

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

DOKTORA TEZİ

**KREDİ RİSKİNİN BELİRLENMESİNDE YAPAY
ZEKÂ YAKLAŞIMLARI VE BİR UYGULAMA**

Gökhan KORKMAZ

2502160279

TEZ DANIŞMANI

Prof. Dr. Ergün EROĞLU

İSTANBUL, 2020

ÖZ

KREDİ RİSKİNİN BELİRLENMESİNDE YAPAY ZEKÂ YAKLAŞIMLARI VE BİR UYGULAMA

GÖKHAN KORKMAZ

Yapay zekâ yöntemleri, emekleme aşamasını hatta yürüme dönemlerini çoktan atlattı ve teknolojinin faydalarından yararlanmak ve böylelikle işlerini daha düzgün, daha verimli ve yaratıcı bir formasyonda yapmak isteyen tüm çevrelerin odak noktası haline geldi. Günden güne hızla gelişen bu yöntem ve yaklaşımların kredi riski değerlendirme gibi kadim problemlere ürettikleri çözümler de en çok merak edilen ve araştırılan konuların başında geliyor. Bu tez çalışmasında, yapay zekâ yöntemlerinden, Yapay Sinir Ağları (YSA) ve Destek Vektör Makineleri (DVM), geleneksel istatistiki yöntemlerden ise Lojistik Regresyon (LR) kullanılmıştır. Yöntemler, bir bankadan alınan, üç farklı gerçek hayat verisine uygulanmıştır. Bunlardan birincisini, bireysel verilerin yer aldığı ve 407 adet gözlemden oluşan "Bireysel Veri Kümesi" oluşturmaktadır. İkincisini ise, kurumsal müşteri verilerinin bulunduğu ve 12921 adet gözlemden oluşan "Kurumsal Veri Kümesi" oluşturmaktadır. Üçüncüsünü ise, yine ikincil veri kümesinden türetilmiş, 1590 tanesi kusurlu ve 1590 tanesi kusursuz müşterilerin verilerinden ve toplam 3180 adet gözlemden oluşan "Dengeli Kurumsal Veri Kümesi" oluşturmaktadır. Her üç veri kümesinde de, DVM en yüksek doğruluk oranını veren yöntem olmuştur. Kurumsal veri kümesinden, dengeli kurumsal veri kümesine geçildiğinde, veri miktarında yaklaşık %75 oranında azalış meydana gelmiştir. Veri miktarındaki bu azalışın, doğruluk oranı bakımından, DVM’ni hiç etkilemediği (%1 arttı), YSA’nı (%22 azaldı) ve LR’u (%30 azaldı) ise oldukça fazla etkilediği görülmüştür. Dolayısıyla sınırlı veri söz konusu olduğunda, DVM’nin diğer yöntemlere kıyasla daha başarılı bir yöntem olduğu, bu tez çalışmasında da görülmüştür.

Anahtar Kavramlar: Kredi riskinin belirlenmesi, yapay sinir ağları, destek vektör makineleri, model karmaşıklığının kontrolü

ABSTRACT

ARTIFICIAL INTELLIGENCE APPROACHES AND AN APPLICATION IN DETERMINING CREDIT RISK

GÖKHAN KORKMAZ

Artificial intelligence methods have already survived the infancy and even the walking periods, and it has now become the focus of all circles who want to take advantage of the benefits of technology and thus do their jobs in a smoother, more efficient and creative formation. The solutions created by these methods and approaches that are rapidly developing day by day to ancient problems such as credit risk valuation are among the most curious and researched topics. In this thesis, among the artificial intelligence methods, Artificial Neural Networks (ANN) and Support Vector Machines (SVM), and from traditional statistical methods, Logistic Regression (LR) were used. The methods are applied to three different real life data from a bank. The first of these is the "Individual Data Set" consisting of individual data and 407 observations. The second is the "Corporate Data Set", which contains corporate customer data and consists of 12921 observations. The third is the "Balanced Corporate Data Set", which is also derived from the secondary data set, consisting of 1590 defective and 1590 flawless customer data and 3180 observations in total. In all three data sets, SVM was the method that gave the highest accuracy. When the transition from the corporate data set to the balanced corporate data set, the amount of data decreased by approximately 75%. It was observed that this decrease in data amount did not affect SVM at all (1% increase), ANN (22% decrease) and LR (30% decrease) in terms of accuracy rate. Therefore, in the case of limited data, it has been seen in this thesis that SVM is a more successful method compared to other methods.

Keywords: Determination of credit risk, artificial neural networks, support vector machines, control of model complexity

ÖNSÖZ

Ekonomilerin üzerine kurulu olduğu “piyasa” sistemlerinin işlerliği, bu piyasaların oyuncularları arasında garantör ve finansör olarak rol alan, kanunlarla kayıtlı, finansal işleyişin ana alterleri diyebileceğimiz ve bu yüzden güçlü sermaye gereksinimleri olan bankaların sağlıklı ve istikrarlı işlerliğine bağlıdır. Bu da onların ana faaliyet konusu olan kredi verme faaliyetinin ve buna bağlı olarak oluşan kredi riskinin doğru ölçülmesi ve yönetilmesi ile mümkündür. Fakat teknik olarak karar verilmesinin çok zor olduğu ve ciddi bir değerlendirme süreci gerektiren, yüksek meblağlı kredilerin yanısıra tipik, küçük meblağlı ve banka çalışanlarını çok uğraştıran ve onları daha nitelikli iş geliştirme süreçlerinden alıkoyan çok sayıda kredinin de değerlendirilmesi gerekmektedir.

Kredilendirme sürecini daha az risk alarak yönetmek isteyen finans kuruluşları, yani büyük ölçüde bankalar, bu işlemleri daha doğru ve daha hızlı bir şekilde yapabilmek için, modern yöntemler diyebileceğimiz yapay zekâ yaklaşımlarını kullanmalarının kaçınılmazlığını, günümüz piyasa ve rekabet şartlarında, giderek daha çok hissettirmektedirler. Yapay zekâ yöntemlerinin etkin olarak kullanımı, kredi değerlendirme süreçlerinin etkinliğini hızla artıracaktır. Bu da bankaların, gelirlerini artırırken kayıplarını minimize etmelerine ve daha güçlü bir bankacılık sistemine sahip olmalarına, bu da daha çok yatırımın finanse etmelerine olanak sağlayacaktır.

Yapay zekâ yöntemlerinden YSA, belki de muhtemel binlerce farklı mimarinin kullanılabileceği, dolayısıyla uygulaması çok zaman alan, yine etkin ve aynı zamanda çeşitli türevleriyle gelişime açık, verimli bir yöntemdir. Bu bakımdan, YSA kullanan bankalar, kendi ağlarını ve mimarilerini oluşturmak suretiyle, özgünleşmektedirler.

Yapay zekâ yöntemlerinden DVM ise, birkaç tane farklı çekirdek fonksiyonu kullanılabilen, yani mimarisine, birkaç işlem sonrasında, çok kolay karar verilebilen ve sonuçları itibarıyla da etkin bir yöntemdir. Son olarak da DVM, matematiksel olarak ispatlanmış güçlü bir alt yapıya ve kapsamlı bir istatikî teoriye sahiptir. Ayrıca DVM eğitim hatasının minimize edilmesinin yanısıra genelleme hatasını da minimize edebilmektedir. Bu özelliklerinden dolayı modelin türevlerine olan ilgi her geçen gün

artmakta ve modeli sadece kendi çalışma alanlarına uygulamak isteyen sosyal bilimciler için değil aynı zamanda modellerin içeriği üzerinde çalışmak isteyen matematikçiler ve mühendisler için de oldukça geniş ve bakir bir çalışma alanı sağlamaktadır.

Yapay zekâ yöntemlerinin sayısı, sezgisel yöntemler de dâhil olmak üzere, günden güne artmakta ve mevcuttaki algoritmalar ise çeşitli parametreler eşliğinde geliştirilmektedir. Veri tabanı yönetim sistemleri sayesinde çok fazla veriyi saklayıp kullanabilen bankalar ise, veri bankalarındaki bu birikimleriyle, yapay zekâ çalışmalarına destek olmakta ve öğrenme düzeylerini de hızla artırmaktadırlar. Kısaca, bugünkü şartlarda bile %90 ve üzeri başarı sağlayan bu yöntemler önümüzdeki birkaç on yılda, %100'e çok yakın, güvenilir, tutarlı ve hızlı sonuçlar üretecektir. Bu yöntemleri iyi kullanabilen piyasa oyuncularını ise, piyasanın önemli aktörleri haline gelebilecektir. Bu konu, ehemmiyeti ve yöntemlerin çeşitliliği ve bu yöntemlerin gelişmeye olan yatkınlıkları dolayısıyla, üzerinde çalışılmayı hakeden konuların başında geliyor.

Bu tezin yazılmasında emeği geçen, yanında iken kendimi inanılmaz rahat hissettiğim, ayrıca beni sayısal yöntemlerde doktora başlarken, çalışma alanı olarak yapay zekâyı belirlerken ve amfide karmaşık sayısal dersleri anlatmak için hazırlarken vb. her konuda cesaretlendiren danışman hocam, sevgili büyüğüm, Prof. Dr. Ergün Eroğlu'na çok teşekkür ederim. Tezimin her aşamasında, engin bilgisi, tecrübesi ve güleryüzlü yardımseverliğiyle, hiçbir karşılık beklemeden, sonuna kadar yardımcı olan, bana kodlamayı öğreten, odamızda çay içerken kendisinden çok şey öğrendiğimiz ve keşke daha önce tanışsaymışız dediğim sevgili hocamız Prof. Dr. Tunçhan Cura'ya çok teşekkür ederim. Beni doktora hazırlayan, sabahlara kadar okulda kalıp yüksek lisans tezimin analizlerini yapan, danışman hocam, muhterem büyüğüm, Prof. Dr. Bahattin Karademir'e sonsuz teşekkürler ederim. Doktora yaparken bize istatistiği öğreten, yeterlilik sürecimizi başarı ve motivasyonla tamamlayan Prof. Dr. Çiğdem Arıcıgil Cilan olmak üzere, bizi öğrenme ve araştırmaya teşvik eden, saygıdeğer hocamız Prof. Dr. Mehpare Timor'a, kitaplarından son derece istifade ettiğimiz, Prof. Dr. Neyran Orhunbilge'ye, arkadaşım Arş. Gör. Uğur

Arcagök'e ve aralarında bir ömür boyu doktora yapabileceğim tüm Sayısal Yöntemler Kürsüsüne çok teşekkür ederim. Dualarıyla her daim arkamda hissettiğim geniş aileme çok teşekkür ederim. Hayatımdaki en değerli insana, eşim Mine Ateş Korkmaz'a ve doktora eğitimim boyunca evimi oradan oraya taşıyan hürmetli ailesine teşekkürü bir borç bilirim. Son olarak ise beni doğuran, büyüten, bugünlere getiren sevgili annem Zekiye Korkmaz'ı ve beni okutan, cesareti ve girişimciliği kendisinden öğrendiğim, rol modelim, babam Nurettin Korkmaz'ı rahmetle anıyor ve faydalı bir birey olma konusunda çaba göstermek suretiyle kendilerine teşekkür etme imkânı yakalamaya çalışıyorum. Sonsöz olarak ise, bu tez çalışmasının konuyla ilgilenen arkadaşlara, bir nebzecek de olsa faydalı olmasını yürekten diliyorum.

İSTANBUL, 2020

GÖKHAN KORKMAZ

İÇİNDEKİLER

ÖZ.....	ii
ABSTRACT	iii
ÖNSÖZ.....	iv
TABLolar LİSTESİ.....	xiii
ŞEKİLLER LİSTESİ.....	xiv
KISALTMALAR LİSTESİ.....	xvii
GİRİŞ	1

BİRİNCİ BÖLÜM

KREDİ RİSKİ

1.1. Risk Ve Kredi Riski Kavramları	8
1.2. Kredi Riski Yönetimi	13
1.3. Kredi Riskinin Belirlenmesi Ve Basel Komitesi Kararları	14
1.3.1. Kredi Riskinin Belirlenmesi.....	14
1.3.2. Basel Komitesi Ve Kredi Risk Yönetimi Konusundaki Tavsiye Kararları	15
1.3.2.1. Basel I Kararları.....	16
1.3.2.2. Basel II Kararları	16
1.3.2.3. Basel III Kararları	18
1.3.3. Basel Ve Kredi Riski Özeti	20
1.3.4. Kredi Riskinin Belirlenmesinde Kullanılan Temel Yaklaşımlar	21
1.3.4.1. Standart Yaklaşım.....	21
1.3.4.2. İçsel Ölçüm Yaklaşımları	23
1.3.5. Kredi İlke Ve Politikaları	25
1.3.6. Kredilerin İzlenmesi Ve Kayıp Kredilerin Önlenmesi.....	26

**1.4. Kredi Riskinin Belirlenmesi Konusunda İstatistik Ve Yapay Zekâ
Özelinde Kullanılan Yöntemlere Dair Bir Literatür Taraması..... 27**

1.4.1.	Diskriminant Analizi.....	28
1.4.2.	Lojistik Regresyon	29
1.4.3.	MARS Modelleri.....	31
1.4.4.	Yapay Sinir Ağları	33
1.4.5.	Destek Vektör Makineleri	40
1.4.6.	Genetik Algoritma.....	42
1.4.7.	Literatür Özeti	43

İKİNCİ BÖLÜM

YAPAY ZEKÂ YÖNTEMLERİ

2.1. Yapay Zekâ	47
2.1.1. Zekâ ve Yapay Zekâ Kavramları	47
2.1.2. Yapay Zekânın Kullanım Alanları	51
2.1.3. Makine Öğrenmesi	52
2.1.4. Makine Öğrenmesi Ve Veri Madenciliği	55
2.1.5. Makine Öğreniminin Teorisi	57
2.1.6. Öğrenme Problemleri	59
2.1.6.1. Regresyon	60
2.1.6.2. Sınıflandırma	62
2.1.6.3. Kümeleme	63
2.1.6.4. Yoğunluk Tahmini	66
2.1.7. Öğrenme Yöntemleri	68
2.1.7.1. Denetimli Öğrenme	68
2.1.7.2. Denetimsiz Öğrenme	69

2.1.7.3.	Yarı Denetimli Öğrenme	71
2.1.7.4.	Güçlendirici Öğrenme.....	71
2.1.7.5.	Aktif Öğrenme	72
2.1.8.	Model Karmaşıklığı ve Kontrolü	72
2.1.8.1.	Occam'ın Usturası	79
2.1.8.2.	Popper'ın Yanlışlanabilirliği.....	79
2.1.8.3.	VC Katsayısı –Teorisi.....	82
2.1.8.4.	Model Karmaşıklığının Optimizasyonu Konusunda Farklı Görüşler	83
2.2.	Yapay Sinir Ağları.....	84
2.2.1.	Yapay Sinir Ağı Yaklaşımı	87
2.2.2.	Yapay Sinir Ağlarının Gelişim Süreci	88
2.2.3.	Sinir Hücrelerinin Yapısı Ve İşleyişi	91
2.2.4.	Yapay Sinir Hücreleri	94
2.2.4.1.	Perseptron	95
2.2.4.2.	ADALINE.....	97
2.2.4.3.	MADALINE	98
2.2.5.	Sayısal Bilgisayarlar Ve Yapay Sinir Ağları	98
2.2.6.	Çok Katmanlı Ağlar	99
2.2.7.	İleri Beslemeli Ve Geri Beslemeli YSA	103
2.2.7.1.	İleri Beslemeli Yapay Sinir Ağları	103
2.2.7.2.	Geri Beslemeli Yapay Sinir Ağları.....	104
2.2.8.	Geriye Yayılmalı Yapay Sinir Ağları	105
2.2.9.	Yapay Sinir Ağlarının İşlevleri	108
2.2.9.1.	Girdi İşlevi	108

2.2.9.2.	Ağırlıkların Oluşturulması İşlevi	109
2.2.9.3.	Toplama İşlevi	109
2.2.9.4.	Etkinlik İşlevi.....	111
2.2.10.	XOR Problemi.....	117
2.2.11.	Yapay Sinir Ağlarının Eğitilmesi.....	120
2.2.11.1.	Hebb Öğrenme Kuralı.....	121
2.2.11.2.	Hopfield Öğrenme Kuralı	123
2.2.11.3.	Delta Öğrenme Kuralı.....	124
2.2.11.4.	Kohonen Öğrenme Kuralı.....	129
2.2.11.5.	Grossberg Öğrenme Kuralı.....	130
2.2.12.	Yapay Sinir Ağları İle İlgili Diğer Kavramlar	131
2.2.12.1.	Yığınlar	131
2.2.12.2.	Yineleme Sayısı	131
2.2.12.3.	Bırakma.....	132
2.2.12.4.	Normalizasyon	132
2.2.12.5.	Model Geçerleme.....	133
2.2.13.	Yapay Sinir Ağlarının Kullanım Alanları	133
2.2.14.	Yapay Sinir Ağlarının Avantajları	137
2.2.15.	Yapay Sinir Ağlarının Dezavantajları	140
2.3.	Destek Vektör Makineleri.....	143
2.3.1.	İstatiksel Öğrenme Teorisi	145
2.3.1.1.	Ampirik Risk Minimizasyonu	147
2.3.1.2.	VC Boyutu	151
2.3.1.3.	Yapısal Risk Minimizasyonu.....	156

2.3.2.	Çekirdek Fonksiyonları	158
2.3.3.	Optimal Ayırıcı Hiperdüzlem	160
2.3.3.1.	Doğrusal Olarak Ayrılabilir Veri Kümeleri için Doğrusal Maksimal Marj Sınıflandırıcı	162
2.3.3.2.	Doğrusal Olarak Ayrılamayan Veri Kümeleri için Doğrusal Maksimal Marj Sınıflandırıcı	169
2.3.4.	Destek Vektör Makinelerinin Avantajları	171
2.3.5.	Destek Vektör Makinelerinin Dezavantajları	173

ÜÇÜNCÜ BÖLÜM

KREDİ RİSKİNİN BELİRLENMESİNDE YAPAY ZEKÂ YAKLAŞIMLARI VE BİR UYGULAMA

3.1.	Uygulamanın Amacı	174
3.2.	Uygulamada Kullanılan Yöntemler	175
3.2.1.	Yapay Zekâ Yöntemleri	177
3.2.1.1.	Yapay Sinir Ağları	177
3.2.1.2.	Destek Vektör Makineleri	177
3.2.2.	İstatistikî Yöntemler	178
3.2.2.1.	Lojistik Regresyon	178
3.3.	Uygulamada Kullanılan Yazılımlar	179
3.4.	Uygulamada Kullanılan Veri Kümeleri	179
3.4.1.	Bireysel Veri Kümesi	179
3.4.2.	Kurumsal Veri Kümesi	180
3.4.3.	Dengeli Kurumsal Veri Kümesi	180
3.5.	Uygulamada Kullanılan Değişkenler	180
3.5.1.	Bireysel Veri Kümesi Değişkenleri	180

3.5.1.1.	Bireysel Veri Kümesi Bağımlı Değişkeni	182
3.5.1.2.	Bireysel Veri Kümesi Açıklayıcı değişkenleri	183
3.5.2.	Kurumsal Veri Kümesi Değişkenleri	183
3.5.2.1.	Kurumsal Veri Kümesi Bağımlı Değişkeni	184
3.5.2.2.	Kurumsal Veri Kümesi Açıklayıcı değişkenleri	185
3.6.	Normalizasyon	185
3.7.	Uygulama Sonuçları	186
3.7.1.	Bireysel Müşteriler İçin Uygulama Sonuçları.....	186
3.7.2.	Kurumsal Müşteriler İçin Uygulama Sonuçları	187
SONUÇLAR		190
KAYNAKÇA		193
ÖZGEÇMİŞ.....		220

TABLÖLAR LİSTESİ

Tablo 2.1: Biyolojik Sinir Ağları İle Yapay Sinir Ağlarının Karşılaştırılması	95
Tablo 2.2: Sayısal Bilgisayarlar İle Yapay Sinir Ağlarının Karşılaştırılması	99
Tablo 2.3: OR, AND ve XOR Problemlerinde x ve y Değerleri	117
Tablo 2.4: Bazı İstatistiki Testlere Karşılık Gelen YSA Modelleri	135
Tablo 2.5: Yapay Sinir Ağı Modelleri Seçim Tablosu	136
Tablo 2.6: Tasarım Parametrelerinin Uç Değerlerinin Eğitim Yakınsaması Ve Ağ Genellemesi Üzerindeki Etkisi.....	142
Tablo 2.7: Çekirdek Fonksiyonları	160
Tablo 3.1: Sınıflandırma Problemlerinde Kullanılan Yöntemler Ve Özellikleri	176
Tablo 3.2: Bireysel Veri Sonuç Sınıfları Ve Sınıf Etiketleri.....	182
Tablo 3.3: Kurumsal Veri Sonuç Sınıfları Ve Sınıf Etiketleri	184
Tablo 3.4: Dengeli Kurumsal Veri Sonuç Sınıfları Ve Sınıf Etiketleri	185
Tablo 3.5: Bireysel Müşteriler Veri Kümesi Uygulama Sonuçları.....	186
Tablo 3.6: Kurumsal Veri Kümesi Uygulama Sonuçları	187
Tablo 3.7: Tüm Uygulamaların Sonuçları	188
Tablo 3.8: Dengeli Kurumsal Veri Kümesi YSA Uygulamasında, Farklı Sayıda Nöron Ve Döğüm Sayısı Kullanıldığında, Ulaşılan Doğruluk Oranları Ve Standart Sapmaları (Maksimum İterasyon: 10000, cv=10, Veri: Dengeli Kurumsal Veri)	189

ŞEKİLLER LİSTESİ

Şekil 2.1: İki Aşamalı Denetimli Makine Öğrenmesi Yöntemi	54
Şekil 2.2: Makine Öğrenmesi Süreci	59
Şekil 2.3: Polinom Dereceleri Sırasıya Bir, İki Ve Dokuz Olan Regresyon Tahminleyicilerinin Performansları	75
Şekil 2.4: Sırasıyla İki ve Altı Gizli Katmana Sahip YSA Modelleri ile Yapılmış Sınıflandırma Sonuçları	76
Şekil 2.5: Kapasite Ve Hata Arasındaki Tipik İlişki	78
Şekil 2.6: Sinir Hücrelerinin Görünümü.....	92
Şekil 2.7: Biyolojik Sinir Hücresi (Nöron).....	93
Şekil 2.8: Perseptronun Yapısı	96
Şekil 2.9: ADALINE (Adaptive Linear Element)	97
Şekil 2.10: Von Neumann Mimarisinde Bilgisayarın Yapısı	98
Şekil 2.11: Yapay Sinir Ağı (Çok Katmanlı) Modeli	100
Şekil 2.12: ÇKA'ların Gösterim Yeteneği.....	102
Şekil 2.13: İleri Beslemeli Yapay Bir Sinir Ağı Modeli	104
Şekil 2.14: Geri Beslemeli Yapay Sinir Ağları Mimarisi.....	105
Şekil 2.15: Genel Bir Geriye Yayılmalı Ağ Mimarisi.....	106
Şekil 2.16: Bir Ağda Geri Yayılımlı Öğrenme Algoritmasının Uygulanması	107
Şekil 2.17: Etkinlik İşlevi İçin Sık Kullanılan Basamak ve Parçalı Doğrusal Fonksiyonlar.....	112
Şekil 2.18: ReLU Etkinlik Fonksiyonu	114
Şekil 2.19: Lojistik Dönüştürme Fonksiyonu.....	115
Şekil 2.20: Gauss Dönüştürme Fonksiyonları	115
Şekil 2.21: Hiperbolik Tanjant Dönüştürme Fonksiyonları	116

Şekil 2.22: Hiperbolik Sekant Dönüştürme Fonksiyonu	116
Şekil 2.23: AND ve OR Probleminde Sınıflandırma	118
Şekil 2.24: XOR Probleminde Sınıflandırma	118
Şekil 2.25: AND ve OR Problemlerinin Tek Katmanlı YSA İle Çözümü	119
Şekil 2.26: XOR Probleminin Çok Katmanlı YSA İle Çözümü	119
Şekil 2.27: Yapay Sinir Ağlarında Denetimli Ve Denetimsiz Öğrenme Mimarileri	120
Şekil 2.28: Hopfield Ağının Yapısı	124
Şekil 2.29: Delta Öğrenme Kuralı ve Gradyan Azaltma	127
Şekil 2.30: YSA’nda Yerel Ve Evrensel Minimum Noktaları	128
Şekil 2.31: Deneysel Risk ve Beklenen Risk İlişkisi.....	129
Şekil 2.32: Yapay Sinir Ağlarında Bırakma Uygulaması Öncesi Ve Sonrası.....	132
Şekil 2.33: Ampirik Riskin, Eğitim Verisi Arttıkça, Beklenen Riske Asimptotik Olarak Yakınsaması	150
Şekil 2.34: İki Boyutlu Uzayda, Üç Ve Dört Gözlemlili Örneklemeler İçin VC Boyutu	152
Şekil 2.35: Büyüme Fonksiyonunun Yapısı	154
Şekil 2.36: VC Boyutu Kullanılarak Beklenen Riskin Minimize Edilmesi	155
Şekil 2.37: Kapasitelerine Göre Belirlenmiş Fonksiyon Kümeleri	158
Şekil 2.38: Girdilerin Doğrusal Olarak Ayrılamadığı Giriş Uzayından, Çekirdek Fonksiyonu Yardımı İle, Doğrusal Olarak Ayrılabildiği Yüksek Boyutlu Nitelik Uzayına Dönüştürülmesi.....	159
Şekil 2.39: Optimal Ayırıcı Hiperdüzlem.....	162
Şekil 2.40: Mümkün İki Farklı Sınıflandırıcı İle Doğrusal Olarak Ayrılabilen İki Sınıflı Bir Örnek.....	163

Şekil 2.41: Birbirinden Farklı Hiperdüzlemlerle Ayrılmış İki Boyutlu Girdi Uzayı	165
Şekil 2.42: Mümkün İki Farklı Sınıflandırıcı İle Doğrusal Olarak Ayrılabilen İki Sınıflı Bir Örnek.....	166
Şekil 2.43: Doğrusal Olarak Ayrılabilir Veriler İçin Doğrusal Maksimal Marj Sınıflandırıcının Matematiksel Gösterimi.....	168
Şekil 2.44: Doğrusal Olarak Ayrılamayan İki Sınıflı Bir Örnek.....	169



KISALTMALAR LİSTESİ

ADALINE	: Adaptive Linear Element Uyarlanabilir Doğrusal Eleman (Nöron)
AIG	: American International Group
ARM	: Ampirik Risk Minimizasyonu Empirical Risk Minimization (ERM)
ART	: Adaptif Rezonans Teorisi Adaptive Resonance Theory (ART)
BDDK	: Bankacılık Düzenleme Ve Denetleme Kurulu
BIS	: The Bank Of International Settlement Uluslararası Takas Bankası
Bkz.	: Bakınız
BP	: Back Propagation Geriye Yayılmalı
BPGD	: Back Propagation Gradient Descent Geriye Yayılmalı Gradyan İnişi
CART	: Classification And Regression Trees Sınıflandırma Ve Regresyon Ağaçları
CBR	: Case Based Reasoning Vaka Bazlı Muhakeme
CDF	: Cumulative Density Function Kümülatif Yoğunluk Fonksiyonu
DR	: Doğrusal Regresyon Linear Regression
DVM	: Destek Vektör Makineleri Support Vector Machine (SVM)

DVR	: Destek Vektör Regresyonu Support Vector Regression (SVR)
DVS	: Destek Vektör Sınıflandırma Support Vector Classification (SVC)
ECA	: Export Credit Agencies İhracat Kredi Kuruluşları
EKK	: En Küçük Kareler Least Squares (LS)
EKK-DVM	: En Küçük Kareler Destek Vektör Makineleri Least Squares Support Vector Machine (LS-SVM)
FSA	: Financial Services Agency İngiltere Finansal Hizmetler Otoritesi
FuzC	: Fuzzy Classifier Bulanık Sınıflandırıcı
GA	: Genetik Algoritma
GCV	: Generalized Cross-Validation Genelleştirilmiş Çapraz Doğrulama
GMLC	: Gaussian Maximum Likelihood Classifier Gaussian Maksimum Olabilirlik Sınıflandırıcısı
GRNN	: General Regresssion Neural Network Genel Regresyon Sinir Ağları
GSYİH	: Gayri Safi Yurt İçi Hâsıla
ID3	: Inductive Dichotomizer Endüktif Dikotomizer
LD	: Lojistik Diskriminant
LDA	: Lineer Diskriminant Analizi

LR	: Lojistik Regresyon
LVQ	: Linear Vector Quantization Doğrusal Vektör Niceleme (LVN)
MADALINE	: Multilayer Adaptive Linear Element Çok Katmanlı Uyarlanabilir Doğrusal Eleman (Nöron)
MARS	: Multivariate Adaptive Regression Splines Çok Değişkenli Uyumlu Regresyon Sınırları / Uzanımları
MO	: Maksimum Olabilirlik Maximum Likelihood (ML)
NEWQ	: Generalization Of Quinlan's Induction Method Quinlan'ın İndüksiyonunun Genelleştirilmesi
PNN	: Probabilistic Neural Network Probabilistik (Olasılıksal) Sinir Ağı
RCSM	: Reassigning Credit Scoring Model Yeniden Atama Kredi Puanlama Modeli
SS	: Standart Sapma
SYR	: Sermaye Yeterliliği Rasyosu
vb.	: Ve Benzeri
VC	: Vapnik-Chervonenkis (Theory), İstatiksel Öğrenme Teorisi Statistical Learning Theory (SLT)
YRM	: Yapısal Risk Minimizasyonu Structural Risk Minimization (SRM)
YSA	: Yapay Sinir Ağları Artificial Neural Network (ANN)

GİRİŞ

Geleceğe hazırlanmak isteyen işletmelerin, planları olan yöneticilerin, bir vizyonu bulunan iş adamlarının, yatırımlarını doğru yönlendirmek isteyen yatırımcıların, doğru rehberlik yapmak isteyen ekonomi danışmanlarının, başarı haritası çizmek isteyen stratejistlerin ve çevresel olayları anlamaya/anlamlandırma çalışan bilim insanlarının sürekli odaklanmış oldukları; “belirsizliği belirli hale getirme” şeklinde ifade edebileceğimiz sorunsal, tarih boyunca çok farklı şekillerde çözülmeye çalışılmıştır.

Belirsizlik sorununu aşabilmek, yani işletmeleri doğrudan veya dolaylı olarak etkileyebilecek içsel veya dışsal olayların işletmeyi nasıl ve ne ölçüde etkileyebileceğini öngörebilmek ve sonrasında oluşacak bu yeni duruma uyum sağlayabilmek ise, günümüzün rekabet koşulları açısından oldukça önemlidir. İşletmeler, parametrelerin sürekli değiştiği piyasa şartlarında, sürekli belirsizlikle yüzleşmekte ve bu belirsizliği doğru yönetebilmek için çok çeşitli yöntemlere başvurumaktadırlar. Kimi zaman elde edilen salt bilgilerle karar süreçleri işletilmeye çalışırken, kimi zaman ise uzun süren yönetim kurulu toplantılarından çıkan ortak kararlara bağlanan güven bu sürecin belirleyici iki unsurunu teşkil etmektedir. Uzman görüşü, profesyonel yardım, tecrübe vb. yöntemlerden sonra, veri tabanı yönetim sistemlerinde büyük veri ambarlarının oluşması ve bu bilgileri anlamlı çıktılara dönüştüren veri madenciliği teknikleri ve bu tekniklerin robotize edilmesini sağlayan makine öğrenmesi yöntemleri de işletmelerin karar süreçlerinde kullanılmaya başlanmıştır.

Modern yüzyılda, İstatistik Biliminin belirsizlik sorunsalı için geliştirdiği nicel tekniklerle, çok farklı bir ufuk yakalandığı aşîkârdır. Son birkaç on yıldır ise, ekonomi çevreleri, istatistik sınavından başarıyla geçen ve nicelliğinin yanı sıra matematiğin katı sınırlarını da eriterek insani bir esnekliği ikame potansiyelini içinde bulunduran yapay zekâyı da kullanarak, kendilerini geleceğe hazırlamanın yollarını araştırmaktadırlar. Bununla birlikte yine aynı çevreler, yapay zekânın işletme süreçlerinde daha çok yetkilendirilmesinin çok boyutlu yansımalarını test etme sürecinden geçmektedirler.

Yapay zekânın ise özellikle işletme çevrelerinde, bilhassa finans sektöründe, kullanılması çok yeni olmamakla birlikte yapay zekâ yöntemlerinin zamanla çok farklı modelleri kullanılarak, karmaşık problemlere doğru cevap verebilme yeteneği, daha önce istatistiki yöntemlerde de olduğu gibi, evrilerek artmaktadır. Yani, standart uygulamalara kıyasla, geliştirilmiş daha kompleks algoritmalar, farklı mimari denemeleri ve muhtelif uygulamalar bu yeni yöntemleri oldukça farklı kılmaktadır.

Yeni yapay zekâ yöntemleri geliştirildikçe, bu yeni yöntemlerin daha önce farklı yöntemlere konu olan bu kadim problemlere verdiği cevapların performansı da ilgili çevrelerce ayrıca merak konusu haline gelmektedir. Her yapay zekâ yönteminin, diğer klasik istatistik ve yapay zekâ yöntemlerine kıyasla, başarılı sonuçları olabildiği gibi daha az verim sağladıkları da sıkça görülebilmektedir. Bunun sebebi, performans düşüklüğü gösteren bu yöntemlerin, aşırı öğrenme veya eksik öğrenme, genelleme (daha önce gözlemlenmemiş girdilerde, yani eğitim verisinin dışındaki veri kümeleri üzerinde, iyi performans gösterme yeteneği olarak tanımlanır¹) hataları ve yerel minimuma yakalanma vb. sorunlarla baş edebilecek teorik yetkinliğe sahip olmayışı olabilmektedir. Sonuç olarak bazı modeller çeşitli sorunları algoritmik altyapısıyla aşabilirlerken diğer bazı yöntemler bu yetkinlikten ve güçlü bir algoritmik yapıdan yoksun olabilmektedirler. Diğer taraftan, uygulama kolaylığı açısından meseleyi ele alacak olursak, bazı yöntemlerin dağılıma bağımlı olmaları ve bunun yanında bir takım katı varsayımları da sağlama gereksinimleri analiz aşamasında uygulayıcıları zorlamaktadır. Diğer bazı yöntemler ise, dağılımdan bağımsız olmaları ve herhangi bir varsayım gerektirmemeleri neticesinde, uygulama aşamasında çok büyük avantaj sağlamakta ve bu da modelin kullanımının daha yaygın olmasını sağlamaktadır.

