritması ve destek vektör makineleri ile birlikte yapay sinir ağları kullanılarak sınıflandırma yöntemi (başarılı, başarısız) uygulanmıştır.

Çalışmanın uygulama bölümünde, 2002-2010 yılları arasında hisse senetleri İMKB’de işlem gören 120 sanayi işletmesi ele alınmıştır. Ele alınan işletmelerin bilanço ve gelir tablolarından belli kriterlere göre “finansal başarılı” ve “finansal başarısız” olarak sınıflandırma yapılmıştır. İşletmelerin başarısız olma durumunu en erken bir yıl olmak üzere dört yıl öncesinden belirleyebilmek için dört model oluşturulmuştur. Dört farklı modelin bir, iki, üç ve dört sene öncesinden C5.0, DVM ve YSA yöntemi ile işletme başarısızlıklarının tahmini için sınıflandırma yapılmış, ayrıca C5.0 algoritması ile karar ağaçları oluşturulmuş ve modellerin sınıflandırma matrisinin ayırım gücünün testi için Press’s Q analizi yapılmış ardından da ileriye yönelik 2010 yılına ait başarısızlık tahmini gerçekleştirilmiştir. YSA yönteminde çok katmanlı perseptron modeli tercih edilmiştir. DVM sınıflandırma yönteminde kernel fonksiyonlarından radyal tabanlı fonksiyon denenmiştir.

Uygulamada analiz öncesi sürecine göre;

- Model 1’de tüm değişkenler yer almıştır. Değişken seçimi yapılmamıştır. 28 değişken ele alınmıştır.
- Model 2’de değişken seçimi için lojistik regresyon analizi yapılarak değişken sayısı 4’e indirilmiştir.
- Model 3’de değişken seçimi için diskriminant analizinden elde edilen 3 değişken belirlenmiştir.
- Model 4’de değişken seçimi için feature selection yönteminden elde edilen 12 değişken ele alınmıştır.

Yöntemler açısından bakıldığında, YSA yönteminin, hem C5.0 hem de DVM yönetimiyle karşılaştırıldığında, tüm modeller için genel anlamda daha iyi sonuçlar verdiği gözlenmiştir. Toplamda doğru sınıflandırma yüzdelerinin doğruluğu açısından sıralama yapıldığında YSA, C5.0 ve DVM olarak tespit edilmiştir.

Her üç yöneme göre başarısızlığın gerçekleştiği yıldan uzaklaştıkça elde edilen tahmin sonuçlarının nispeten azaldığı, benzer şekilde başarısızlığın gerçekleştiği yıla yaklaştıkça elde edilen tahmin sonuçlarının daha yüksek tahmin sonuçlarına ulaştığını söylemek mümkündür. Her üç yöneme göre başarısızlıktan bir yıl öncesine ait tahmin sonuçlarının, sırasıyla iki, üç ve dört yıl öncesine ait sonuçlara göre daha yüksek tahmin değerlerini elde ettiği gözlenmiştir.

Modeller arasında kıyaslama yapıldığında Model 3'ün doğru sınıflandırma yüzdesiyle en iyi tahmin modeli olduğu ortaya çıkmaktadır. Başarısızlıktan bir yıl öncesi verilerle yapılan analizde % 95,8'lik doğru sınıflandırma yüzdesi elde edilmiştir. Buradan hareketle ileriye yönelik 2010 yılına ait finansal başarısızlık tahmini gerçekleştirildiğinde YSA ile % 86, C5.0 ile % 81 ve DVM yöntemi ile % 74 doğrulukla sınıflandırma sonuçları elde edilmiştir.

Sınıflandırma matrislerinin ayırım gücünü test etmek için kullanılan Press's Q istatistiğini hesaplanan değer bir serbestlik dereceli ki-kare değeri ile karşılaştırıldığında sınıflandırmaların şansa göre istatistiki olarak anlamlı olduğu gözlenmiştir.

Yukarıda yapılan tüm bu açıklama ve model karşılaştırmalarından bir sıralama ve model tercihi yapmak gerekirse başarısızlıktan bir yıl öncesi model en iyi model, iki yıl öncesi model ikinci en iyi, üç yıl öncesi model ise üçüncü ve dört yıl öncesi model ise dördüncü ve son sıradaki model konumundadır. Bu iki, üç ve dört yıl öncesi modellerin dikkate alınmaması şeklinde algılanmamalıdır.

Yapılan analizler neticesinde, C5.0 algoritmasının karar ağacından elde edilen kurallardan hareketle işletmelere öneride bulunmak gerekirse:

- İşletmelerin finansal başarısızlığa uğramaması adına aktif kârlılık oranlarını  $X_{21} > -0,009$  üzerinde tutması gerekir. Bu durum işletmeleri %89,4 oranında finansal anlamda başarılı kılacaktır.

- İşletmeler aktif kârlılık oranlarını  $X_{21} > 0,015$  üzerinde ve borç oranlarını  $X_9 \leq 0,714$  altında gerçekleştirebilirlerse finansal anlamda başarılı olabilmek için avantaj sağlayabileceklerdir.
- İşletmeler faizleri karşılama oranlarını  $X_{23} > 0,865$ , cari oranlarının  $X_1 > 0,917$  ve vergi öncesi kârdaki büyüme oranlarını  $X_{22} > -1,002$  gerçekleştirebilirlerse finansal anlamda başarılı olabilmek için avantaj sağlayabileceklerdir.
- İşletmeler asit test oranlarını  $X_2 > -0,009$ , faizleri karşılama oranlarını  $X_{23} > 0,005$  ve stok/satışlar oranlarını  $X_{15} \leq 4,198$  gerçekleştirebilirlerse finansal anlamda başarılı olabilmek için avantaj sağlayabileceklerdir.

Bu çalışma ile birlikte işletmelerin gelecekteki faaliyetlerini sürdürebilirliği konusunda önemli bilgiler ortaya çıkardığı görülmektedir. Elde edilen sonuçların diğer çalışma sonuçları ile karşılaştırılması uygun olacaktır. Bundan sonraki veri setleri geniş tutularak fuzzy logic (bulanık mantık) ile araştırma yapılabilir ve işletmeler başarısızlığa düşmeden birkaç sene öncesinden hem işletme yöneticilerine hem de yatırımcılara ve kredi kurumlarına önerilerde bulunabilir.

## KAYNAKÇA

- Adıyaman, F., *Talep Tahmininde Yapay Sinir Ağlarının Kullanılması*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Teknik Üniversitesi Fen Bilimler Enstitüsü, İstanbul 2007.
- Adriaans, Pieter, and Zantinge, Dolf, *Data Mining* [Elektronik Sürüm], Addison-Wesley, England Boston 1997.
- Akdağ, Y., *Vadeli İşlem ve Opsiyon Borsası'nda TL/Dolar Vadeli İşlem Sözleşmelerinin Gün Sonu Uzlaşma Fiyatının Yapay Sinir Ağları ile Tahmini*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi, Bankacılık ve Sigortacılık Enstitüsü, İstanbul 2010.
- Akgüç, Öztin, *Finansal Yönetim*, (Yenilenmiş 7. Baskı), Avcıol Basım-Yayın, İstanbul 1998.
- Akkoç, S., *Finansal Başarısızlığın Öngörülmesinde Sinirsel Bulanık Ağ Modelinin Kullanımı ve Amprik Bir Çalışma*, (Yayımlanmamış Doktora Tezi), Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü, Kütahya 2007.
- Akpınar, Haldun, "Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği," *İstanbul Üniversitesi İktisadi Fakültesi Dergisi*, 29(1), 2000, 1-22.
- Aktaş, R., Doğanay, M., ve Yıldız, B., "Mali Başarısızlığın Öngörülmesi İstatistiksel Yöntemler ve Yapay Sinir Ağı Karşılaştırması", *Ankara Üniversitesi Siyasal Bilimler Fakültesi Dergisi*, 58(4), 2003, 1-24.
- Aktaş, R., Endüstri İşletmeleri İçin Mali Başarısızlık Tahmini (Çok Boyutlu Model Uygulaması), *Türkiye İş Bankası Kültür Yayınları*, Yayın No: 323, Ankara 1993.
- Al-Hudaİry, Hazem, H.M. A.A., *Data Mining and Decision Making Support in the Governmental Sector*, (Yayımlanmamış Yüksek Lisans Tezi), University of Louisville Department of Computer Engineering and Computer Science, USA 2004.
- Alpaydın, E. *Introduction to Machine Learning*, MIT, US 2004.
- Anderson, D. ve McNeill, G., *Artificial Neural Networks Technology*, Kaman Science Corporation, USA New York 1992.

- Atiya, A. F., "Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results", [Yapay Sinir Ağlarını Kullanarak Kredi Riski İçin Başarısızlık Tahmini, Bir Uygulama ve Sonuçlar], *IEEE Transaction on Neural Networks*, 12(4), 2001, 929-935.
- Aydoğan, Ü., *Destek Vektör Makinelerinde Kullanılan Çekirdek Fonksiyonların Sınıflama Performanslarının Karşılaştırılması*, (Yayımlanmamış Yüksek Lisans Tezi), Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara 2010.
- Baş, N., *Yapay Sinir Ağları Yaklaşımı ve Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2006.
- Bilgehan, M., and Turgut, P., "The Use of Neural Networks In Concrete Compressive Strength Estimation" [Beton Mukavemet Tahmininde Yapay Sinir Ağları Kullanımı], *Computers and Concrete*, 7(3), 2010, 271-283.
- Burges, C. J. C., Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining Knowledge Discovery*, 2(2), 1998, 121-167.
- Burmaoğlu, S., *Birleşmiş Milletler Kalkınma Programı Beşeri Kalkınma Endeksi Verilerini Kullanarak Diskriminant Analizi, Lojistik Regresyon Analizi ve Yapay Sinir Ağlarının Sınıflandırma Başarılarının Değerlendirilmesi*, (Yayımlanmamış Doktora Tezi), Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Erzurum 2009.
- Büker, S., Aşıkoğlu, R., ve Sevil, G., *Finansal Yönetim*, (2. Baskı), Anadolu Üniversitesi, Eskişehir 1997.
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., ve Zanasi, A., *Discovering Data Mining: From Concept to Implementation*, (1 edition), Prentice Hall PTR, NJ Upper Saddle River USA, 1998.
- Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Shearer, Colin ve Wirth, Rüdiger, *CRISP-DM 1.0 Step-by-step data mining guide*, 2000.
- Chen, Ming-Syan, Han, Jiawei, and Yu, Philip S., "Data Mining: An Overview from Database Perspective", [Veri Madenciliği: Veritabanı Penceresinden Bir Bakış], *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 1996, 866-883.