Yapılan aştırmalarda, kullanılan veri setlerinin hacmindeki çok büyük artışla birlikte ve kullanılan yöntemlerin mimarisinde yapılacak bazı düzenlemelerle, eskiye oranla daha iyi sonuçlar veren modeller oluşturulmaya çalışılmaktadır. Günümüzde bilgisayarların ve sensörlerin, oldukça uygun maliyetli hale gelmesi neticesinde, iş süreçlerinde kullanımı hızla yaygınlaşmakta bu da kişilerin ve kurumların

¹ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **Deep Learning**, 1. Baskı, Cambridge: The MIT Press, 2017, s. 110.

yorumlanabilir büyük veri bankalarına sahip olmalarını sağlamaktadır. Dolayısıyla, bilinen bir problemin çözümüne dönük olarak, beş yıl önce yapılan bir çalışmanın sonuçlarıyla, bugün yapılacak bir çalışmanın sonuçları, hem veri hacmindeki iyileşmenin sağlayacağı fayda bakımından hemde yöntemlerin algoritmik içeriklerinin güçlenmesinden dolayı, oldukça farklılık gösterecektir.

Bir modelin performansının iyi olup olmadığına, bu modelin sonuçlarının bir başka modelin sonuçlarıyla mukayese edilmesi yoluyla karar verilmektedir. Burada modellere dönük olarak, geleneksel yani istatistiki olanlarla, yapay zekâ temelli olanların mukayesesi de ayrıca merak konusudur. Diğer bir konu da bu iki yaklaşımın yani istatistik ve yapay zekâ yöntemlerinin artık girift bir hale gelmesidir. Çünkü günümüzde, bir yöntemin diğeriyle optimize edildiği, bir yöntemle diğer yöntemin parametresine karar verildiği ya da melez diyebileceğimiz hem istatiksel hem yapay zekâ yöntemlerinin kullanıldığı birçok hibrit model bulunmaktadır. Bu hibrit yöntemlerin sonuçları da güncel yöntemlerle karşılaştırılmaktadır. Bu da zaman içinde model yaklaşımlarını bir doğal seleksiyona tabi tutmakta ve sonuçta başarılı olan modeller uygulanmaya ve üzerinde çalışılmaya devam etmektedir.

Yapay zekâ yöntemlerinden biri olan Yapay Sinir Ağları (YSA), insandaki sinir hücrelerinin yapısından esinlenilerek geliştirilmiş, sonrasında ise muazzam sayıdaki türevleriyle, çok çeşitli problemlere etkin çözümler üreten ve hâlihazırda, birçok sahada, etkin bir şekilde kullanılan karar verme aracı haline gelmiştir.

YSA'nın birçok avantajının yanı sıra bilinen en önemli dezavantajı ise, karar verirken içinde kullandığı fonksiyonların parametlerinin bilinmemesidir. YSA'nın bu özelliği "kara kutular" olarak adlandırılmaktadır. Bu bilinmezlik, etkili sonuçların yanı sıra, tüm süreci algoritmaya terk etme konusunda, kullanıcıların kafasında sürekli soru işaretleri bırakmakta ve işletmelerin şeffaflık, hesap verebilirlik gibi kurumsal bazı ilkeleriyle çelişmektedir. O yüzden YSA sonuçlarının sürekli olarak bir başka yöntemin sonuçlarıyla kıyaslanması ve böylelikle oluşturulan YSA mimarisinin uygulama sonuçlarının kalitesinin farkında olunması önem taşımaktadır.

Bir diğ er yapay zek a y ntemi olan Destek Vekt r Makineleri (DVM) ise, g zlem uzayında yer alan sınıflandırma birimlerinin arasına, bu birimleri iki ayrı sınıfa atayacak bir aşırı d zlem (hiperd zlem)  izilirken bu hattın kendine en yakın komşu g zlem birimleriyle olan mesafesini maksimize etme d ş ncesine dayanan bir y ntemdir. Denetimli ya da yarı denetimli  alıřabilen, sınıflandırma ve regresyon yapabilen ve YSA'na kıyasla daha yeni ve ilgili  evrelerce kabul g ren,  nemli bir yapay zek a yaklařımıdır.

DVM, son zamanlarda finansal bařarsızlıktan kredi riskini belirlemeye, hisse senetlerinin y n n  tahmin etmekten gayrimenkul deęerlemeye ve y z tanımadan parmak izi ayırmaya kadar, bir ok sahada arařtırma ve uygulama alanı bulmakta,  zellikte  rneklem sayısının az fakat boyut sayısının fazla olduęu veri k melerinde, YSA'na alternatif bir y ntem olarak  n plana  ıkmaktadır.

Yapay zek nın belki de en  nemli uygulama alanlarından birisi olan finans sekt r , “ekonomilerin mikroorganizmaları” diyebileceğimiz iřletmeler a ısından, ulusal ve uluslararası ticarete, olduk a b y k  nem arzetmektedir. řirketler, faaliyetlerinin devamlılıęı, rekabet avantajı saęlama, yatırımlarını  eřitlendirme ve b y me gibi stratejik hedeflerini ger ekleřtirebilmek i in s rekli olarak finansman ihtiya ı duymaktadırlar. Bug n n ekonomilerinde řirketler, finansman ihtiya larını kredilerden, hisse senedi ve tahvil gibi sermaye piyasası ara larına, factoringden, forfaiting ve finansal kiralamaya kadar  ok  eřitli finans y ntemleri kullanarak karřılayabilmektedirler. Bu y ntemler ise genel hatları ile ele alındıęında finansal sistemler, banka temelli ve piyasa temelli finansal sistemler olmak  zere iki ana gruba ayrılabilir². Bu y ntemler arasında ise en pratięi, en  ok kullanılanı ve en hızlısı yine kredi kullanma se eneęidir.   nk  kredi kullanımını k  k, orta ve b y k  aplı řirketler, hatta řahıřlar i in bile ge erli olabilen bir y ntemken, factoring ve benzeri y ntemlere ise her  apta iřletmenin ya da bireylerin bařvurabilmesi m mk n ya da pratik g r nmemektedir.

² Mustafa Tevfik Kartal, **Bankaların Finans Sekt r ndeki  nemi**, Finansal İktisat, Orion Yayıncılık, 2018, s.5-27.

Bankalar, aracı olmak, teminat oluşturmak, yatırımları finanse etmek, vb. birçok şekillerde ekonomilere can veren, çok önemli aktörlerdir. Dolayısıyla ekonomilerin güçlü ve güvenilir olmasının arkasında sağlam bir bankacılık sisteminin olduğunu söylemek abartılı olmayacaktır. Çünkü günümüz piyasalarında yaşanan birçok ekonomik krizin arka planında bankacılık sisteminin zaafa uğramış olduğu bilgisi yer almaktadır. Meseleye tersinden bakıldığında ise ekonomik krizlerden çıkışın yine bankacılık sisteminin güçlendirilmesi ile mümkün olabildiği bilinmektedir. Çeşitli finansal krizler incelendiğinde, ortak paydanın bankacılık sistemindeki yolsuzluklar ya da etkin olmayan kredi risk yönetiminin olduğu söylenebilir. Özetle bankaların kredi kullandırma işlevi, ne kadar verimli ve kapsamlı olursa, ekonomilerde gerçekleştirilen girişimlerin de o denli verimli ve geniş ufuklu olması beklenebilir.

Bankaların temel faaliyetlerine bakıldığında; finansal aracılık, dış ticareti fonlama ve ihracatı teşvik, uluslararası mali piyasalarda risk yönetimi, kaynakların etkin kullanımı, gelir ve servet dağılımını etkileme, kaydi para yatırma, para ve maliye politikalarının yürütülmesine yardımcı olma, ekonomiyi yönlendirme, enflasyonu kontrol altına alma ve tam istihdam sağlanması gibi birçok fonksiyonlarının olduğunu görmekteyiz³. Bu çalışma da bankaların temel faaliyet konularından finansal aracılık hizmeti, yani genel itibarıyla kredi kullandırma fonksiyonu üzerinde durulacaktır.

Kredi kullandırma fonksiyonu bankaların, piyasalarla olan en geniş tabanlı etkileşim alanını oluşturmaktadır. Dolayısıyla bankaların karlılıkları, bu etkileşimin sağlıklı yürütülmesiyle mümkün olabilmektedir. Bir diğer ifadeyle, bireysel ve kurumsal müşterilere kullandırılan kredilerin geri dönüşlerinin tam olarak ve zamanında sağlanabilmesiyle mümkün olmaktadır. Bu da, kredilerin geri ödenmemesi ya da geç ödenmesi gibi ihtimallerden doğan “kredi riski” olgusunun makul seviyelerde tutulabilmesi ile yakından alakalıdır. Kredi riskine dönük bu optimize yaklaşım, doğru kredi tahsisleri ve etkin kredi yönetimi süreçleri yardımıyla gerçekleştirilmekte ve bu

³ Mikail Altan, “**Fonksiyonlar ve İşlemler Açısından Bankacılık**, Beta Basım Yayın, 2001, s.66.

temel faaliyet alanında; kar maksimizasyonu, geri dönmeyen kredilerin en aza indirgenmesi suretiyle sağlanabilmektedir⁴.

Piyasalara işlerlik kazandıracak olan bankalara ait kredi verme fonksiyonunda doğru kredi tahsislerinin nasıl yapılacağı, önemli bir problem olup, bu konuda bir hayli akademik çalışma ve pratikte kullanıla gelen bankacılık uygulamaları vardır. Fakat yeni yöntemler geliştirildikçe, geliştirilen bu yöntemlerin, bu kadim sorulara vereceği cevaplar yeniden araştırma ve merak konusu haline gelmektedir.

Bu tezde, bankaların kredi riskini belirlemeye dönük olarak, yapay zekâ yöntemlerinden YSA ve DVM kullanılmış ve sonuçları klasik yöntemlerden Lojistik Regresyon (LR) yönteminin sonuçları ile karşılaştırılmıştır. Analizler bir bankadan alınan bireysel ve kurumsal kredi verileri üzerinde gerçekleştirilmiştir.

Kullanımı, dünya genelinde, birçok disiplinde, teoride ve pratikte giderek artan DVM uygulamasının, kredi riskini belirleme konusunda, YSA'na göre performansı, dengeli ve dengesiz dağılmış veri üzerinde yapılacak çeşitli denemelerle, ayrıntılı bir şekilde test edilmiştir. Araştırmanın, kredi riskini belirleme konusunda, ilgili yöntemlerin özgün ve büyük gerçek hayat verisi üzerinde uygulanması sayesinde, literatüre katkı sağlaması beklenmektedir.

“Kredi riskinin belirlenmesinde yapay zekâ yaklaşımları ve bir uygulama” konulu tezin birinci bölümünde, kredi riski kapsamında risk ve kredi riski kavramı, kredi riskinin yönetimi ve bu yönetimin yöntemleri olan içsel ve dışsal yaklaşımlar ve Basel süreçlerinin kredi riski konusuyla olan etkileşimi ele alınmıştır. Sonrasında ise kredi riskinin ölçülmesi konusunda kullanılan istatistik ve yapay zekâ yöntemlerinin bir literatür taraması ve son olarak da bu literatür taramasının özeti yer almaktadır. Tezin ikinci bölümünde ise tezde kullanılan yapay zekâ yöntemlerinden YSA ve DVM, kavram ve yöntem bakımından, ayrıntılı olarak incelenmiştir. Ayrıca bu yöntemlerin, esin kaynakları, gelişim süreçleri, literatür taramaları, son olarak da avantaj ve

⁴ Erol Muzır, “Basel II Düzenlemeleri Doğrultusunda Kredi Riski Analizi Ve Ölçümü: Geleneksel Ekonometrik Modellerin Yapay Sinir Ağı Ve Mars Modelleriyle Karşılaştırılmasına Yönelik Ampirik Bir Çalışma”, Doktora Tezi, İstanbul Üniversitesi, 2011, s.10.

dezavantajları başlıklar halinde tartışılmıştır. Tezin üçüncü bölümü ise uygulama kısmı olup, problem, kullanılan yöntemler (YSA, DVM ve LR), kullanılan yazılımlar ve veri kümeleri ve uygulama sonuçları bu bölümde ayrıntılı olarak raporlanmıştır.



BİRİNCİ BÖLÜM

KREDİ RİSKİ

1.1. Risk Ve Kredi Riski Kavramları

Risk kelimesi, Fransızca “riscue” kökünden gelmekte olup, “Zarara uğrama tehlikesi, riziko” olarak tanımlanmaktadır⁵.

Bunun dışında Aven ve Renn, risk konulu çalışmalarında, önemli buldukları risk tanımlarını şu şekilde özetlemektedirler⁶:

- Risk, olumsuz bir sonucun oluşmasıdır.
- Risk, beklenen zararlar eşittir.
- Risk, bir olayın olasılığının ve sonuçlarının birleşimidir.
- Risk, olumsuz etkilerin olasılığının ve ciddiyetinin bir ölçüsüdür.
- Risk, sonucun, eylemlerin ve olayların belirsizliğini ifade eder.
- Risk, insani değeri olan bir şeyin (insanların kendileri dâhil) tehlikede olduğu ve sonucun belirsiz olduğu bir durum veya olaydır.

Yukarıdaki tanımlara ilave olarak risk kelimesi, belirsizliğin ve zarar ihtimalinin olasılığını tanımlarken aynı zamanda, oluşabilecek yeni durumun, mevcuttaki durumdan daha iyi olma olasılığını, yani tehditlerin yanı sıra fırsatları da içeren, insani bir sürece de işaret etmektedir. Çünkü kişi ve kurumların, çok istekli olmasalar da, almış oldukları risklerin makul bir sebebi var. O da, risk alınmama durumunun olumlu veya olumsuz getirileri ile risk almanın getireceği olumlu ya da olumsuz getiriler arasında bir tercih yapmaktır. Risk almaktan umulan ise mevcut durumdan ya da risk alınmama durumundaki beklentiden daha iyi bir duruma erişme düşüncesidir. Risk almak, kazanımları mı yoksa yıkımları mı getireceği sorusunun cevabı, birazda mevcut

⁵ Türk Dil Kurumu, **Genel Türkçe Sözlük**, Çevrimiçi, <https://sozluk.gov.tr/?kelime=risk>, 28/10/2019.

⁶ Aven, T. Ve Renn, O. “On risk defined as an event where the outcome is uncertain”, **Journal of Risk Research**, Cilt: 12, Sayı: 1, 2009, s.2.

riskini iyi yönetilebilmesi ile alakalıdır. Sonuç olarak risk; tam olarak ölçülemeyen, zararlı sonuçları olabilen ve de yönetilebilen bir iş fonksiyonudur⁷.

Kredi riskini tanımlamadan önce, ilk olarak banka ve sonrasında kredi kavramını tanımlamak gerekmektedir.

Banka kelimesi, köken itibarıyla İtalyanca, “banco” kökünden gelmekte olup, sıra ya da tezgâh anlamına gelmektedir. Kavramsal olarak bankalar, diğer adıyla mevduat kurumları, fon sahibi çevrelerden toplamış oldukları fonları (öz sermayesi de dâhil olmak üzere) fon talebinde bulunan çevrelere kullandıran, yatırım fırsatları yaratan ve kullandıkları finansal sistemlerle içinde bulundukları ekonomilerin etkinliğini artırmak için önemli roller üstlenen kurumlar olarak kabul edilmektedir. Ticari bankaların temel misyonu, ticaretin ve üretimin fonlanmasıdır⁸.

Türkiye’de mevduat bankalarının yapabilecekleri faaliyetler, 5411 Sayılı Bankacılık Kanunun 4. Madde’sinde aşağıdaki gibi düzenlenmiştir⁹:

- a) Mevduat kabulü,
- b) Katılım fonu kabulü,
- c) Nakdî, gayrinakdî her cins ve surette kredi verme işlemleri,
- d) Nakdî ve kaydî ödeme ve fon transferi işlemleri, muhabir bankacılık veya çek hesaplarının kullanılması dâhil her türlü ödeme ve tahsilat işlemleri,
- e) Çek ve diğer kambiyo senetlerinin iştirası işlemleri,
- f) Saklama hizmetleri,
- g) Kredi kartları, banka kartları ve seyahat çekleri gibi ödeme vasıtalarının ihracı ve bunlarla ilgili faaliyetlerin yürütülmesi işlemleri,
- h) Efektif dâhil kambiyo işlemleri; para piyasası araçlarının alım ve satımı; kıymetli maden ve taşların alımı, satımı veya bunların emanete alınması işlemleri,

⁷ Ender Aykut Yılmaz, “Basel II Kapsamında Kredi Riski Ve Temerrüt Olasılığı Hesaplama Yolu İle Tespiti”, Doktora Tezi, Gazi Üniversitesi, 2010, s.4.

⁸ Birgül Şakar, **Banka Kredileri Ve Yönetimi**, İstanbul: Beta Yayınları, 6. Baskı, 2015, s.1.

⁹ Mevzuat Dergisi, “5411 Sayılı Bankacılık Kanunu”, Çevrimiçi, <https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5411.pdf>, 05.09.2020.

i) Ekonomik ve finansal göstergelere, sermaye piyasası araçlarına, mala, kıymetli madenlere ve dövizde dayalı; vadeli işlem sözleşmelerinin, opsiyon sözleşmelerinin, birden fazla türev aracı içeren basit veya karmaşık yapıdaki finansal araçların alımı, satımı ve aracılık işlemleri,

j) Sermaye piyasası araçlarının alım ve satımı ile geri alım veya tekrar satım taahhüdü işlemleri,

k) Sermaye piyasası araçlarının ihraç veya halka arz yoluyla satışına aracılık işlemleri,

l) Daha önce ihraç edilmiş olan sermaye piyasası araçlarının aracılık maksadıyla alım satımının yürütülmesi işlemleri,

m) Başkaları lehine teminat, garanti ve sair yükümlülüklerin üstlenilmesi işlemleri gibi garanti işleri,

n) Yatırım danışmanlığı işlemleri,

o) Portföy işletmeciliği ve yönetimi,

p) Hazine Müsteşarlığı ve/veya Merkez Bankası ve kuruluş birlikleri nezdinde oluşturulan bir sözleşme kapsamında üstlenilen yükümlülükler çerçevesinde alım satım işlemlerine ilişkin piyasa yapıcılığı,

r) Faktöring ve forfaiting işlemleri,

s) Bankalararası piyasada para alım satımı işlemlerine aracılık,

t) Finansal kiralama işlemleri,

u) Sigorta acenteliği ve bireysel emeklilik aracılık hizmetleri,

v) Kurulca belirlenecek diğer faaliyetler.

Mevduat bankaları birinci fıkranın (b) ve (t), katılım bankaları (a), kalkınma ve yatırım bankaları (a) ve (b) bentlerinde belirtilen faaliyetleri gerçekleştirememektedirler.

Kredi (credit) kavramı, Latince’de itimat etmek, inanmak, güvenmek anlamına gelen “credere” kelimesinden türetilmiştir. Türkçede ise, günlük dilde çeşitli, saygınlık

ve itibarlılık anlamında kullanımları mevcuttur. Kredi, geniş anlamda “hazır bir servetin ileride doğacak bir servetle değişimi” veya “bir malın bedeli olan paranın, hemen verilmeyerek ileride ödenmesi vaadi” olarak da tanımlanabilir. Ekonomik açıdan ise kredi, “hazır bir satın alma gücünün kullanılma imkânından, belli bir süreyle, başka bir kişi adına vazgeçilmesi veya bu gücün ona, yine belli bir süreyle, kullandırılması”dır¹⁰. Bankaların çeşitli yöntemlerle oluşturdukları parasal kaynak stoğunun, nakit ihtiyacı olan bireysel ya da kurumsal müşterilere, belli bir zaman ve miktar ölçüsünde, kullandırılması genel olarak “kredi” kavramıyla açıklanmaktadır. Bankacılık kanununda kredinin tanımı, “Bir bankanın, yapacağı istihbarat sonucunda, gerçek veya tüzel kişilere yasaları, iç politikaları ve kendi kaynaklarını da göz önünde bulundurarak teminat karşılığı ve teminatsız olarak para, teminat veya kefalet vermek şeklinde tanınmış olan olanak veya limit” şeklindedir. Krediyi veren kredi kuruluşu, yani bankalar ya da özel finans kurumları, belli bir oranda faiz ödeyerek topladıkları bu parasal kaynakları, müşterinin kredi kullanım sürelerine karşılık, biraz daha yüksek bir faiz oranı belirleyerek ihtiyaç duyan çevrelere kullandırma yoluyla kar elde etmeye çalışmaktadırlar. Bankaların kredi verme fonksiyonunun doğru bir şekilde yürütülebilmesi, onların karlılıkları ve dolayısıyla devamlılıkları açısından oldukça önem taşımaktadır¹¹.

Kredi riski ise, “kredi müşterisinin yapılan sözleşme gereklerine uymayarak yükümlülüğünü, kısmen veya tamamen, zamanında, yerine getirememesinden dolayı bankanın maruz kalabileceği zarar olasılığı şeklinde tanımlanmaktadır”¹². Diğer bir tanımla “kredi riski, en basit anlamıyla, bir bankanın kredi müşterisinin (borrower) ya da kendisiyle bir anlaşmaya taraf olanın (counterparty) anlaşma koşullarına uygun biçimde yükümlülüklerini karşılayamama olasılığıdır”¹³.

¹⁰ Mevzuat Dergisi, “Kredi Sözleşmelerinin Hukuki Niteliği Ve Türleri”, Çevrimiçi, <https://www.mevzuatdergisi.com/2002/01a/01.htm>, 11/03/2020.

¹¹ Kezban Öztürk, **Kredi Politikası Ve Değerlendirmesi**, Ankara: Siyasal Kitabevi, 2015, s.13.

¹² “Bankaların İç Sistemleri Hakkındaki Yönetmelik”, Resmi Gazete, 26333, 01.11.2006. Çevrimiçi, http://www.bddk.org.tr/WebSitesi/turkce/Mevzuat/Bankacilik_Kanununa_Iliskin_Duzenlemeler/1678_bankalarin_ic_sistemleri_hakkinda_yonetmelik_24_02_2011.pdf.

¹³ Türkiye Bankalar Birliği, **Kredi Riskinin Yönetimine İlişkin İlkeler**, Çevrimiçi, https://www.tbb.org.tr/Dosyalar/Arastirma_ve_Raporlar/risk_yonetim.doc, 28/10/2019.

Bireylerin veya kurumların, içsel veya dışsal olarak herhangi bir nedenden ötürü, yükümlülüklerini yerine getirememesi olarak bilinen operasyonel risk, bankaların karşılaştıkları, en eski risk türüdür. Etkin kredi risk yönetiminin sağlıklı bir şekilde hayata geçirilebilmesi, bunun için gerekli ve yeterli önlemlerin alınabilmesi, yasal yükümlülüklerin yerine getirilebilmesi, risk odaklı performans sisteminin geliştirilebilmesi gerekir. Bunun dışında sağlıklı risk yönetimi için; kurumun, risk durumu hakkında bilgi almak isteyen üçüncü kişi ya da kurumlara zamanında ve doğru cevap verilebilmesi ve bankaların kredi riskini doğru belirleyebilmeleri için operasyonel riski de, kesinlikle göz önünde bulundurmaları gerekmektedir¹⁴.

Kredi kalemi, bankalar her ne kadar çok farklı amaçlar için borç verseler de, genelde aktiflerin içinde en büyük ağırlığa sahiptir ve birçok bankacı, bankacılığın en önemli faaliyetinin kredi vermek olduğunu düşünmektedir. Örnek vermek gerekirse, bankacılık kurumunun en çok geliştiği ABD’nde bankaların kredi verme faaliyetlerinden elde ettiği gelirler, bankaların toplam gelirinin üçte ikisine (2/3) karşılık gelmektedir. Bu durum ülkemizde de, diğer birçok ülkede de çok farklı değildir. Dolayısıyla kredi riski, bankacılık risk grubunun en önemli kalemini oluşturmaktadır. Yani kredi verme, karlı olduğu kadar riskli de bir faaliyet alanıdır. Sonuç olarak kredi riski yönetiminin etkinliği, bankaların en kritik yönetim süreçlerinden birisini oluşturmaktadır. Çünkü üçüncü kişilere ait olan fon kaynaklarını kullandıkları için, kredilerin geri dönmemesi, belirli kişi ve kuruluşlarda donması, bankaların varlıklarını sürdürememeleri sonucunu doğuracak ve bu da ekonominin etkinliğine belki de en üst zararı verebilecek potansiyele sahip, tehlikeli bir yönetim sürecidir¹⁵. Yine örnek vermek gerekirse, 2008 yılında ABD’nde gerçekleşen mortgage krizi, dünya genelindeki birçok ülkeyi etkileyen, küresel bir finansal krize dönüşmüştür. Mortgage krizi (konut kredilerindeki kontrolsüz ve aşırı büyümeden kaynaklanan finansal kriz), ABD’ni çok büyük zarara uğratmış, temerrüde düşen birçok kişi ve kurum da evlerini satışa çıkararak borç ödeme yoluna gitmiş, bu da

¹⁴ Erdinç Altay, **Bankacılıkta Risk, Piyasa Riski Kredi Riski Ve Operasyonel Riskin Ölçümü Ve Yönetimi**, İstanbul: Derin Yayınları, 2015, s.429.

¹⁵ Birgül Şakar, **a.g.e.** s.1-2.

evlerin fiyatlarının düşmesine neden olarak krizi daha da derinleştirmiştir. Bu krizin sebebi de, iyi yönetilemeyen bir kredilendirme süreci sonucudur denilebilir.

1.2. Kredi Riski Yönetimi

Bankacılıkta kredi riski, “belli bir tarihte, hali hazırda kullanılmakta olan tüm nakit ve nakit dışı kredilerin toplamı” olarak kabul edilir. Daha geniş anlamıyla bankacılıkta risk, “gerçek ya da tüzel kişiye ait nakit, mal, kefalet ve teminat şeklinde kullanılan kredilerin toplamıdır, yani kredi olarak verilen paranın geri alınmasına ve verilen garantiye konu olan taahhüdün yerine getirilmesine kadar olan sürede oluşabilecek tehlikelerin toplamıdır” şeklinde tanımlanmaktadır. Kredi işlemlerinin sonucunda çok çeşitli risk kaynakları oluşabilir. Bu riskler; ödenmeme, likidite, teminat kaynaklı, faizde değişim, satın alma gücü, firma yöneticilerinin hatalarından kaynaklı, çevresel, sektörel, ürün/pazar, politik, yönetim kalitesinden kaynaklı, döviz kurunda dalgalanma ve politik istikrarsızlık gibi risklerdir. Ayrıca, kredi talebi iyi değerlendirilse bile, kredi yanlış yapılandırılırsa, bu da onun geri ödeme riskini artırıcı bir etki doğurabilmektedir. *Kredi risk yönetimi*, kredi kayıplarını en aza indirmek için uygulanan yöntem ve sistemleri içermektedir ¹⁶.

Risk, serbest piyasaların “olmazsa olmaz!” koşuludur. Riskin, yönetimi, rehabilitasyonu ve optimizasyonu mümkündür. Fakat riskin hiç olmadığı bir durum; düşünülebilse bile, çoğunlukla “sıfır risk, sıfır getiri” anlamına geleceği için, gelir elde etmek isteyen serbest piyasa oyuncularının tercih etmek isteyeceği bir yol değildir. Az ya da çok muhakkak çeşitli oranlarda, farklı türden risklere maruz kalınan piyasa koşullarında önemli olan nokta; bankaların sahip oldukları sermaye imkânlarını titizlikle dikkate alarak, almak istedikleri riskin sınırlarını iyi belirleyebilmeleri ve bu doğrultuda sermaye yeterliliklerini oluşturmalarıdır.

Kar maksimizasyonunu gerçekleştirmeye çalışan bankalar, doğal olarak, maksimum kapasitede kredi kullandırma seçeneği üzerine yoğunlaşacaklardır. Fakat burada söz konusu olan, sadece kullanılan kredi miktarının maksimizasyonu olsaydı; komplike

¹⁶ Birgül Şakar, **a.g.e.** s.4-39.

herhangi bir y nteme gerek kalmaksızın, bankalar y ksek getiri saėladıkları y ksek riskli m  terilerle kar maksimizasyonunu ger ekle tirirlerdi ve kaynaklarını bu y ne kanalize ederlerdi. Fakat kredi risk y netiminde s z konusu olan; kullandırılan kredi miktarını maksimize ederken, bu kredilerden dolayı maruz kalınan kredi riskinin de minimize edilebilmesidir. Geni   er evede k r maksimizasyonu, ancak bu  ekilde saėlanabilir.   te bu noktada, etkin bir kredi risk y netimi i in kapsamlı analizler gerekmektedir.

Kredi riskinin  l  lmesinin ardından, kredi kararının bir kurul tarafından verilmesi, tahsil tının takibi, kontrol  ve s recin saėlıklı bir  ekilde tamamlanması ve bankanın, en az yılda bir kere, kredi durumunun kapsamlı bir  ekilde g zden ge irilmesi, gerekli  nlemlerin alınması; kredi risk y netiminin unsurlarını te kil etmektedir¹⁷.

Son yıllarda  ok b y k bankaların  nemli bir kısmı, i  kollarından kaynaklanan kredi risklerini deėerlendirmek amacıyla karma ık i sel sistemler geli tirmi tir. “Riskin  l  m  ve y netimine y nelik model tabanlı yakla ımın faydaları nelerdir?” sorusunun  ne  ıkan cevabı; en  nemlisi, daha iyi risk  l  m  ve y netiminin daha verimli bir sermaye tahsisini saėlamasıdır. Risk daha iyi y netildiėinde, krediler daha doėru bir  ekilde fiyatlandırılabilir ve daha fazla sayıda piyasa oyuncusu i in daha kolay yaygınla abilir. Kredi riski modellemesindeki iyile me, kredi t revleri ve teminatlı bor  y k ml l kleri gibi kredi riski transferi i in yeni pazarların olu masına yol a acaktır. Bu yeni pazarlar, piyasa katılımcılarının kredi riskini payla ma y ntemlerini geli tirecek ve bu da riskin daha etkin fiyatlandırılmasını saėlayacaktır¹⁸.

1.3. Kredi Riskinin Belirlenmesi Ve Basel Komitesi Kararları

1.3.1. Kredi Riskinin Belirlenmesi

Kredi riskinin belirlenmesinde kullanılan iki temel yakla ım olan standart ve i sel  l  m yakla ımlarını a ıklamadan  nce, bu yakla ımları tavsiye niteliėinde

¹⁷ Ender Aykut Yılmaz, **a.g.e.** s.4.

¹⁸ Eliana Angelini, Giacomo Di Tollo, Andrea Roli, “A Neural Network Approach For Credit Risk Evaluation”, **The Quarterly Review of Economics and Finance**, Cilt: 48, 2008, s.734.

yayınlayan (Basel II) ve uluslararası bir akreditasyon kriteri haline gelen, “Basel Kararları” ve bu kriterleri oluşturan Basel Komitesinin gelişim sürecine ve işlevlerine bakmak, bugünkü yaygın kredi riski politikalarını anlama açısından önemlidir.

1.3.2. Basel Komitesi Ve Kredi Risk Yönetimi Konusundaki Tavsiye Kararları

Uluslararası Takas Bankası, BIS (Bank of International Settlements), 1930 yılında İsviçre’nin Basel şehrinde, Belçika, Kanada, Fransa, Almanya, İtalya, Japonya, Lüksemburg, Hollanda, İspanya, İsveç, İsviçre, İngiltere ve ABD’nin de dâhil olduğu toplamda 13 ülkenin merkez bankalarının ve bankacılık denetim otoritelerinin yetkilileri tarafından kurulmuştur. BIS, “merkez bankalarına parasal ve finansal istikrar arayışı içinde hizmet etmek, bu alanlarda uluslararası işbirliğini teşvik etmek” amacını taşımaktadır¹⁹. BIS, parasal ve finansal istikrarı sağlamak için düzenli olarak politika analiz sonuçları ve uluslararası bankacılık ve finans istatistikleri yayınlamaktadır. 1970’li yılların başında, sabit kur sisteminin terk edilmesi, 1974 petrol krizi vb. sebeplerle uluslararası bankacılık ve döviz alanında yaşanan aşırı dalgalanma neticesinde BIS tarafından, “Bankacılık Düzenleme ve Denetleme Kurulu”, yani bilinen adıyla “Basel Komitesi” oluşturulmuştur (1974). Basel Komitesi’nin aldığı kararların, bankalar üzerinde, herhangi bir yaptırım olmayıp kararlar tamamen tavsiye niteliğindedir. Fakat bu komite tarafından alınan kararlara uymak, özellikle üye ülkelerin parasal ve finansal güvenlikleri için, piyasa koşulları ve rakabet edebilirlikleri açısından kaçınılması zor bir hal almıştır. Türkiye’nin de üyesi olduğu komite, bugün dünya genelindeki GSYİH’nın yaklaşık % 95’ine karşılık gelen bir üyelik portföyüne²⁰ ulaşmıştır. Komite tarafından yapılan standartlaştırma çabalarıyla uluslararası denetim farklılıklarının ve boşluklarının giderilmesi hedeflenmektedir²¹.

¹⁹ İbrahim Arslan, “Basel Kriterleri Ve Türk Bankacılık Sektörüne Etkileri”, **Selçuk Üniversitesi Sosyal Bilimler Dergisi**, 2007, s.49.

²⁰ Bank of International Settlements, About BIS Overview, Çevrimiçi, <https://www.bis.org/about/index.htm>, 28/10/2019.

²¹ Ahmet Turgut, “Bankalar Ve Kobi’ler İçin Dönüm Noktası: Basel II Standartları”, **Karamanoğlu Mehmetbey Üniversitesi Sosyal Ve Ekonomik Araştırmalar Dergisi**, Cilt: 2007, Sayı: 2, s.31.

1.3.2.1. Basel I Kararları

Basel Komitesi, 1974 yılındaki petrol krizinin ardından kurulmuş olmasına rağmen ilk bildirgesini 1988 yılında yayınlamıştır. Basel I olarak bilinen bu bildiride, Sermaye Yeterliliği Rasyosu (SYR), “Toplam Sermaye/Kredi Riski” olarak belirlenmiş ve bankalar için bu oranın %8 olması tavsiye edilmiştir. Sonra bu rasyonun formülü, 1996 yılında aşağıdaki şekilde (1.1) güncellenmiş olup, oranı ise değiştirilmemiştir²².

$$SYR = \text{Toplam} \frac{\text{Sermaye}}{\text{Kredi Riski} + \text{Piyasa Riski}} \quad (0.1)$$

1.3.2.2. Basel II Kararları

2004 yılında ilk kez yayınlanan, 2005’de yeniden güncellenen ve 2006’da en son şeklini alan Basel II kararları ise, riskleri kayıtlı öz sermaye ihtiyaçları ile daha uyumlu hale getirmiş, riskin belirlenmesinde ve yönetilmesindeki gelişmeleri dikkate alarak daha kapsamlı bir finansal yaklaşım inşa etmiştir. Basel II kararlarıyla, finansal sistemlerin daha güvenli olması gerektiğini tavsiye eden ve rekabet eşitliği ilkesini teşvik eden komite, bu kararların yayınlanmasından itibaren, çeşitli düzeylerde karmaşıklık sergileyen uluslararası bankalara odaklanmıştır. Basel II bankaların, kurumsal müşterilerine kredi verirken, onları tasnif etmelerini ve oluşturacakları bu yeni sınıfları ayrı ayrı değerlendirmeye tabi tutmalarını (içsel ölçüm yaklaşımı) ve değerlendirme süreçlerinde, derecelendirme kuruluşlarının vereceği kredi notlarını dikkate alarak (standart yaklaşım) kredi taleplerini cevaplandırmaları gerektiğini belirtmiştir. Yani kredi notları aracılığıyla, müşterinin kredi alıp alamayacağı veya hangi oranlarla kredi alabileceğinin kararına varılması (standart yaklaşım) gerektiği tavsiye edilmiştir. Dolayısıyla, Basel II’nin finans dünyasında getirdiği temel yenilik, müşterinin kendi not durumuna göre, kredilendirmenin farkı şekillerde uygulanabileceği esası olmuştur²³.

²² Ömer Faruk Demirkol, Emel ABA, “Basel II Sermaye Yeterliliği Uzlaşısı Süreci İçinde Gelen Ek Düzenlemeler Seti: Basel III Kriterleri, **Niğde Üniversitesi İİBF Dergisi**”, 2012, Cilt: 5, Sayı: 2, s.51.

²³ İbrahim Yılmaz, **a.g.e.** s.53.

Basel-II'nin Basel-I'den farkları ise:

- Basel-I'de yer alan, kredi riski açısından sermaye yükümlülüğünün OECD ülkesi olup olmama kriterine göre belirlenmesi prensibine dayanan “külüp kuralı (club rule)” Basel-II ile birlikte kaldırılmıştır.
- Basel-II'de kredi riski, krediyi alan tarafların derecelendirme notlarına göre belirlenmektedir. Basel-II'de yer alan bazı yöntemler bağımsız derecelendirme şirketleri (ör. Standard&Poors, Fitch, Moodys, vb) tarafından verilen derecelendirme notlarını kullanırken, bazı ileri yöntemlerde -bankacılık denetim otoritesi iznine tabi olmak üzere- bankaların kendi değerlendirmelerine (içsel ölçüm yaklaşımı) dayanarak verdikleri derecelendirme notları dikkate alınmaktadır.
- Basel-I'de sadece kredi ve piyasa riskleri için sermaye zorunluluğu bulunmaktadır. Basel-II'de ise bu risklere ilaveten operasyonel risk sermaye yükümlülüğü de eklenmiştir. Basel-II'de operasyonel risk, yetersiz veya aksayan iç süreçler, insanlar ya da sistemler ya da harici olaylar sonucu ortaya çıkan zarar riski olarak tanımlanmakta (örneğin zimmet, deprem, bilgi işlem arızası sonucu ortaya çıkan zararlar) ve bankalardan bu riskler için de sermaye bulundurmaları istenmektedir.
- Basel-II içerisinde bankaların sermaye yeterliliklerini kendilerinin değerlendirmesi istenmekte ve hem sermaye yeterliliği hem de bankanın kendisini değerlendirme sürecinin bankacılık denetim otoritesi tarafından denetlenmesi ve değerlendirilmesi istenmektedir.
- Basel-II kapsamında sermaye yeterliliğine ilişkin olarak Basel-II'ye özgü detaylı bilgilerin kamuya açıklanması zorunluluğu getirilmektedir. Bu husus Basel-I'de yer almamaktadır²⁴.

Bunun dışında Basel II'nin bankacılık ve finans dünyasına katkılarını aşağıdaki gibi özetlemek mümkündür²⁵:

²⁴ Bankacılık Düzenleme Ve Denetleme Kurumu, “10 Soruda Yeni Basel Sermaye Uzlaşısı (Basel-II)”, 2005, s.2, Çevrimiçi, https://www.bddk.org.tr/ContentBddk/dokuman/duyuru_basel_0001_38.pdf, 07.09.2020

²⁵ Ömer Faruk Demirkol, Emel ABA, **a.g.e.** s.261.

- Etkin ve disipline edilmiş bir bankacılık sisteminin oluşması,
- Çağdaş risk yönetimi tekniklerinin geliştirilmesi,
- Bankaların mevcut sermayelerinin, maruz kalabilecekleri riski karşılayabilmesi,
- Basel II hükümleri gereği, bankaların güçlü bir özkaynak yapısına sahip olma koşulu sayesinde, yabancı bankalarla birleşme sürecini hızlandırması,
- Şeffaflık ilkesi gereği piyasa disiplini için gerekli özverinin sağlanması ve çeşitli kriz senaryolarına uygun risk yönetimi anlayışının piyasalarda oluşması,
- Basel II hükümleri gereği bankalardaki problemlerin etkin çözümüne yönelik insan kaynakları ve bilgi işlem gibi alt yapı unsurlarının, bankaların müşteri portföy tercihlerini, kredibilitesi yüksek ve derecelendirmesi iyi firmalardan yana kullanmaları ve söz konusu firmalara kredi faiz oranından daha düşük oranda kredi kullandırmaları.

Basel II kararlarının, Basel I'e göre getirdiği, birçok yenilik bulunmakla birlikte bu yenilikler arasında en büyük etkiyi oluşturan; kullandırılan kredilerin risk ağırlıklarının belirlenmesinde, kredi borçlusunun derece notlarının esas alınmasıdır²⁶.

1.3.2.3. Basel III Kararları

Basel II düzenlemelerinden sonra, 2008 yılında, büyük küresel finansal kriz sürecinde Lehman Brothers iflas etti. Bunun dışında, ABD'deki büyük yatırım bankaları holding şirketlerine dönüştürüldü. Federal Ulusal Mortgage Derneği (Fannie Mae, Federal National Mortgage Association, FNMA) ve Federal Ev Kredisi Mortgage Kurumu (Federal Home Loan Mortgage Corporation, Fredi Mac), bu kriz neticesinde, ulusallaştırıldı. Ayrıca dünyanın en büyük sigorta şirketlerinden olan, American International Group'un (AIG) çöküşün eşiğine geldi. Bunun dışında Hollanda-Belçika Finans Grubu (Fortis) dağıldı ve satıldı. Yine bunun gibi, İzlanda'nın en büyük ticari bankası battı ve ardından ülkenin bankacılık sisteminin tamamı çöktü. Birçok ülkenin kendi bankalarına ciddi destekler vermesini gerektiren bu gibi olaylar, kriz durumları

²⁶ Hasan Candan, Alper Özün, **Bankalarda Risk Yönetimi Ve Basel II**, İstanbul: Türkiye İş Bankası Kültür Yayınları, 2006, s.136.

için, yeterince önlemin henüz alınmamış olduğunu ve mevcut finansal sistemlerin çok ciddi boşluklar içerdiğini göstermiştir. Tüm bunların yanı sıra, küresel finansal kriz, reel sektörü de etkilemiş refah seviyesinde çok ciddi kayıplar görülmüş ve yüksek seviyelerde iş kaybı meydana gelmiştir. Tüm bu sebeplerle, gelecekte karşılaşılabilecek krizlere karşı, bankacılık ve finans sisteminin, daha dirençli olmasını sağlamak amacıyla likidite ve sermaye kalitesinin artırılması, ekonomik döngünün dikkate alınması ve sermaye yükümlülüğünün artırılması gibi önemli reformların gerçekleştirilmesi adına, Basel Komitesi tekrar toplanmış ve Basel III kriterleri olarak bilinen reformlar, 12 Eylül 2010 tarihinde, komitenin kendisi tarafından kamuoyuna bir basın açıklaması ile deklare edilmiştir²⁷.

Şüphesiz ki yukarıda da sayılan ve ciddi kayıplara neden olan bu gibi ekonomik krizlerin önüne geçebilmek, ekonomik çevrelerin temel sorunudur. Bu temel sorunun önemli bir nedeni ise daha önce geçerli olan kriterlerin (Basel II) eksik yönleridir. Basel III ile Basel II kriterlerinin eksik yönleri giderilmeye çalışılmış ve yeni yaklaşımlar ve yeni tedbirler ortaya konmuş böylece yeni finansal krizlerin önüne geçilmeye çalışılmıştır.

Basel III deklarasyonunu ortaya çıkaran nedenleri ve halihazırda yürütülen çalışmalar şu şekilde özetlenebilir²⁸:

- Olumsuz piyasa koşullarında ani düşüş gösterebilen sermaye tamponlarının güçlendirilmesi,
- Banka sermayelerinin kalitesinin artırılması,
- Basel II'ye destek olmak üzere bir kaldıraç oranı uygulamasının getirilmesi,
- Asgari sermaye gereksinimindeki döngüsellik azaltma ve karşılık ayrılması,
- Bankacılık sektörünün kuvvetlendirilmesi amacıyla ortaya koyduğu sermaye ve likidite düzenleme önerileri sunması,

²⁷ BDDK, Sorularla Basel III, Risk Yönetimi Dairesi, Aralık, 2010, s.3-4.

²⁸ Eymen Gürel, Esra Burcu Bulgurcu Gürel ve Neslihan Demir “Basel III Kriterleri”, **BSAD Bankacılık ve Sigortacılık Araştırmaları Dergisi**, Cilt: 1, Sayı: 3-4, 2012, s.17.

- Risk yönetiminin geliştirilmesine ek olarak, bankaların stres ortamlarına karşı dayanıklılığının artırılması.

Basel III ile ulaşılmak istenen hedefler ve halihazırda devam eden çalışmalar:

- Kaynağı ne olursa olsun finansal ve ekonomik şoklara karşı bankacılık sisteminin dayanıklılığının artırılması,
- Kurumsal yönetim ve risk yönetimi uygulamalarının geliştirilmesi,
- Bankaların şeffaflığının ve kamuya bilgi verme özelliklerinin artırılması,
- Mikro bazda yapılan düzenlemelerle bireysel olarak bankaların dayanıklılığının artırılması,
- Makro bazda düzenlemelerle finansal sistemin şoklara karşı direncinin artırılması.

Devam eden çalışmalar ise:

- Alım satım hesaplarının (trading book) temelden gözden geçirilmesi,
- Dışsal derecelendirme notlarının seküritizasyon işlemlerinde (sermaye hesaplamaları çerçevesinde) kullanılması ve etkileri,
- Sistemik olarak öneme haiz finansal kuruluşlara ilişkin politikalar geliştirilmesi, Büyük risklere ilişkin düzenlemeler,
- Sınır ötesi (cross-border) bankacılık ilkelerinin geliştirilmesi,
- “Etkin Bankacılık Denetiminin Temel İlkeleri”nin krizinden çıkarılan dersler sonrasında gözden geçirilmesi,
- Standart uygulamaların ve güçlü işbirliğinin banka denetçileri ve denetim otoriteleri arasında geliştirilmesi şeklindedir²⁹.

1.3.3. Basel Ve Kredi Riski Özeti

Kredi riskinin değerlendirilmesi konusunda BIS’in fonksiyonunu özetleyecek olursak, bankaların almaları gereken önlemleri belirlemenin yanı sıra, kredi verme konusunda, kurumsal müşteriler için standart yaklaşımı tavsiye etmektedir. Bunun dışında BIS,

²⁹ Eymen Gürel, Esra Burcu Bulgurcu Gürel ve Neslihan Demir, **a.g.e.** s.17.

yine Basel Komitesi aracılığıyla, 2000 yılında yayımlanan “Kredi riski yönetim prensipleri” adlı bildirisinde, bankalara kredi riski yönetimi için, tavsiye niteliğinde, on yedi adet prensip yayınlamıştır. Bu prensiplerden onuncusu; uygun bir kredi yönetimi ölçüm ve izleme süreci için, “bankalar kredi riskini yönetmede, bankanın icraatlarını, karmaşıklığını, büyüklüğünü ve doğasını dikkate alarak bir iç risk ölçüm sistemi geliştirmeleri ve bunu faydalı bir hale getirmeleri teşvik edilir” şeklindedir³⁰. Yani Basel, standart yaklaşımın yanı sıra, bankalar için bir içsel risk analiz yöntemi oluşturmalarını tavsiye etmektedir.

Özetle finans kuruluşları için, hem kredi verme faaliyetlerinde tavsiye kriterler edinmek ve böylelikle kendilerini operasyonel olarak sağlama almak hem de dünya çapında geçerli bu kriterlere akreditasyon sağlayarak uluslararası işlemlerde ivme ve güven kazanmak için, Basel kriterlerine uyum oldukça önem arz etmektedir. Fakat daha fazla kar elde etmek isteyen piyasa oyuncularını için yüksek riskli finansal hedefler her zaman var olmaya devam edecektir.

1.3.4. Kredi Riskinin Belirlenmesinde Kullanılan Temel Yaklaşımlar

Kredi riskinin belirlenmesinde kullanılan yöntemleri, iki temel başlıkta özetlemek mümkündür. Bunlardan birincisi, bankaların dışsal puan ve sınırları kullanmalarına dönük kredi değerlendirme stratejisi diyebileceğimiz “standart yaklaşım”, ikincisi ise, bankaların içsel bilgi ve sistemlerle karar verme yöntemi diyebileceğimiz, “içsel ölçüm yaklaşımı”dır.

1.3.4.1. Standart Yaklaşım

Standart yaklaşımla, finans kurumlarının, kredilendirme süreci için; Basel II doğrultusunda, Standart & Poors, Moody’s, Fitch, Dış Kredi Değerlendirme Kuruluşları (External Credit Assessment Institutions, ECAI) vb. açıklayıcı kredi derecelendirme kuruluşlarının ülkeler, bankalar ve şirketlere verdikleri notlar değerlendirilerek karar verilmesi kast edilmektedir. Bunun, daha basiti olan,

³⁰ Bank of International Settlements, , Principles for the Management of Credit Risk, 2000, Çevrimiçi, <http://www.bis.org/publ/bcbs75.pdf>, 29/06/2011, s.4.

basitleştirilmiş standart yaklaşımdan farkı; risk ağırlıklarının belirlenmesinde ihracat kredi kuruluşları (Export Credit Agencies), kısa adıyla ECA tarafından verilen notların kullanılmasıdır³¹.

Standart yaklaşımda alacaklar çeşitli varlık sınıflarına bölünmüş ve her bir sınıfın risk ağırlığı ayrı ayrı belirlenmiştir. Bu varlık sınıfları:

- Hazine ve merkez bankalarına kullandırılan krediler,
- Bankalara kullandırılan krediler,
- Çok taraflı kalkınma bankalarına kullandırılan krediler,
- Merkezi hükümet içinde yer almayan kamu kurum ve kuruluşlarına kullandırılan krediler,
- Menkul kıymetler şirketlerine kullandırılan krediler,
- Kurumsal krediler,
- Perakende krediler,
- Konut ipoteği karşılığı krediler,
- Ticari gayrimenkul ipoteği karşılığı krediler,
- Tahsili gecikmiş alacaklar,
- Yüksek risk kategorileri,
- Diğer

Sınıflandırılmayan konuların “diğer” başlığı altında sınıflandırılması, sınıflandırma içeren tüm düzenlemelerde yer almaktadır. Teamülde ise bu başlıkların en büyük varlık sınıflarını kapsadığı görülmektedir. Basel II kararlarında, her ne kadar sermaye yeterliliğinin hesaplanmasında kullanılan yöntemler ayrıntılı bir şekilde açıklansa da, belli bazı hususlarda tercih hakkı ulusal denetim otoritelerine bırakılmıştır³².

Standart yaklaşım, anlaşılması ve ölçülmesi özellikle içsel ölçüm yaklaşımlarına kıyasla çok daha kolay olması nedeniyle, bu yaklaşımı uygulamak isteyen kuruluşların sayısı da diğer yönteme göre daha fazla olmaktadır. Basel II'ye yeni geçmiş ya da geçmeyi planlayan, gelişmekte olan ülkelerde faaliyet gösteren kurumların/bankaların

³¹ BDDK, Basel II Sayısal etki Çalışması (QIS-TR) Değerlendirme Raporu, Aralık, 2004, s.18.

³² Hasan Candan, Alper Özün, **a.g.e.** s.136-137.

büyük çoğunluğu ve gelişmiş ülkelerdeki küçük ölçekteki bankalar standart yaklaşımı kullanmak istemektedirler. Öte yandan uygulamak isteyen ancak yetersiz veri ya da yüksek maliyet gibi nedenlerle içsel değerlendirme yöntemini uygulama imkânını pratikte yakalayamayan bankaların da, en azından bazı kredi türleri bazında, standart yaklaşımı kullanmak istedikleri görülmektedir. Fakat bazı ülkeler de, örneğin İngiltere’de, FSA (Financial Services Agency) Temmuz 2003, tarihli bir çalışmasında “İçsel değerlendirme yaklaşımının, ölçeğine bakılmaksızın tüm şirketlerce uygulanabileceğine ve uygulanması gerektiğine inanıyoruz” şeklinde bankaların, içsel değerlendirme yaklaşımını , faaliyet göstermelerini isteyen açıklamalar yapan kuruluşların da bulunduğu görülmektedir³³.

1.3.4.2. İçsel Ölçüm Yaklaşımları

İçsel ölçüm yaklaşımlarının temeli, banka karar vericilerinin de onayı doğrultusunda, bankaların kendi oluşturdukları kredi riski ölçüm modelleriyle kendi sermaye gereksinimlerini hesaplayabilmeleri ve bu şekilde risk yönetimini, kendi içsel, operasyonel, ihtiyaçlarına dönük olarak ve içinde bulunduğu piyasa şartlarına uyumlu olarak risk yönetimini gerçekleştirebilmeleridir. Literatüre bakıldığında, perakende çalışan bankaların, içsel ölçüm yaklaşımını seçmeleri durumunda, %30'lara varabilen önemli miktarlarda sermaye indirimi sağlayabileceği, bu yöntemin kabul görmüş bir avantajıdır. Buna karşın finansal kuruluşun, kurumsal inisiyatif olarak, içsel ölçüm yaklaşımını kullanması durumunda risk, standart yaklaşıma kıyasla daha yüksek olup, bu da sistemi kullanan bankaların dezavantajı olarak kabul görmektedir. İçsel ölçüm yaklaşımlarının standart yaklaşıma nazaran bir diğer dezavantajı da çok yoğun bir çalışma ve adaptasyon süreci gerektirmesidir. Bu süreçte diğer iş süreçlerine uyumlu ve de kapsamlı risk ölçüm yöntemlerine ve bunları doğru bir şekilde uygulayabilecek uzman ekiplere ihtiyaç duyulmaktadır. Bu yöntemler ise sürekli, doğru ve etkin veri bankaları gerektirmektedir. Veri bankaları da birçok gelişmekte ola ülke için, bankanın kendi verisinden ibaret olduğu için, bankaların kısmen deneme yanılma yoluyla, bunu

³³ Ozan CANGÜREL, “Basel II Kapsamında Kredi Riskinin Ölçümünde Otorite Etkinliği: Türkiye İçin Alternatif Bir Öneri”, Doktora Tezi, Marmara Üniversitesi, 2011, s.49.

gerçekleştirebildiği söylenebilir³⁴. Bankalar içsel ölçüm yaklaşımları kullanarak elde ettikleri marjinal gelirin, yani sadece standart yaklaşımla elde edilen karın üzerinde elde ettikleri edindiklerin kazancın, içsel ölçüm yaklaşımlarının maliyetinin üzerinde olmasına özen göstermelidirler. Aksi takdirde, yoğun emeklerin ve yüksek maliyetlerin karşılığı, bankaya gelir olarak yansımayacaktır.

İçsel ölçüm yaklaşımlarını kullanan bankalar, kredi derecelendirme sistemlerini belirledikten sonra, vermiş oldukları kredilerin dereceler bazında dağılımını, dereceler arası geçiş istatistiklerini, beklenen ile gerçekleşen kayıp sıklığının karşılaştırılmasını ve beklenen ile gerçekleşen temerrüt kaybı ve bakiyesinin karşılaştırmalarını içeren raporlar oluşturmalıdır. Bununla birlikte içsel ölçüm yaklaşımlarının, kredi risk yönetiminin odağını teşkil etmesinden dolayı, bankaların üretecekleri veri tabanlarının temerrüt olasılığı, temerrüt halinde kayıp oranı, temerrüt tutarı ve efektif vade gibi risk bileşenlerini de hesaplayabilmeleri gerekmektedir. Bu sayede gerekli olan karşılık tutarlarının hesaplanması için beklenmeyen kayıpların tahmin edilmesi, mümkün hale gelebilecektir³⁵.

Kredi riskini içsel olarak ölçebilmek için, günümüz koşullarında, bankalar bireysel başvurularda kişinin gelir beyanı, gelirinin üzerinde herhangi bir haciz kesintisinin olup olmadığını, üzerine kayıtlı mal varlıklarının durumu gibi ölçütleri değerlendirmeye almaktadırlar. Bunun yanı sıra Türkiye’de dokuz bankanın geçmişe dönük veri birikimini tek bir havuzda toplayarak, kişi ya da kurumların kredi puanına karşılık gelen bir “findeks puanı” belirlenmekte ve bu puan kredilendirme sürecinde oldukça etkin rol oynamaktadır. Findeks puanı, kişinin kredi çekme, geri ödeme, fatura ödeme, mevduat ve borç ödeme alışkanlıkları gibi özellikleri dikkate alınarak ve çeşitli algoritmalar kullanılarak hesaplanmaktadır³⁶.

Bankalar kredi başvurularını sonuçlandırırken tüm bunları, yani kişinin mal varlığı, borç stoğu, findeks puanı ve standart yaklaşım çerçevesinde kullanılan kredi puanı gibi verilerin hepsini birer karar değişkeni olarak göz önünde bulundurabildikleri gibi

³⁴ Ozan CANGÜREL, **a.g.e.** s.54-55.

³⁵ Erdiñç Altay, **a.g.e.** s.425.

³⁶ Findeks, Çevrimiçi, <https://www.findeks.com/urunler/findeks-kredi-notu>, 28/10/2019.

sadece uluslararası kurumlardan ya da findeks gibi yerel kurumlardan aldıkları not bilgisi ile de hareket edebilmektedirler.

Sonuç olarak, kredi riskinin belirlenmesinde temel unsurların soyut olarak tanımlanması kolaydır. Ancak oluşabilecek binlerce farklı duruma göre uyarlanması çok daha zordur. Her durum, bireylerin rolleri ve yetenekleri, teknoloji sistemi, kurumların faaliyetleri ve hedefleri etrafında inşa edilmiş olup, benzersizdir. Sonuç olarak, rekabetçi kalabilmek için kurumların kredi riski belirleme tekniklerini, piyasa şartlarına ve kendi yeterlilik ve eksiklerini dikkate alarak kredilendirme sürecine uyum sağlamalı ve sürekli olarak sistemlerini güncellemeleri ve geliştirmeleri gerekmektedir³⁷.

1.3.5. Kredi İlke Ve Politikaları

Bankalar bir yandan içsel ölçüm yaklaşımlarına dönük olarak, karmaşık bir sistem uygulayarak bir yandan da bu sistemlerini niceliksel ve niteliksel olarak geliştirerek, zamanla kendi “kredi ilke ve politikalarını” belirlemektedirler.

Kredi politikası ise, “bankanın temel yönetim anlayışının kredilere ilişkin uygulamalarıyla uyumlaştırılması” olarak tanımlanmaktadır. Yazılı olması halinde, kredi ilke ve politikaları, banka çalışanlarının, yetki ve sorumluluk alanının farkındalığına ve bu çerçevede hareket edebilmesine olanak sağlar. Yani karar aşamasında, politikalar bir referans noktası ve standart uygulama fikirleri içerirler. Yönetime çalışmasını rapor ederek onay alma ihtiyacı hisseden personel, belirgin kredi ilke ve politikaları aracılığıyla, kredi talebini, kendi yetki alanı çerçevesinde, rahatlıkla cevaplandırabilmektedir. Yani politikalar, modern yönetim anlayışlarından olan personel güçlendirme (empowerment) fonksiyonunu da olanaklı hale getirmektedir. Sonuç olarak, bankaların kredi yönetiminde sağladıkları etkinlik seviyesi, bir anlamda onların uzun süredir uygulayageldikleri kredi politikalarının uzun vadeli bir sonucudur. Bankaların kredi politikaları, onların kredi felsefelerini ve bu sürece dair

³⁷ Eliana Angelini, Giacomo Di Tollo Ve Andrea Roli, **a.g.e.** s.741.

tüm prosedürlerini içerir³⁸. Yani bu ilke ve politikalar, bir anlamda, fona ihtiyaç duyan çevreler için, bankanın kişiliğini temsil ederler.

Konumuz açısından düşünüldüğünde ise, bankalar bundan sonra tüm kredi kararlarını, insani bir sürece tabii tutmaksızın, tamamen yapay zekâ yöntemlerine teslim edebilmektedirler. Etmeseler bile, söz konusu olan bu yapay zekâ yöntemlerini kredi politikalarının etkin bir aracı haline getirebilirler.

1.3.6. Kredilerin İzlenmesi Ve Kayıp Kredilerin Önlenmesi

Bankaların kullandırmış oldukları kredilerin etkin bir şekilde izlenebilmesi ve böylelikle kredi kayıplarının en aza indirgenmesi için aşağıdaki listede belirlenmiş unsurların takibi önem taşımaktadır:

- Bankanın, kredi müşterisinin güncel finansal durumu hakkında, ayrıntılı bilgi sahibi olması,
- Verilen tüm kredilerin yürürlükteki sözleşmelere uygunluğu,
- Onaylanmış kredilerin müşteriler tarafından kullanımının takibi,
- Büyük kredilerle ilgili olarak planlanmış nakit alışverişinin hizmet gerekliliklerini karşılayabilmesi,
- Kredi alan kişinin, güncel finansal durumuna ilişkin teminat yeterliliği,
- Problemlili kredilerin zamanında fark edilmesi ve ilgili sınıfa alınması,
- TCMB memzuçları (kredi limiti ve riski bulunan şirketlerin toplam limit risklerinin ayrıntılı olarak derlendiği bir tablo³⁹),
- Çek, icra, protesto kayıtları,
- Devre faizlerinin tahsilat durumu,
- İstihbarat raporları,
- Firma ziyaretleri,
- Firmanın mevduat ortalamaları,
- Firmanın keşide çekleri (ödemelerin düzenli yapılıp yapılmadığı),

³⁸ Birgül Şakar, **a.g.e.** s.6.

³⁹ Yılmaz Velioğlu, “Memzuç Nedir? Memzuç Kayıtlarının İçeriği ve Finans Dünyasındaki Yeri”, Çevrimiçi, <http://www.muhasabetr.com/yazarlarimiz/yilmazvelioglu/011/>, 11/03/2020.

- Sistem üzerinden şubenin firma ile yapılan işlemlerin kayıtları,
- Bilanço ve gelir tablosu gibi finansal tablolar,
- Ticari firmalardan toplanan düzenli istihbarat⁴⁰.

Yukarıda belirtilen prosedürler, yapay zekâya dair, yöntem ve uygulamalar geliştirildikçe, kredilendirme süreci beşeri bir süreç olmaktan çıkarılıp, robotize edilerek, personelin mesaisinden tasarruf edilebilir ve müşterilerin menfaati açısından da daha kısa sürede sonuç alınabilir. Ayrıca yeni oluşan kredi risklerini azaltmak adına, banka yönetimi, daha hızlı ve sistematik bir biçimde, bilinçlendirilebilir ve ilgili süreçlerin takibi anlık olarak yapılabilir.

1.4. Kredi Riskinin Belirlenmesi Konusunda İstatistik Ve Yapay Zekâ Özelinde Kullanılan Yöntemlere Dair Bir Literatür Taraması

Bilginin hızlı üretilmesi ve tüketilmesi neticesinde, birçok alanda da yaşandığı gibi, kredi pazarlarında rekabet oldukça şiddetli bir hal almıştır. Bu durum banka vb. finans kurumlarını daha agresif hareket etmeye ve dolayısıyla daha fazla risk almaya zorlamaktadır. Rekabet bu seviyede değilken, risksiz müşterilere sağlanan krediler, rekabetin ağır baskısıyla, artık çeşitli risk faktörlerinden bir ya da birkaçını taşıyan müşterilere de sunulmaya başlanmıştır. Ayrıca yüksek risk taşıyan sınıflara, yüksek faizli (getirili) krediler kullandırılmasından dolayı, bu sınıfı daha cazip gören finansal kurumlar da, bu sınıfa dönük, kredi arzı gerçekleştirmeye çalışmaktadırlar. Yaşanan bu gelişmeler, finans kurumları üzerindeki baskıyı artırırken, bir yandan daha çok kredi müşterisine ulaşmaya çalışırken, bir yandan da, kendilerini güvence altına almak adına, oluşacak kredi riski balonunu küçültmeye çalışmaktadırlar. Dolayısıyla bu durum da, kredi başvuru riski değerlendirme modellerine olan talebi giderek artırdı⁴¹. Bu konuda her geçen gün yeni modeller üretilmekte ve kullanıma sunulmaktadır. Aşağıda yer alan literatür

⁴⁰ Birgül Şakar, **a.g.e.** s.46-47.

⁴¹ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, "Credit Scoring With A Data Mining Approach Based On Support Vector Machines", **Expert Systems With Applications**, Cilt: 33, 2007, s.847.

taramasında, kredi riski belirlemeye dönük, kabul görmüş yöntemleri kullanan makalelere yer verilmiştir.

Bu literatür taramasında, Tian-Shyug Lee ve I-Fei Chen'in⁴² ve Huang ve diğerlerinin⁴³ ve Lyn Thomas'ın⁴⁴ çalışmalarında yer alan literatür özetinden de yararlanılmıştır. Ayrıca, Vellido ve diğerlerinin⁴⁵, kredi riskini belirleme konusunda, 1992 yılından 1998 yılına kadar olan ve muhtelif YSA algoritmaları kullanarak geliştirilen model çalışmaları, araştırılan konu, kullanılan algoritma, onaylayan kurum ve veri boyutu gibi bilgiler açısından sınıflandırmış oldukları, makalelerinden yararlanılmıştır.

1.4.1. Diskriminant Analizi

Diskriminant analizi (DA) ilk olarak Fisher tarafından 1936 yılında ayırma modellerinin oluşturulması için önerilmiştir ve bu konuda kullanılan ilk istatistiksel metodolojidir. Bardos⁴⁶, Desai ve diğerleri⁴⁷, Martell ve Fitts⁴⁸, Overstreet ve diğerleri⁴⁹ kredi riskini belirlemek için, model olarak, DA'ni önermişlerdir.

DA'nin geçerli olabilmesi için, ilk olarak, ayırmacı değişkenlerin dağılımları birlikte normal dağılıma uymalıdır. İkinci olarak da, ayırt edici fonksiyonun hata teriminin, normal olarak eş varyanslı (homoscedastic) dağıldığı varsayımına uymalıdır. Bu varsayımlar, ampirik tahmin sonuçlarını yorumlama ve güvenilir tahminler üretme

⁴² Tian-Shyug Lee, I-Fei Chen, "A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks And Multivariate Adaptive Regression Splines", **Expert Systems With Applications**, Cilt: 28, 2005, s.745.

⁴³ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, **a.g.e.** s.850.

⁴⁴ Lyn. C. Thomas, "A Survey Of Credit And Behavioural Scoring: Forecasting Financial Risk Of Lending To Consumers", **International Journal Of Forecasting**, Cilt: 16, Sayı: 2, 2000, s.151.

⁴⁵ A. Vellido, S.J.G. Lisboa, J. Vaughan, "Neural Networks In Business: A Survey Of Applications (1992–1998)", **Expert Systems With Applications**, Cilt: 17, 1999, s.53.

⁴⁶ Mireille Bardos, "Detecting The Risk Of Company Failure At The Banque De France", **Journal Of Banking & Finance**, Cilt: 22, Sayı: 10–11, October 1998, S.1407.

⁴⁷ Vijay S.Desai, Jonathan N. Crook B, George A. Overstreet, Jr., "A Comparison Of Neural Networks And Linear Scoring Models In The Credit Union Environment", **European Journal Of Operational Research**, Cilt: 95, 1996, s.3.

⁴⁸ TF Martell, RL Fitts, "A Quadratic Discriminant Analysis Of Bank Credit Card User Characteristics", **Journal Of Economics And Business**, Cilt: 33, 1981, s.15.

⁴⁹ G.A. Overstreet Jr., E.L. Bradley Jr. ve R.S.Kemp, "The Flat-Maximum Effect And Generic Linear Scoring Model: A Test", **IMA Journal Of Mathematics Applied In Business And Industry**, Cilt: 4, 1992, s.100.

konusunda, bazı sorunlara neden olabilmektedir. Ayrıca açıklayıcı değişkenler, ağırlıklı olarak, bilanço verileriyle sınırlı kalabilmektedir⁵⁰.

DA, Altman'ın öncülük ettiği⁵¹, çok sayıda, ayırmacı kredi puanlama analizidir. Model, “iyi firmalar” ı, “kötü firmalar” dan ayırmada istatistiksel açıklayıcı güce sahip finansal değişkenleri (açıklayıcı değişkenleri) tanımlar. Modelin parametreleri elde edildikten sonra, kredi adaylarına bir “Z puanı” verilir. Bazı durumlarda, puan temerrüt olasılığı olarak da yorumlanabilir; diğerlerinde, puan bir sınıflandırma sistemi olarak da kullanılabilir: potansiyel bir borçlu, bir puan ve bir kesme noktasına dayanarak iyi veya kötü bir gruba yerleştirilir⁵².