- Civalek, Ömer, ve Calayır, Yusuf, “İnce Dikdörtgen Plakların Titreşim Frekanslarının Yapay Sinir Ağları ile Tahmini”, *İMO Teknik Dergi*, Yazı: 275, 2007, 4161 – 4176.
- Çakır, Özgür, *Veri Madenciliğinde Sınıflandırma Yöntemlerinin Karşılaştırılması: Bankacılık Müşteri Veri Tabanı Üzerinde Bir Uygulama*, (Yayımlanmamış Doktora Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2008.
- Çelik, B., *Yapay Sinir Ağları Metodolojisi ile Zaman Serisi: Teori ve Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul 2008.
- Çomak, E., *Destek Vektör Makineleri Çoklu Sınıf Problemleri İçin Çözüm Önerileri*, (Yayımlanmamış Yüksek Lisans Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2004.
- Çomak, E., *Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar*, (Yayımlanmamış Doktora Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2008.
- Demirci, D. A., *Destek Vektör Makineleri ile Karakter Tanıma*, (Yayımlanmamış Yüksek Lisans Tezi), Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, Sakarya 2007.
- Detienne K. B., Detienne D. H., Joshi S. A. “Neural Networks As Statistical Tools For Business Researchers”, [İşletmeciler Araştırmacıları için İstatistiksel Araç Olarak Yapay Sinir Ağları], *Organizational Research Methods*, 6(2), 2003, 236-265.
- Efe, M. Ö., ve Okyay K., *Yapay Sinir Ağları ve Uygulamaları*, Boğaziçi Üniversitesi Yayınevi, İstanbul 2000.
- Elman, J.L., “Finding structure in time”, [Zamanında Yapı Keşfi], *Cognitive Science*, 14(2), 1990, 179-211.
- Elmas, Ç., *Yapay Zeka Uygulamaları*, Seçkin Yayınevi, Ankara 2011.
- Fausett, L.V., *Fundamentals of Neural Networks* [Elektronik Sürüm], Prentice Hall, Canada 1993.
- Freeman, J.A., and Skapura, D.M., *Computation and Neural Systems Series* [Elektronik Sürüm], Addison-Wesley Publishing Company, USA 1991.
- Fu, Li Min., *Neural Networks in Computer Intelligence* [Elektronik Sürüm], 2nd Edition, McGraw Hill, New York 1994.

- Giudici, P., *Applied Data Mining: Statistical Methods for Business and Industry* [Elektronik Sürüm], West Sussex, John Wiley& Sons, England 2003.
- Gratzer, Karl., “Business Failure and The New Economy” [İşletme Başarısızlığı ve Yeni Ekonomi], *EBHA Conference 2001, Business and Knowledge: July 2001*, 20p, Oslo Norway 2001.
- Hair, Joseph F., Tahtam, Ronald L., Anderson, Rolph E., and Black, William, *Multivariate Data Analysis* [Elektronik Sürüm], Prentice Hall, New Jersey 1998.
- Ham, J., and Kamber, M., *Data Mining: Concepts and Techniques* [Elektronik Sürüm], Morgan Kaufmann Publisher, San Fransisco 2000.
- Hand, David., Mannila, Heikki, ve Smyth, Padhraic, *Principles of Data Mining* [Elektronik Sürüm], The Mit Press, England 2001.
- Hawley, Delvin.D., Johson, J.D., and Raina, D., “Artificial Neural Systems: A new tool for Finacial Decision Making” [Yapay Sinir Ağları Sistemleri: Yeni Bir Finansal Karar Verme Aracı], *Financial Anaylyst Journal*, 46(6), 1990, 63-72.
- Haykin S., *Neural Networks: A Comprehensive Foundation* [Elektronik Sürüm], Prentice Hall Inc, New Jersey 1999.
- Hopfied, J.J., “Neural Networks and Physical Systems with Emergent Collective Computational Abilities” [Acil Toplu Hesaplamalı Yeteneklerle Birlikte Fiziksel Sistemler ve Yapay Sinir Ağları], *Proceedings of National Academy of Science*, 79(8), 1982, 2554-2558,
- Hopgood, A.A., *Intelligent Systems for Engineers and Scientists* [Elektronik Sürüm], CRC Press, (2nd Edition), Florida 2001.
- Hristev, R.M., *The ANN Book* [Elektronik Sürüm], GNU Public License, (Edition 1), Boston USA 1998.
- Hsu, C.-C., Huang, Y.-P., Chang, K.-W., “Extended Naive Bayes Classifier for Mixed Data” [Karışık Veri için Genişletilmiş Naive Bayes Sınıflandırıcıları], *Expert Systems with Applications*, 35(3), 2008, 1080-1083.
- Hu, Y.H., and Hwang, J.N., *Handbook of Neural Network Signal Processing* [Elektronik Sürüm], CRC Press, Florida 2002.

- Huang, W., Lai, K.K., Nakamori, Y., and Wang, S., "Forecasting Foreign Exchange Rates With Artificial Neural Network: A Review" [Yapay Sinir Ağlarıyla Döviz Kurlarının Tahmini: Bir Uygulama], *International Journal of Information Technology & Decision Making*, 3(1), 2004, 145-165.
- Irmak, S., *Veri Madenciliği Yöntemleri ile Sağlık Sektörü Veritabanlarında Bilgi Keşfi: Tanımlayıcı ve Kestirimci Model Uygulamaları*, (Yayımlanmamış Doktora Tezi), Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü, Antalya 2009.
- Jain., A.K., and Mao J., "Artificial Neural Networks: A Tutorial" [Yapay Sinir Ağları: Bir Eğitimi], *IEEE Computer Special Issue on Neural Computing*, IEEE Computer Society, 1996, 31-44.
- Jang, J.S.R., Sun, C.T., and Mizutani, E., *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence* [Elektronik Sürüm], Printice Hall, Upper Saddle River NJ 1997.
- Jiawei, Han and Micheline, Kamber, *Data Mining Concepts and Techniques* [Elektronik Sürüm], Morgan Kaufman Publishers, USA 2001.
- Kaastra, Iebling and Milton S. Boyd, "Designing a Neural Network for Forecasting Financial and Economic Time Series" [Finansal ve Ekonomik Zaman Serileri Tahmini için Bir Yapay Sinir Ağı Dizaynı], *Neurocomputing*, 10(3), 1996, 215-236.
- Kantardzic, M., *Data Mining Concepts, Models, Methods and Algorithms* [Elektronik Sürüm], IEEE Press, Piscataway- NJ 2001.
- Kaplan, M., ve Tekeli, R., *Ekonomide Bekleyişler ve Tahmin: Yapay Sinir Ağları Uygulamaları*, Tablet Yayınları, Konya 2008.
- Karadeniz, Nur, *Müşteri İlişkileri Yönetimi Açısından Veri Madenciliği Yöntemi ve Hizmet Sektörü Üzerine Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2008.
- Kecman, V., *Learning and Soft Computing: Support Vector Machines. Neural Networks and Fuzzy Logic Models* [Elektronik Sürüm], MIT Press, Cambridge-Massachusetts 2001.
- Keskin, Yasemin, *İşletmelerde Finansal Başarısızlığın Tahmini, Çok Boyutlu Model Önerisi ve Uygulaması*, (Yayımlanmamış Doktora Tezi), Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara 2002.

- Kılıç, Y., *Finansal Başarısızlık Tahmininde Veri Madenciliğinin Kullanılması: İMKB’de Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Gaziantep Üniversitesi Sosyal Bilimler Enstitüsü, Gaziantep 2011.
- Kiremitçi, B., *Veri Ambarlarında Veri Madenciliği ve Ulaştırma-Lojistik Sektöründe Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2005.
- Koç M. L., Balas C. E., Arslan A., “Taş Dolgu Dalgakıranların Yapay Sinir Ağları İle Ön Tasarımı”, *İMO Teknik Dergi*, 15(225), 2004, 3351-3375.
- Kohonen, T., *Self-Organizing Maps* [Elektronik Sürüm], Springer Series in Information Sciences, (3rd Edition), New York 2001.
- Koyuncugil, Ali Serhan, *Bulanık Veri Madenciliği ve Sermaye Piyasalarına Uygulanması*, (Yayımlanmamış Doktora Tezi), Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara 2006.
- Kröse, B., Van Der Smagt P., *An Introduction to Neural Networks* [Elektronik Sürüm], The University of Amsterdam Press, Amsterdam 1996.
- Krysztof J., and Horis, W. P., *Data Mining A Knowledge Discovery Approach* [Elektronik Sürüm], Springer Sciences Business Media, New York 2007.
- Küçükönder, H., *Yapay Sinir Ağları ve Tarımda Bir Uygulama*, (Yayımlanmamış Doktora Tezi), Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş 2011.
- Larose, Daniel T., *Discovering Knowledge in Data: An Introduction to Data Mining* [Elektronik Sürüm], Wiley and Sons publishing, USA 2005.
- Lippmann, R.P., “An introduction to computing with neural nets” [Yapay Sinir Ağlarıyla Bir Hesaplama Giriş], *IEEE ASSP Magazine*, 4(2), 1987, 4-22.
- Luan, Jing, “Data Mining and Knowledge Management in Higher Education,” *Workshop and Presentation at 42nd Associate of Institutional Research International Conference*, Toronto, Canada 2002.
- Melgani, Farid, and Bruzzone, Lorenze, “Classification of Hyperspectral Remote Sensing Images with Support Vector Machines” [Destek Vektör Makineleriyle Hiperspektral Uzaktan Algılama Resimlerinin Sınıflandırması], *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 2004, 1778-1790.