DA, istatistiksel olarak, dağılıma uyum ve varsayım gerektirmesi, sonuçların kullanılan yeni yöntemlere kıyasla geride kalması, gibi nedenlerle artık, diğer yöntemlere kıyasla, önemini kaybetmiş olsa da, literatür geçmişine bakıldığında bugüne dek, kredi riskini belirleme probleminde, en çok kullanılan yöntem olarak rapor edilmektedir⁵³. DA doğrusal ilişkileri gözlemlemek için geliştirilen, doğrusal bir yöntem olduğu için, girdiler arasındaki doğrusal olmayan (nonlinear) ilişkileri tespit etme konusunda, YSA gibi doğrusal olmayan içerikli yöntemlere kıyasla daha düşük performans göstermektedir. DA'nin YSA ve DVM gibi, hatta LR gibi diğer yöntemlerin henüz kullanılmadığı dönemlerde, elde edilen sonuçlar açısından, önemli bir risk değerlendirme uygulaması olmasına rağmen, doğrusal bir model olmasından kaynaklı olarak, ilişkili risk hakkında herhangi bir tahmin sağlayamadığı düşünülmektedir⁵⁴.

1.4.2. Lojistik Regresyon

Doğrusal regresyon (DR) modelleri, seçilen bir örnekte bazı rastgele değişkenleri tanımlarlar. Bu değişkenler, önemli bilgileri yansıtır ve bağımlı değişkenin Z değişkeni ile temsil edildiği doğrusal bir regresyonda, açıklayıcı değişkenler olarak

⁵⁰ Eliana Angelini, Giacomo Di Tollo, Andrea Roli, **a.g.e.** S.740.

⁵¹ Edward Altman, “Financial Ratios, Discriminant Analysis And The Prediction Of Corporate Bankruptcy”, **Journal Of Finance**, Cilt: 13, 1968, s.589.

⁵² Eliana Angelini, Giacomo Di Tollo Ve Andrea Roli, **a.g.e.** s.739.

⁵³ Lyn. C. Thomas, **a.g.e.** s.151.

⁵⁴ Tian-Shyug Lee, I-Fei Chen, **a.g.e.** s.745.

kullanılır. Bu şekilde yapılan değerlendirmede, model istatistiksel olarak anlamlı değişkenleri belirler ve regresyon katsayılarını (sistemin davranışını kontrol eden değerleri yani parametreleri), da tahmin eder. Sonuçta, bildiğimiz, o analitik regresyon formülü (bakınız denklem 2.2) oluşur.

DR yaklaşımında, yeni bir borçlunun temerrüde düşme olasılığı, sıfır ila bir (0, 1) aralığında değerler kazandığı için, “katsayılar hangi değer üzerinde olduğunda kredi müşterisi olumlu yanıt alabilecek?” gibi önemli bir seçim sorununu da beraberinde getirmektedir. Logit ve probit modeli, (0, 1) aralığında sınırlı olmayan çıktıya ilişkin doğrusal model problemi ise, LR modeli (logit) ile çözülür; üstel bir dönüşüm kullanır ve regresyon analizinin sonuçları bu aralığa dâhil edilerek çıktılar kategorize edilir. Yani müşterinin ödeme ve ödememe olasılıkları hesaplanarak müşteri, bu olasılıklardan hangisi daha yüksekse o sınıfa, atanır. Böylelikle DR kullanımı sonucu oluşacak bu karar verme problemi de, bu şekilde, çözülmüş olur⁵⁵.

LR, kategorik sonuçlarla bir açıklayıcı veri kümesinin ilişkilendirildiği, yaygın olarak bilinen ve birçok bilim sahasında; kategorize etme ve sınıflandırma problemlerinde kullanılan⁵⁶, önemli bir istatistiki yöntemdir.

Harrell ve Lee⁵⁷, çalışmalarında LR’un DA kadar etkin sonuçlar verdiğini ve DA varsayımlarının ise kullanılan modeli oldukça zora soktuğunu makalelerinde özellikle belirtmişlerdir. Yapılan çalışmalar neticesinde, LR’un, kredi müşterisinin, kredi durumunun ilişkili olasılığını tahmin edebildiği ve finansal riskin dağılımının DA’ndan daha iyi anlaşılmasını sağladığı tespit edilmiştir⁵⁸. West ise LR’un, geleneksel yöntemler arasında, en iyi sınıflandırma metodu olduğunu belirtmiştir⁵⁹.

⁵⁵ Eliana Angelini, Giacomo Di Tollo Ve Andrea Roli, **a.g.e.** S.739.

⁵⁶ Maja Pohar, Mateja Blas, And Sandra Turk, “Comparison Of Logistic Regression And Linear Discriminant Analysis: A Simulation Study”, **Metodološki Zvezki**, Cilt: 1, Sayı: 1, 2004, s.143.

⁵⁷ Frank E. Harrell, Jr. Kerry L. Lee, “A Comparison Of The Discrimination Of Discriminant Analysis And Logistic Regresyon Under Multivarite Normality”, **S.K. Se (Ed.), Biostatistics: Statistics In Biomedical, Public Health, And Environmental Sciences**, North-Holland, Amsterdam, Cilt: 1985, s.333.

⁵⁸ Tian-Shyug Lee, I-Fei Chen, **a.g.e.** s.745.

⁵⁹ David West, “Neural Network Credit Scoring Models”, **Computers & Operations Research**, Cilt: 27, 2000, s.1131.

Laitinen⁶⁰, Westgaard ve van der Wijst⁶¹, Wiginton'ın⁶² çalışmaları başta olmak üzere, LR, kredi riskinin belirlenmesinde bugüne dek sıklıkla kullanılmıştır.

1.4.3. MARS Modelleri

Çok Değişkenli Uyumlu Regresyon Uzanımları⁶³ (Multivariate Adaptive Regression Splines, MARS), ilk olarak Friedman tarafından 1991 yılında, yüksek boyutlu verilerin esnek regresyon modellemesi için yeni bir yöntem olarak sunulmuştur. En fazla birkaç değişkene bağlı olan nonlinear ve nonparametrik modelin, ilişkilerinde regresyona kıyasla daha fazla güç ve esnekliğe sahip olduğu, ayrıca veri kümesindeki ek katkıları ayrı ayrı saptadığı ve onları farklı çok değişkenli etkileşimlerle ilişkilendiren bir formda temsil ettiği, yöntemin geliştiricisi Friedman tarafından belirtilmektedir⁶⁴.

Optimal MARS modeli iki aşamalı bir süreçle uygulanır. İlk aşamada, MARS, başlangıçta verileri örtmek için çok sayıda temel işlev oluşturur; burada değişkenlerin sürekli, kategorik veya sıralı olarak girmesine izin verilir. İkinci aşamada, Genelleştirilmiş Çapraz Doğrulama (Generalized Cross-Validation, GCV) kriteri kullanılarak, en az katkı sırasına göre, temel fonksiyonlar silinir. Modelden bir değişken çıkarıldığında hesaplanan GCV'deki düşüş gözlemlenerek değişken öneminin bir ölçüsü daha sonra değerlendirilebilir. Bu süreç, kalan temel işlevlerin tümü önceden belirlenmiş gereksinimleri karşılayana kadar devam eder⁶⁵.

⁶⁰ Erkki K. Laitinen, Teija Laitinen, "Bankruptcy Prediction Application Of The Taylor's Expansion In Logistic Regression", **International Review Of Financial Analysis**, Cilt: 9, 2000, s.327.

⁶¹ Sjur Westgaard, Nico Van Der Wijst, "Default Probabilities In A Corporate Bank Portfolio: A Logistic Model Approach", **European Journal Of Operational Research**, Cilt: 135, 2001, S.339.

⁶² John C. Wiginton, "A Note On The Comparison Of Logit And Discriminant Models Of Consumer Credit Behavior", Cilt: 15, Sayı: 3, 1980, s.757.

⁶³ K. Batu Tunay, "Türkiye'de Paranın Gelir Dolaşım Hızlarının MARS Yöntemiyle Tahmini", **ODTÜ Gelişme Dergisi**, Cilt: 28, Sayı: 3-4, 2001, s.176.

⁶⁴ Jerome H. Friedman, "Multivariate Adaptive Regression Splines", **Annals Of Statistics**, Cilt: 19, Sayı: 1, 1991, s.1.

⁶⁵ Chun-Ling Chuang, Rong-Ho Lin, "Constructing a reassigning credit scoring model", **Expert Systems with Applications**, Cilt: 36, Sayı: 2, 2009, ss. 1689.

MARS modelleri, tahmin ve sınıflandırma alanlarındaki problemlerin modellenmesinde De Gooijer ve diğerleri⁶⁶, Friedman ve Roosen⁶⁷, Kuhnert ve diğerleri⁶⁸, Griffin ve diğerleri⁶⁹, Nguyeng-Cong ve diğerleri⁷⁰, Ohmann ve diğerleri⁷¹ vb. birçok araştırmacı tarafından, muhtelif disiplinlerde kullanılmıştır⁷².

Lee ve Chen, YSA ve MARS ile iki aşamalı hibrit modelleme prosedürü kullanarak kredi puanlamasının performansını araştırmış ve sonuçta, oluşturulan hibrit yaklaşım, Doğrusal Diskriminant Analizi (Linear Discriminant Analysis, LDA), LR, YSA ve MARS kullanarak elde edilen sonuçlardan daha iyi sonuçlar üretmiştir. Analizlerin arka planındaki yaklaşım, ilk olarak kredi puanlama modelinin oluşturulmasında MARS kullanılması, ikinci aşamada ise elde edilen önemli değişkenler daha sonra YSA modeline giriş düğümleri olarak girilmesidir. Sonuçların iyi olması neticesinde geliştirilen yeni yaklaşımın, kredi puanlama görevlerini yerine getirmede, iyi bir alternatif olacağı şeklindedir⁷³.

Lee ve diğerleri, iki veri madenciği yöntemi olan Sınıflandırma Ve Regresyon Ağacı (Classification And Regression Trees, CART) ve MARS kullanarak bir kredi puanlama modeli oluşturmuşlardır. Sonuçlarının kıyaslayabilmek için ise modelleri, bir bankadan edinmiş oldukları kredi kartı verileri üzerinde denemiş ve bu iki modelin kredi puanlama doğruluğu açısından geleneksel DA, LR, YSA ve DVM'den daha iyi

⁶⁶ Jan G. De Gooijer, U., Bonnie K. Ray, Horst Krager, "Forecasting Exchange Rates Using TSMARS", **Journal of International Money and Finance**, Cilt: 17, 1998, s.513.

⁶⁷ J.H. Friedman, C.B. Roosen, "An Introduction To Multivariate Adaptive Regression Splines" **Statistical Methods in Medical Research**, Cilt: 4, 1995, s.197.

⁶⁸ Petra M. Kuhnert, Kim-Anh Do, Rod McClure, "Combining Non-Parametric Models With Logistic Regression: An Application To Motor Vehicle Injury Data", **Computational Statistics & Data Analysis**, Cilt: 34, 2000, s.371.

⁶⁹ W.L. Griffin, N.I. Fisher, J.H. Friedman, C.G. Ryan, "Statistical Techniques For The Classification Of Chromites In Diamond Exploration Samples", **Journal of Geochemical Exploration**, Cilt: 59, 1997, s.233.

⁷⁰ V. Nguyen-Cong, G. Van Dang, B. M. Rode, "Using Multivariate Adaptive Regression Splines To QSAR Studies Of Dihydroartemisinin Derivatives", **Eur J Med Chem**, Cilt: 3, Sayı 1, 1996, s.797.

⁷¹ C. Ohmann, V. Moustakis, Q. Yang, K. Lang, "Evaluation Of Automatic Knowledge Acquisition Techniques In The Diagnosis Of Acute Abdominal Pain", **Artificial Intelligence in Medicine**, Cilt: 8, 1996, s.23.

⁷² Tian-Shyug Lee, I-Fei Chen, **a.g.e.** s.746.

⁷³ Tian Shyug Lee, I-Fei Chen, "A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks And Multivariate Adaptive Regression Splines", **Expert Systems With Applications: An International Journal**, Cilt: 28, Sayı: 4, 2005, s.589.

performans gösterdiğini, bu nedenle söz konusu yöntemlerin, kredi puanlamada, etkili alternatifler olduğunu belirtmişlerdir⁷⁴.

Chuang ve Lin, kredi puanlama probleminin, kredi piyasalarındaki aşırı rekabet sonucu, çok önemli bir hale geldiğini ve bu yüzden son yıllardaki tüm yöntemlerin bu konuda önemli bir sınav verdiklerini ifade etmişlerdir. Chuang ve Lin, kredi sınıflandırma problemini çözmek ve kredi puanlama modelinin Tip I hatasını azaltmak için, iki aşamalı MARS ve YSA tabanlı bir Yeniden Atama Kredi Puanlama Modeli (Reassigning Credit Scoring Model, RCSM) sunmuşlardır. Bu atama aşamasını oluşturmaktadır. Yeniden atama aşamasında ise, Vaka Bazlı Muhakeme (Case Based Reasoning, CBR) tabanlı sınıflandırma tekniğini kullanarak reddedilen iyi kredi başvuru sahiplerini koşullu sınıfa yeniden atayarak Tip I hatasını azaltmaya çalışmışlardır. Önerdikleri modelin etkinliğini göstermek için ise, RCSM modelini, makine öğrenimi veritabanlarından elde edilen, bir akademik veri tabanından alınan, bir kredi kartı veri kümesi üzerinde denemişlerdir. Sonuçların ise, yaygın olan diğer dört yaklaşımdan, daha doğru kredi puanlaması gerçekleştirdiğini ve aynı zamanda modelin puanlama sisteminin Tip I ve Tip II hatasını azaltarak iş gelirini de artırmaya yardımcı olacağı fikrini savunmuşlardır⁷⁵.

1.4.4. Yapay Sinir Ağları

YSA, evrensel (global) yaklaşım özelliği itibarıyla, ilgiyi üzerine çekmektedir. Bununla birlikte, sinir ağlarının karar vermede kullanımıyla ilgili önemli bir dezavantajı bulunmaktadır. Bu da girdiler ile çıktılar arasındaki bağlamın, regresyonda olduğu gibi analitik bir şekilde, açıklanamamasıdır. Tahminlerinde oldukça yüksek bir doğruluk oranına rağmen, YSA'nın sonuç çıktılarını nasıl oluşturduklarının anlaşılması pek mümkün görünmemektedir. Bu durumun sonuçlara yansımaları, bir örnekle açıklanabilir: kredi talebi kabul edilen ya da reddedilen bir müşterinin talebinin niçin

⁷⁴Tian-Shyug Lee, Chih-Chou Chiu, Yu-Chao Chou ve Chi-Jie Lu, "Mining The Customer Credit Using Classification And Regression Tree And Multivariate Adaptive Regression Splines", **Computational Statistics & Data Analysis**, Cilt: 50, Sayı: 4, 2006, ss. 1113-1130.

⁷⁵ Chun-Ling Chuang, Rong-Ho Lin, "Constructing a reassigning credit scoring model", **Expert Systems with Applications**, Cilt: 36, Sayı: 2, 2009, ss. 1685.

kabul edildiği ya da reddedildiğini açıklamak mümkün değildir. Sistem bir çıktı veriyor kabul ya da red şeklinde, fakat bunun sebebi olarak gösterilebilecek bir açıklayıcı değişken öne sürmek mümkün değildir. Dolayısıyla müşteri ya da üst yönetim bu çıktı karşısında “niçin kabul/red edildi?” şeklinde sorusunun somut bir cevabı modele bakılarak verilememektedir. Bu da, şeffaflık ve hesap verebilirlik ilkeleri açısından, olumsuz bir durum olarak kabul edilmektedir⁷⁶.

Geleneksel istatistiki yöntemlerin geçerliliği ve etkinliği, büyük ölçüde; doğrusallık, normallik, açıklayıcı değişkenlerin birbirinden bağımsız olması ve ölçüt değişkeni ve açıklayıcı değişkenlerle ilgili önceden var olan fonksiyonel form gibi, bazı kısıtlayıcı varsayımlara dayanmaktadır. Bu geleneksel yöntemler, en iyi sonucu, ancak istatistiksel varsayımların tamamı ya da önemli bir çoğunluğu mümkün olduğunda iyi sonuçlar verebilmektedir⁷⁷.

YSA, finans alanında, ilk olarak 1990 yılında, iflas tahmini problemlerinin çözümü için kullanılmaya başlandı ve bugün finansal süreçlerin belki de tamamında aktif olarak kullanılmaktadır. YSA gibi doğrusal olmayan modellerin doğrusal olanlardan, genel olarak daha üstün olmasının birkaç nedeni bulunmaktadır. İlk olarak finansal oranlar ile temerrüt kazançlar/toplam varlıklar 0,1'de iken 0,2 miktarında değişiyorsa, bu oran 1,0'dan 1,2'ye değiştiğinden çok daha büyük bir etkiye sahip olacaktır. Ayrıca çarpımsal faktörlerin de olduğu iddia edilebilir. Örneğin negatif nakit akışı olan bir firma, büyük yükümlülükleri varsa, kredi riski artacaktır. Bunun nedeni yüksek kaldıraçlı firmaların açıklarını finanse etmek için borç almakta, diğer firmalara kıyasla, çok daha zorlanmalarındır. Sonuç olarak doğrusal olmayan hesaplama içeriği ile YSA, mevcuttaki diğer doğrusal yöntemlerden daha iyi performans göstermektedir. Şu anda, büyük ve ticari kredi riski belirleme ürünlerinden bazıları sinir ağlarına dayanmaktadır. Örneğin, Moody'nin Kamu Şirketi Risk Modeli, temel teknoloji

⁷⁶ Bart Baesens, Rudy Setiono, Christophe Mues, Jan Vanthienen, “Using Neural Network Rule Extraction And Decision Tables For Credit-Risk Evaluation”, **Management Science**, Cilt: 49, Sayı: 3, 2003, s.312.

⁷⁷ Guoqiang Zhang, Michael Hu, B. Eddy Patuwo, Daniel C. Indro, “Artificial Neural Networks In Bankruptcy Prediction: General Framework And Cross-Validation Analysis,” **European J. Operational Research**, Cilt: 116, 1999, s.16-32.

olarak YSA'na dayanmaktadır. Birçok banka da kendine ait, tescilli sinir ağı tahmin modelini oluşturmuştur ve hali hazırda aktif olarak kullanmaktadır⁷⁸.

YSA, geleneksel yöntemlere kıyasla, herhangi bir varsayım gerektirmemesinden, yüksek genelleme kabiliyetine sahip olmasından ve özellikle durağan olmayan süreçlerin modellenmesinde etkin bir yöntem olmasından dolayı finansal iş çevrelerinde kullanımı giderek artmaktadır⁷⁹. Angelini ve diğerleri, Atiya, Pang ve diğerleri, Rong, Wu vb. birçok araştırmacı bir kredi müşterisinin temerrüde düşme olasılığını hesaplamada YSA'nın çok başarılı bir yöntem olduğunu belirtmişlerdir. Ayrıca, Barney, Glorfeld, Piramuthu, Huang, vb. birçok araştırmacı, YSA'nın, kredi risk modellemesinde diğer modellerden daha doğru, uyarlanabilir ve güçlü bir metod olduğunu belirtmişlerdir⁸⁰.

Angelini ve diğerleri, biri standart ve biri ileri beslemeli, diğeri ise özel amaçlı bir mimariye sahip iki farklı sinirsel ağ sistemi geliştirmiş ve bu mimarileri İtalya'daki küçük işletmeler üzerinde uygulamışlardır. Sonuç olarak da sinir ağlarının, dikkatli veri analizinin, veri ön işleminin (verilerin alfa-nümerik hale getirilmesi) ve eğitiminin yapılması koşuluyla, bir borçlunun temerrüt eğilimini öğrenme ve tahmin etmede, çok başarılı bir yöntem olduğu sonucuna varmışlardır⁸¹.

Atiya, Merton tarafından geliştirilen geleneksel kredi riski modellerinden esinlenerek, sinir ağları sistemi için yeni göstergeler önermiş olduğu çalışmasında, geleneksel finansal oran göstergelerine ek olarak bu göstergelerin kullanılmasının tahmin doğruluğunda önemli bir iyileşme sağladığı (üç yıllık bir tahmin için, % 81,46'dan % 85,5'e) sonucuna varmıştır⁸².

⁷⁸ Amir F. Atiya, "Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results", **IEEE Transactions On Neural Networks**, Cilt: 12, Sayı: 4, 2001, s.930.

⁷⁹ Bing Cheng, D. M. Titterington, "Neural Networks: A Review From A Statistical Perspective", **Statistical Science**, Cilt: 9, Sayı: 1, 1994, s.26.

⁸⁰ David West, **a.g.e.** s.1131.

⁸¹ Eliana Angelini, Giacomo di Tollo, Andrea Roli, **a.g.e.** s.733.

⁸² Amir F. Atiya, **a.g.e.** s.929.

Rong ve diğerleri, 120 müşteri verisini iyi, orta ve kötü şeklinde üç gruba bölerek yaptıkları çalışmada kullandıkları YSA modeliyle, geleneksel olan LDA yönteminden çok daha yüksek bir doğruluk oranına ulaşmışlardır⁸³.

Wu ve diğerleri, YSA'nın geleneksel yöntemlere göre başarılı bir yöntem olduğu ve parametrelerini genetik algoritma ile belirledikleri bir DVM algoritmasının performansını YSA ile test etmişler ve geliştirdikleri modelin en iyi doğruluk oranını veren model olduğu sonucuna varmışlardır⁸⁴.

Barney ve diğerleri, çiftçi müşteriler üzerinde yaptıkları bir kredi risk belirleme çalışmasında, En Küçük Kareler (EKK) kullanan çok değişkenli istatistiksel bir model ve LR, GA ve YSA modellerini kullanmışlar ve modellerin tamamının, kredi veren kurumun geliştirdikleri sistematikten daha başarılı olduğu ve YSA modelinin diğerlerinden daha başarılı olduğu sonucuna varmışlardır. Buna ek olarak, Barney ve diğerleri, YSA'nın, hatalar arasındaki kesme noktasını yayması ve Tip I ve Tip II hataları arasındaki dengede iyileşmeye izin vermesi nedeniyle, kamu politikası açısından, diğer çok değişkenli modellerden daha üstün olduğunu belirtmişlerdir⁸⁵.

Desai ve diğerleri, küçük boyutlu müşteri verileri üzerinde Çok Katmanlı Algılayıcı (ÇKA), (Multi Layer Perceptron, MLP) ve modüler YSA gibi, YSA ve LDA ve LR benzeri geleneksel tekniklerin kredi puanlama modeli oluşturma yeteneklerini araştırmıştır. Çalışmada ayrıca genel (generic) modellerle, özelleştirilmiş (customized) modellerin bir sonuç karşılaştırması yapılmıştır. Sonuçlar ise iki farklı şekilde yorumlanmıştır. Birincisi eğer, performansın ölçüsü doğru sınıflandırılan kötü kredilerin yüzdesi ise özelleştirilmiş sinir ağlarının çok umut verici olduğu sonucuna, ikincisi ise eğer performansın ölçüsü doğru sınıflandırılan iyi ve kötü kredilerin yüzdesi ise LR modelinin sinir ağları gibi bir sonuç verebildiği sonucuna ulaşmışlardır.

⁸³ L. Rong-Zhou, S.Su-Lin, X. Jian-Min, "Neural network credit-risk evaluation model based on back-propagation algorithm", **In Proceedings of the first international conference on machine learning and cybernetics**, 2002, s.1702.

⁸⁴ C.Wu, X. M. Wang, "A Neural Network Approach For Analyzing Small Business Lending Decisions", **Review Of Quantitative Finance And Accounting**, Cilt: 15, 2000, s.397.

⁸⁵ Douglas K. Barney, O. Finley Graves, John D. Johnson, "The Farmers Home Administration And Farm Debt Failure Prediction", **Journal of Accounting and Public Policy**, Cilt: 18, 1999, s.99.

Özellikle kötü krediler söz konusu olduğunda genel modellerin özelleştirilmişler kadar iyi olmadığı sonucuna varılan çalışmada, iyi sonuçlar için daha iyi mimarilerin oluşturulması gerektiği gibi, bulgularını da mevcut YSA modellerinin yetersiz olduğu şeklinde yorumlamışlardır. Çünkü oluşturdıkları özelleştirilmiş modüler sinir ağının üç farklı kredi verisi arasındaki farklılıkları ayırt edemediğini belirtmişlerdir⁸⁶.

Glorfeld, öncelikle 17 açıklayıcı değişkenle, %88 doğruluk oranı elde etmiş, fakat değişkenlerini “değişken seçim prosedürü” ile yediye (7) düşürünce %94,1 doğruluk oranına ulaşmıştır. Glorfeld, çalışmasının sonucunda, sinir ağlarının kara kutular olarak görülmesinin gereksiz bir yaklaşım olduğunu ve burada sinir ağlarının başarısını, geliştirdiği değişken seçim prosedürüne bağlamış ve bu prosedürün başarılı olması halinde, iyi sonuçlar veren sinir ağı mimarilerinin geliştirilmesinin kolay olduğunu belirtmiştir. Ayrıca Glorfeld çalışmasında, LDA, LR ve CART yöntemlerini sinir ağlarının karşılaştırmasında kullanmıştır⁸⁷.

Piramithu, üç farklı gerçek dünya verisi üzerinde, nörofuzzy sistemlerle karşılaştırdığı sinir ağları modelinin, nörofuzzy sistemlerden daha iyi sonuç verdiği sonucuna varmıştır⁸⁸.

Huang ve diğerleri, yapay zekâ temelli yaklaşımların artık geleneksel istatistiki yöntemlerden daha iyi sonuçlar verdiğini belirtmişlerdir. ABD ve Tayvan pazarlarında yaptıkları çalışmada referans olarak geri yayılım ağını kullanmışlar ve sonuçta hem YSA hem de DVM ile %80 doğruluk oranına ulaşmışlardır. Ayrıca çalışmada yapay zekâ tabanlı modellerin yorumlanabilirliğini geliştirmek adına YSA modeli yorumlamasında, son araştırma sonuçlarını uygulayarak, YSA modellerinin girdi finansal değişkenlerinin göreceli önemini kullanmışlardır⁸⁹.

⁸⁶ Vijay S.Desai, Jonathan N. Crook B, George A. Overstreet, Jr., **a.g.e.** s.24.

⁸⁷ Louis W. Glorfeld, : “A Methodology for Simplification and Interpretation of Backpropagation-Based Neural Network Models”, **Expert Systems With Applications**, Cilt: 10, Sayı: 1, 1996, s.51.

⁸⁸ Selwyn Piramuthu, “Financial Credit-Risk Evaluation With Neural And Neurofuzzy Systems”, **European Journal of Operational Research**, Cilt: 112, 1999, s.320.

⁸⁹ Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun-Hwa Chen, Soushan Wu, “Credit Rating Analysis With Support Vector Machines And Neural Networks: A Market Comparative Study”, **Decision Support Systems**, Cilt: 37, 2004, s.543.

Vellido ve diğerleri, kredi risk değerlendirme konusunda, 1992 yılından 1998 yılına kadar olan ve muhtelif YSA algoritmaları kullanılarak model geliştirilen çalışmaları, araştırılan konu, kullanılan algoritma, onaylayan kurum ve veri boyutu gibi bilgiler açısından sınıflandırmışlardır. Bu çalışmaya göre Jensen⁹⁰, Richeson, Zimmermann, ve Barnett⁹¹, Desai ve diğerleri⁹², Torsun⁹³, Arminge, Enache ve Bonne⁹⁴, Williamson⁹⁵, bireysel kredi değerlendirme probleminde geriye yayılmalı gradyan iniş (back propagation gradient descent, BPGD) yöntemini, kullanmışlardır. Bunlardan geriye yayılmalı gradyan iniş metodunun sonuçlarını, LDA ile kıyaslamışlardır. Torsun, yöntemin sonuçlarını Gaussian Maksimum Olabilirlik Sınıflandırıcısı (Gaussian maximum likelihood classifier, GMLC) ve Bulanık Sınıflandırıcı (Fuzzy Classifier FuzC) ile mukayese etmiştir. Arminge, Enache ve Bonne ise ilgili yöntemin sonuçlarını LDA ve CART yönteminin sonuçlarıyla kıyaslamıştır⁹⁶.

Kurumsal kredi müşteri verileri üzerinde, Deng⁹⁷, Glorfeld ve Hardgrave, geriye yayılmalı gradyan iniş yöntemini kullanmıştır. Bunlardan Piramuthu, Shaw ve Gentry yöntemin sonuçlarını kıyaslamak için probit regresyon, endüktif dikotomizer (ID3) ve Quinlan'ın indüksiyonunun geliştirilmesi (NEWQ) yöntemlerini kullanmıştır. Glorfeld ise LDA, LR ve CART yöntemlerini karşılaştırmasında kullanmıştır⁹⁸.

Glorfeld and Hardgrave, YSA'nın, özellikle tahmin ve sınıflandırma problemleri için, diskriminant ve regresyon analizi gibi geleneksel istatistiksel tekniklere iyi bir

⁹⁰ H. L. Jensen, „Using Neural Networks For Credit Scoring. **Managerial Finance**, Cilt: 18, Sayı 6, s.15-26.

⁹¹ Richeson, L., Zimmermann, R. A., Barnett, K. G.. “Predicting Consumer Credit Performance: Can Neural Networks Outperform Traditional Statistical Methods?” **International Journal Of Applied Expert Systems**, Cilt: 2, Sayı 2, 1994, s.116-130.

⁹² V. S.Desai, J. N., Crook, G. A. Overstreet Jr., “A Comparison Of Neural Networks And Linear Scoring Models In The Credit Union Environment” **European Journal Of Operational Research**, Cilt: 95, 1996, s.24-37.

⁹³ I. S.Torsun, “A Neural Network For A Loan Application Scoring System” **The New Review Of Applied Expert Systems**, Cilt: 2, 1996, S.47-62.

⁹⁴ G. Arminge, D. Enache, T. Bonne, “Analyzing Credit Risk Data: A Comparison Of Logistic Discrimination Classification Tree Analysis And Feedforward Networks”, **Computational Statistics**, Cilt: 12, 1997, S.293-310.

⁹⁵ A. G. Williamson, “Refining A Neural Network Credit Application Vetting System With A Genetic Algorithm” **Journal Of Microcomputer Applications**, Cilt: 18, 1995, S.261-277.

⁹⁶ A. Vellido, S.J.G. Lisboa, J. Vaughan, **a.g.e.** s.53

⁹⁷ Deng, P-S.. “Automatic Knowledge Acquisition And Refinement For Decision Support: A Connectionist Inductive Inference Model”, **Decision Sciences**, Cilt: 24, Sayı 2, 1993, s.371-393.

⁹⁸ A. Vellido, S.J.G. Lisboa, J. Vaughan, **a.g.e.** s.53.

alternatif olduğunu kanıtlamış olduğunu belirtmektedir. Ayrıca Glorfeld ve Hardgrave, yapılan çalışmalarda kullanılan sinir ağlarında mimarilerin geleneksel yöntemlerle seçildiğini, bunun da yöntemin performansını azalttığını belirtmişlerdir. İstatistiksel tekniklerin, kredi riski belirlemede yetersiz kaldığını ve sinir ağı mimarilerinin, doğru oluşturulması halinde, önünün açık olduğunu belirtmişlerdir⁹⁹.

Piramuthu, Shaw ve Gentry, karar destek sistemlerinde karar kuralları oluşturmak için tümevarımsal öğrenme yöntemlerini kullanmışlardır. AQ-Star yöntemi, Ağaç indüksiyon yöntemi ve olasılıksal öğrenme yöntemi yöntemleri ile YSA ile algoritmik özellikleri ve karar destek performansları karşılaştırmışlar ve bunun içinde kredi değerlendirme örnekleri ve ampirik veriler kullanmışlardır¹⁰⁰.

Baesens ve diğerleri, gerçek hayattaki üç farklı kredi riski veri kümesini analiz etmişler, YSA kararlarını, ağlara yerleştirilmiş, öğrenilmiş bilgileri yakalayan açıklayıcı kurallarla açıklığa kavuşturmaya çalışarak, kredi riski yöneticisinin belirli bir kredi müşterisinin neden kötü ya da iyi olarak sınıflandırıldığını açıklamaya çalışmışlardır. Sonuç olarak, YSA'nın, kredi riski değerlemesi için gelişmiş ve kullanıcı dostu karar destek sistemleri oluşturulmasını sağlayan, güçlü birer yönetim araçları olduğu sonucuna varmışlardır¹⁰¹.

West, "YSA modellerinin kredi puanlama doğruluğunu araştırdığı çalışmasında beş farklı ÇKA, uzman karışımı, Radyal Tabanlı Fonksiyon (Radial Based Function, RBF), Doğrusal Vektör Niceleme (LVN), (Linear Vector Quantization, LVQ) öğrenme ve bulanık uyarlamalı rezonans şeklindeki beş farklı sinir ağını kullanmıştır. Sonuçlarını ise, LDA, LR, K en yakın komşu, çekirdek yoğunluğu tahmini ve Karar Ağaçları (Decision Trees) dâhil olmak üzere ticari uygulamalar için daha geleneksel yöntemlerle karşılaştırmıştır. Sonuçlara bakıldığında; ÇKA'nın en doğru sonucu vermeyebileceği, YSA modeli ve kredi puanlama uygulamaları için hem uzmanların

⁹⁹ L. W. Glorfeld, B. C. Hardgrave, "An Improved Method For Developing Neural Networks: The Case Of Evaluating Commercial Loan Creditworthiness", **Computers & Operations Research**, Cilt: 23, Sayı 10, 1996, s.933.

¹⁰⁰ Selwyn Piramuthu Michael J. Shaw, James A. Gentry, "Inductive Learning Methods for Knowledge-Based Decision Support: A Comparative Analysis", **Computer Science in Economics and Management**, Cilt: 3, 1990, s.147–165.

¹⁰¹ Bart Baesens, Rudy Setiono, Christophe Mues, Jan Vanthienen, **a.g.e.** 312.

karışımı hem de radyal temel işlevine sahip YSA modellerinin göz önünde bulundurulması gerektiği ve LR'un geleneksel yöntemler içinde en doğru sonucu verdiğini belirtmiştir¹⁰².

Basu ve diğerleri, YSA ile ilgili çalışmaları kronolojik olarak incelediklerinde, her bir YSA uygulamasının bazı avantajları ve dezavantajları olmasının yanısıra, her örüntü tanıma probleminde kullanılan YSA uygulamasının daha iyi sonuç verdiği sonucuna varmışlardır¹⁰³.

YSA'nın gelişim sürecinde öne çıka yayınlar yukarıda sıralanmış olup, YSA literatürde, kredi risk belirleme konusunda alternatif bir metot olarak kabul gördükten sonra, modelin bu konuda kullanımına yönelik yapılan araştırmalar sayılamayacak kadar çoktur.

1.4.5. Destek Vektör Makineleri

DVM, şu anda sınıflandırmada aktif olarak kullanılan önemli bir araştırma ve uygulama alanıdır. DVM, birçok disiplinde, sınıflandırma problemlerini başarıyla çözebilmektedir¹⁰⁴. Bu niteliği sayesinde DVM, sınıflandırma konusunda, birçok geleneksel yöntem ve yapay zekâ özelinde ise YSA'na alternatif olarak ön plana çıkmaktadır.

Huang ve diğerleri, DVM'ni yeni bir kredi riski belirleme yöntemi olarak incelerken bir kıyaslama olarak geri yayılım sinir ağını, ABD ve Tayvan pazarları için kullanmışlardır. Hem geri yayılım sinir ağı için hem de DVM yöntemleri için yaklaşık % 80 tahmin doğruluğu elde ettikleri çalışmada DVM'in, YSA'ya göre, çok hafif bir ilerleme gösterdiği bulgusuna ulaşmışlardır. Araştırmanın diğer bir yönünün, "yapay zekâ tabanlı modellerin yorumlanabilirliğini geliştirmek" olduğu belirtilen çalışmada, YSA modellerinin yorumlamasında, son araştırma sonuçlarını uygulayarak ve YSA modellerinden girdi finansal değişkenlerinin göreceli önemini alarak bunu

¹⁰² David West, **a.g.e.** s.1131.

¹⁰³ Jayanta Kumar Basu, Debnath Bhattacharyya, Tai-hoon Kim, "Use of Artificial Neural Network in Pattern Recognition", **International Journal of Software Engineering and Its Applications**, Cilt: 4, Sayı: 2, 2010, s.32.

¹⁰⁴ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, **a.g.e.** s.847.

gerçekleştirmeye çalışmışlar fakat bunu gerçekleştirebilmenin daha derin pazar analiziye mümkün olabileceğini savunmuşlardır¹⁰⁵.

Baesens ve diğerleri, çalışmalarında, bazı önemli veri setleri büyük Benelüks ve İngiltere finans kuruluşlarından gelen sekiz gerçek hayat kredi puanlama verisine, LR, LDA, K-en yakın komşu, YSA ve karar ağaçları gibi yöntemleri uygulamışlardır. O yıllarda bu bilinen yöntemlere ek olarak yeni yeni kullanılmaya başlanan DVM ve En Küçük Kareler Destek Vektör Makineleri (EKK-DVM, Least Squares Support Vector Machine, LS-SVM) gibi gelişmiş çekirdek tabanlı sınıflandırma algoritmasını da çalışmalarında kullanmışlardır. Çalışmanın sonuç performansları ise, sınıflandırma doğruluğu ve alıcının çalışma karakteristik eğrisinin altındaki alan kullanılarak değerlendirilmiştir. Performans farklılıklarının anlamı, uygun test istatistikleri kullanılarak, istatistiksel olarak yorumlanan çalışmada, hem EKK-DVM hem de YSA sınıflandırıcılarının çok iyi bir performans gösterdiği, ancak LR ve LDA gibi basit sınıflandırıcıların da kredi puanlaması için çok iyi performans gösterdiği sonucuna varmışlardır¹⁰⁶.

Huang ve diğerleri, iki farklı veri kümesi üzerinde, geliştirdikleri üç farklı DVM modelini uygulamışlar ve bu modelleri YSA, Genetik Algoritma (GA) ve karar ağacı sınıflandırıcıları ile karşılaştırmışlardır. Sonuç olarak, kurdukları DVM modelinin, az girdi ile özdeş ve en iyi sonucu veren model olduğu sonucuna varmışlardır. Çalışmada ek olarak, genetik algoritmayı DVM sınıflandırıcısı ile birleştiren, kendi önerdikleri hibrit GA-DVM modelinin, eşzamanlı olarak özellik seçme görevi ve model parametreleri optimizasyonunu gerçekleştirebildiğini ve deneysel sonuçlarının, DVM'nin mevcut veri madenciliği yöntemlerine umut verici bir yöntem olduğunu gösterdiği sonucuna varmışlardır¹⁰⁷.

¹⁰⁵ Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun-Hwa Chen, Soushan Wu, **a.g.e.** s.544-556.

¹⁰⁶ B Baesens, T. Van Gestel, S.Viaene1, M. Stepanova, J. Suykens, J. Vanthienen, "Benchmarking state-of-the-art classification algorithms for credit scoring", **Journal of the Operational Research Society**, Cilt: 54, 2003, s.627-635.

¹⁰⁷ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, **a.g.e.** s.847.

1.4.6. Genetik Algoritma

Bir arama metodu olan GA¹⁰⁸, kredi risk belirlemede çok yaygın bir yöntem olmamakla birlikte yinede bu konuda gerçekleştirilen bazı önemli uygulamalar mevcuttur. Ong ve diğerleri¹⁰⁹, kredi puanlama modellerinin doğruluğunu önemli ölçüde iyileştirmek için, daha karmaşık bir model önerilmesi gerektiğini belirtmişlerdir. GA kullandıkları bu çalışmada Ong vd. hata oranını, iki farklı sayısal örnek üzerinde YSA, karar ağaçları, kaba kümeler ve LR dâhil olmak üzere diğer kredi puanlama modelleri ile karşılaştırmışlar ve sonuç olarak da GA'nın diğer modellerden daha iyi performans sağlayabileceği sonucuna varmışlardır.

Chen ve Huang, iki ilginç kredi analizi problemi için YSA ve GA modellerini kullanmışlar ve YSA'nın kredi skorlama için oldukça uygun bir model olduğu kanaatine varmışlardır. İlk olarak YSA ile kredileri kabul ve ret şeklinde sınıflandırdıktan sonra reddedilen kredileri daha iyi anlamak ve GA tabanlı ters sınıflandırma tekniğini kullanarak, bunları tercih edilen kabul edilebilir sınıfa yeniden atama yapmışlardır. Sonuç olarak da Chen ve Huang, GA tabanlı ters sınıflandırmayı kullanarak; alacaklılar, sözkonusu kredi müşterilerine şartlı kabul önerebilir ve reddettikleri başvuru sahiplerine kabul koşullarını açıklayabilirler ve reddedilen kredi müşterileri ise kabul alabilmeleri için, şartlarındaki minimum değişiklik seçeneklerini öğrenebilecekleri faydalı bir model önerisinde bulunmuşlardır¹¹⁰.

Fogarty ve Ireson ise, GA ve Bayesian sınıflandırıcı ile geliştirdikleri hibrit modelin (IDIOMS) karar ağaçları ve Bayesian sınıflandırma ve K en yakın komşu vb. diğer modellerden daha iyi sonuç verdiği sonucuna varmışlardır. Bu uygulamayı, 51023 adet müşteri verisi üzerinde yapan Fogarty ve Ireson, zaten % 94 doğru sonuca ulaşılmış verilerde hibrid modelin bu zorlu görevden başarıyla çıktığını belirtmiş ve sonuç

¹⁰⁸ Ergün Eroğlu, Timur Kesintürk : “Warehouse Location Problem With Genetic Algorithm”, **35 th International Conference On Computers And Industrial Engineering**, 2005, s.656.

¹⁰⁹ Chorng-Shyong Ong, Jih-Jeng Huang, Gwo-Hshiung Tzengb, “Building Credit Scoring Models Using Genetic Programming”, **Expert Systems with Applications**, Cilt: 29, 2005, s.41.

¹¹⁰ Mu-Chen Chen, Shih-Hsien Huang, “Credit Scoring And Rejected Instances Reassigning Through Evolutionary Computation Techniques”, **Expert Systems with Applications** Cilt: 24, 2003, s.435.

olarak, “bir sınıflandırıcı yalnızca onu eğitmek için kullanılan veriler kadar iyi olabilir” görüşünü vurgulamışlardır¹¹¹.

1.4.7. Literatür Özeti

Literatürde, kredi riski belirleme konusunda kullanılan yöntemler elbette yukarıda bahsedilenlerle sınırlı değildir. Bunların dışında, Henley ve Hand¹¹², K-en yakın komşu modelini kredi riskini belirleme probleminde kullandılar. Ayrıca sınıflandırma ağaçları veya Özyinelemeli Bölümleme Algoritmaları (Classification trees or recursive partitioning algorithms, RPA) istatistikte Breiman ve diğerleri, yapay zekâda Safavian ve Landgrebe¹¹³, makine öğrenmesi alanında ise Quinlan¹¹⁴, tarafından geliştirilen model sonrasında ilk kez Makowski¹¹⁵ tarafından kredi skortlama probleminde uyarlanmıştır. Daha sonra Coffman¹¹⁶, karar ağaçlarını DA ile karşılaştırdı. Mehta(1968), carlet (1987) ve Boyle (1992) karar ağaçlarının kredi skortlama kullanımını tartışan diğer araştırmacılarıdır¹¹⁷. Davis ve diğerleri¹¹⁸ tarafından yapılan çalışma ise kredi riski belirlemede karar ağaçlarının kullanımı konusunda ön plana çıkan bir diğer çalışmadır¹¹⁹.

Yukarıda bahsedilen hibrit ve hibrit olmayan modellerin dışında, daha farklı hibrit modeller de geliştirilmiştir. Hsieh kümeleme ve sinir ağı tekniklerinden oluşan bir hibrit sistem önerdiler¹²⁰. Lee, Chiu, Lu ve Chen, geri beslemeli sinir ağlarını

¹¹¹ Terence C. Fogarty, Neil S.Ireson, “Evolving Bayesian Classifiers For Credit Control—A Comparison With Other Machine-Learning Methods”, **IMA Journal Of Mathematics Applied In Business & Industry**, Cilt: 5, Sayı: 4, 1993, s.63-75.

¹¹² Henley, W. E., & Hand, D. J. “A k-nearest neighbor classifier for assessing consumer credit risk”, **Statistician**, Cilt: 44, Sayı: 1, 1996, s.77.

¹¹³ Safavian, S.R., Landgrebe, D. A “Survey Of Decision Tree Classifier Methodology”. **IEEE Transactions On Systems, Man And Cybernetics**, Cilt: 21, 1991, s.660–674.

¹¹⁴ Quinlan, J. R. “**C4.5: Programs For Machine Learning Discriminating**”, Morgan Kaufman, San Mateo, CA, 1993.

¹¹⁵ S. Makowski, “Credit Scoring Branches Out”, **The Credit World**, Cilt: 75, 1985, s.30–37.

¹¹⁶ J. Y. Coffman, “The Proper Role Of Tree Analysis In Forecasting The Risk Behaviour Of Borrowers”, **Management Decision Systems**, Atlanta, MDS Reports 3,4,7 ve 9, 1986.

¹¹⁷ Lyn. C. Thomas, **a.g.e.** s.151.

¹¹⁸ Davis, R. H., Edelman, D. B., Gammerman, A. J., “Machine Learning Algorithms For Credit-Card Applications”, **Journal Of Mathematics Applied In Business And Industry**, Cilt: 4, 1992, s.43–51.

¹¹⁹ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, **a.g.e.** s.848.

¹²⁰ Nan Chen Hsieh, “Hybrid Mining Approach In The Design Of Credit Scoring Models”. **Expert Systems With Applications**, Cilt: 28, Sayı: 4, 2005, s.655.

geleneksel DA yaklaşımıyla birleştirmeyi önerdiler. Chen ve Huang, iki farklı kredi analizi problemi içeren bir çalışma sunmuş ve YSA ile GA yöntemlerini kullanarak¹²¹ bu problemleri çözdüler. Ayrıca Malhotra vd.¹²² sinirsel bulanık (nörofuzzy) modellerin geleneksel modellere göre avantajlarını araştırdılar. Hoffmann, Baesens, Martens, Put ve Vanthienen, 2002 yılında yaptıkları çalışmada, kredi riskini belirleme problemi için bir genetik bulanık ve bir nöro-bulanık sınıflandırıcı uyguladılar. Baesens ve diğerleri, kredi puanlaması için karşılaştırmalı en gelişmiş sınıflandırma algoritmalarını kullandılar¹²³. Back ve diğerleri girdi seçimi için ÇKA birlikte kullanılacak GA'nın kullanılmasını önerdiler.¹²⁴

Literatüre genel bir bakışla, sözkonusu yöntemlerden herhangi birinin, kredi riskini belirleme probleminde, hemen her veri kümesi için, doğruluk oranı olarak, diğer bütün yöntemlere salt bir üstünlük sağladığını, diğer bir deyişle, evrensel yöntem olduğunu söylemek mümkün görünmemektedir. Genel itibarıyla, konuyla alakalı, yukarıda da künyesi geçen makalelerde de belirtildiği gibi, geleneksel yöntemlerden LR, hâlâ kredi riski belirleme konusunda, etkin bir şekilde kullanılmakla birlikte, istatistiki yöntemlerden daha önce yine aynı konuda kullanılan DA'ne bariz üstünlüğü literatürde kabul görmüş bir konudur. Analizlerde kullanılan veri kümesinin içeriği özellikle nonlinear yöntemleri gerektiren, karmaşık bir yapı arzettiği durumlarda, YSA ve DVM gibi nonlinear genelleme yeteneğine sahip yapay zekâ yöntemleri ön plana çıkmaktadır.

YSA, özellikle kredi riski belirleme konusunda, hem finasta hem diğer birçok disiplinde oldukça etkin bir biçimde kullanılmaktadır. Doğruluk oranı bakımından oldukça iyi sonuçlar veren YSA; finans, tıp ve mühendislik¹²⁵ gibi birçok alanda kullanılmaktadır. Ayrıca sonuçlarına güvenilerek, bankacılıkta kredi verilen, tıpta tedavi uygulanan YSA'nın, kara kutu (black box) denilecek kadar, içeriğinin

¹²¹ Mu-Chen Chen, Cheng-Lung Huang, **a.g.e.** s.433.

¹²² Malhotra, R., Malhotra, D. K., "Differentiating Between Good Credits And Bad Credits Using Neuro-Fuzzy Systems", **European Journal Of Operational Research**, Cilt: 136, Sayı: 1, 2002, s.190-211.

¹²³ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, **a.g.e.** s. 848.

¹²⁴ Amir F. Atiya, **a.g.e.** s.931.

¹²⁵ A. Vellido, Paulo J. G. Lisboa, J. Vaughan, **a.g.e.** s.51.

bilinemeyişi ve uygun mimarinin (uygun sayıda gizli katman ve düğüm sayısının) oluşturulmasının çok zaman alması, çok sayıda veri gerektirmesi, modele yöneltilen ve hâlâ geçerliliğini ciddi ölçüde koruyan, önemli eleştiriler arasındadır. YSA, kredi riski belirleme konusunda çok fazla kullanılması ve oldukça geniş kabul görmesi ve birçok araştırma probleminde en iyi sonucu vermesine rağmen, bazı araştırmada ise diğer yöntemlere göre geri kaldığı görülmektedir.

DVM ise birçok disiplinde kullanım alanı bulabilen ve hızla türevleri geliştirilen bir yeni bir yaklaşımdır. DVM, araştırma problemlerinden sınıflandırma ve regresyon konusunda, sonuçlarına mutlaka başvurulmak istenen, önemli bir yapay zekâ yöntemi olarak ön plana çıkmaktadır¹²⁶. DVM'nın, diğer geleneksel yöntemlerin tamamından, YSA'ndan ise hafifte olsa daha iyi sonuçlar verdiği birçok kaynakta belirtilmektedir.

Genetik algoritma, çok yaygın ve önemli bir yapay zekâ yöntemi olmasının yanı sıra kredi skorlama konusunda kullanımı kesinlikle YSA ve DVM gibi yaygın olmamakla birlikte, YSA'ndan daha iyi sonuçlar verdiğini belirten çalışmalar olmuşsa da, literatürün geneline bakıldığında ise, böyle bir genel kabul görülmemektedir.

Hibrit yöntemlerin ise genellikle, kullanılan bu yaygın algoritmalarından daha iyi netice verdiğini belirten birçok çalışma bulunmaktadır. Fakat bu hibrit algoritmaların, YSA ve DVM gibi bilinen ve yaygın bir formuna literatürde rastlanmamaktadır.

Genel itibariyle, yapay zekâ yöntemlerinin sınıflandırma, örüntü tanıma, tahmin, optimizasyon vb. birçok konuda, geleneksel istatistiki yöntemlerden geri kalmadığı ve birçok çalışmada daha iyi sonuçlar da verdiği bilinmektedir. Ayrıca bu yapay zekâ yöntemleri, herhangi bir dağılıma bağlı olmamaları ve varsayım gerektirmemeleri gibi pratik özellikleriyle de oldukça yaygın bir kullanıma ulaşan, kabul görmüş yöntemlerdir¹²⁷. Yapay zekâ yöntemleri arasında da DVM'nın YSA'na alternatif teşkil eden bir yöntem olduğu da literatürde ayrıca mükerreren dile getirilmektedir.

¹²⁶ Vladimir N. Vapnik, "An Overview of Statistical Learning Theory", **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.988.

¹²⁷ Cheng-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, **a.g.e.** s.848.

Kredi riski belirleme konusunda, YSA ve özellikle DVM'nın ve geleneksel yöntemlerden de LR'un, güçlü sınıflandırma özellikleriyle ön plana çıkması neticesinde, bu tezde de, kredi riskinin belirlenmesi konusunda, yöntem olarak; yapay zekâ özelinde YSA ve DVM, istatistiksel olarak da, LR kullanılması uygun bulunmuştur.



İKİNCİ BÖLÜM

YAPAY ZEKÂ YÖNTEMLERİ

2.1. Yapay Zekâ

2.1.1. Zekâ ve Yapay Zekâ Kavramları

Yapay zekânın ne olduğunu, nasıl işlediğini, gelişim potansiyelini ve sınırlarını tartışmaya başlamadan önce, zekâ kavramını etraflıca tartışmak gerekmektedir. Bugüne dek, zekânın farklı birçok tanımı yapılmıştır. Fakat “insani bir olgu” olan zekânın, soyut bir kavram olması nedeniyle, üzerinde fikir birliğine varılmış bir tanımı bulunmamaktadır.

Mevcut durumda zekâyı, “bir varlığın kendi ortamında, doğru düzgün ve olan biteni öngörerek hayatını devam ettirmesini sağlayan bir işlev, bir özellik, bir yetenek ya da bir öğrenme, soyutlama ve uyum kabiliyeti” şeklinde¹²⁸ ya da, daha kısa bir tanımla, “insanın dış çevreye uyum sağlama yeteneğidir” şeklinde tanımlayabiliriz. Dış çevreye uyum sağlama ise çok çeşitli yeteneklere sahip olmayı gerektirmektedir. Genel olarak zekâyı, “algılama, muhakeme (aklî süreç), planlama, problem çözme, soyut düşünme, karmaşık fikirleri anlama, hızlıca düşünme ve tecrübelerden öğrenme yeteneklerinin tamamı” şeklinde de tanımlamak mümkündür¹²⁹.

Üzerinde ittifak edilmiş bir tanımı bulunmayan, soyut bir kavramı tanımlamaya çalışmak yerine, özelliklerine ve çıktılarına yoğunlaşarak da bu kavram daha iyi anlaşılabilir. Bu bakımdan Robinson vd. zekâ ile ilgili birçok kuramı analiz etmişler ve bu analizlerine dayanarak zekânın en çok başvurulmuş üç özelliğini tespit etmişlerdir. Bunlar¹³⁰:

1. Öğrenme kapasitesi: Kişinin verilen eğitimden yararlanabilme kapasitesi.

¹²⁸ Nils J. Nilson, **Yapay Zekâ Geçmişi ve Geleceği**, 1. Baskı, İstanbul: Boğaziçi Üniversitesi Yayınevi, 2019, s.13.

¹²⁹ S.Leng ve M. Hutter, “A Collections Of Definitions Of Intelligence”, **Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms**, s.18.

¹³⁰ Mustafa Özkes, “Peabody Resim Kelime Testi 3.01-3.12 Yaş Aralığı İzmir Bölgesi Standardizasyonu Çalışması”, **Ege Eğitim Dergisi**, 2013, Cilt: 14, No: 1, 92.

2. Öğrenilmiş bilginin toplamı: Kişinin kendi yetenekleri içinde öğrendiği tüm kavram ve bilgiler.
3. Çevrenin istemlerine uyabilme: Kişinin kendisini çevresine adapte edebilmesi ve çevresinde görülen değişikliklere başarılı bir şekilde uyum sağlayabilmesi.

Yapay zekâ ise, yukarıda belirtilen yeteneklerin işlevlerini yerine getirmek için geliştirilen yazılımların bizlere sağladığı bir faydadır. Yapay zekâ, kabaca, “makinelere zekâ kazandırmaya adanmış etkinlik”¹³¹ şeklinde de tanımlanabilir. Yani, yapay zekâ bir nevi, insan zekâsının; dolayısıyla, “insanın bilgi işleme ve karar verme süreçlerinin yazılım ve tümleşik yongalar yoluyla taklit edilmesi” olup, insan eliyle üretilmiş bu türden bir zekâyı, “yapay zekâ” denilmesi son derece yerinde bir kavramsal yakıştırmadır. Burada, akıl ile zekânın bir ayrımını yapmak gerekirse; akıl gerçek ile yalanı, doğru ile yanlış ayırabilme, bir konuda muhakeme yapabilme ve görüş bildirme yeteneği olarak tanımlanabilir ve aklın herhangi bir yazılımla veya başka bir yolla taklit edilmesi de mümkün değildir. Akıl genetik yolla intikal eder ve sevgi, korku, kıskançlık, doğal savunma güdülerinin yanı sıra çevresel etkileşimlerle de gelişir ve değişir¹³². Akıl, zekânın algılamadan sonraki muhakeme kısmını temsil etmektedir. Yani zekâ, aklın çıktılarını kullanan ve bunu algılamayla ilişkilendiren ve eylemle neticelendiren kapsamlı bir bilgi işleme sürecini temsil etmektedir. Bir benzetme de bulunmak gerekirse aklı; kıymetli ve hayatî bir cevhere, zekâyı ise aklın işlenmesini ve kullanılmasını sağlayan bir makineye benzetebiliriz.

Bir bilim olarak yapay zekâ, “İnsan aklının nasıl çalıştığını göstermeye çalışan bir kuram” şeklinde de tanımlanmaktadır. Bunun dışında yapay zekâ için yapılan; “insan tarafından yapıldığında zeki olarak adlandırılan davranışların makine tarafından yapılması” veya “düşünme, anlama, faaliyete geçmeyi sağlayacak bilgi işleme çalışması” şeklinde de tanımlanabilir. “İnsan olmaksızın, kompleks işleri yapabilen makineler inşa etmek için araştırma yapan bilişsel bir bilim dalı” gibi farklı açıklayıcı yaklaşımlar da, yapay zekânın farklı yönlerine atıf yapan birer tanım olarak göz önüne alınabilir¹³³.

¹³¹ Nils J. Nilson, **a.g.e.** s.13.

¹³² Çetin Elmas, **Yapay Zekâ Uygulamaları**, 3. Baskı, Ankara: Seçkin Yayıncılık, 2016, s.21.

¹³³ Harun Pirim, “Yapay Zekâ”, **Journal Of Yasar University**, Cilt: 1, Sayı: 1, s. 84.

Meseleye tersinden bakacak olursak, yazılımların, daha önce hiç karşılaşılmamış iş süreçlerine bile, kayıtlı tecrübeler çerçevesinde, bilgi işlem sürecini işleterek, mümkün olan en makul kararı verebilme özelliği, insandaki; “zekâ” kavramına karşılık gelmektedir.

İnsan zekâsı (doğal zekâ) ve yapay zekâ arasındaki etkileşimi bir örnekle açıklamak gerekirse, bir insan çok haneli iki sayının çarpımı gibi karmaşık matematiksel işlemleri, bir bilgisayar kadar hızlı ve doğru yapamaz. Fakat bir arabayla yolda giderken, kaza ihtimali bulunan bir anda, çevredeki tüm araçları, hızlarını, yolun ve havanın durumu vb. birçok faktörü hızla algılayıp, değerlendirip (muhakeme) gerekli müdahale emrini vermek (davranış) gibi, çoklu bir idrak yeteneği söz konusu olduğunda, insan zekâsı yine rakipsiz olmaktadır¹³⁴. Tüm bu yetenekler otonom araçlara, yazılımlar aracılığıyla, kazandırılabilir. Fakat, programlanmamış, yani yazılımsal olarak işlenmemiş yeni bir faktörün varlığı söz konusu olduğunda, yine insan zekâsına ihtiyaç duyulacaktır. Yani yapay zekâ, doğal zekâyâ tâbidir.

Yapay zekâ, matematik, istatistik, yöneylem, geometri, psikoloji, biyoloji, fizik, kimya, tıp, zooloji, genetik, mühendislik ve bilgisayar bilimleri vb. hemen her disiplini içine alan bir çalışma alanıdır. Fakat yapay zekâyı disiplinlerarası bir yaklaşım olarak görmek, eksik bir bakış açısı olabilecektir. Çünkü bilgisayarların gelişerek tüm iş süreçlerine dâhil olduğu zamanları (daha çok 1990’lı ve 2000’li yıllar) hatırlayacak olursak, yapay zekânın bizi getireceği süreç de bundan çok farklı olmayacak ve hemen her işte yapay zekâ araçlarına başvurulur olacaktır. Yani yapay zekâ yalnızca disiplinler arası bir yaklaşım değil, insanın zeki yeteneklerini üstlendiği için, insanın ulaştığı tüm bilimsel birikimi süreçlerine dâhil edebilecek, kapsamlı, tutarlı, tertipli ve öğrenme bakımından sınırsız potansiyelli bir bilim öğrencisi ve hatası minimize edilmiş bir uygulama makinesi olarak görmek, daha kapsayıcı bir bakış açısı olacaktır.

Yapay zekânın en önemli kullanım alanlarından biri olan “bir dilden bir başka dile tercüme (doğal dil işleme)” sahasından bir örnek getirecek olursak, herhangi bir gramer kuralını dahi bilmeyen söz konusu zeki yazılımlar, yaptıkları bu tercümeyi

¹³⁴ Çetin Elmas, **a.g.e.** s.21.

şimdiye kadar yapılmış ve sanal ortama aktarılmış tercüme metinlerinin tamamını göz önünde bulundurarak yapmaktadırlar. Bu seviyede bir mütercimlik ise hiçbir insanın tek başına güç yetirebileceği bir seviye değildir. Yani şimdiye kadar ki yapılmış tüm tercüme metinlerine vakıf olan bir tercüme yaklaşımı, bir insanın hatta (Oxford sözlüğünün, çok fazla insanın katkısıyla, onlarca yılda oluşturulabildiğini anımsayacak olursak), bir milletin dahi takatinin çok üzerindedir. Buradan şu şekilde bir sonuca varmak mümkün olabilir: yapay zekâ, insanlığın kolektif zekâsını temsil etmektedir ve insanlığın ulaşabildiği en son noktaya ulaşacaktır. Bilgisayar, hesap makinesi gibi araçlar nasıl ki tüm iş süreçlerinin önemli bir parçası ve bu cihazları kullanamayan kişi ve kurumların artık varlıklarını sürdüremedikleri ya da çok dar bir çerçeveye sıkıştıkları bilinmektedir. Çok yakın bir gelecekte, insanların yine nihai karar vericiler olmasının yanı sıra yapay zekâyı etkin kullanamayan kişi ve kurumların uzmanlığından, stratejik yönetim süreçlerinden, karlılıklarından ve nihayet sürdürülebilirliklerinden bahsetmek de pek mümkün görünmemektedir.

İnsanın öğrenme yeteneğinin bir yazılımdan çok daha üstün olduğunu da unutmamak gerekir. Örneğin bir insan, yeni tanıştığı iki kişiyi, bir kez gördükten sonra çok rahat ayırt edebilir. Fakat bunu bir yapay zekâ yazılımının yapabilmesi için, bu iki kişinin ortalama olarak yüz ila iki yüz adet fotoğrafının modele girilerek, modelin eğitilmesi gerekmektedir. Doğal zekâ, yaratıcılık bakımından hiçbir yazılımla mukayese edilemeyecek bir kapasiteye, inovasyon kabiliyetine, çağrışımlı ve uyarlanabilir hafıza derinliğine, sınırsız bir algılama yeteneğine ve en önemlisi ise hayal gücüne sahip eşsiz bir varlıktır.

Günümüze kadar ki teknolojik gelişmelerle yapay zekânın bir mukayesesini yapacak olursak, örneğin bir makineye veya bir otomobile bakıldığında, birçok mekanizmasının icadının ve geliştirilmesinin arka planında, çeşitli insan uzuvlarının taklit edilmesi olduğu görülecektir. Örnek vermek gerekirse fotoğraf makinesi, insan gözünden esinlenilerek geliştirilmiştir. Arabaların cam sileceklerinin rahatsız edici bir şekilde sürekli değil de fasılalı bir biçimde çalıştırılması, ortalama altı saniyede bir gözü temizlemek için açılıp kapanan göz kapaklarından esinlenerek icat edilmiştir. Fakat yapay zekâ söz konusunda olduğunda, bu kez şimdiye kadarkilerden farklı olarak, insanın fiziksel işlevlerinin değil, zihinsel yeteneklerinin taklidi söz konusudur.

Yapay zekâ ile normal zekâyı, gerçek anlamda kıyaslamak pek mümkün değildir. Fakat konumuz çerçevesinde ele alınacak olursa, yapay zekâ ile doğal zekâ arasındaki farklar şu şekilde sıralanabilir¹³⁵:

- Doğal zekâda bilgiler zamanla unutulabilir. Fakat yapay zekâda bilgiler sistem içerisinde depolandığı için unutulmaz ve olduğu gibi korunur.
- Doğal zekâda bilgi ve tecrübenin aktarılması çok zaman alır. Ayrıca tecrübenin bütün yönleriyle, tamamen aktarımı çoğu zaman mümkün olmaz. Ama yapay zekâ da bilgi aktarımı oldukça kolaydır.
- Doğal zekânın eğitimi çok uzun zaman alırken, yapay zekânın eğitimi ise artık gelişen bilgisayarlar ve uygun maliyetli sensörlerin toplamış olduğu bilgilerle oldukça kolay ve hızlıdır.
- Doğal zekâda kişiye özgü tecrübelerle hareket edildiği için, belli bir zaman sonra neden böyle bir karara varıldığı sorusu cevapsız kalmaktadır. Oysa yapay zekâ belli bir sistematik içinde hareket ettiği için raporlama ve belgeleme kolaydır.
- Yapay zekâda herhangi bir yaratıcılık söz konusu değildir. Yapay zekânın tüm yaratıcılığı, kendini tasarlayan doğal zekâdan kaynaklanır.
- Doğal zekâ da, yapay zekâ da geçmiş tecrübelerden öğrenir. Fakat yeni karşılaşılan durumları, tecrübelerinden hareketle yeni duruma uyarlama ve sonuç alma doğal zekânın yetisidir. Yapay zekâ, bunu doğal zekâdan öğrenir. Yani bilinen, kurumsallaşmış eylemlerde doğal zekâdan daha hızlı ve doğru ama bilinmeyende, doğal zekâyı taklit eder. Kısaca taklit eden yapay zekâ, taklit edilen ise doğal zekâdır.

2.1.2. Yapay Zekânın Kullanım Alanları

Yapay zekâ, hemen her disiplinde, yönetim süreçlerinin tamamında, etkin olarak kullanılmaktadır. Yapay zekânın kullanıldığı alanlar, her ne kadar sayması zor olsa da, başlıklandırmak gerekirse, şu şekilde özetlenebilir:

¹³⁵ Atıncı Yılmaz, Yapay Zekâ, 2. Baskı, İstanbul: Kodlab Yayınevi, 2017, s.7.

- Çıkarım (tümdengelim),
- Genelleme (tümevarım),
- Akıl yürütme (muhakeme),
- Problem çözme,
- Veriden sonuç çıkarma, veri toplama, veri birleştirme, veri düzenleme
- Sınıflandırma, kümeleme, tahmin, optimizasyon,
- Yönetim (planlama, gerçekleştirme, hata payı oranında düzeltme),
- Makine öğrenimi, veriden öğrenme, karar destek sistemleri oluşturma,
- Doğal dil işleme (Natural Language Processing, NLP),
- Robotik,
- Yaratıcılık (teorik ve pratik)

olarak özetlenebilir¹³⁶.

Yukarıda sıralanan kullanım alanları içinde konumuzla alakalı olan “makine öğrenmesi”, yapay zekânın merkezi bir alt dalıdır ve bu tez çalışmasının da konusudur.

2.1.3. Makine Öğrenmesi

Makine öğrenmesi kavramını tanımlamadan önce, bu kavramın geliştirilmesine zemin oluşturan evrimsel süreci gözlemleyebilmek için, ilgili kavramları sırayla tanımlayacak olursak, *veri (data)*; insanların daha sonra çeşitli amaçlar için kullanmak amacıyla saklama ihtiyacı duydukları, salt, yorumsuz ve ham bilgilerdir. Bir başka tanımla veri, “bir problemde bilinen anlatımlardan yola çıkarak bilinmeyeni bulmaya yarayan şey, bilgi, data” şeklinde tarif edilmektedir¹³⁷.

Öğrenmek, “görevi yerine getirme becerisine sahip olma yolumuz” olarak tanımlanmaktadır. *Makine öğrenmesi (machine learning)*, esasen, “karmaşık işlevleri istatistiksel olarak tahmin etmek için bilgisayar kullanımına daha fazla vurgu yapan ve bu işlevler etrafındaki güven aralıklarını kanıtlamaya daha az vurgu yapan (analitik

¹³⁶ Huimin Lu, Yujie Li, Min Chen, Hyoungeop Kim ve Seiichi Serikawa, “Brain Intelligence: Go beyond Artificial Intelligence”, **Mobile Netw Appl**, Cilt: 23, 2018, s.368–375.

¹³⁷ TDK, Genel Türkçe Sözlük, Çevrimiçi, <https://sozluk.gov.tr/?kelime=>, 11/11/2019.