- Michie, D., Spiegelhalter, D.J., Taylor, C.C., Machine Learning, *Neural and Statistical Classification* [Elektronik Sürüm], Elis Horwood Publisher, London 1994.
- Nabiyev, V.V., *Yapay Zeka : Problemler-Yöntemler - Algoritma*, Seçkin Yayıncılık, Ankara 2010.
- Ntungo, Chrispin, ve Milton, Boyd, “Commodity Futures Trading performance Using Neural Network Models Versus Arima Models” [Arıma Modellerine Karşı Yapay Sinir Ağları Modellerini Kullanarak Emtia Malların Ticari Performansı], *Journal Of Futures Markets*, 18(8), 1998, 965-983.
- Oğuzlar, Ayşe, “Veri Madenciliğinde Birliktelik Kuralları,” *Marmara Üniversitesi Sosyal Bilimler Enstitüsü Hakemli Dergisi Öneri*, 6(22), 2004, pp. 315-321.
- Okka, O., *Analitik Finansal Yönetim*, Nobel Yayın Dağıtım, Ankara 2009.
- Öz Bülent, *Türkiye’de Ticari Bankaların Başarısızlığında Etkili Olan Faktörlerin Çok Değişkenli İstatistik Yöntemlerle İncelenmesi*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2005.
- Özdamar, Kazım, *Paket Programlar ile İstatistiksel Veri Analizi*, Kaan Kitabevi, Eskişehir 2004.
- Özekes, Serhat, *Veri Madenciliği Uygulaması*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2002.
- Özkan, Yalçın, *Veri Madenciliği Yöntemleri*, Papatya Yayıncılık, İstanbul 2008.
- Özkaya, A.U., *Intelligent Arrhythmia Classification Based On Support Vector Machines*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2003.
- Özmen, Şule, *Ağ-Ekonomisinde Yeni Ticaret Yolu: e-Ticaret*, İstanbul Bilgi Üniversitesi Yayınları, İstanbul 2003.
- Özşahin, M., *Türkiye Otomotiv Sektörünün Kendini Örgütleyen Haritalar İle Finansal Analizi*, (Yayımlanmamış Yüksek Lisans Tezi), Çukurova Üniv., Fen Bilimleri Enstitüsü Endüstri Mühendisliği ABD, Adana 2009.
- Öztemel E., *Yapay Sinir Ağları*, (1. Baskı), Papatya Yayıncılık, İstanbul 2003.
- Özyirmidokuz, E.K., *Veri Madenciliği Tekniklerini Kullanarak İmalat Verilerinin Modellenmesi ve Analizi*, (Yayımlanmamış Doktora Tezi), Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, Kayseri 2009.

- Pavel Berkhin, *Survey of Clustering Data Mining Techniques* [Elektronik Sürüm], Accrue Software, California 2002.
- Piramuthu, Selwyn, “Evaluating feature selection methods for learning in data mining applications” [Veri Madenciliği Uygulamalarında Öğrenme Yöntemleri için Değerlendirme Özelliği Seçimi], *European Journal of Operational Research*, 156(2), 2004, 483–494.
- Sağiroğlu Ş., Beşdok E., Erler M., *Mühendislikte Yapay Zeka Uygulamaları I: Yapay Sinir Ağları*, Ufuk Kitap Kıratsiye-Yayıncılık, Kayseri 2003.
- Saraç, T., *Yapay Sinir Ağları*, Basılmamış Seminer projesi Gazi Üniversitesi, Ankara 2004.
- Scott, Dick., Meeks, Aleksandra, Last, Mark, Bunke, Horst, ve Kandel, Abraham, “Data mining in software metrics databases,” [Yazılım Ölçümleri Veritabanında Veri Madenciliği], *Fuzzy Sets and Systems*, 145 (1), 2004, 81-110.
- Scott, Paul D., “How Size Matters: The Role of Sampling in Data Mining” [Veri Madenciliğinde Örnekleme Yöntemi: Örnek Hacmi Ne kadar Önemli], R.A. Sarker (Ed.), *Heuristic and Optimization for Knowledge Discovery*, Idea Group USA, 2002, 122-141.
- Sezer, E.A., Bozkır, A.S., Yağız, S., ve Gökçeoğlu, C., “Karar Ağacı Derinliğinin CART Algoritmasında Kestirim Kapasitesine Etkisi: Bir Tünel Açma Makinasının İlerleme Hızı Üzerinde Uygulama”, *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu, Kayseri: ASYU 21- 4 Haziran 2010*.
- Shahnaz, Farial, *Decision Tree Based Algorithms* [Elektronik Sürüm], Michael W. Berry (Ed.), *Lecture Notes in Data Mining*, World Scientific Publisher, USA 2006.
- Sharma, Subhash, Vijay, Mahajan, “Early Warning Indicators of Business Failure” [İşletme Başarısızlığının Erken Uyarıcı Göstergeleri], *Journal of Marketing*, Vol. 44, Fall 1980, pp.80-89., Erişim Tarihi: 24 Şubat 2012, Web of Science
- Shmueli, G., Patel, N. R., ve Bruce, P. C., *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner*, John Wiley & Sons, Hoboken, NJ, USA 2007.
- Silahtaroglu, Gökhan, *Veri Madenciliği*, Papatya Yayıncılık Eğitim, İstanbul 2008.

- Sim, Jaesung, *Critical Success Factors in Data Mining Projects*, (Yayımlanmamış Doktora Tezi), University of North Texas, Texas 2003.
- Simpson, P.K., *Neural Networks Theory, Technology and Applications* [Elektronik Sürüm], The Institute of Electrical and Electronics Engineers Press, New York 1990.
- Sund, Reijo., "Utilization of Administrative Registers using Statistical Knowledge Discovery," International Workshop on Mining Official Data, *National Research and Development Centre for Welfare and Health*, Helsinki- Finland 2002.
- Şen Z., *Yapay Sinir Ağları İlkeleri*, Su Vakfı Yayınları, İstanbul 2004.
- Şentürk, Aysan, *Veri Madenciliği, Kavram Ve Teknikler*, Ekin Yayınevi, Bursa 2006.
- Şimşek, U.T., *Veri Madenciliği ve Müşteri İlişkileri Yönetiminde (CRM ) Bir Uygulama*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2006.
- Tan Wang, J., *Encyclopedia of Data Warehousing and Mining* [Elektronik Sürüm], Idea Group, USA 2006.
- Tan, P.N., Steinbach, M., ve Kumar, V., *Introduction to Data Mining*, Pearson, Addison- Wesley, Boston -MA -USA 2006.
- Tang, Z. H., ve MacLennan, J., *Data Mining with SQL Server 2005*, Wiley Publishing Inc, Indianapolis - USA 2005.
- Terlemez, Levent, *Eş İşlem Stratejisi Yöntemiyle İMKB'de Portföy Oluşturmada Veri Madenciliği Uygulaması*, (Yayımlanmamış Doktora Tezi), Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir 2008.
- Tokan, Türker, *Destek Vektör Makinelerinin Mikrodalga Teori ve Tekniğindeki Uygulamaları*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2009.
- Torun, T., *Finansal Başarısızlık Tahmininde Geleneksel İstatistikî Yöntemlerle Yapay Sinir Ağlarının Karşılaştırılması ve Sanayi İşletmeleri Üzerinde Uygulama*, (Yayımlanmamış Doktora Tezi), Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, Kayseri 2007.
- Türko, R. Metin., *Finansal Yönetim*, Alfa Yayın, İstanbul 1999.
- Two Crows Corp, *Introduction to Data Mining and Knowledge Discovery*, USA 2005.

- U.S. Defense Advanced Research Projects Agency(DARPA), *DARPA Neural Network Study*. 2.Basım. AFCEA International Press, USA 1988. Widrow, Bernard, DARPA Neural Network Study [Elektronik Sürüm], AFCEA International Press, USA 1988, ISBN 0-916159-17-5.
- Vapnik, V. N., *The Nature of Statistical Learning Theory* [Elektronik Sürüm], Springer-Verlag, NewYork 1995.
- Wasserman, P.D., *Advanced Methods in Neural Computing* [Elektronik Sürüm], Van Nostrand Reinhold, New York 1993.
- Witten, I. H., ve Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques* [Elektronik Sürüm], Second Edition, Morgan Kaufmann Publishers, San Francisco 2005.
- Yalama, A., Birol, Y., ve Ulaş Ünlü. “İlk Halka Arzlarda Uzun Dönem Getirilerinin Tahmini: Yapay Sinir Ağları ile İMKB İçin Ampirik Bir Uygulama”, *İstanbul Üni. İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, Sayı:10, 2009, 29-47.
- Yurtoğlu, H., *Yapay Sinir Ağları Metodolojisi ile Öngörü Modellemesi: Bazı Makroekonomik Değişkenler İçin Türkiye Örneği*, DPT yayım 2683, Ekonomik Modeller ve Stratejik Araştırmalar Genel Müdürlüğü, Ankara 2005.
- Zhang G., Patuwo B. E., Hu M. Y., “Forecasting With Artificial Neural Networks: The State Of The Art” [Yapay Sinir Ağlarıyla Tahmin: ], *International Journal Of Forecasting*, 14(1), 1998, 35-62.



**EKLER**

**Ek-1 Uygulama için Örnek Setlerini Oluşturan İşletmeler**

BORSA KODU	ŞİRKETİN ADI	BAŞARISIZ/ BAŞARILI
ADANA	Adana Çimento	1
AFYON	Afyon Çimento	1
AKALT	Akal Tekstil	0
AKENR	Ak Enerji	0
ALARK	Alarko Holding	1
ALCAR	Alarko Carrier	1
ALCTL	Alcatel Telekomikasyon	0
ALKIM	Alkim Kimya	1
ALYAG	Altınyag	0
ANACM	Anadolu Cam	1
ARCLK	Arçelik	1
ARSAN	Arsan Tekstil	0
ATEKS	Akın Tekstil	0
BERDN	Berdan Tekstil Sanayi ve Ticaret A.Ş	0
BFREN	Bosch Fren Sistemleri Sanayi ve Ticaret A.Ş	0
BISAS	Bisaş Tekstil Sanayi ve Ticaret A.Ş	0
BOLUC	Bolu Çimento	1
BOSSA	Bossa	1
BRISA	Brisa	1
BRMEN	Birlik Mensucat Ticaret ve Sanayi İşletmeleri A.Ş.	0
BROVA	Brova Yapı Endüstrisi A.Ş	0
BRSAN	Borusan Mannesman	1
BSOKE	Batisöke Çimento	1
BURCE	Burçelik Bursa Çelik Döküm Sanayii A.Ş	0
BURVA	Burçelik Vana Sanayii ve Ticaret A.Ş	0
BYSAN	Boyosan Tekstil	0
CBSBO	ÇBS Boya Kimya	0
CEMTS	Çemtaş	1
CEYLN	Ceylan Giyim	0
CIMSA	Çimsa	1
CMBTN	Çimbeton	1
COMDO	Componenta Döküm	1
DARDL	Dardanel	0
DENCM	Denizli Cam	0
DENTA	Dentaş Ambalaj	1
DERIM	Derimod	1
DITAS	Ditaş Doğan	0
DOBUR	Doğan Burda	1
DOGUB	Doğusan	0
DUROF	Duran Ofset	0
ECILC	Eczabaşı İlaç	1
ECYAP	Eczabaşı Yapı	0

## Ek-1 Devamı

EDIP	Edip İplik	0
EGGUB	Ege Gübre	1
EGPRO	Ege Profil	1
EGSER	Ege Seramik	1
EMNIS	Eminiş Ambalaj	0
ENKAI	Enka İnşaat	1
EPLAS	Ege Plastik	0
EREGL	Ereğli Demir Çelik	1
FRIGO	Friigo Pak Gıda	0
GEDIZ	Gediz İplik	0
GENTS	Gentaş	1
GOLDS	Goldas Kuyumculuk	1
GOLTS	Göлтаş Çimento	1
KARSN	Karsan Otomotiv	0
KENT	Kent Gıda	1
KERT	Kerevitaş Gıda	1
KLBM	Kelebek Mobilya	0
KONYA	Konya Çimento	1
KOZAA	Koza Madencilik	1
KRSTL	Kristal Kola	0
MAKTK	Makine Takım	0
MRDIN	Mardin Çimento	1
MRSHL	Marshall	1
MTESK	Metem Enerji ve Tekstil	0
NERGS	Nergis Holding	0
NUHCM	Nuh Çimento	1
OLMKS	Olmuksa	1
OTKAR	Otokar	1
PETKM	Petkim	1
PIMAS	Pimaş	1
PINSU	Pınar Su	1
POLYL	Polyl	0
PRTAS	ÇBS Printaş	0
PTOFS	Petrol Ofisi	1
RAKSE	Raks Elektronik	0
SASA	Advansa SASA	0
SELGD	Selçuk Gıda	0
SIFAS	Sifaş	0
SKTAS	Söktaş	1
TATKS	Tat Konserve	0
TBORG	Tuborg	0
TIRE	Mondi Tire Kutsan	0
TRCAS	Turcas Petrol	1

**Ek-1 Devamı**

TRKCM	Trakya Cam	1
TUDDF	T.Demir Döküm	0
TUMTK	Tümteks	0
TUPRS	Tupraş	1
UKIM	Uki Konfeksiyon	0
UNYEC	Ünye Çimento	1
UZEL	Uzel Makina	0
VAKKO	Vakko Tekstil	0
VKING	Viking Kağıt	0
YATAS	Yataş	1
YUNSA	Yünsa	0

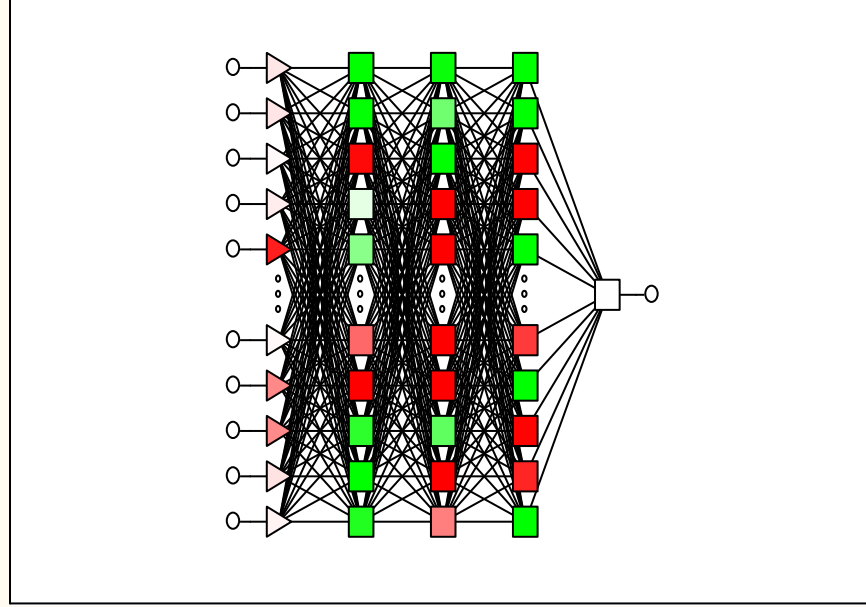
**Ek-1 Devam Test Seti**

Borsa Kodu	Şirketin Adı	Başarısız/ Başarılı
AEFES	Anadolu Efes	1
AKCNS	Akçansa	1
ASELS	Aselsan	1
BSPRO	Bosch Profilo	1
BTCIM	Batı Çimento	1
BUCIM	Bursa Çimento	1
CELHA	Çelik Halat	1
CMENT	Çimentaş	1
DMSAS	Demisaş Döküm	1
DYOBY	DYO Boya	0
ERSU	Ersu	0
ESEMS	Esem Spor	0
IZOCM	İzocam	1
MEMSA	Mensa	0
MERKO	Merko Gıda	0
PINSUT	Pınar Süt	1
PKART	Plastikkart A.Ş.	0
RKSEV	Raks Ev Aletleri	0
SERVE	Serve Kırtasiye	0
SKPLC	Şeker Piliç	0
SONME	Sönmez Filament	0
TOASO	Tofaş	1
VESTL	Vestel	1
ZOREN	Zoren	0

## Ek-2 Model 1-2-3-4 için Yapay Sinir Ağının Genel Mimari Yapısı

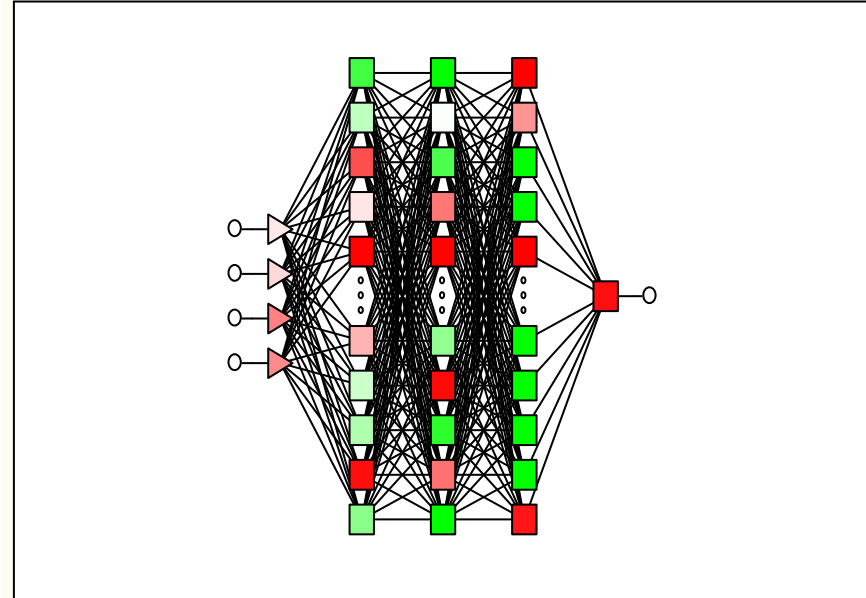
### Model 1 için YSA Genel Mimari Yapısı

Profile : MLP 28:28-12-15-17-1:1 , Index = 5  
 Train Perf. = 1,000000 , Select Perf. = 0,000000 , Test Perf. = 0,875000



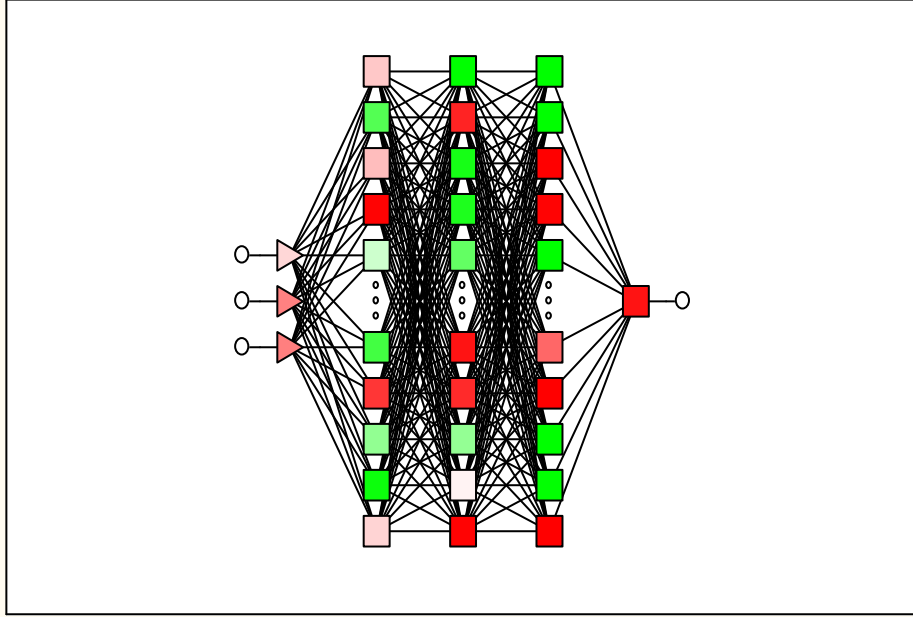
### Model 2 için YSA Genel Mimari Yapısı

Profile : MLP 4:4-12-15-17-1:1 , Index = 3  
 Train Perf. = 0,989583 , Select Perf. = 0,000000 , Test Perf. = 0,958333