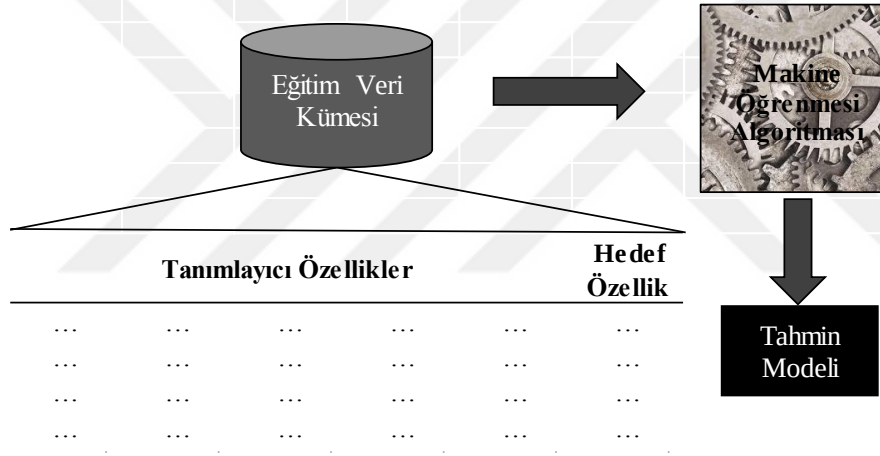
değil, pratik) bir uygulamalı istatistik biçimidir”. Makine öğrenimi, insanlar tarafından yazılmış hazır programlarla çözülmesi çok zor olan görevlerin bilgisayarlara, makinelere yaptırılmasının imkânlarını araştırır. Bilimsel ve felsefi açıdan bakıldığında, makine öğrenimi, ilgi çekicidir çünkü makine öğrenimi kullanımını geliştirmek demek, zekânın altında yatan, insani ilkelere ilişkin süreçlerin anlaşılması demektir. Yani makine öğrenimi süreçleri, bir bakıma, “insanın yine insanı etraflica ve daha yakından tanıma süreci”ne vurgu yapmaktadır. *Makine öğrenmesi algoritması (machine learning algorithm)* ise “veriden öğrenebilen algoritma” şeklindedir. *Veri kümesi (dataset)*, “özelliklerin sıralı bir koleksiyonundan oluşan örneklerin bir koleksiyonu” şeklinde tanımlanabilir. En genel şekliyle *veri kümesi (dataset)* bir “dizayn matrisi” olarak bilinir. Dizayn matrisinin her bir satırı bir örneğe karşılık gelir ve sütunlardan herbiri ise bir özelliğe karşılık gelmektedir. Bir de bu matristeki her bir örnek, geometrik anlamda, bir “vektör” olarak tanımlanabilir. *Makine öğrenmesi görevleri (machine learning tasks)* ise genellikle “makine öğrenmesi sisteminin bir örneği nasıl işlemesi gerektiği” şeklinde tanımlanmaktadır ve makine öğrenmesi sistemleri ile gerçekleştirilebilen görevler şu şekilde listelenebilir¹³⁸:

- Sınıflandırma
- Eksik girişli sınıflandırma
- Regresyon
- Uyarlama (transcription)
- Makine tecümesi
- Ayırıştırma-eşleme (parcing-mapping)
- Anomali tespiti
- Sentez ve örnekleme
- Eksik değerlerin bulunması
- Gürültü arındırma
- Yoğunluk tahmini veya olasılık kütle fonksiyonu tahmini

¹³⁸ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 98-106.

Makine öğrenmesi, “veriden örüntüleri (desenleri) bulup çıkararak, otomatize edilmiş bir süreç” olarak da tanımlanabilir. Tahminleyici veri analitiği uygulamalarında, modelleri inşa etmek için denetimli makine öğrenmesi kullanılır. Denetimli makine öğrenmesi, daha önceden elde edilmiş örneklerle veya vakalara dayanan, açıklayıcı değişkenlerden oluşan tanımlayıcı özellikler (descriptive features) ve bağımlı değişken dediğimiz hedef özellik (target feature) arasındaki ilişkilerin bir modelini öğrenen, otomatik bir tekniktir. Sonrasında geçmiş vakalarla eğitilen (tümevarım) bu model, yeni vakalar için de kullanılabilir (tümdengelim) bir araç niteliği kazanır. Bu iki aşamalı süreç, Şekil 2.1.’de gösterilmektedir¹³⁹:

Şekil 0.1: İki Aşamalı Denetimli Makine Öğrenmesi Yöntemi



(a) Geçmiş tecrübelerden öğrenen bir model



(b) Tahminler yapmak için kullanılan bir model

Kaynak: John D. Kelleher, Brian Mac Namee, Aoife D’arcy, **Fundamentals Of Machine Learning For Predictive Data Analytics, Algorithms, Worked Examples, And Case Studies**, 1. Baskı, America: MIT Press, 2015, s.3.

¹³⁹ John D. Kelleher, Brian Mac Namee, Aoife D’arcy, **Fundamentals Of Machine Learning For Predictive Data Analytics, Algorithms, Worked Examples, And Case Studies**, 1. Baskı, America: MIT Press, 2015, s.3.

2.1.4. Makine Öğrenmesi Ve Veri Madenciliği

Makine öğrenmesi temelde bir veri madenciliği yöntemidir. Konuyla alakalı tanımlara, yukarıda kaldığımız yerden devam edecek olursak, *veri tabanı (database)*, “insanların ya da kullanılan hafızalı cihazların ürettiği verilerin belirli bir düzen içerisinde, çeşitli uygulamalar tarafından saklandığı/depolandığı bir bilgi sistemi, düzenli bilgiler topluluğudur”. Veri tabanı kavramı, bilgisayar terminolojisinde daha detaylı bir şekilde, “sistemik erişim imkânı olan, yönetilebilir, güncellenebilir, taşınabilir, birbirleri arasında tanımlı ilişkiler bulunabilen bilgiler kümesi” olarak tanımlanmaktadır. Yine bir başka tanım ile veri tabanı, “bilgisayarda sistemik şekilde saklanmış, programlarca istenebilecek ham veri yığını” olarak tanımlanmaktadır¹⁴⁰. Yine veri tabanı, “bir kuruluştaki kullanılan uygulama programlarının kullandığı verilerin bütünüdür” denilebilir. *Veri tabanı sistemi (database system)* ise, “çeşitli amaçlar için verileri düzenli bir şekilde saklayan ve bu verileri kullanarak yararlı bilgiler üreten bilgisayarlı kayıt tutma sistemidir”. Bu doğrultuda, *Veri tabanı yönetim sistemi (database management system)*, “kullanıcı ile veri tabanı arasındaki iletişimi sağlayan yazılımların bütünüdür”. Veri tabanı yönetim sistemleri, veri tabanlarının oluşturulmasını, yönetilmesini, yedeklenmesini, kullanıcı erişimlerinin düzenlenmesini kapsamaktadır. Bu sistemler, kişi ve kurumların veriyi etkin bir şekilde kullanmalarına ve programcılarının kaydedilen bu verilere kolay erişimine imkân tanıyan bir tür yazılımdır. Ayrıca veri tabanı yönetim sistemleri, farklı türden bilgilerin, aynı ortamda tutulmalarını, bu bilgilere erişimin hızlı ve kolay olmasını, bilgi tekrarlarının önlenmesini, sürekli yapılması gereken işlemlerin hızlı ve doğru yapılmasını, aynı bilgi kümelerinin birden fazla kişi ya da kurum tarafından, aynı anda kullanılabilmelerini ve verilerin güvenliğini sağlayan yazılımlardır¹⁴¹.

Veri madenciliği (data mining), büyük miktarda veri içerisinde, gelecekle alakalı tahminler yapılabilmesini sağlayacak bağlantıların, kuralların veya örüntülerin çeşitli

¹⁴⁰ Kaan Kocamaz, “Hastane bilgi yönetim sistemlerinde veri madenciliği ve Konya Meram Tıp Fakültesi hastanesindeki hastane bilgi yönetim sistemi uygulamasının incelenmesi”, Yüksek Lisans Tezi, Selçuk Üniversitesi 2007, s.22.

¹⁴¹ Selçuk Alp, Savaş Özdemir, Arzu Kilitçi, **Veri Tabanı Yönetim Sistemleri**, Türkmen Kitabevi, İstanbul, 2011, s.12.

bilgisayar programları yardımıyla ortaya çıkarılması işlemidir. Veri madenciliği, büyük miktarlarda ve oldukça hızlı toplanan verilerin, çeşitli analizler neticesinde anlamlı ve yararlı çıktılara dönüştürme süreci olarak da tanımlanabilir. Veri madenciliği tanımlarına genel olarak bakıldığında, bu tanımların hepsinde yer alan iki unsurdan birincisinin, çok fazla miktarlarda verinin veri tabanlarında kullanışlı bir şekilde depolanması (veri ambarlarının oluşturulması), ikincisinin ise bu verilerden anlamlı ve yararlı tespitlerin yapılması olduğu görülmektedir¹⁴².

Veri madenciliği, şirketlerde müşterilerin satın alma alışkanlıklarının tespitinde, satış tahmininde, müşteri segmentasyonunda, çeşitli müşteri grupları için ürün geliştirme faaliyetlerinde, çeşitli risklerin analizi gibi pazarlama ve finans başta olmak üzere birçok sektörde kullanılmaktadır ve karar vericilerin en uygun kararı vermelerine yardımcı olmaktadır.

Veri madenciliği yöntemleri, “veriden öğrenme (learning from data)” diyebileceğimiz bu süreçleri ilgili veri kümesinde yer alan değişkenler arasındaki ilişkileri, çeşitli yöntemlerle, keşfederek (öğrenerek) bir model geliştirir. Geliştirilen bu model, kullanılan veri kümesini haritalandırır, özetler ve veriden anlamlı sonuçlar üreten bir makine görevi görür. Model kullanarak tahminler yapma, eldeki verilerin bir özetini çıkarma anlamına geleceği için, çarpma işleminin nasıl yapıldığını matematiksel olarak modelleyip, tüm sayıların birbiriyle çarpımını ezberlemeye ya da sürekli olarak yanında taşımaya gerek kalınmaması örneğine benzetilebilir. Modellerin geçerlilik seviyesi ve sınırları içerisinde artık bu makine, yeni bilgiler için de sonuç hesaplamaları yapabilir. Modelin üretildiği veri kümesi ise, bir bakıma modelin tecrübesini yansıtmaktadır¹⁴³. İlgili veri kümesi büyüdükçe modelin kararlılığı da artar. Tüm bu süreçler kavramsal olarak makine öğrenmesi olarak ifade edilmektedir.

¹⁴² Umman Tuğba Şimşek Gürsoy, **Veri Madenciliği Ve Bilgi Keşfi**, 1. Baskı, Ankara: Pegem Yayınevi, 2009, s.27.

¹⁴³ Şenol Emir, “Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması: Borsa Endeks Yönünün Tahmini Üzerine Bir Uygulama”, Doktora Tezi, İstanbul Üniversitesi, 2013, s.41-42.

2.1.5. Makine Öğreniminin Teorisi

Makine öğrenmesi süreci Vapnik tarafından, aşağıdaki şekilde özetlenmektedir: “Makine öğrenmesi, denetleyicinin yanıtını mümkün olan en iyi şekilde öngören bir dizi fonksiyonlardan birini seçmektir ve aşağıdaki üç bileşenden oluşur¹⁴⁴:

- 1- Her birine x denilen tesadüfi vektörler, $P(x)$ gibi sabit fakat bilinmeyen bir dağılımdan açıklayıcı bir şekilde çizilir,
- 2- Denetmen, her giriş vektörü (x) için, sabit de olan fakat bilinmeyen bir şartlı dağılım fonksiyonuna göre bir çıkış vektörü (y) döndürür. Bu seçim yine sabit ve aynı dağılımlı gözlemleri ($P(x, y) = P(x)P(y|x)$) oluşturur.
- 3- Fonksiyonların bir dizi uygulama yeteneğine sahip bir öğrenme makinesi ($f(x, \alpha) \alpha \in \Lambda$) oluşturulur”.

Öğrenme problemi, aslında bir fonksiyon ($f(x, \alpha) \alpha \in \Lambda$) tahmini problemidir. Denetim otoritesinin cevabına en uygun yaklaşımı seçmek için, denetim otoritesinin verilen bir cevaba verdiği cevap ile öğrenme makinesi tarafından verilen cevap arasındaki kayıp veya tutarsızlık, yani yanlış tahmin riski ($L(y, f(x, \alpha))$) ölçülür ve bu tutarsızlığı kullanarak bir risk fonksiyonu;

$$R(\alpha) = \int L(y, f(x, \alpha)) dP(x, y) \quad (0.1)$$

oluşturulur. Burada amaç, risk fonksiyonunu ($R(\alpha)$) minimize ederek, optimum fonksiyonu ($f(x, \alpha_0)$) bulmaktır. Fakat bu, o kadar kolay olmayacaktır. Çünkü eğitim ve test veri kümeleri iki ayrı veri kümesidir ve oldukça da sınırlıdır. Bunlardan birisinin üzerinde geliştirilen bir modelin diğer veri kümesini ayıklaması pratikte pek mümkün görünmemektedir. Bu sorunu aşabilmek için, bu iki veri kümesi ile ilgili bazı varsayımlara ihtiyaç vardır. Makine öğrenmesinde, eğitim veri kümesi üzerinde geliştirilmiş bir modelin bundan farklı bir test verisi üzerinde başarılı bir değerlendirme yapabilmesi için eğitim verisi ile test verisinin aynı olasılık dağılımına sahip veri kümeleri olduğu varsayılır. Bununla birlikte veri kümeleri içinde yer alan

¹⁴⁴ Vladimir N. Vapnik, “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.988.

her bir gözlemin, yani girdi vektörlerinin birbirinden bağımsız olduğu varsayımı yapılır ve bu varsayımlar “bağımsız ve özdeş dağılımlı (independent identically distributed, i.i.d.)” şeklinde ifade edilir. Bu varsayımlar, veri oluşturma sürecinin tek bir örnek üzerinden olasılık dağılımıyla tanımlanmasına olanak sağlar. Aynı dağıtım daha sonra her eğitim örneğini ve her test örneğini oluşturmak için kullanılır. Buna altta yatan dağılım, p olarak gösterilen, dağıtımı oluşturan veri denilmektedir. Bu olasılık çerçevesi ve ilgili varsayımlar, eğitim hatası ile test hatası arasındaki ilişkiyi matematiksel olarak incelememize izin verir. Eğitim ve test hatası arasında hemen gözlemleyebileceğimiz bir ilişki, rastgele seçilen bir modelin beklenen eğitim hatasının, o modelin beklenen genelleme hatasına eşit olmasıdır. Bir olasılık dağılımımız olduğunu varsayalım $p(x, y)$ ve eğitim veri setini ve test veri setini oluşturmak için ondan tekrar tekrar örneklediğimizi varsayalım. Bazı sabit değerler (w) için, beklenen eğitim seti hatası, beklenen test seti hatasıyla tamamen aynıdır, çünkü her iki beklenti de aynı veri seti örnekleme süreci kullanılarak oluşturulur. İki koşul arasındaki tek fark, örneklediğimiz veri kümesine atadığımız isimdir. Elbette, bir makine öğrenimi algoritması kullanıldığında, parametreler önceden sabitlenmez, ardından her iki veri kümesi de örneklenir. Eğitim seti örneklenir, hemen ardından eğitim seti hatasını azaltacak parametreleri seçmek için kullanılır ve ardından test seti örneklenir. Bu süreç altında, beklenen genelleme (test) hatası, beklenen eğitim hatası değerine eşit veya ondan büyüktür. Bir makine öğrenimi algoritmasının ne kadar iyi performans göstereceğini belirleyen faktörler, aşağıda belirtilen iki koşulu yerine getirme yeteneğidir. Bunlar:

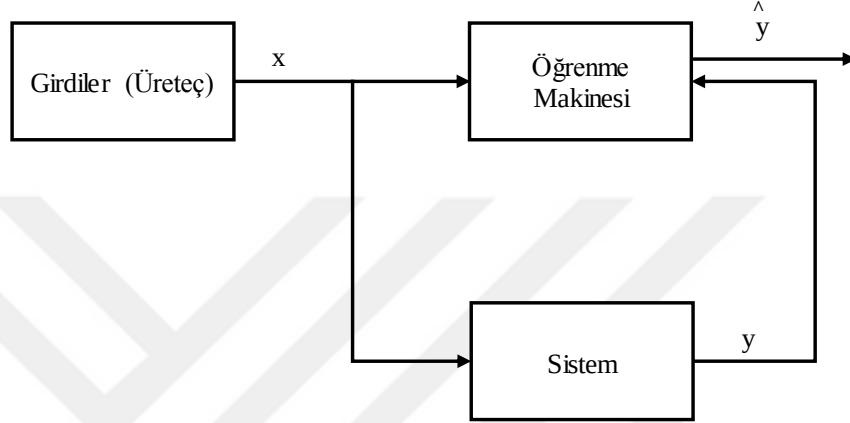
1. Eğitim hatasını küçük yapın.
2. Eğitim ve test hatası arasındaki boşluğu küçük tutun.

Bu iki koşul, makine öğrenimindeki iki temel probleme karşılık gelir. Bunlardan birincisi eksik uyum (underfitting), ikincisi ise aşırı uyum (overfitting). Eksik uyum, model eğitim setinde yeterince düşük bir hata değeri elde edemediğinde meydana gelir.

Aşırı uyum ise, eğitim hatası ile test hatası arasındaki boşluk çok büyük olduğunda meydana gelir¹⁴⁵.

Makine öğrenmesinin tipik bir gösterimi Şekil 2.2'deki gibidir:

Şekil 0.2: Makine Öğrenmesi Süreci



Kaynak: Vladimir Cherkassky, Xuhui Shao, Filip Mulier ve Vladimir Vapnik, “Model Complexity Control for Regression Using VC Generalization Bounds”, IEEE Transactions On Neural Networks, Cilt: 10, Sayı: 5, 1999, s. 1078.

2.1.6. Öğrenme Problemleri

Gerçek hayat problemlerinde, bir takım sorulara, cevaplar bulunmaya çalışılmaktadır. Bunlar bazen bir gözlemin mevcuttaki sonuç sınıflarından hangisine dâhil edileceği (sınıflandırma ya da kategorilendirme), ya da düzenlenmek istenen bir gözlem grubunun kaç homojen altküme ile temsil edilebileceği (kümeleme) olabilir. Bir başka örnek vermek gerekirse, “bir hastanın diyabet ya da kanser olup olmadığının bilgisini veren kan değerleri kompozisyonunun ne olması gerekir?” gibi sorular olabilir. Daha önceki yıllarda tüm bu sorulara kişi ve kurumlar tecrübeleri doğrultusunda, sezgilerini kullanarak karar verirlerken, bugün veri tabanlarından elde edilen çok büyük veri kümeleri üzerinde yapılan modelleme çalışmalarıyla, bir diğer ifadeyle makine öğrenmesiyle, sonuçlarda çok yüksek doğruluk seviyelerine ulaşılabilmektedir.

¹⁴⁵ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 111.

Makine öğrenmesinin uygulanabildiği problemler, genel olarak, dört ana başlıkta toplanabilir. Bunlar, regresyon (haritalama), sınıflandırma (örüntü tanıma), kümeleme ve yoğunluk tahmini problemleridir.

2.1.6.1. Regresyon

Regresyon, herhangi bir değişkenin, (bağımlı değişken) bir veya birden fazla açıklayıcı (bağımsız) değişkenle arasındaki ilişkinin matematiksel bir fonksiyon yardımıyla ifade edilmesidir. Bu fonksiyonel gösterime “regresyon denklemi” adı verilmektedir. Regresyon denklemi yardımıyla açıklayıcı değişkenlerin çeşitli değerlerine karşılık, bağımlı değişkenlerin değerinin ne olacağı tahmin edilir. Bağımlı değişkeni etkileyen açıklayıcı değişkenlerin belirlenmiş olması da, bağımlı değişken konusunda alınacak kararlarda hangi değişkenlerin önem kazandığını saptamaktır¹⁴⁶. Regresyon analizi, bir takım varsayımların sağlanması halinde geçerli olan bir çözüm yöntemidir. Regresyon denkleminin genel gösterimi aşağıda verilmiştir:

$$y_i = a + b_i x_i + \varepsilon_i \quad (0.2)$$

Burada y_i bağımlı değişkenin her bir değeri için yapılan tahmini, x_i her bir açıklayıcı değişkenin değerlerini, a modelin sabitini, b_i her bir açıklayıcı değişkenin katsayısını ve ε_i ise bağımlı değişken için yapılan her bir tahminin sonuç değerinden sapmasını, yani modelin hatasını göstermektedir. Regresyon, bir (basit regresyon) ve birden fazla açıklayıcı değişkene sahip (çoklu regresyon) problemlere uygulanabilir.

x_i açıklayıcı değişkenlerinin üslerinin en fazla bir olduğu durumlar “doğrusal en küçük kareler regresyonu” olarak adlandırılırken, tahminlerin Hata Kareleri Toplamını (HKT) minimize edecek bir doğru çizilerek değişkenler arasındaki etkileşim sayısallaştırılır. Regresyon modelleri, (x_i) açıklayıcı değişkenlerin üsleri birden fazla olduğu durumlarda, “doğrusal olmayan en küçük kareler regresyonu” olarak

¹⁴⁶ Neyran Orhunbilge, **Uygulamalı Regresyon Ve Korelasyon Analizi**, 2. Baskı, İstanbul: İşletme Fakültesi Yayınları, 2002, s.12.

adlandırılır ve bu sefer bazı dönüşüm formülleriyle model yine doğrusal hale getirilerek çözüme gidilir¹⁴⁷.

$$HKT = (x + a)^n = \sum_i^n e_i^2 = \sum_i^n (y_i - a - b_i x_i)^2 \quad (0.3)$$

Yukarıda formülü verilen, hata kareleri toplamı minimize edilerek, modelin hatası minimize edilmiş olur ve böylelikle sonucu en iyi tahminleyen açıklayıcı değişken katsayı takımı (b_i) belirlenmiş ve model kurulmuş olur.

Regresyon günümüzde, çok değişkenli yaklaşımların temelini oluşturmaktadır. Geliştirilen diğer birçok çok değişkenli çözüm yöntemi, regresyonun eksik kalan taraflarını telafi etmek, güçlendirmek veya söz konusu varsayımların telafi edilmesi için geliştirilmiştir.

Regresyon varsayımlarının sağlanması gerekliliği ve bu varsayımlardan ilki olan veri kümesinin normal dağılıma uyması zorunluluğu, bu yöntemin araştırmacıyı zorlayan yanını teşkil ederken (gerçek hayat problemlerinde çoğu veri kümesi normal dağılmamaktadır), bu metot araştırmacıya iki önemli kolaylık da sağlamaktadır. Bunlar, regresyonun formülüne bakıldığında, açıklayıcı değişken katsayıları (b_i) yardımıyla, birincisi; her bir açıklayıcı değişkenin bağımlı değişkeni ne kadar etkilediği görülebilir. Katsayısı, matematiksel olarak mutlak anlamda, en büyük olan açıklayıcı değişken bağımlı değişkeni en çok, yine mutlak anlamda en küçük olan değişken ise en az etkileyecektir. Ayrıca katsayının işareti ile de ilişkinin yönü belirlenebilecektir. İkincisi ise, bir açıklayıcı değişkendeki bir birimlik değişimin, ilgili açıklayıcı değişkenin katsayısına bakılmak suretiyle, bağımlı değişkeni ne kadar etkileyeceği de görülebilir.

Denklem 2.2'ye geri dönelim. Denklemde geçen ve a terimi, girdi vektörülerinden (x_i) bağımsız bir sapma (orjinden) parametresidir. Genellikle girdi vektöründen çıktı vektörüne dönüşümün (afin dönüşüm) sapma parametresi olarak kullanılır. Bu

¹⁴⁷ Emre Çomak, “Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar”, Doktora Tezi, Selçuk Üniversitesi, 2008, s.77.

terminoloji, herhangi bir girdi olmadığında bu afin dönüşümün çıktısının a olmaya doğru meyilli olduğu bakış açısıyla ortaya çıkar. Bu terim, istatistiksel bir tahmin algoritmasının beklenen miktar tahmininin gerçek miktara eşit olmadığı istatistiksel sapma (ε_i) fikrinden farklıdır. Sadece iki sınıfı birbirinden ayıran regresyon doğrusunun orjinden sapma miktarına karşılık gelmektedir. DR elbette son derece basit ve sınırlı bir öğrenme algoritmasıdır, ancak bir öğrenme algoritmasının nasıl çalışabileceğine dair iyi bir örnek sağlar. Sonuç olarak regresyonda, bazı girdiler verildiğinde, sonuca dair sayısal bir değeri tahmin etmesi beklenir. Dolayısıyla bu görevi çözmek için, öğrenme algoritmasından bir $f: \mathbb{R}^n \rightarrow \mathbb{R}$ fonksiyonunu çıkarması beklenir. Bu tür bir analiz, çıktı formatının farklı olması dışında, sınıflandırmaya benzemektedir¹⁴⁸.

2.1.6.2. Sınıflandırma

Sınıflandırma problemleri, gözlemin atanacağı böğelerin (sınıflandırmada sınıf ve kümelemede küme olarak özelleşir) etiketlerinin, uygulama öncesinde, uygulayıcı tarafından belirlendiği ayırma problemleridir. Sınıflandırmada, örneğin kişinin diyabet hastası olup olmadığı araştırılıyorsa, sınıf etiketlemeleri, uygulama öncesinde, mesela “diyabet” ve ya da “diyabet değil” şeklinde belirlenir. Bir başka örnekle, bankaya kredi için başvuran müşterinin yükümlülüklerini yerine getirip getiremeyeceği gibi, sınıf etiketleri (isimleri) uygulama öncesinde belirlenerek çözümü gerçekleştirilen problemler, sınıflandırma problemleridir.

Sınıflandırma problemlerinde de, regresyonda olduğu gibi, en iyi sonucu veren yani sabiti ve katsayıları sınıflandırma hatasını minimize eden fonksiyonu bulabilmek için, (sonuç çıktılarının da yer aldığı) eğitim verisine ihtiyaç duyulur.

Sınıflandırma, yine ideal bir fonksiyon bulma şeklinde gerçekleştirilir. Bu fonksiyonların kümesi ikili sınıflandırma problemi için 2.4’deki şeklini alır.

$$f(\mathbf{x}, \omega), \omega \in \Omega \quad (0.4)$$

¹⁴⁸ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 101-110.

Bu fonksiyon “gösterge (indicator)” fonksiyonlarının bir kümesini temsil etmektedir. Bu ikili sınıflandırma probleminin hatasını bulabilmek için Denklem 2.5 (kayıp fonksiyonu) yaygın olarak kullanılır.

$$L(y, f(x, \omega)) = \begin{cases} 0, & \text{eğer } y = f(x, \omega), \\ 1, & \text{eğer } y \neq f(x, \omega). \end{cases} \quad (0.5)$$

bu kayıp fonksiyonu kullanıldığında, risk fonksiyonu;

$$R(\omega) = \int L(y, f(x, \omega)) p(x, y) d_x d_y \quad (0.6)$$

hâlini alır¹⁴⁹.

Bilgisayar programları, modele girilen tüm gözlemlerin, her zaman, her analizde, sonuç değerlerini sağlayamazlarsa (yarı denetimli öğrenme) sınıflandırma işlemi daha zor bir hale gelecektir. Veriler eksik olduğunda, kullanılan sınıflandırma algoritması bu eksik veriyi telafi edebilmek için bir takım “öğrenme” işlevlerini yerine getirmesi gerekir. Bu soruna özellikle tıp sahasında rastlanılır. Çünkü çoğu testler ya çok maliyetli ya da hızlı müdahale gereğinden dolayı oluşan zaman yetersizliğinden dolayı yapılamamaktadır. Öğrenme makinesi, böylesine büyük bir fonksiyon setini verimli bir şekilde tanımlama sorununu genellikle, verinin olasılık dağılımını tahminleyip eksik veriyi marjinalleştirme yoluna giderek aşmaya çalışmaktadır. Eksik girdilerle sınıflandırma, makine öğreniminin neler yapabileceğinin yalnızca bir örneğidir¹⁵⁰.

2.1.6.3. Kümeleme

Büyük problemler, küçük parçalara ayrılarak daha kolay çözülebilir. Fakat bu ayırma işleminin sistematik yapılması gerekmektedir. “Sınıflamaları hakkında açık bilgi (sonuç çıktısı) bulunmayan durumlarda, topluluğa ilişkin tahminlerin yapılmasında yararlanılan yöntemler” olarak da tanımlayabileceğimiz kümeleme analizi, bir araştırmacıya “bir veri kümesindeki benzer(homojen) birey gruplarını bulma, kendi

¹⁴⁹ Vladimir Cherkassky, Filip Mulier, **Lerning From Data: Concepts, Theory and Methods**, 2. Baskı, New Jersey, John Wiley & Sons, Inc., 2007, s.26.

¹⁵⁰ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 100.

içinde türdeş, fakat diğerlerinden farklı olacak biçimde, kümelere ayırma olanağı tanıyan çok değişkenli istatistiksel analiz tekniği”dir. *Segment analizi* veya *taksonomi analizi* olarak da adlandırılan kümeleme analizinin genel amacı, araştırmacıya verisini daha iyi tanıyarak, segmentasyon odaklı strateji geliştirme faaliyetlerine yardımcı olmaktır. *Küme (cluster)*, “birbirine yakın (benzer) nesnelerin çok boyutlu uzayda oluşturdukları bulutlar benzetmesi” şeklinde tanımlanabilir. *Kümeleme analizi (cluster analysis)*, “X veri matrisinde yer alan ve doğal grupları kesin olarak bilinmeyen birimleri, değişkenleri ya da birim ve değişkenleri birbirleri ile benzer olan alt kümelere ayırmaya yardımcı olan yöntemler topluluğu”dur. Kümeleme analizi, “birimleri değişkenler arası benzerlik ya da uzaklıklara dayalı olarak hesaplanan bazı ölçülerden yararlanarak homojen gruplar oluşturmaya çalışır”. Kümeleme analizi değişkenleri bağımlı ve bağımsız değişkenler olarak ayırmaz ve birini diğerine üstün tutmaz. Sonuç olarak birbirine bağımlı tüm değişkenleri inceler ve bu değişkenler arasındaki ilişkileri tespit eder¹⁵¹.

Sınıflandırma problemlerinde sınıf etiketleri önceden bellidir. Kümeleme analizinde ise küme etiketleri önceden belli değildir. Her iki sınıflandırma yönteminde de sınıf ve küme (her ikisi için ise “bölge”) sayısı, uygulamadan önce araştırmacı tarafından belirlenir. Örneğin sınıflandırmada hastaları diyabet/diyabet değil şeklinde sınıflandırırken, kümeleme de yine hastaların iki kümeye ayrılması istenir. Fakat öncesinde diyabet/diyabet değil şeklinde bir etiketleme yapılmaz. Kümeleme metodunda, uygulama yapıldıktan sonra oluşan kümeler, temel bileşenler analizinde olduğu gibi, oluşan kümelerde yer alan gözlemlerin ortak özellikleri belirlenerek etiketleme yapılır. Yani etiketleme, analiz öncesinde yapılırsa problemin adı sınıflandırma, sonrasında yapılırsa kümeleme olur. Her ikisinde, uygulama öncesinde bölge sayısı bellidir.

Sınıflandırma da gözlem sonuçları, tamamen (denetimli öğrenme) ya da kısmen (yarı denetimli öğrenme) bellidir. Model, bu sonuçlarla kendini kıyaslayarak hata fonksiyonunu, yani tahminlerindeki sapmayı, minimize ederek öğrenir ve yeni

¹⁵¹ Şebnem Koltan Yılmaz Ve Said Patir, “Kümeleme Analizi Ve Pazarlamada Kullanımı Kümeleme Analizi Ve Pazarlamada Kullanımı”, *Akademik Yaklaşımlar Dergisi*, Cilt: 2, Sayı: 1, 2011, s.91-113.

gözlemleri bu tecrübe doğrultusunda tasnif eder. Kümelemede gözlemlerin sonuçları belli değildir. Açıklayıcı değişkenler yoluyla birbirine benzeyen gözlemleri aynı kümeye, benzemeyenleri ise farklı bir kümeye ayırma mantığıyla çalışır. Gözlemler, daha çok modelde yer alan değişkenler arasındaki korelasyonlar dikkate alınarak, kümelere bölgelere atanır.

Sınıflandırmada araştırmacı örneğin hastaları diyabet ya da diyabet değil şeklinde sınıflandırırken, kümelemede iki kümeye ayırma işleminin ardından, kümelere düşen gözlemlerin ortak özellikleri üzerinden isimlendirme yapılır. Bu kez kümelerin isimleri; yatış gerektiren/gerektirmeyen hastalar, takip gerektiren/gerektirmeyen hastalar, riskli/risksiz hastalar, aynı zamanda tansiyonlu/tansiyonsuz hastalar, A/B tipi diyabetli hastalar şeklinde etiketlenmeleri de muhtemeldir. Yani, kümeleme analizi söz konusu olduğunda, gözlemleri iki ya da daha fazla gruba bölme konusunda, inisiyatif, sınıflandırmaya nispeten, biraz daha modelin kendisindedir.

Kümeleme analizi, temel bileşenler analizi gibi bir “veri azaltma (data reduction)” yöntemidir. Kümeleme de Maksimum Olabilirlik (MO) (Maximum Likelihood, ML) yöntemi kullanılır.

$f(x, \omega), \omega \in \Omega$ şeklindeki gösterge fonksiyonunu kümeleme analizinde bu kez vektör niceleyicileri (vector quantizers) olarak adlandırılır ve haritalamayı sağlar.

$$x \xrightarrow{f(x, \omega)} c(x). \quad (0.7)$$

Bu durumda hata kareleri toplamı;

$$L(y, f(x, \omega)) = (x - f(x, \omega))^2 p(x) dx \quad (0.8)$$

olur. Risk fonksiyonunun minimizasyonu ise;

$$R(\omega) = \int L(y, f(x, \omega)) = (x - f(x, \omega))^2 p(x) dx \quad (0.9)$$

formülüyle elde edilir ve gözlemlenen verilere dayanarak optimal bir vektör niceleyicisini verir. Yukarıdaki risk fonksiyonunun değerini minimize eden, yani

optimal vektör niceleyicisi yeni gözlemlerin atamasında kullanılır. Bu risk fonksiyonunu en aza indiren vektör niceleyicinin, gelecekteki verileri en uygun şekilde ölçmek için tasarlandığını unutmamak gerekir. Bu bağlamda, vektör nicelemesi bir öğrenme problemidir ve amaç veri kümesini sıkıştırmak değil, veri azaltmaktır¹⁵².