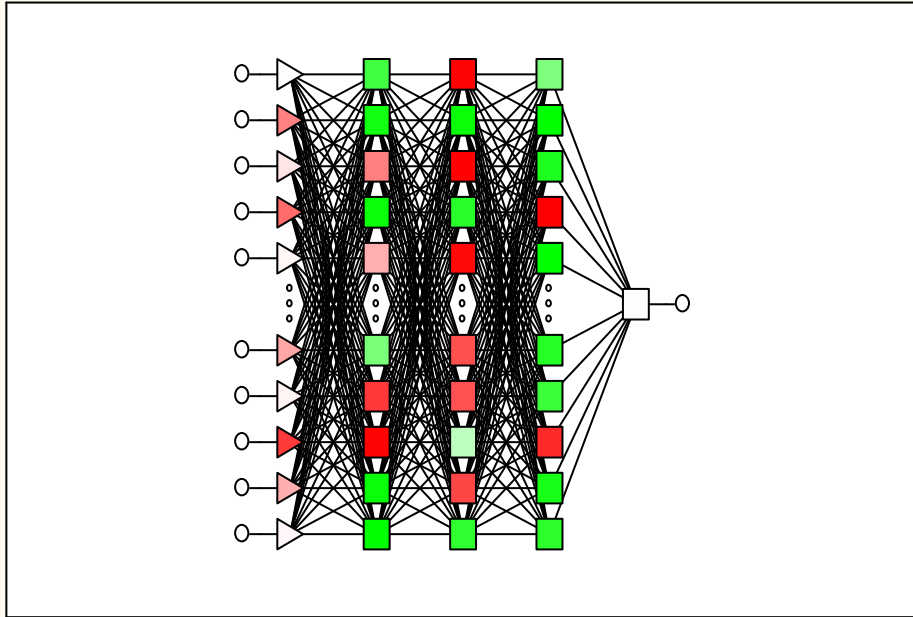
### Model 3 için YSA Genel Mimari Yapısı

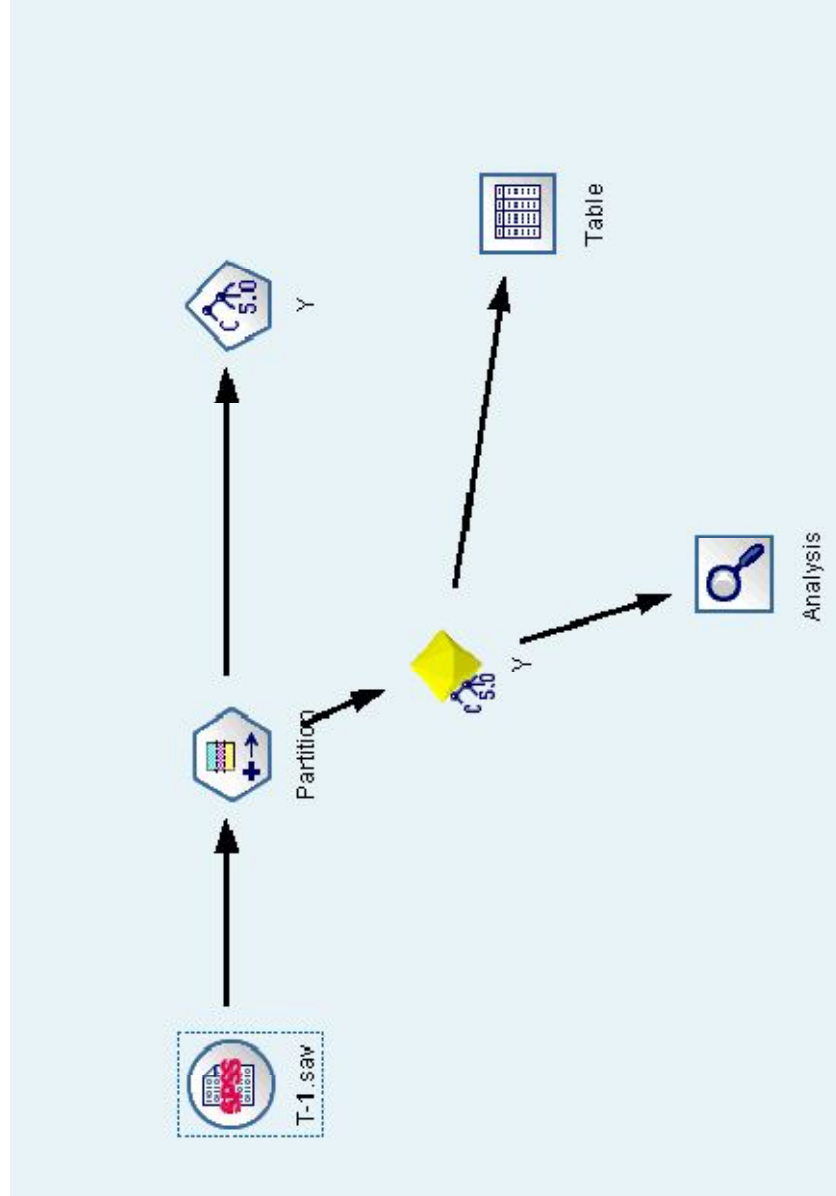
Profile : MLP 3:3-12-15-17-1:1 , Index = 1  
 Train Perf. = 0,968750 , Select Perf. = 0,000000 , Test Perf. = 0,875000



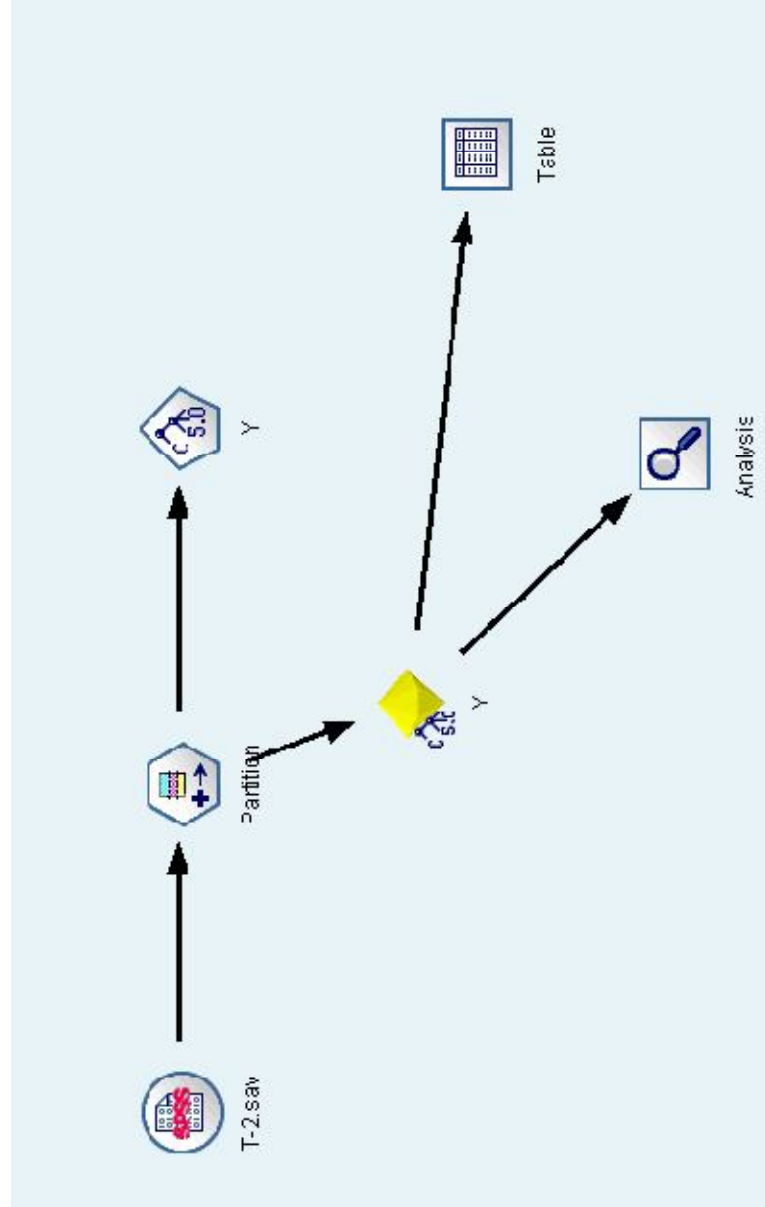
### Model 4 için YSA Genel Mimari Yapısı

Profile : MLP 12:12-12-15-17-1:1 , Index = 3  
 Train Perf. = 0,947917 , Select Perf. = 0,000000 , Test Perf. = 0,875000



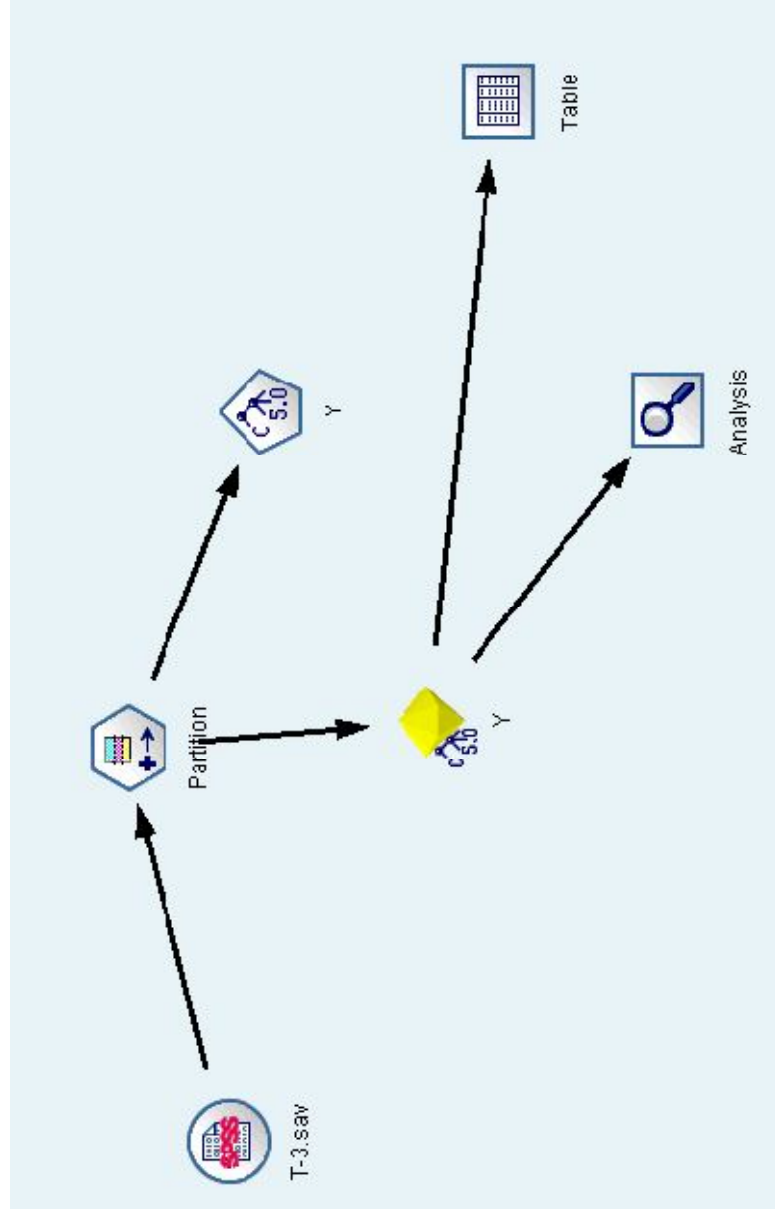
**Ek-3 T-1 Dönemi için C5.0 Algoritması Veri Akışı**

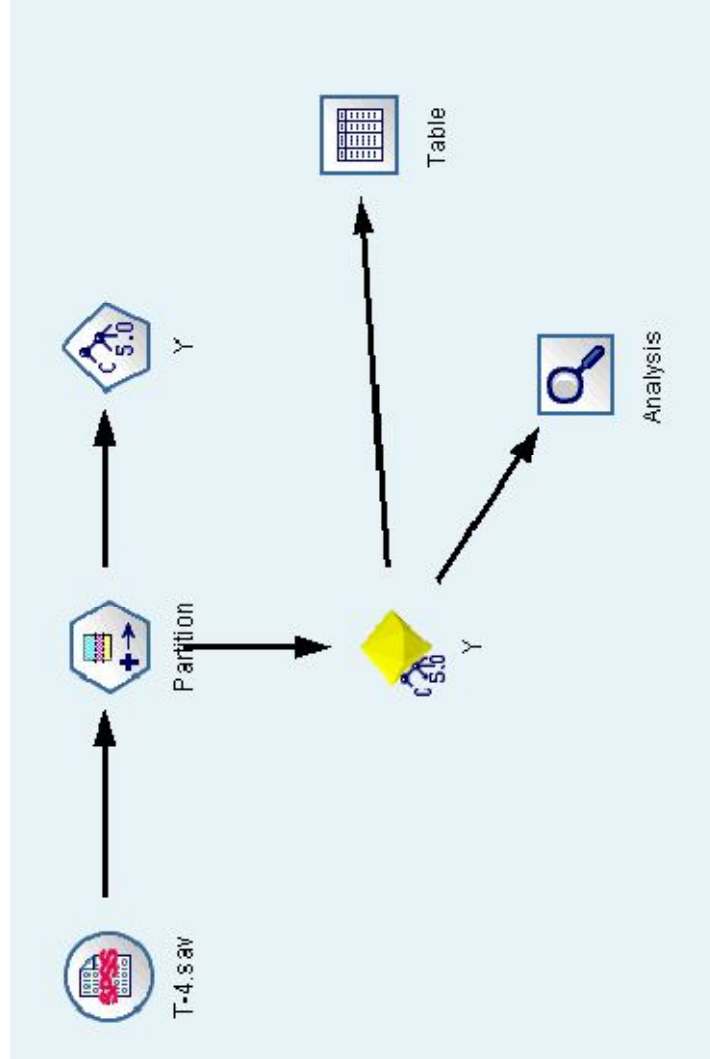
Ek-4 T-2 Dönemi için C5.0 Algoritması Veri Akışı