2.1.6.4. Yoğunluk Tahmini

Yoğunluk tahmini (density estimation) probleminde, makine öğrenmesi algoritmasından bir $p: R^n \rightarrow R$ fonksiyonunu öğrenmesi istenir. Burada $p(x)$, sürekli ise bir olasılık yoğunluk fonksiyonunu (probability density function), kesikli ise örneklerin çizildiği uzaydaki olasılık kütle fonksiyonunu (probability mass function) gösterir. Böyle bir işlevi yerine getirebilmek için modelin, algoritmaya sunulan verilerin genel yapısını öğrenmesi gerekir. Yani örneklerin nerede yoğun bir şekilde kümелendiğini ve nerede oluşma ihtimalinin düşük olduğunu öğrenmelidir. Sınıflandırma, kümeleme, regresyon vb. görevlerin çoğu, öğrenme algoritmasının olasılık dağılımının yapısını, en azından bulanık ya da kısmi olarak, yakalamasını gerektirir. Yoğunluk tahmini, bu dağılımı açıkça yakalamamıza izin verir. Pratikte, yoğunluk tahmini her zaman bu ilgili tüm görevleri çözmemize izin vermez, çünkü çoğu durumda $p(x)$ üzerinde gerekli işlemler hesaplama açısından zordur. Fakat bu dağılım algoritma tarafından kestirilebildiğinde, prensip olarak diğer görevleri de çözmek için, bu dağılım temel alınarak, hesaplamalar yapılabilir. Örneğin, bir makine öğrenmesi algoritması bir yoğunluk tahmini görevini yerine getirmiş ve bir $p(x)$ olasılık dağılımı elde etmiş ise, bu dağılım eksik değer bulunan diğer öğrenme görevlerini yerine getirmek için kullanılabilir. Bir x_i olarak gösterilen tüm değerler verilmiş ve yalnızca tek bir x_i değeri eksikse o zaman bu veri seti üzerindeki dağılımın $p(x_i | x_i)$ tarafından verildiği bilinir ve eksik değer çok rahat bir şekilde bulunabilir¹⁵³.

Bilinmeyen bir olasılık yoğunluk işlevini tahmin etme yöntemleri, genellikle basit temel fikirlere dayanmaktadır. Ancak tahminlerin yakınsadığı keskin gösterimler,

¹⁵² Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.28.

¹⁵³ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 103.

büyük bir dikkat gerektirmektedir. Bu temel teknik, bir x vektörünün R bölgesine düşme olasılığı, Denklem 2.10'da gösterildiği gibi¹⁵⁴;

$$P = \int_R p(x') dx' \quad (0.10)$$

hesaplanır. Dolayısıyla P , $p(x)$ yoğunluk fonksiyonunun yuvarlanmış veya ortalaması alınmış bir sürümüdür ve P olasılığını tahmin ederek bu yuvarlanmış p değerini tahmin edebiliriz. n adet $X_1, \dots, X_n$ örneklerinin, $p(x)$ olasılık yasasına göre, bağımsız ve özdeş dağılımlı (independently and identically distributed, i.i.d.) olduğunu varsayalım. Açıkçası, bu n adet verinin k tanesinin R bölgesine, aşağıda verilen binom yasasına (2.11) göre, düşme olasılığı;

$$P_k = \binom{n}{k} P^k (1-P)^{n-k}, \quad (0.11)$$

olur ve k 'nın beklenen değeri aşağıdaki gibi (2.12) olur;

$$\varepsilon[k] = np \quad (0.12)$$

Ayrıca, k için bu binom dağılımı, ortalamayla ilgili olarak, çok keskin bir şekilde yükselir, böylece k/n oranının P olasılığı ve dolayısıyla yuvarlatılmış yoğunluk fonksiyonu için çok iyi bir tahmin olması beklenir. Bu tahmin özellikle gözlem sayısı çok büyük olduğunda, iyi bir doğruluk oranı sağlar. Şimdi $p(x)$ 'in sürekli olduğunu ve R bölgesinin o kadar küçük olduğunu varsayarsak, p içinde önemli ölçüde değişiklik göstermezse, aşağıdaki bağlantı (2.13) yazılabilir;

$$= p(x)V \quad (0.13)$$

burada x , R içindeki bir noktadır ve V , R tarafından çevrelenen bir hacimdir. Buraya kadar ki eşitlikleri (2.10, 2.11, 2.12 ve 2.13) birleştirerek, $p(x)$, k için aşağıdaki açık bir tahminde (2.14) bulunabiliriz;

¹⁵⁴ Richard O. Duda, Peter E. Hart, David G. Stork, **Pattern Classification**, 2. Edition, Canada: John Wiley & Sons, INC, 2000, s. 161.

$$p(x) \simeq \frac{\frac{k}{n}}{V} \quad (0.14)$$

Eğer V sabitlenir ve çok daha fazla sayıda gözlem modele dâhil edilirse k/n oranı, olasılıksal olarak beklendiği gibi yakınsayacak ve biz bir $p(x)$ 'in bölgeye göre ortalaması alınmış bir değeri (2.15) edinilmiş olur¹⁵⁵;

$$\frac{P}{V} = \frac{\int_R p(x') dx'}{\int_R dx'} \quad (0.15)$$

Yoğunluk tahmini için benimsenen klasik yaklaşım; öğrenme makinesi tarafından desteklenen yoğunluk fonksiyonlarının sınıfı parametrik bir kümeyle sınırlanır. Ampirik Risk Minimizasyonu (ARM) tümevarım (inductive) prensibini ve eğitim verilerini kullanarak risk fonksiyonunu, ampirik olarak, tahmin eder. Ampirik risk, eğitim verileri için ortalama risktir. Ampirik risk adı verilen bu tahmin, uygun parametrelerin seçilmesiyle en aza indirgenir¹⁵⁶. ARM, DVM yönteminin altında, istatistiksel öğrenme teorisi başlığı altında ayrıntılı olarak incelenmiştir.

2.1.7. Öğrenme Yöntemleri

Makine öğrenmesi, verinin özünü elde etmeye çalışırken, öğrenme işlevini çeşitli şekillerde yerine getirebilir. Öğrenme kabaca; denetimli, denetimsiz, yarı denetimli, güçlendirici ve aktif öğrenme şeklinde sınıflandırılabilir. Bunların dışında da öğrenme yöntemleri bulunmaktadır.

2.1.7.1. Denetimli Öğrenme

Denetimli (supervised) öğrenme, bir veri kümesinin ampirik çıktısı ile hedef çıktı olarak da bilinen, istenen çıktı arasındaki doğrudan karşılaştırmaya dayanır. Yani sisteme tanıtılan sonuç verileri, örneğin bir hastanın açıklayıcı kan değerlerine karşılık aynı hastanın kanser olup olmadığının bilgisi, yani gözleme ait gerçek hayat çıktısı da

¹⁵⁵ Richard O. Duda, Peter E. Hart, David G. Stork, **a.g.e.** s.163.

¹⁵⁶ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.29.

modele tanıtılır. Dolayısıyla model, eğer kendisinin bulduğu sonuç bu çıktı ile örtüşmüyorsa, fonksiyon parametrelerini güncelleyerek hatasını minimize etmeye, böylelikle de bu tür veri kümeleri için ideal olan fonksiyonu bulmaya çalışır. Model çıktıları, sonuç çıktıları ile benzeştiğinde ise modelin eğitimi durdurulur. Yani denetimli öğrenmede, eğitim faaliyetinin durdurulmasına yine modelin içinde oluşturulan bu denetim mekanizması karar vermektedir. Bu yöntemin en önemli özelliği, gerçek hayattan alınmış girdi ve çıktı değerlerinin modele tanıtılmasıdır¹⁵⁷.

Genellikle, gerçek çıktı ile istenen çıktı arasındaki tüm ortalama veriler toplam hata kare gibi bir hata fonksiyonunun en aza indirilmesiyle formüle edilir. Regresyon ve sınıflandırma problemleri, denetimli öğrenme yöntemiyle çözülebilen problemlerdir. Yani bu iki tarz problemin çözümünde, veri kümesinde sonuç çıktılarına ihtiyaç duyulmaktadır.

2.1.7.2. Denetimsiz Öğrenme

Denetimsiz (unsupervised) öğrenmede öğrenme işlevinin gerçekleştirilebilmesi için “doğru çıktı” ile ilgili bilgi yoktur¹⁵⁸. Denetimsiz (unsupervised) öğrenme, yalnızca girdi verileri arasındaki korelasyonlara dayanmaktadır. Boyut indirgeme, özellik çıkarımı ve kümeleme yöntemleri genellikle denetimsiz öğrenir. Denetimsiz öğrenen sistemler, şimdiki bulgular üzerinden gelecekteki muhtemel doğru seçenekleri tahmin etmeyi öğrenebilirler ancak belli bir fayda fonksiyonu olmadan hangi eylemde bulunacaklarını öğrenemezler. Daha çok kümeleme problemlerinde kullanılan bu yöntemin sonuç kümeleri, sınıflandırma yönteminin aksine, eğitimden sonra etiketlenir. Denetimsiz öğrenmeye Yarışmacı Öğrenme (Competitive Learning), Kendi Kendini Düzenleyen Haritalar (Self-Organizing Maps, SOM), Hebbian

¹⁵⁷ Emre Yakut, “Veri Madenciliği Tekniklerinden C5.0 Algoritması Ve Destek Vektör Makineleri İle Yapay Sinir Ağlarının Sınıflandırma Başarılarının Karşılaştırılması: İmalat Sektöründe Bir Uygulama”, Doktora Tezi, Atatürk Üniversitesi, 2012, s.63.

¹⁵⁸ Xin Yao, “Evolving Artificial Neural Networks”, **Proceedings Of The IEEE**, Cilt: 87, Sayı: 9, 1999, s.1424.

Öğrenme, Grossberg öğrenme, Adaptif Rezonans Teorisi (ART) (Adaptive Resonance Theory, ART) gibi öğrenme kuralları örnek olarak verilebilir¹⁵⁹.

Denetimsiz öğrenme algoritmaları, birçok özelliği içinde barındıran, bir veri kümesini tecrübe eder, sonrasında bu veri kümesinde bulunan yararlı özellikleri öğrenir. Derin öğrenme söz konusu olduğunda, yoğunluk tahmininde olduğu gibi ister açık bir şekilde; isterse birleştirme veya gürültü azaltma yöntemlerinde olduğu gibi dolaylı olsun, genellikle bir veri setini oluşturan olasılık dağılımının tamamını bilmek isteriz. Dolayısıyla, bu denetimsiz öğrenme yönteminde, verilerden öğrenme yolunu takip etmeyi benimseyen yapay zekâ yöntemlerinin öğrenecek bir çıktıdan mahrum olmaları, çıktıları tahmin etme ya da sınıflara ayırma gibi görevleri yerine getiremez. Özet olarak, denetimsiz öğrenme, rastgele bir vektörün birkaç örneğini gözlemlemeyi ve olasılık dağılımını $p()$ veya bu dağılımın bazı ilginç özelliklerini zımnî veya açık bir şekilde öğrenmeye çalışmayı içerirken, denetimli öğrenme rastgele bir vektörün birkaç örneğini ve ilişkili bir değer veya vektör ve genellikle $p()$ öngörerek tahmin etmeyi öğrenme şeklindedir. Denetimsiz öğrenme ve denetimli öğrenme, resmi olarak tanımlanmış bir tanif değildir. Yani iki yöntem arasındaki ayrım çok net değildir. Her iki görevi de gerçekleştirmek için birçok makine öğrenimi teknolojisi kullanılabilir. Denetimli ve denetimsiz öğrenmenin resmi bir tanımı olmadığı gibi, veri kümelerinin veya deneyimlerin katı bir sınıflandırması da yoktur. Makine öğrenmesi kapsamında kullanılan yöntemler çoğu durumu kapsar, ancak yeni uygulamalar için yeni algoritmalar tasarlamak her zaman mümkündür¹⁶⁰.

Denetimsiz öğrenme, tamamen insan denetiminden uzak, veriden öğrenip, karar verip uygulayan bir bakış açısıyla, gerçek anlamda yapay zekâyı temsil etmekte olduğunu ve yapay zekâsı yüksek bir geleceğin potansiyelini içinde barındıran bir öğrenme yöntemi olduğunu ileri sürenler bulunmaktadır¹⁶¹.

¹⁵⁹ Tolun, Seda. “Destek Vektör Makineleri: Banka Başarısızlığının Tahmini Üzerine Bir Uygulama”, Doktora Tezi, İstanbul Üniversitesi, 2008, s.15.

¹⁶⁰ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 105.

¹⁶¹ Çetin Elmas, **a.g.e.** s.47.

2.1.7.3. Yarı Denetimli Öğrenme

Yarı denetimli (semisupervised) öğrenme yöntemi, sonuç çıktılarının bir kısmının belli olduğu öğrenme problemleridir. Eğer sonuç etiketlerinin bir kısmı mevcutsa, denetimli öğrenmeye kıyasla daha iyi sonuç vereceği beklentisiyle yarı denetimli öğrenme yöntemi uygulanabilmektedir. Regresyon ve sınıflandırma problemleri bu yöntemle de çözülebilmektedir.

2.1.7.4. Güçlendirici Öğrenme

Güçlendirilmiş (Reinforcement) Öğrenme, ağ yanıtı sonuçlarını bir fonksiyona bağlar ve bu fonksiyonu ödüller veya cezalar vererek eğitir. Yani bir sinir ağını eğitmedeki temel yaklaşım olan ağırlıkları değiştirme işlevini, ödül veya ceza verme yoluyla gerçekleştirir. Güçlendirilmiş öğrenme algoritmaları, uyarlanabilir sistemleri eğitmek için, bir dizi işlemde oluşan bir görevi yerine getirerek uygulanmaktadır. Nihai sonuç, bu diziden elde edilen sonuçtur. Bu nedenle her bir eylemin katkısı, üretilen eylem zinciri kapsamında değerlendirilmelidir¹⁶².

Genel olarak bir sınıflandırıcıyı eğitmenin en tipik yolu, bir girdi sunmak, geçici kategori etiketini hesaplamak ve sınıflandırıcıyı geliştirmek için bilinen hedef kategori etiketini kullanmaktır.

Diğer adları “pekiştirmeli öğrenme” ve “ takviyeli öğrenme” olan güçlendirici öğrenmede, istenen kategori sinyali verilmemektedir. Dolayısıyla öğretim için yapılan tek geri bildirim geçici kategori bilgisinin doğru ya da yanlış olduğudur. Bu, sadece bir tespitin doğru veya yanlış olduğunu belirten ancak neden yanlış olduğunu veya doğrusunun nasıl elde edilebileceğini belirtmeyen bir eleştirmen gibidir. Yani ödül sadece, cevap tam doğru olduğunda verilir. Örüntü tanımda, bu güçlendirmenin ikili bir şekilde yapılması en yaygın olanıdır. Yani modelin verdiği geçici karar doğrudur ya da yanlıştır. Örneğin satranç oynayan bir yazılım hamlenin doğru olup olmadığına bakmaksızın oynar. Oyunun sonunda yenilip yenilmemesine göre hamleleri doğru veya yanlış hamleler olarak belirler ve bu şekilde satranç oynamasını öğrenir. LVN,

¹⁶² Eliana Angelini, Giacomo Di Tollo Ve Andrea Roli, **a.g.e.** s.743.

bu tür ağırlara örnek olarak verilebilir¹⁶³. Hinton ve Sejnowski'nin geliştirdiği Boltzmann kuralı ve GA yine güçlendirici öğrenmeye örnek olarak verilebilir¹⁶⁴.

2.1.7.5. Aktif Öğrenme

Aktif öğrenme, üyelik sorguları yoluyla kavram öğrenmenin basit ve sezgisel bir yoludur. Bir üyelik sorgusunda, öğrenci girdi etki alanındaki bir noktayı sorgular ve bir uzman bu noktanın sınıflandırmasını döndürür. Biçimsel öğrenme teorisindeki çoğu çalışma, sorgulamaların incelenmesine yönlendirilmiştir, genelleme davranışını geliştirmede önemli role sahiptirler¹⁶⁵.

2.1.8. Model Karmaşıklığı ve Kontrolü

Veriden öğrenme yöntemlerinin genel yaklaşımı; değişkenler arasındaki ilişkilerin tespiti yani örüntü tanıma yoluyla bir model geliştirmek, sonra bu modeli yeni verilere uygulayarak muhtelif gözlemlerin sonuçlarını doğru bir şekilde tahmin etmektir. Oluşturulan bu modeli ise; ya oluşturulan bir fonksiyon ya da çok sayıdaki fonksiyonlar kümesinden en iyi sonucu vermesi için seçilmiş bir fonksiyon temsil eder. Bu fonksiyonun (yani modelin) tahmin ettiği sonuçlar, gerçek hayattaki sonuçlardan bir miktar sapma yani hata verecektir. Bu sapma bir kayıp fonksiyonuyla hesaplanır. Model olması için geliştirilen fonksiyon ise katsayıları, kayıp fonksiyonunu minimize edecek şekilde, optimize edilmiş ya da yine aynı amaç için muhtemel sonsuz sayıdaki fonksiyonların arasından seçilmiş bir fonksiyondur.

Farklı farklı yöntemlerle çok çeşitli modeller geliştirilebilmektedir. Geliştirilen bu modellerin hangisinin daha başarılı olduğunu gösteren en önemli ölçütün, modelin test verisi, yani uygulamanın başında eğitim verisinin dışında tutulmuş, modelin bilmediği bir veri kümesi söz konusu olduğunda elde edilen doğruluk oranı (accuracy rate) olduğu şimdilik nettir. Modelin başarısını sınarken aslında hangi miktarı ölçmek

¹⁶³ Richard O. Duda, Peter E. Hart, David G. Stork, **a.g.e.** s.161.

¹⁶⁴ Emre Yakut, **a.g.e.** s.64.

¹⁶⁵ David Cohn, Les Atlas Ve Richard Ladner, "Improving Generalization with Active Learning", **Machine Learning**, Cilt: 15, 1994, s.202.

istediğimizi biliriz, ancak ölçme işlemi pratik değildir. Örneğin, bu sorun yoğunluk tahmini bağlamında sıklıkla ortaya çıkmaktadır. En iyi olasılık modellerinin çoğu, olasılık dağılımlarını yalnızca zımnî ya da dolaylı olarak temsil edebilir. Bu tür modellerde, gerçek olasılık değerinin hesaplanması oldukça zordur. Bu durumlarda, tasarım hedeflerini dikkate alan bir kriter tasarlanmalı veya istenen kritere uygun bir yaklaşım tasarlanmalıdır¹⁶⁶.

Konumuza dönecek olursak birçok model, eğitim verisi dışındaki bir veri kümesine uygulandığında, doğruluk oranı ciddi şekilde düşmektedir. Genel olarak, bunun sebebi modellerin karmaşıklık düzeylerinin doğru ayarlanamamasıdır. Yani ya çok basit ya da aşırı karmaşık bir model oluşturulmuş olmasıdır. Bu da eğitim aşamasının uzunluğuyla alakalı bir durumdur. Aşırı öğrenen modeller aşırı karmaşık bir fonksiyon oluştururlarken eksik öğrenen model ise örüntüyü yeterince tanımayan basit bir fonksiyon geliştirirler.

İyi bir yapay zekâ modelinin eldeki eğitim verisine olan odaklanma kabiliyeti ile bunun dışındaki veri kümelerine genelleme kabiliyeti arasındaki dengeyi iyi kurabilmelidir. Bunun için oluşturulan model, eğitim verisinin örüntülerini mümkün olan en iyi seviyede, verinin içindeki gürültüyü de modele dâhil etmeyecek şekilde, tanımalı ki bu eğitim hatasını minimize edilecektir. Diğer yandan eğitim verisinin dışındaki (test verisi, geçerleme verisi veya tamamen bağımsız) veri kümelerine uyarlanabilecek esnekliğe sahip olmalıdır ki bu da modelin genelleme hatasının düşük tutulabilmesi ile mümkündür. Sonuç olarak, bir yapay zekâ modelinin başarısını büyük ölçüde, bu eğitim ve genelleme hatasının birlikte minimize edilebilmesi belirlemektedir. Model az karmaşık olursa, eksik öğrenme (under fitting) problemi, aşırı karmaşık olursa da bu kez aşırı öğrenme (over fitting) problemiyle karşılaşılır. Eksik öğrenen modellerin varyansı düşük, aşırı öğrenenlerin yüksektir. Eksik öğrenen bir modelin sonuçlarının değişmesi için çok fazla sayıda yeni gözlem gerekirken, aşırı öğrenen bir model için az sayıda yeni gözlem yeterlidir. Ayrıca aşırı öğrenme problemi, “ezberleme” olarak, yani bir bakıma gerçekte dersi öğrenmeyip ezberleyen,

¹⁶⁶ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 103.

dolayısıyla gösterilen dersin dışında bir soru söz konusunda olduğunda cevap veremeyen bir öğrenci profiline benzetilebilir. Sonuçta aşırı öğrenen modeller, eğitim verisine “bağımlı” hale gelirler.

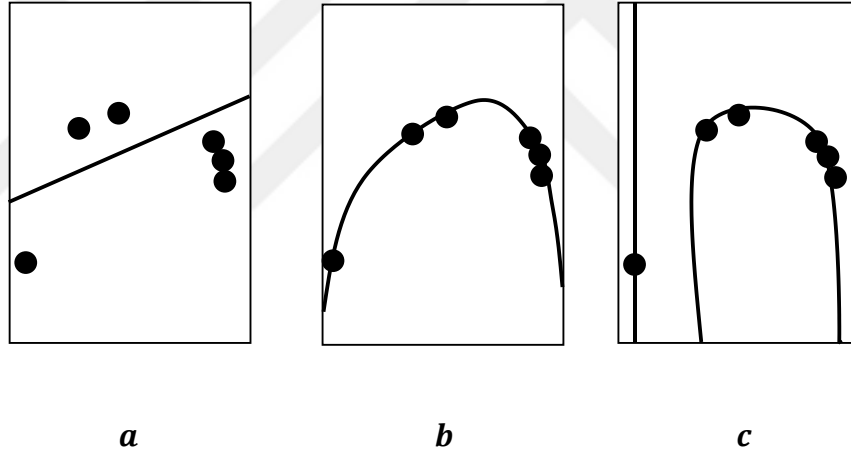
Konuyu örneklendirecek olursak, ağaç fotoğraflarıyla eğitilen ama biri eksik öğrenmiş diğeri aşırı öğrenmiş iki ayrı örüntü tanıma modeli düşünelim. Burada eksik öğrenen model, her gördüğü yeşili ağaç olarak kabul ederken, aşırı öğrenen model ise sırf yaprak sayısı farklı diye gördüğü resmin ağaç olduğunu reddedecektir¹⁶⁷. Başka bir örnek vermek gerekirse; iki ayrı emlak şirketinin (A ve B), ev değerlendirme konusunda kendilerine özgü birer model geliştirdiklerini düşünelim. A şirketinin, evin fiyatını tespit ederken sadece evin metrekaresine ve muhitine göre sınıflandırma yapan basit bir modele, B şirketinin ise evin metrekaresine ve muhitine ek olarak katı, oda sayısı, banyo sayısı, balkon sayısı, yer döşemesi, mutfak mobilyası, alışveriş merkezine, metroya, üniversiteye, merkeze mesafesi vs birçok özelliğin dâhil edildiği çok karmaşık bir modele sahip olduğunu varsayalım. Sonuç olarak A şirketi eksik öğrenecek, B şirketi ise aşırı öğrenecektir. A şirketinde varyanslar düşük, yani fiyatlandırma aralığı dar, B şirketinde ise geniş olacaktır. A şirketinin fiyatlandırma politikasının değişmesi için çok fazla sayıda yeni evin fiyatlandırması söz konusu olması gerekirken B şirketinde bu sayı çok küçük olacaktır. Bu niteliklere sahip iki şirketin birleştiğini ve fiyat değerlendirme konusunda birlikte karar verebilecekleri bir model ürettiklerini varsayalım. Oluşturulan yeni model, muhtemelen iki tür hatadan da uzak, optimal bir model olacaktır.

Bir öğrenme algoritmasının kapasitesini (karmaşıklığını) kontrol etmenin bir yolu, onun hipotez alanını, yani öğrenme algoritmasının çözüm olarak seçmesine izin verilen işlevler kümesini seçmek yani sınırlamaktır. Örneğin DR algoritması, hipotez uzayı olarak girdisinin tüm doğrusal fonksiyonlarının kümesine sahiptir. DR’u, hipotez uzayında sadece doğrusal fonksiyonlar yerine, polinomları da içerecek şekilde genelleştirebiliriz. Bunu yapmak modelin kapasitesini artıracaktır. Makine öğrenimi algoritmaları, gerçekleştirmeleri gereken görevin gerçek karmaşıklığı ve sağladıkları

¹⁶⁷ Christopher J.C. Burges, “A Tutorial On Support Vector Machines For Pattern Recognition”, **Data Mining and Knowledge Discovery**, Cilt: 2, No:2, 1998, s.122.

eğitim verilerinin miktarı açısından kapasiteleri uygun olduğunda genellikle en iyi performansı gösterir. Yetersiz kapasiteye sahip modeller karmaşık görevleri çözemez. Yüksek kapasiteye sahip modeller karmaşık görevleri çözebilir, ancak kapasiteleri mevcut görevi çözmek için gerekenden daha yüksek olduğunda gereğinden fazla çalışabilirler. Şekil 2.3'te doğrusal (a), kuadratik (b) ve dokuzuncu dereceden (c) tahminleyicilerin performansları görülmektedir. Burada a'da kullanılan modelin eksik uyum sağladığı, b'de kullanılan modelin en uygun karmaşıklık seviyesine sahip olduğu ve son olarak c'dekinin ise aşırı uyum sağladığı net bir şekilde görülmektedir¹⁶⁸.

Şekil 0.3: Polinom Dereceleri Sırasıyla Bir, İki Ve Dokuz Olan Regresyon Tahminleyicilerinin Performansları



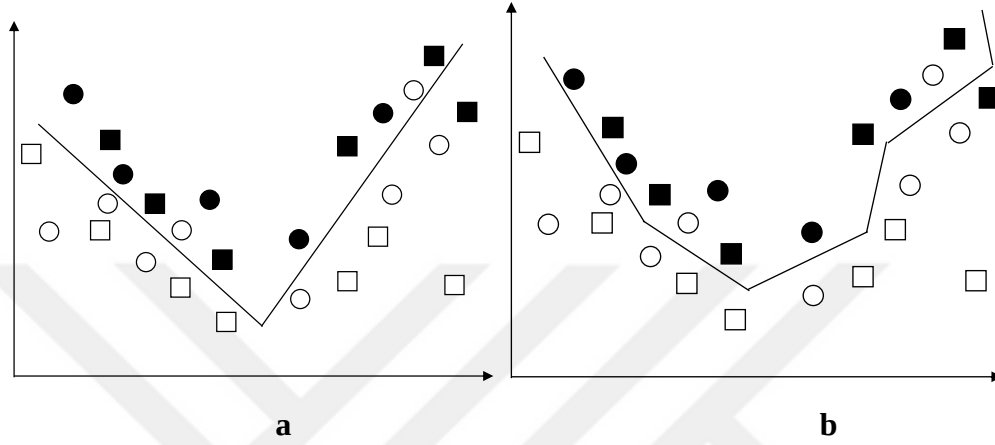
Kaynak: Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **Deep Learning**, 1. Baskı, Cambridge: The MIT Press, 2017, s. 113.

Model karmaşıklığı konusunda, daha önce Tosh ve Ruxton tarafından yapılmış bir başka uygulama, YSA uygulaması ve bu uygulamanın Şekil 2.4'teki grafiği bize normal bir model ile aşırı öğrenmiş bir modelin yapılarını yanyana göstermektedir. Şekil 2.4a'da görülen grafikte iki gizli katmanlı bir sinir ağı kullanılmış ve sadece iki tane gözlem yanlış sınıflandırılmıştır. 2.4b'de ise altı gizli katmanlı bir ağ kullanılmış ve model, veri kümesinin altında yatan fonksiyonel ilişkiyi yani örüntüyü de dikkate

¹⁶⁸ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 112.

almadan aşırı öğrenmiş ve bu onun sonuç performansına olumlu olarak yansımamış ve yine iki gözlemi hatalı olarak sınıflandırmıştır¹⁶⁹.

Şekil 0.4: Sırasıyla İki ve Altı Gizli Katmana Sahip YSA Modelleri ile Yapılmış Sınıflandırma Sonuçları



Kaynak: Colin R. Tosh ve Graeme Di Ruxton, **Modelling Perception With Artificial Neural Networks**, 1. Baskı, Cambridge: Cambridge University Press, 2010, s.17.

Eğitim verisinin boyutu değiştikçe eğitim ve genelleme hatası da değişir. Eğitim örneklerinin sayısı arttıkça, beklenen genelleme hatası ise asla artmamalıdır (asimptotik tutarlılık). Parametrik olmayan modeller için, mümkün olan en iyi hata elde edilene kadar daha fazla veri daha iyi genelleme sağlamalıdır. Optimal kapasiteden daha düşük kapasiteye sahip, sabitlenmiş herhangi bir parametrik model, Bayes hatasını aşan bir hata değerine asimptotik olarak yakınsayacaktır. Örnek için bkz. Şekil 2.5. Modelin optimum kapasiteye sahip olmasının mümkün olduğunu ve yine de eğitim ve genelleme hataları arasında büyük bir boşluk olduğu unutulmamalıdır. Bu durumda, daha fazla eğitim örneği toplayarak bu boşluk azaltılabilir¹⁷⁰.

Aşırı öğrenme problemi yaşanan uygulamalarda model, eğitim verisini çok iyi analiz edip çok iyi öğrenir adeta ezberler. Fakat yeni bir veri kümesi söz konusu olduğunda, model aşırı nüfuz ettiği eğitim verisinden yeni veri kümesi için genelleme

¹⁶⁹ Colin R. Tosh ve Graeme Di Ruxton, **Modelling Perception With Artificial Neural Networks**, 1. Baskı, Cambridge: Cambridge University Press, 2010.

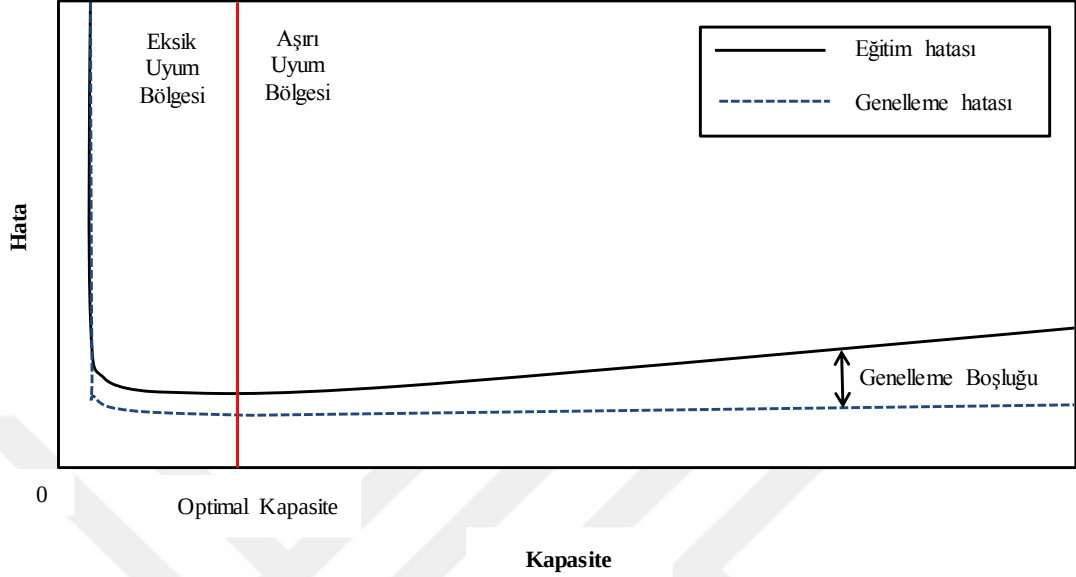
¹⁷⁰ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, a.g.e. s. 114.

yapamayacaktır. Eksik öğrenen modellerde ise, evet genelleme yeteneği çok yüksek olacaktır fakat modelin uygulanan yeni veri setleri üzerindeki odaklanma yeteneği ise zayıf olacaktır. Bu yüzden modellerin karmaşıklık düzeyi iyi ayarlanmalıdır.

Modelin kapasitesinin (modelin hipotezinin karmaşıklık seviyesinin) ölçülmesi, istatistiksel öğrenme teorisinin nicel tahminler yapmasını sağlar. İstatistiksel öğrenme teorisindeki en önemli özellik, eğitim hatası ile genelleme hatası arasındaki tutarsızlığın yukarıdan modelin karmaşıklığı (kapasitesi) büyüdükçe büyüyen ancak eğitim örneklerinin sayısı arttıkça küçülen bir miktarla sınırlandığını göstermesidir. Bu sınırlar, makine öğrenimi algoritmalarının çalışabileceğine dair bilgiye dayalı bir gerekçe sağlar, ancak öğrenme algoritmalarıyla çalışırken pratikte nadiren kullanılmazlar. Tipik olarak, eğitim hatası, model kapasitesi arttıkça olası minimum hata değerine asimptot olana kadar azalır (hata ölçüsünün minimum bir değere sahip olduğu varsayılarak). Tipik olarak genelleme hatası, model kapasitesinin bir fonksiyonu olarak U şeklinde bir eğriye sahiptir. Bu, Şekil 2.5'te gösterilmektedir. Parametrik modeller, herhangi bir veri gözlemlenmeden önce boyutu sonsuz ve sabitlenmiş bir parametre vektörü tarafından tanımlanan bir işlevi öğrenir. Parametrik olmayan modellerin böyle bir sınırlaması yoktur. Bazen, parametrik olmayan modeller, yalnızca teorik soyutlamalardır (olası tüm olasılık dağılımlarını araştıran bir algoritma gibi) ve pratikte uygulanmaktadır. Şekil 2.5'e bakıldığında eğitim ve test hatasının farklı seyrettiği görülmektedir. Grafiğin sol ucunda eğitim hatası ve genelleme hatası yüksektir. Bu, en düşük verimdir. Kapasiteyi artırdıkça hatalar azalır, ancak eğitim ile genelleme hatası arasındaki boşluk artar. Sonunda, bu boşluğun boyutu, eğitim hatasındaki azalmadan ağır basar ve kapasitenin çok büyük olduğu, optimum kapasitenin üzerinde aşırı uyum bölgesine girilir.¹⁷¹ Dolayısıyla bu kapasite artışını maksimize etmemek ve hem eğitim hem de genelleme hatası arasındaki makas artmaya başlayınca durarak optimize etmek gerekmektedir. Bu şekilde problemin karmaşıklık seviyesine en uygun karmaşıklıkta modelin oluşturulmasına bir adım daha yaklaşılabacaktır.