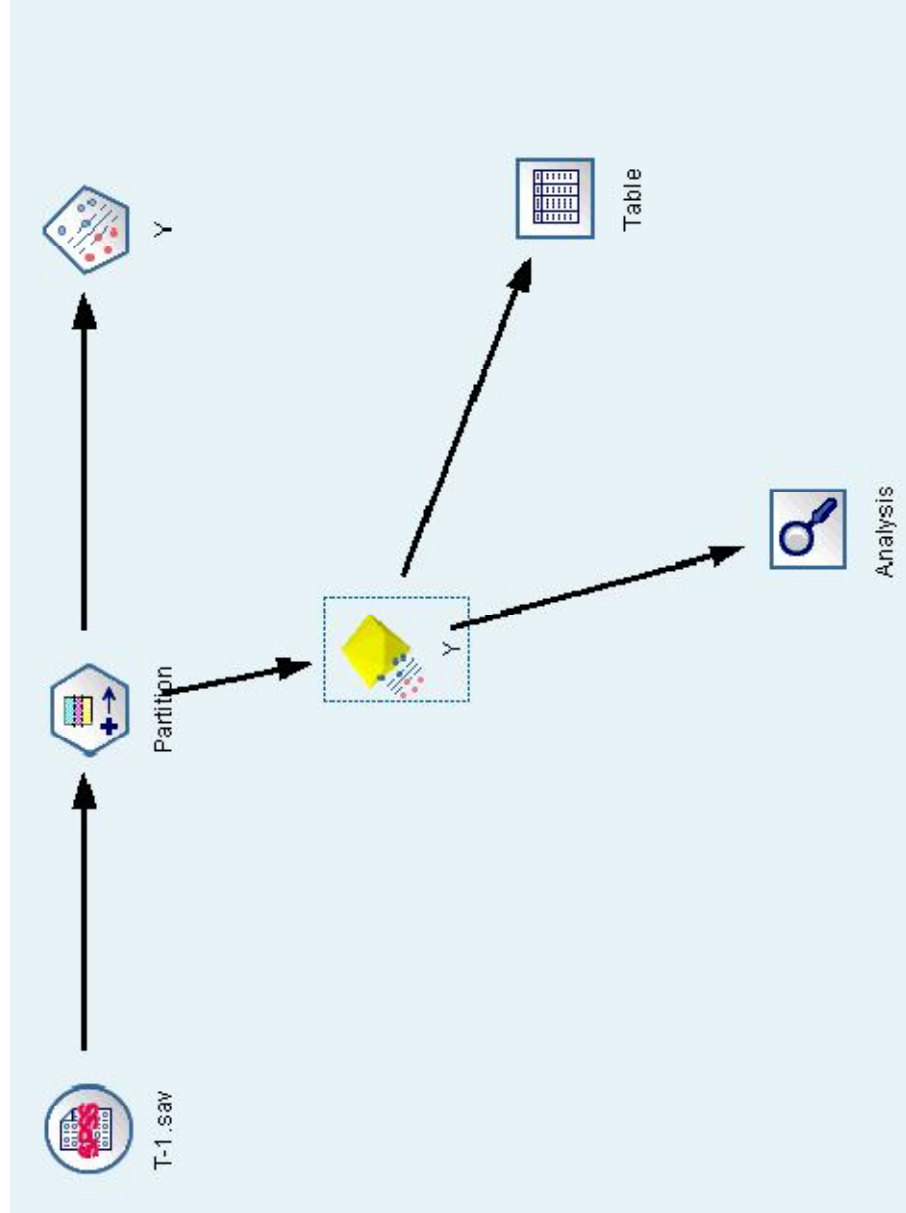


Ek-5 T-3 Dönemi için C5.0 Algoritması Veri Akışı

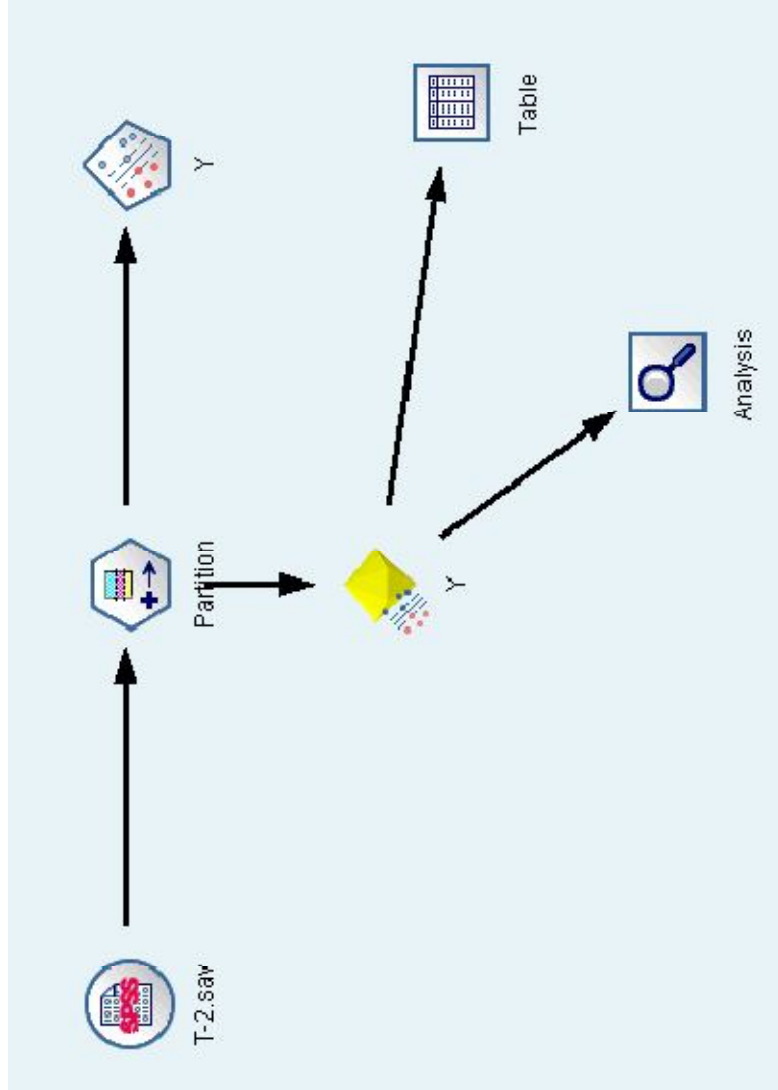


**Ek-6 T-4 Dönemi için C5.0 Algoritması Veri Akışı**

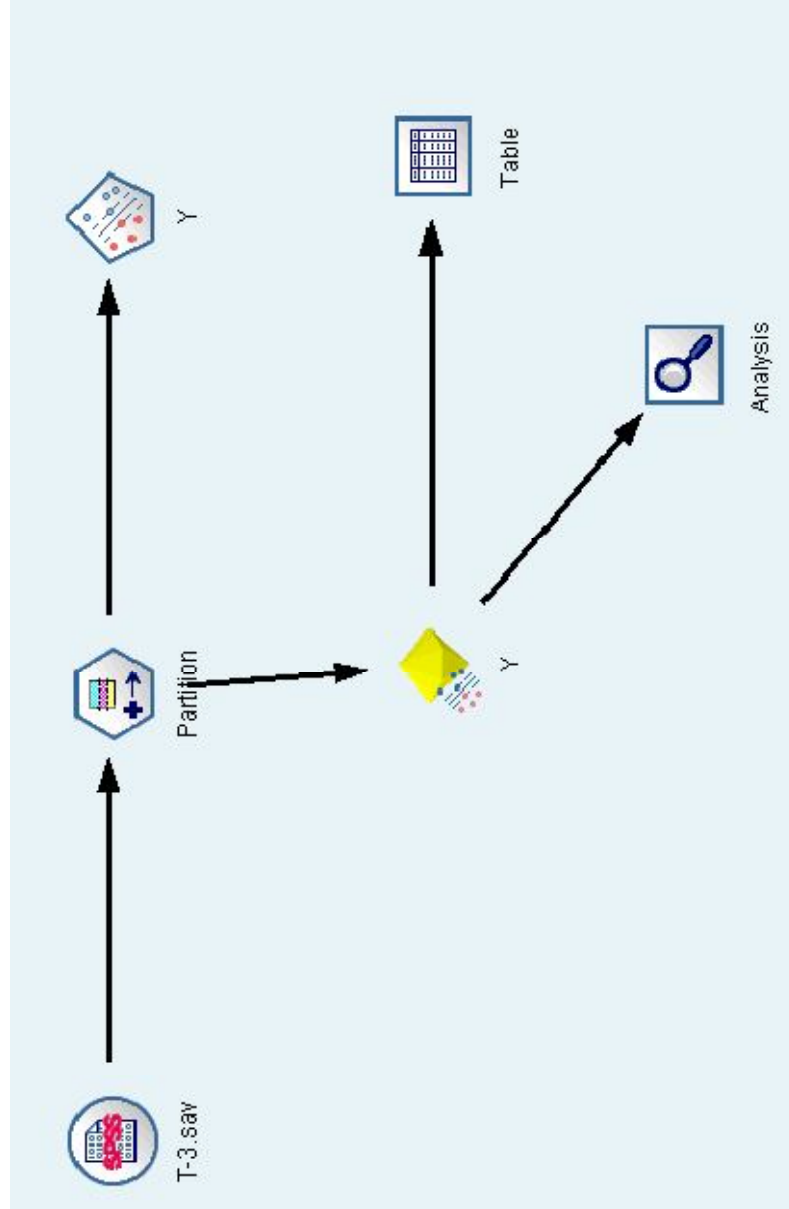
## Ek-7 T-1 Dönemi için DVM Veri Akışı



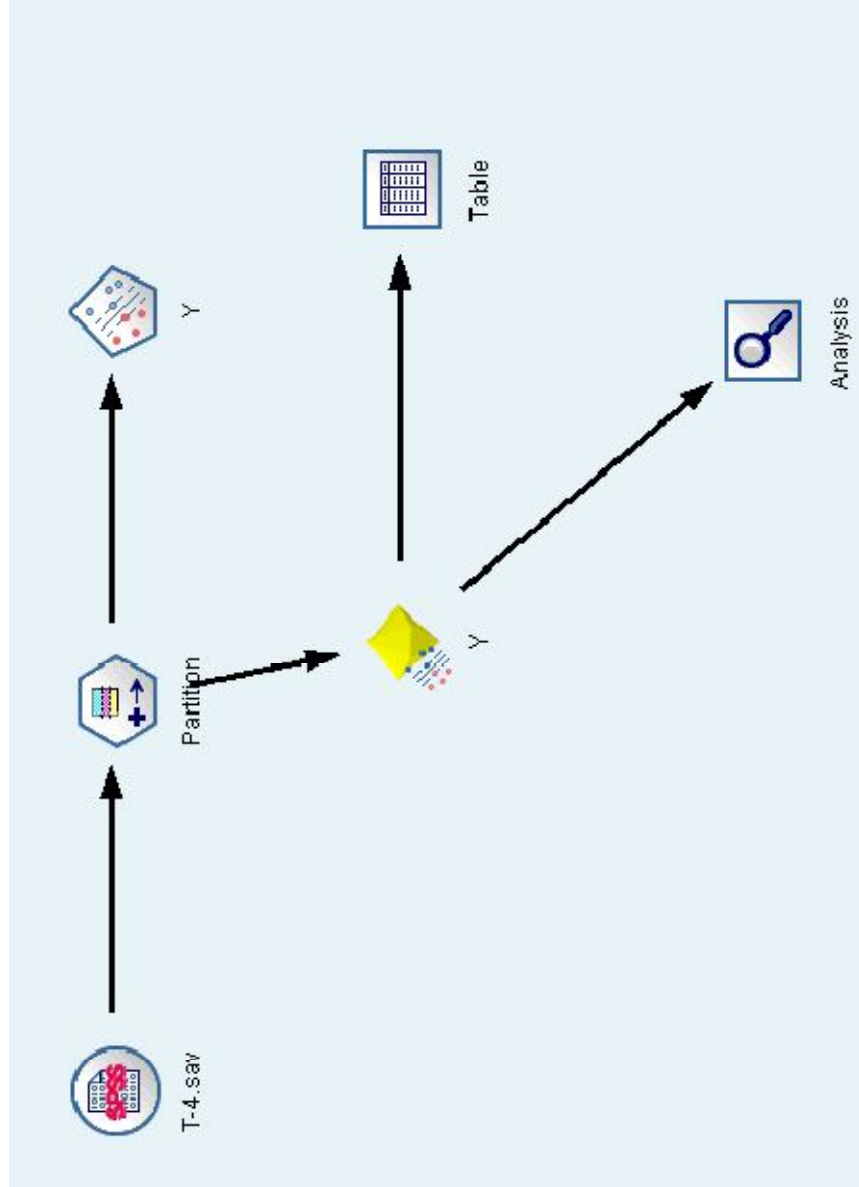
Ek-8 T-2 Dönemi için DVM Veri Akışı



## Ek-9 T-3 Dönemi için DVM Veri Akışı



Ek-10 T-4 Dönemi için DVM Veri Akışı



## Ek-11 Başarısızlıktan 1, 2, 3 ve 4 Yıl Öncesi İçin YSA Analizinin Tahmin Sonuçları

Şirket Kodu	Başarı Durumu	MODEL 1 Tahmin				MODEL 1 Tahmin				MODEL 3 Tahmin				MODEL 4 Tahmin			
		T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4
ADANA	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AEFES	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AFYON	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AKALT	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1
AKCNS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AKENR	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
ALARK	1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	1
ALCAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALCTL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALKIM	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALYAG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ANACM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ARCLK	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ARSAN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASELS	1	1	0	1	0	1	1	1	1	1	1	1	0	0	1	1	0
ATEKS	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
BERDN	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1
BFREN	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0
BISAS	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
BOLUC	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1
BOSSA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
BRISA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BRMEN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BROVA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
BRSAN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BSOKE	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
BSPRO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BTCIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BUCIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BURCE	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BURVA	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
BYSAN	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0	0	1
CBSBO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CELHA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CEMTS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CEYLN	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0
CIMSA	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CMBTN	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1
CMEN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
COMDO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DARDL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DENCM	0	0	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1
DENTA	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1
DERIM	1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	0	1
DITAS	0	1	0	1	1	0	1	1	1	1	0	1	1	0	1	1	1
DMSAS	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
DOBUR	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1
DOGUB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DUROF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DYOBY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ECILC	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ECYAP	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1
EDIP	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
EGGUB	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0
EGPRO	1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1
EGSER	1	0	1	1	1	0	0	1	1	0	0	1	1	0	0	0	1
EMNIS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ENKAI	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
EPLAS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EREGL	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
ERSU	0	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	1

## Ek-11 Devamı

ESEMS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRIGO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GEDIZ	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
GENTS	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
GOLDS	1	0	1	0	1	1	1	1	0	1	0	1	1	0	0	1
GOLTS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
IZOCM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
KARSN	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	1
KENT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
KERVT	1	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0
KLBMO	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
KONYA	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
KOZAA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
KRSTL	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
MAKTK	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1
MEMSA	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MERKO	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
MRDIN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
MRSHL	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
MTESK	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
NERGS	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
NUHCM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
OLMKS	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
OTKAR	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1
PETKM	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1
PIMAS	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0
PINSU	1	0	1	0	0	1	1	1	0	1	1	0	0	0	1	0
PINSUT	1	1	1	0	0	1	1	1	0	1	1	0	0	1	1	0
PKART	0	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1
POLYL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
PRTAS	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
PTOFS	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0
RAKSE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RKSEV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SASA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SELGD	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
SERVE	0	0	0	1	1	0	1	1	1	0	1	1	1	0	1	1
SIFAS	0	1	1	0	0	0	0	0	1	0	0	0	1	0	0	1
SKPLC	0	0	0	1	0	0	0	1	1	0	0	1	1	0	0	1
SKTAS	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
SONME	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
TATKS	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1
TBORG	0	0	1	0	1	0	0	0	1	0	0	0	1	0	0	1
TIRE	0	0	1	1	1	0	0	1	1	0	1	1	1	0	1	1
TOASO	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0
TRCAS	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
TRKCM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TUDDF	0	0	0	1	1	0	0	1	1	0	0	0	1	0	1	1
TUMTK	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
TUPRS	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
UKIM	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
UNYEC	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1
UZEL	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0
VAKKO	0	1	0	1	1	0	1	1	1	0	1	0	0	0	1	1
VESTL	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1
VKING	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	1
YATAS	1	1	1	0	0	1	0	1	0	1	0	0	0	0	0	1
YUNSA	0	1	0	1	0	0	0	1	1	0	1	1	1	0	1	1
ZOREN	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1



**Ek-12 Başarısızlıktan 1, 2, 3 ve 4 Yıl Öncesi İçin C5.0 Algoritmasının Tahmin Sonuçları**

Şirket Kodu	Başarı Durumu	MODEL 1 Tahmin				MODEL 2 Tahmin				MODEL 3 Tahmin				MODEL 4 Tahmin			
		T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4
ADANA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AEFES	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AFYON	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AKALT	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
AKCNS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AKENR	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
ALARK	1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
ALCAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALCTL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALKIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALYAG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ANACM	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1
ARCLK	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0
ARSAN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASELS	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
ATEKS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BERDN	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1
BFREN	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1
BISAS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BOLUC	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
BOSSA	1	0	0	0	1	1	0	1	0	1	0	1	0	0	0	0	1
BRISA	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0	1	0
BRMEN	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
BROVA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BRSAN	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
BSOKE	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1
BSPRO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BTCIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BUCIM	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1
BURCE	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
BURVA	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
BYSAN	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0	0	1
CBSBO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CELHA	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
CEMTS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CEYLN	0	0	1	1	1	0	1	0	1	0	1	0	1	0	1	1	1
CIMSA	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
CMBTN	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	0
CMENT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
COMDO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DARDL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DENCM	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
DENTA	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0
DERIM	1	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1	1
DITAS	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
DMSAS	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	0
DOBUR	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
DOGUB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DUROF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DYOBY	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0
ECILC	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	1	1
ECYAP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EDIP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EGGUB	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
EGPRO	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1
EGSER	1	0	1	1	1	0	0	1	1	0	0	1	1	0	1	1	1
EMNIS	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
ENKAI	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1
EPLAS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EREGL	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
ERSU	0	1	1	0	1	0	0	1	0	0	0	1	0	1	1	0	1

## Ek-12 Devamı

ESEMS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRIGO	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
GEDIZ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GENTS	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1
GOLDS	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1
GOLTS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
IZOCM	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1
KARSN	0	0	1	1	0	0	0	1	0	0	0	1	0	0	1	1
KENT	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1
KERVT	1	1	1	1	1	0	1	0	1	0	0	0	0	1	1	1
KLBM0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
KONYA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
KOZAA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
KRSTL	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1
MAKTK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MEMSA	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
MERKO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MRDIN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
MRSHL	1	0	0	0	0	1	1	0	1	1	1	0	1	0	0	0
MTESK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NERGS	0	0	0	1	1	0	0	1	1	0	0	0	0	0	1	1
NUHCM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
OLMKS	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1
OTKAR	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1
PETKM	1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
PIMAS	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1
PINSU	1	1	0	0	0	1	0	0	0	1	1	0	0	1	0	0
PINSUT	1	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1
PKART	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
POLYL	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1
PRTAS	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
PTOFS	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	0
RAKSE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RKSEV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SASA	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
SELGD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SERVE	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
SIFAS	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
SKPLC	0	0	0	1	0	0	0	1	1	0	0	1	1	0	0	1
SKTAS	1	1	1	1	1	0	0	1	1	0	0	1	1	1	1	1
SONME	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
TATKS	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0
TBORG	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
TIRE	0	1	1	1	1	0	0	1	1	0	0	1	1	1	1	1
TOASO	1	0	0	1	0	1	1	0	0	1	1	0	0	0	1	0
TRCAS	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	0
TRKCM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TUDDF	0	0	0	1	1	0	0	1	1	0	0	0	1	0	0	1
TUMTK	0	0	0	1	1	0	0	1	1	0	0	0	0	0	1	1
TUPRS	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
UKIM	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
UNYEC	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	1
UZEL	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1
VAKKO	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
VESTL	1	1	1	1	1	0	0	1	0	0	0	1	0	1	1	1
VKING	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
YATAS	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
YUNSA	0	0	1	1	1	0	0	1	1	0	0	1	1	0	1	1
ZOREN	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1

**Ek-13 Başarısızlıktan 1, 2, 3 ve 4 Yıl Öncesi İçin Dvm Yönteminin Tahmin Sonuçları**

Şirket Kodu	Başarı Durumu	MODEL 1 Tahmin				MODEL 2 Tahmin				MODEL 3 Tahmin				MODEL 4 Tahmin			
		T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4	T-1	T-2	T-3	T-4
ADANA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AEFES	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AFYON	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AKALT	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
AKCNS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AKENR	0	0	0	1	1	0	0	1	1	0	0	0	1	0	0	0	1
ALARK	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALCAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALCTL	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	1
ALKIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALYAG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ANACM	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ARCLK	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ARSAN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASELS	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ATEKS	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
BERDN	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
BFREN	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1
BISAS	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
BOLUC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BOSSA	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BRISA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BRMEN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BROVA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BRSAN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BSOKE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BSPRO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BTCIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BUCIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BURCE	0	1	0	0	0	1	0	0	1	1	1	0	1	1	0	0	1
BURVA	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1
BYSAN	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1
CBSBO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CELHA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CEMTS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CEYLN	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1
CIMSA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CMBTN	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1
CMENT	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
COMDO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DARDL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DENCM	0	0	0	1	1	0	1	1	1	0	1	1	1	0	0	1	1
DENTA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DERIM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DITAS	0	1	1	1	1	1	0	1	1	0	1	1	1	0	0	1	1
DMSAS	1	1	0	1	1	1	0	1	0	1	0	1	0	1	0	0	1
DOBUR	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1
DOGUB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DUOF	0	0	1	0	1	0	1	0	1	0	1	0	1	0	0	0	1
DYOBY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ECILC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ECYAP	0	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1
EDIP	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
EGGUB	1	0	0	0	1	1	1	1	0	1	1	1	1	1	0	0	1
EGPRO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EGSER	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
EMNIS	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
ENKAI	1	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1
EPLAS	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
EREGL	1	1	0	1	1	1	0	1	0	1	0	1	0	1	0	1	1
ERSU	0	1	0	0	1	1	0	0	0	1	0	0	0	1	0	0	1



## ÖZGEÇMİŞ

<b>Kişisel Bilgiler</b>	
Adı ve Soyadı	Emre YAKUT
Doğum Yeri ve Tarihi	Gaziantep – 09.09.1982
<b>Eğitim Durumu</b>	
Lisans Öğrenimi	Kocaeli Üniversitesi Mühendislik Fakültesi Endüstri Mühendisliği Bölümü
Y.Lisans Öğrenimi	Gaziantep Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı
Bildiği Yabancı Diller	İngilizce
Uluslararası bilimsel toplantılarda sunulan ve bildiri kitabında ( <i>Proceedings</i> ) basılan Bildiriler	BAKIRCI, F., YAKUT, E., DEMİRCİ, A., ve GÜNDÜZ, M., (2012), Türkiye'deki Kömür İşletmeciliğinin Veri Zarflama Analizi ve Veri Madenciliği ile Etkinlik Ölçümü, 13th International Conference on Econometrics, Operations Research and Statistics, 24-26 Mayıs 2012, Kıbrıs".  ELMAS, B., ALKAN, Ö., YAKUT, E., "Finansal Başarısızlığın Tahmin Edilmesinde Yapay Sinir Ağları Ve Lojistik Regresyon Modelinin Karşılaştırılması, 12. International Symposium On Econometrics Statistics and Operations Research, 26-29.05.2011 Denizli."
Ulusal hakemli dergilerde yayınlanan makaleler	ÖZÇOMAK, M.S., GÜNDÜZ, M., DEMİRCİ, A., YAKUT, E., (2012), "Çeşitli İklim Ve Ürün Verileri Arasındaki İlişkinin Kanonik Korelasyon Analizi Ve Veri Zarflama Analizi Yöntemleri İle İncelenmesi", Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 2012; 26(1):111-131  GEREK, İ.H., ERDİŞ, E., YAKUT, E., "Finansal Performansa Dayalı Etkinlik Ölçümü: Çimento Sektörü Uygulaması", NWSA-5813-2704, Yayın No:1A0303, s:311-321  ELMAS, B., YAKUT, E., ALKAN, Ö., (2011) "İşletmelerin Mali Başarısızlığının Yapay Sinir Ağları ve Lojistik Regresyon Modeli ile Tahmin Edilmesi". <i>Finans Politik &amp; Ekonomik Yorumlar</i> , 48(560), 45-56  ELMAS, B., ve YAKUT, E., (2010).Türk İmalat Sektörünün Etkinliğine Krizlerin Etkileri 1998-2008, <i>Finans Politik ve Ekonomik Yorumlar Dergisi</i> , 2010; 47(546):0-0

	ATA, A., ve YAKUT, E., Finansal Performansa Dayalı Etkinlik Ölçümü: İmalat Sektörü Uygulaması, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2009; 18:80-100.
Ulusal bilimsel toplantılarda sunulan ve bildiri kitabında basılan bildiriler	GEREK, İ.H., ERDİŞ, E., YAKUT, E., Finansal Performansa Dayalı Etkinlik Ölçümü: Çimento Sektörü Uygulaması, 6. İnşaat Yönetim Kongresi, 2011, Bursa  YAKUT, E., ELMAS, B., (2011), İşletmelerin Mali Başarısızlığının Veri madenciliği ve Diskriminant Analizi Modelleri ile Tahmin Edilmesi, 15. Finans Sempozyumu, Malatya  ATA, A., YAKUT, E., (2008). Çalışma Sermayesi Yöntemi ve Karlılık İlişkisi: İmalat Sektörü Uygulaması, 12. Ulusal Finans Sempozyumu Kayseri
<b>İş Deneyimi</b>	
Stajlar	
Projeler	
Çalıştığı Kurumlar	
<b>İletişim</b>	
E-Posta Adresi	emreyakut82@yahoo.com
<b>Tarih</b>	06.12.2012