¹⁷¹ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, a.g.e. s. 113.

Şekil 0.5: Kapasite Ve Hata Arasındaki Tipik İlişki



Kaynak: Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **Deep Learning**, 1. Baskı, Cambridge: The MIT Press, 2017, s. 113.

Modellerin karmaşıklık düzeyinin ayarlanmasında, bugüne dek geliştirilen yöntemlerde felsefi açıdan iki temel yaklaşım, Occam'ın usturası ve Popper'ın yanlışlanabilirliği etkili olmuştur. Modellerin karmaşıklığının optimize edilmesi konusunda şimdiye dek, budama, bırakma (Bkz. Şekil 2.24) ceza puanı, Bayesçi çıkarım vs. yöntemler, uygulanmış fakat bu yöntemler matematiksel olarak kapsamlı bir teoriden yoksun olup, optimum sonucu garanti edememektedir. Tabii ki Vapnik ve Chervonenkis tarafından geliştirilen İstatistiksel Öğrenme Teorisi (Statistical Learning Theory, SLT) bundan hariç tutulmalıdır. Çünkü istatistiksel kapsamı ve ispatlanmış matematiksel arka planı ile model karmaşıklığını optimize edebilen ve bu sayede eğitim ve genelleme hatasını birlikte minimize edebilen bir yöntem olarak geçtiğimiz bir-iki on yıldır sınıflandırma ve regresyon uygulamalarında sıklıkla başvurulmakta ve literatür taramasından da anlaşılacağı üzere iyi sonuçlar vermektedir. Model karmaşıklığı ve bunun optimizasyonu ile alakalı fikirlere, ilk iki başlıkta (Occam'ın usturası ve Popper'ın yanlışlanabilirliği) tartıştıktan sonra bu felsefi yaklaşımlara matematiksel zemin kazandıran istatistiksel öğrenme teorisinin yaklaşımına ve daha sonraki bölümlerde içeriğine bakmak tüm bu konuların anlaşılması adına takip edilmesi gereken bir yol haritası sunacaktır.

2.1.8.1. Occam'ın Usturası

Ockhamlı William, 1287 ve 1347 yılları arasında yaşamış, skolastik, İngiliz filozoftur. Ockham, genel olarak mantık, fizik, teoloji ve poitika alanında önemli çalışmalar yapmış olmasına rağmen, daha çok geliştirdiği Occam'ın Usturası isimli metodolojik prensiple anılmaktadır¹⁷².

Occam'ın usturası, “bir model eğer daha az karmaşık bir modelle çözülebiliyorsa, modele ait unsurlar, gerekliliğin ötesinde çoğaltılamaz” görüşünü benimsemektedir¹⁷³.

İşletmelerin yönetim ve organizasyonunda, “kullanılmayan veya lazım olmayan her şeyi sistemin dışında bırakma ve sistemi daha kompakt ve daha işler hale hale getirmek için daha yalın bir hale getirme” mantığını taşıyan yalın organizasyon modern yönetim yaklaşımını anımsatmaktadır.

Bu teori, başka bir benzetme ile işe elverişli olmayan, yağlı ve hantal bir insan vücudunu, sağlıktan ödün vermeden, daha atik ve kaslı bir yapıya dönüştürme çabası olarak yorumlanabilir. Böylelikle yapılan işlerde hız, esneklik ve işlevsellik kazanma amaçlanmaktadır.

2.1.8.2. Popper'ın Yanlışlanabilirliği

Popper'ın yanlışlanabilirliği (Popper's falsificationism), VC Boyutu (Vapnik – Chervonenkis) ya da buna benzer matematiksel öngörülerden çok, genel olarak bilimsel keşiflerin, metodolojik tutarlılık sınırlarını test etme noktasında önemli bir bilim felsefesi yaklaşımı içermektedir. Hatta Popper'a göre “yanlışlanabilirliğin tedavisi” olarak isimlendirilen bu yaklaşım, zamanla bilim çevrelerinde popülerleşmiştir. Popper bu yaklaşımını, 1934 tarihli Araştırma Mantığı (Logik der Forschung) adlı kitabında, temellendirmiştir¹⁷⁴.

¹⁷² Wikipedia, Çevrimiçi, https://tr.wikipedia.org/wiki/Ockhaml%C4%B1_William, 1004/2020.

¹⁷³ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.147.

¹⁷⁴ David Corfield, Bernhard Schölkopf, Vladimir Vapnik, “Popper, Falsification And The VC-Dimension”, **Max Planck Institute for Biological Cybernetics**, Teknik Rapor, Sayı: 145, 2005, s.1.

1959'da ilk kez İngilizceye çevrilen kitabında, Popper; bilimi diğer iddialardan ayıran tek faktörün, yapılacak tahrifata, dayanıklılık olduğunu belirtmiştir. Yani öne sürülen bilimsel keşif eğer doğruysa, tahrif edilme riskine karşı koyma yetisine sahip olması gerektiğini söylemiştir. Popper bunu yaparak Viyana Çemberi (Viyana Çevresi olarak da adlandırılan ve mantıksal pozitivizmi savunan ve 20. YY.'da çok önemli etkileri olan bir bilimsel kurul olup iki amacı vardır. Bunlardan birincisi, dünyanın bilimsel kavranışında metafizik öğelerin ve teolojik unsurların kuramsal olarak doktrinden arındırılmasıdır. İkincisi ise felsefeye bilimsel bir nitelik kazandırılmasıdır) üyelerinden (Hans Hahn, Kurt Gödel, Felix Kaufmann, Philip Frank, Bertnard Russel, Whitehead, A.J. Ayer, Wittgenstein)¹⁷⁵ farkını ortaya koydu. Einstein'ın, test edilebilir tahminlerle, cesur varsayımlar yapma isteğinden etkilenen Karl Popper, Einstein'ı paradigmatik, yani modelci bir bilim adamı olarak görmektedir. Örneğin, başta güneşten teleskoplarımıza gelen yıldız ışığının, güneşin yakınından geçerken, belli bir miktar büküldüğü bulundu. Deneysel hata söz konusuysa, yani bükülmenin gerçekleşmediği kesin olarak tespit edilirse, bu durumda Einstein, gözlemlerin düzgün bir şekilde yürütüldüğünden emin olduğu müddetçe, genel görelilik teorisinden vazgeçmeye, doğal olarak hazır olacaktır. Öte yandan, belli bir miktarda bükülme gözlenirse, bu Einstein'ın söz konusu teorisini doğrulamak için kesinlikle yeterli olmayacaktır. Yani Popper'a göre, bilimsel keşfi çürütmek için keşfin aleyhine tek bir bulgu yeterliyken, lehine tek bir bulgu ile keşfin doğruluğu ispatlanamaz. O'nun için hayati olan temel düşünce şuydu: *“evrensel ifadeler (varsayımlar) tahrif edilebilir, fakat doğrulanamaz”*. Popper bir tahmin doğrulandığında, bilimsel bir teorinin daha olası hale geldiği fikrini içselleştiremiyordu. Yani tek bir bulguya dayanarak bilimsel bir keşfin genel-geçer bir doğruluğa sahip olamayacağı düşüncesindeydi. Dolayısıyla, Popper'a göre, söyleyebileceğimiz tek şey; “teorinin ciddi şekilde test edilmiş veya iyi bir şekilde doğrulanmış” olabileceğiydi. Bu arada Viyana Çemberinde yer alan kişiler ve konuyla ilgilenen mantıkçı ampiristler, doğrulama ve tahrifatın bu simetrik olmayan çözümüne katılmadılar ve olasılıksal bir çerçevede bir doğrulama teorisi geliştirdiler. “Bilimsel Keşif Mantığı (The Logic of Scientific Discovery)” adı verilen bu teorinin önemli bir çoğunluğu, bilginin söz konusu olduğu durumlarda, yapılan keşfin doğruluk

¹⁷⁵ Wikipedia, Çevrimiçi, https://tr.wikipedia.org/wiki/Mant%C4%B1ksal_pozitivizm, 07/04/2020.

olasılıklarını göz önüne alma yaklaşımıyla ele alınmaktaydı. Böylece Popper zaman içinde, Bayes karşıtı bilim filozoflarının toplanma noktası haline geldi¹⁷⁶.

Bir dizi sınıflama hipotezinin VC (Vapnik-Chervonenkis) boyutu, en büyük doğal sayıdır. n , bir dizi n noktası olacak, ancak '+' veya '-', olarak etiketlenecek bu etiketlemeyi kabul eden bir hipotez var. Nokta kümesinin hipotez sınıfı tarafından parçalandığı söylenir. Örneğin, yarım düzlem kümesinin VC-boyutu üçtür. Çünkü üç doğrusal olmayan nokta kümesinden herhangi biri bu sınıflandırıcılar tarafından parçalanırken, dört nokta kümesinden herhangi biri aynı sınıf tarafından parçalanamaz (Bakınız Şekil 20). Bir boyut daha yüksek, dolayısıyla VC boyutu dört olur. VC boyutu aşağıdaki şekilde kullanılır. Belirli bir popülasyondaki büyük bir insan örneğinin ağırlık, boy ve bel boyutunun ölçüldüğünü hayal edin. Bir araştırmacı, yukarıdaki ölçümleri girdi olarak kullanarak erkekleri bu popülasyondaki kadınlardan doğru bir şekilde ayırabilen doğrusal bir sınıflandırıcı (ayırma düzlemi) oluşturmak istemektedir. Bu sınıflandırıcı sınıfının VC boyutu dördür. Veriler üç boyutlu uzayda çizilirse ve aynı düzlemde yer almayan dört nokta seçilirse, o zaman bu noktalara karşılık gelen insanların cinsiyetleri ne olursa olsun, erkekleri dişilerden ayıran bir düzlem bulunabilir. Şimdi, büyük örneğimizi alır ve doğru bir doğrusal sınıflandırıcı bulursak, bir dizi hipotez tarafından parçalanmış bu nokta fikri, Karl Popper'ın yanlışlanamazlık fikrini akla getirebilir. Doğrusal bir sınıflandırıcının, etiketlenmiş olmalarına rağmen, dört boyutlu bir örneğe bakarak cinsiyetleri ayırabileceği teorisini tahrif edemediği görülmektedir. Sonuç olarak Popper'ın yanlışlamacılığa getirdiği çözüm, istatistiksel öğrenme teorisi için önemli bir ilham kaynağı olmuştur. Şöyle ki "tümevarım" ilkesini bilimsel kabul etmeyen Popper'ın yanlışlanabilirliği açısından olaya baktığımızda, Vapnik vd. VC boyutunun "genelleme hatası" mantığıyla adeta Popper için bu katsayı, bir çeşit, "yanlışlanabilirliğin ölçüsü" olmuştur. Yani Popper'a göre tümevarımı bilimsellikten uzaklaştıran tümevarım hatasını, genelleme hatasını belirleyerek, kayıt altına aldığı yine Vapnik vd. tarafından belirtilmektedir¹⁷⁷.

¹⁷⁶ David Corfield, Bernhard Schölkopf, Vladimir Vapnik, **a.g.e.** s.1.

¹⁷⁷ David Corfield, Bernhard Schölkopf, Vladimir Vapnik, **a.g.e.** s.1.

2.1.8.3. VC Katsayısı –Teorisi

Son on yılda, makine öğrenimi alanında bir paradigma değişikliği meydana geldi. Bu süre zarfında, geleneksel istatistiksel yöntemlerin tipik olarak başaramadığı, son derece yüksek boyutlu verilerin tümevarımsal (endüktif) olarak işlenmesine izin veren, bir dizi yeni teknik ortaya atılmıştır. Geliştirilen modellerin, geleneksel yaklaşımda olduğu gibi verileri temsil etmek için bir model oluşturmaktan ziyade, bu makine öğrenimi tarzı, girdileri doğru bir şekilde etiketlemek için girdiler arasında ayrım yapabilmeyi amaçlamaktadır. Bu tür sınıflandırıcıların üretimini tümevarımsal öğrenme örnekleri olarak düşünürsek, bu terimin kapsamı, uzun zamandır çalışması felsefi çözümlerin merkezi olan bilimsel yasaların ve modellerin üretiminin ötesine genişlemiştir. Bu ayrımcı sınıflandırıcıların teorik temeli “İstatistiksel Öğrenme Teorisi” olarak bilinir. Burada, tümevarımsal öğrenmenin sağladığı en önemli fayda olan, “doğruluk ve aşırı uyum arasında önemli bir denge kurulması”, teorinin var olma sebebidir. Aşırı vukufiyet, bir veri kümesini temsil etmek için, çok zengin bir hipotez alanı kullanıldığında ortaya çıkar. Bir örnek vermek gerekirse, az sayıda veri noktasına yüksek derecede bir eğrinin yerleştirildiği regresyon örneğidir. O zaman bu eğrinin yeni veri noktaları üzerinde iyi tahminler yapması beklenirken, daha serbestçe oturan düz bir çizgiye daha çok güvenilebilir. Şimdi, istatistiksel öğrenme teorisinde hipotez alanının zenginliği bir eğrinin derecesi veya bir hipotezin parametre sayısı ile değil, VC boyutu olarak bilinen bir yapı ile kontrol edilir. Bu boyutun, göreceğimiz gibi, bir yanlışlanabilirlik derecesini ölçmesi düşünülebilir. O zaman Karl Popper’ın isminin istatistiksel öğrenme teorisi tartışmalarında ortaya çıkması şaşırtıcı olmamaktadır. Tam olarak bu konuda bilimsel ve makine öğrenimi arasındaki benzerliklerin ortaya çıkmasını beklenmelidir. Son olarak istatistiksel öğrenme teorisi, Popper’ın yaklaşımını desteklemektedir fakat Popper’ın, istatistiksel öğrenme teorisinin gelişimine, bakış açısı nedeniyle, bir katkısının olduğunu söylemek doğru değildir¹⁷⁸.

¹⁷⁸ David Corfield, Bernhard Schölkopf, Vladimir Vapnik, **a.g.e.** s.1.

VC boyutu, hesaplanma yöntemi, model karmaşıklığına etkisi vb. konunun birçok yönü DVM başlığı altında, ayrıntılı olarak tartışıldığı için matematiksel hesaplamalar buraya alınmamıştır.

2.1.8.4. Model Karmaşıklığının Optimizasyonu Konusunda Farklı Görüşler

Aşırı uyumdan kaçınma, birçok araştırmacı için önemli bir problemdir. Bu problem Occam'ın Usturası gibi, teori niteliği kazanmadan önce de; düzenlileştirme, budama, Bayesçi çıkarım, ceza şartlarının dâhil edilmesi ve bir açıklama uzunluğunun en iyi hale getirilmesi vb. çeşitli yöntemlerle aşılmaya çalışılmıştır. Uygulama sahasında, kullanılacak olan algoritmayı sınırlama veya detaylarını azaltma gibi bir evrensel eğilim hep söz konusu olmuştur. Fakat unutmamak gerekir ki, modelin detaylarını budamak, sonuçlarda bir miktar sapmaya da sebep olacaktır. Bir algoritmayı diğerine tercih etmek için probleme özgü açıklayıcı nedenler yoksa modelin karmaşıklığından kaçınmanın da makul bir sebebi kalmamaktadır. “Ampirik başarıyı belirleyen, öğrenme algoritmasının problemle eşleşmesidir; aşırı uyumdan kaçınmanın dayatılması değil”. Ayrıca her model önce en basit şekliyle geliştirilir. Örneğin YSA ilk başta, perseptron dediğimiz basit algılayıcı şeklinde geliştirildi. Fakat perseptron, örüntü tanıma problemini çözemez. Hoparlörden açıklayıcı bir ses tanıma işlemi, basit tek katmanlı bir YSA ile çözülemez. Yani karmaşıklık ihtiyaç neticesinde oluşmuş bir sonuç olup, minimize edilmesi gereken değil, bundan farklı olarak optimize edilmesi gereken bir süreçtir. Fazla kompleks olduğu için mevcut problemi iyi çözemeyen algoritmalar da olmuştur. Sonuç olarak burada yapılması gereken bir modeli budamaktan çok, probleme en uygun modelin ve mimarinin oluşturulmaya çalışılmasıdır. Yani bir bakıma kilolu bir gövdeyi, kaslı bir vücuda dönüştürmek, evet istenen bir durumdur. Fakat sağlıktan ve işlevsellik kapasitesinden ödün vermeden yapılırsa, bu sadeleştirme faydalı olacaktır. Öğrenme algoritmaları yoluyla birçok makine öğrenmesi modeli geliştirilebilmektedir. Bu algoritmalarından bazıları, diğerlerine kıyasla, daha çok tercih edilebilmektedir. Bunun önemli bir sebebi de bu algoritmalarından bazıları hem model olarak uygulaması çok basit, yani anlaşılabilirliği çok kolay, hem de bilgisayarlara getirdiği zaman maliyetinin

düşüklüğüdür. Diğer yandan karmaşık algoritmalar ise, diğer verinin içinde yer alan ilişkileri daha basit olan algoritmalara nazaran daha iyi keşfedebilmekte, bunlar da modelin ürettiği sonuçların doğruluk oranına katkıda bulunmaktadır. Örneğin sınıflandırma konusunda “hangi sınıflandırıcı, neden tercih edilmelidir?”, sorusuna verilecek cevap; yöntemin, gereksiz karmaşadan uzak ama doğruluk oranı yüksek, boyut lanetinden uzak ama öğrenme kabiliyeti yüksek, az zaman alan, güvenilir, tutarlı ve uyarlanabilir olması gibi unsurları kapsamalıdır¹⁷⁹. Tüm bunların yanı sıra bir modelin anlaşılması ve uygulanması ne kadar kolay olursa, kullanımı da o denli yaygın olacaktır.

Makine öğrenimi için, Wolpert tarafından 1996 yılında ileri sürülen, “bedavaya öğle yemeği yok (The no free lunch)” teoremi genel olası veri üreten dağılımların ortalamasını alarak, her sınıflandırma algoritmasının daha önce gözlemlenmemiş noktaları sınıflandırırken aynı hata oranına sahip olduğunu belirtir. Başka bir deyişle, bir anlamda, hiçbir makine öğrenimi algoritması evrensel olarak diğerlerinden daha iyi değildir. Düşünebildiğimiz en gelişmiş algoritma, yalnızca her noktanın aynı sınıfa ait olduğunu tahmin etmekle aynı ortalama performansa (tüm olası görevlere göre) sahiptir. Gerçek dünya uygulamalarında karşılaştığımız olasılık dağılımlarının türleri hakkında varsayımlar yaparsak, bu dağılımlarda iyi performans gösteren, öğrenme algoritmaları tasarlayabiliriz. Bu, makine öğrenimi araştırmasının amacının evrensel bir öğrenme algoritması veya mutlak en iyi öğrenme algoritması aramak olmadığı anlamına gelir. Bunun yerine amacımız, bir yapay zekâ temsilcisinin deneyimlediği "gerçek dünya" ile ne tür dağıtımların alakalı olduğunu ve ilgilendiğimiz veri üreten dağıtım türlerinden alınan verilerde ne tür makine öğrenimi algoritmalarının iyi performans gösterdiğini anlamaktır¹⁸⁰.

2.2. Yapay Sinir Ağları

Gerekli bilgileri doğrudan ilgili bir uzman veya uzman grubundan çıkarmak, oldukça maliyetli, hatalara açık ve çok zaman alıcı olabilmektedir. Hatta çoğu zaman uzman

¹⁷⁹ Richard O. Duda, Peter E. Hart, David G. Stork, **a.g.e.** s.453.

¹⁸⁰ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 114-116.

sistemlerin zamanında uygulanması mümkün olmayabilir. Bu bilgi mühendisliği sorunsalı, bilgi edinme sürecini otomatikleştiren uzman sistemlerin inşasını gündeme getirmiştir. Uzman görüşü denilince esasında, muhtelif uzmanlar, problemleri farklı farklı tanımlayıp çözümü için farklı değişkenleri referans alsalar da, ortalama olarak uzman görüşleri arasındaki farklar azdır. Örnek referans verilerinden hareketle, otomatik bilgi edinimi için yaygın olarak kullanılan bir yaklaşım, tümevarım (induction) olarak adlandırılır. Tümevarım yaklaşımında, bir karar örneği modelleme sistemi oluşturulur. Sistem daha sonra, örnek verilere dayanarak yaklaşık bir uzman modeli oluşturur. Sinirsel ağ tabanlı, kullanımı kolay ve gittikçe daha fazla bilgi işlem gücü elde edilebilen uzman sistem uygulamalarının kullanımı zamanla artmakta ve muhtelif disiplinlerde yapılan çalışmaların sayısı hızla yayılmaktadır. Bu artışın bir sebebi de, anlaşılabileceği üzere sinir ağlarının, tümevarımsal (inductive) öğrenme stratejisini, yani genelleme fonksiyonunu temsil ediyor olmasıdır¹⁸¹. YSA da bir makine öğrenmesi algoritması olup, kullanımı birçok alanda yaygın olan, son derece etkili bir öğrenme yöntemidir.

YSA, 1940'lı yıllardan itibaren geliştirilmekte olan, kapsamlı bir yöntem olup; “diğer yapay zekâ yöntemlerinde olduğu gibi, örneklerden öğrenen ve bu öğrenme olayını düğümler arası bağlantı ağırlıklarını ayarlayarak gerçekleştirebilen ve sigmoid ile Gaussian fonksiyonlarında olduğu gibi, nonlinear ve iteratif bir yöntem”¹⁸² şeklinde tanımlanabilir.

YSA, “beynin sinir hücrelerinin yapısına dayanan, nispeten kaba elektronik modellerdir. Beyin gibi, temelde deneyimlerden öğrenir. Mevcut bilgisayarların kapsamı dışındaki bazı sorunların, gerçekten de küçük enerji verimli paketlerle çözülebileceğinin doğal bir kanıtıdır. Bu beyin modellemesi, aynı zamanda makine çözümleri geliştirmek için daha az teknik olan (sade) bir yol vaadediyor”.

¹⁸¹ Louis W. Glorfeld, **a.g.e.**, s.37.

¹⁸² Xin Yao, **a.g.e.** s.1423-1424.

Hesaplamaya yönelik bu yeni yaklaşım, sistemin aşırı yüklenmesi sırasında, daha geleneksel olan muadillerinden daha zarif bir bozulma sağlamaktadır¹⁸³.

Bunun dışında kullanılan bazı YSA tanımları şu şekildedir¹⁸⁴:

- “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”.
- “Yapay sinir ağı, basit hesaplama elemanlarının yoğun bağlantılarından meydana gelmiş bir bilgi işleme elemanıdır”.
- “YSA insan idrakinin ve biyolojik sinirlerinin matematiksel modelinin genelleştirilmesi olarak tanımlanabilir”.
- “Bir sinir ağı, basit işlem birimlerinden oluşan, deneyimsel bilgileri biriktirmeye yönelik doğal bir eğilimi olan ve bunların kullanılmasını sağlayan yoğun bir şekilde paralel dağıtılmış bir işlemcidir. Bu işlemcinin insan beyni ile ilişkisi iki şekilde açıklanabilir:
 1. İşlemci tarafından öğrenilen bilgi, belirli bir öğrenme işlemi sayesinde edinilir,
 2. İşlemciyi oluşturan yapay nöronlar arasındaki bağlantı ağırlıkları, edinilen bilginin saklanması için kullanılır”.
- “YSA, insan beyninin özelliklerinden olan, öğrenme yolu ile yeni bilgiler türetebilme, yeni bilgiler oluşturabilme ve keşfedebilme gibi yetenekleri herhangi bir yardım almadan otomatik olarak gerçekleştirmek amacıyla geliştirilen bilgisayar sistemleridir”
- “YSA kavramı beynin çalışma ilkelerinin sayısal bilgisayarlar üzerinde taklit edilmesi fikri ile ortaya çıkmış ve ilk çalışmalar beyni oluşturan biyolojik hücrelerin, ya da literatürdeki ismi ile nöronların matematiksel olarak modellenmesi üzerinde yoğunlaşır”.

¹⁸³ Dave Anderson and George McNeill, “Artificial Neural Networks Technology”, A DACS State-of-the-Art Report, 1992, s. 2.

¹⁸⁴ Emre Yakut, **a.g.e.** s.52-53.

Bilgisayarlar sayısal hesaplamalar söz konusu olduğunda, insan beyninden çok daha hızlıdır ve doğrulukları açısından da bir o kadar güvenilirlerdir. Fakat sayısal bilgisayarlar, yani günlük kullandığımız bilgisayarlar, insan beyni gibi, idrak edebilmeyi gerektiren, olgular arası bağlantılar kurabilme ve en etkin kararı verebilme gibi özel bazı yeteneklerden mahrumdur. Bunlar; çabucak tanıma, verinin çok açık olmadığı durumlarda bile anlama ve yorumlama, geçmiş tecrübelerle dayanarak muhakemesini yapabilme ve bunları daha önce hiç karşılaşılmamış olan durumlarla ilişkilendirme ve çıkarsamalar yapabilme yetenekleridir. YSA, beyni taklit etmek suretiyle, bilgisayarların bu eksikliklerinin üstesinden gelmeye çalışmaktadırlar. YSA'nın, kendi içinde kurduğu modelin, doğrusal olmayı ya da parametrik olmayı gerektirmemesi, YSA'nı çok boyutlu veri setlerindeki karmaşık ilişkileri ortaya çıkarmada son derece başarılı bir yaklaşım yapmaktadır. YSA, ilişkilerin özelliğiyle ilgili herhangi bir ön varsayımda bulunmayı gerektirmemektedir. Dolayısıyla bu özellikleriyle YSA istatistiki yöntemlerden ayrılmakta ve istatistik yöntemlere nispeten daha kullanışlı hale gelmektedir. YSA'nın en güçlü özelliği, verilerden, bu veriler arasındaki ilişkileri, önemlerine nispeten ağırlıklandırarak, öğrenebilme yetisinin olmasıdır¹⁸⁵.

2.2.1. Yapay Sinir Ağı Yaklaşımı

YSA, “kendisine sunulan verinin doğasında bulunan ilişkileri öğrenir, çünkü bu yaklaşım, geleneksel istatistiksel model geliştirme, doğrusal olmayan ilişkileri arama ve yordayıcı değişkenler arasındaki etkileşimleri saptama gibi zaman alıcı manuel veri inceleme etkinliklerini içemektedir”¹⁸⁶.

YSA, bir bakıma “paralel bilgi işleme sistemleri” olarak da adlandırılabilir. YSA'nın işlediği bu bilgiler, ilgili olaya ait, daha önceden gerçekleşmiş gözlemler aktararak, YSA sistemi eğitilir. Böylece, bilinen örneklerle tespit edilmiş bazı önemli özellikler üzerinde genelleştirmeler yapılarak, daha sonra sonuçları bilinmeyen fakat bu önemli

¹⁸⁵ Talip Torun, “Finansal Başarısızlık Tahmininde Geleneksel İstatistiksel Yöntemlerle Yapay Sinir Ağlarının Karşılaştırılması Ve Sanayi İşletmeleri Üzerinde Uygulama”, Doktora Tezi, Erciyes Üniversitesi, 2007, s.47-49

¹⁸⁶ Desai, Vijay S., N. Crook Jonathan ve A. George Overstreet, a.g.e. s. 25.

özellikleri tespit edilmiş örneklerin sonuçlarının ne olacağı tahmin edilir¹⁸⁷. Yani, YSA, normalde bir insanın yaşayarak veya deneyerek öğrendiği gibi, tahmin yapma yeteneğini sahip olduğu öğrenme yeteneğiyle, her yeni örnekte artırarak, gerçekleştirirler¹⁸⁸.

YSA'nın bugünkü modellerinin üretilmesi noktasında bir çıkış noktası olan, Frenk Rosenblatt'ın, ilk olarak 1958 yılında ortaya attığı perseptrondur. Perseptornun yapısının anlaşılması için, önce sinir hücrelerinin yapısının anlaşılması gerekir. YSA'nın gelişim sürecini ele aldıktan sonra, yöntemin temel yapı taşları olan biyolojik sinir hücrelerinin yapısına ve sinir hücrelerinin birer taklidi olan basit algılayıcıların çalışma prensiplerine değinilecektir.

2.2.2. Yapay Sinir Ağlarının Gelişim Süreci

YSA'nın gelişim sürecinde bazı önemli olaylar, yöntemin bugüne gelmesinde önemli dönüm noktalarını teşkil etmektedir. Bunlar¹⁸⁹:

- 1890 – İnsan beyninin yapısı ve fonksiyonları ile ilgili ilk yayın. 1890 yılında Alman bilim adamı Heinrich Wilhelm Gottfried von Waldeyer-Hartz beynin sinir hücrelerinden oluştuğunu ve sinir hücresinin bugünkü bilinen yapısını kabaca belirledi. Ayrıca nöron (neuron) ve kromozom (chromosome) kelimelerinin, tarihte ilk kez, bu yayında kullanılması,
- 1911 – İnsan beyninin bileşenlerinin belirli bir düzenek ile sinir hücrelerinden (nöronlardan) oluştuğu fikrinin benimsenmesi,
- 1943 – Yapay sinir hücrelerine dayalı hesaplama teorisinin ortaya atılması ve eşik değerli mantıksal devrelerin (threshold logic device) geliştirilmesi (Mcculloch ve Pits),
- 1949 - Hebbian Öğrenme (Donald Hebb),

¹⁸⁷ Elmas Çetin. **a.g.e.** s.22.

¹⁸⁸ Esen Ersoy, ve Ömer Karal, “Yapay Sinir Ağları Ve İnsan Beyni”, **İnsan ve Toplum Bilimleri Araştırmaları Dergisi**, 2012, Cilt: 1, Sayı, 2.

¹⁸⁹ Dave Anderson and George McNeill, “Artificial Neural Networks Technology”, A DACS State-of-the-Art Report, Contract Number F30602-89-C-0082, 1992, s.17-18.

- İlk YSA modelleri diyebileceğimiz, basit algılayıcılar olan, Perseptronun 1958 yılında geliştirilmesi, (Frank Rosenblatt), ardından ADALINE (Adaptive Linear Neuron) ve MADALINE (Multi Adaptive Linear Neuron) modellerinin geliştirilmesi,
- 1969- Perseptron'un XOR problemini çözemediğinin ispatlanması (Minsky, Pappert) ve DARPA'nın YSA çalışmalarını desteklemeyi bırakması,
- 1982- ÇKA'nın geliştirilmesi (Rumelhart) ve Geriye yayımlı ÇKA'nın (Rumelhart ve öğrencisi) geliştirilmesi
- 1985 – XOR probleminin ÇKA ile çözülmesi (Parker, Werbos).

Ayrıntılı olarak ise YSA'nın tarihçesini Öztemel¹⁹⁰, kronolojik olarak, şu başlıklar altında toplamıştır.

- 1890 : İnsan beyninin yer alan sinir hücrelerinin yapısının ve çalışma prensiplerinin keşfedilmesinin ardından sinir ağlarına olan ilgi arttı.
- 1911 : İnsan beyninin bileşenlerinin belirli bir düzenek ile sinir hücrelerinden (nöronlardan) oluştuğu fikri benimsendi.
- 1943 : Beyindeki sinir hücrelerini taklit eden bir hesaplama teorisi, YSA fikri, ortaya atıldı ve eşik değerli mantıksal devreler (threshold logic device) geliştirildi.
- 1956-1962 : Rosenblatt'ın algılayıcı modeli ile benzer nitelikler taşıyan ve YSA'nın ilk mühendislik uygulaması olan ADALINE/MADALINE (Widrow, Hoff/Widrow ve Lehr) uygulamaları geliştirildi.
- 1957-1962 : Perceptron (tek katmanlı algılayıcı) geliştirildi.
- 1965 : İlk makine öğrenmesi kitabı yayınlandı.
- 1967-1969 : Grosberg vb. öğrenme algoritmaları geliştirildi.
- 1969 : Yapay zekâ biliminin önde gelen araştırmacılarından Minsky ve Pappert, Algılayıcılar (Perceptrons) isimli kitapla, perceptronların bilinen XOR problemini çözemediğini kanıtladılar. Sonrasında perceptronların doğrusal olmayan problemleri çözemeyeceğini belirttiler. Bu kitap ve bu ispat

¹⁹⁰ Ercan Öztemel, **Yapay Sinir Ağları**, İstanbul, Papatya Yayıncılık, 2012, s.37-42.

çok ses getirdi ve YSA ile ilgili çalışmalar bitme noktasına geldi ve “yapay sinir ağı yapmak mümkün değildir” inancı hızla yayıldı. Bu çalışma bir devrim niteliğinde olup YSA ile ilgili çalışmaların ve gelişmelerin 1970 öncesi ve sonrası şeklinde iki ana başlıkta incelenmesine neden olmuştur.

- 1972 : 1969’dan itibaren bütün ilgiyi kaybeden YSA ile ilgili çalışmalarına devam eden Amari, Anderson, Cooper, Fukushima, Grossberg, Kohonen ve Hopfield gibi araştırmacılar çalışmalarına devam ettiler. Özellikle elektrik mühendisi Kohonen ve nöropsikolojist Anderson “çağrışimli bellek” (associative memory) konusundaki, birbirlerinin hemen hemen aynı çalışmaları, daha sonraki denetimsiz öğrenme yaklaşımının temellerini oluşturdu. Ayrıca 1972’de de korelasyon matris belleği geliştirildi.
- 1974 : ÇKA modeli geliştirildi.
- 1978 : 1960’ların sonlarında YSA’nın psikolojik mantıksallığını ve mühendislik uygulamalarındaki kolaylığını gösteren Grossberg, Carpenter ile birlikte ART (Adaptif Rezonans Teorisi) modelini geliştirdi. ART, denetimsiz öğrenme konusunda gelmiş geçmiş en kompleks YSA’nı temsil etmektedir.
- 1978-1980 : Fukushima görsel şekil ve örüntü tanıma amaçlı, NEOCOGNITRON modelini tanıttı. Bu modelle artık, kullanılan YSA modellerinin biyolojik sinir ağlarına benzerliğinden çok sonuçlarının kullanılabilirliği ön plana çıkmaktaydı.
- 1982-1984 : Hopfield tarafından, “Hebbian Öğrenme” olarak da bilinen, “Hopfield Ağı” geliştirildi ve ilk kez “gezgin satıcı problemi” bu modelle çözüldü. Hinton ve arkadaşları Boltzman makinesini geliştirdiler. Rumelhart vd. paralel programlama çalışmaları yaptılar. Paralel programlama çalışmaları, ÇKA’nın geliştirilmesini beraberinde getirdi. Böylelikle tek katmanlı algılayıcılarla çözülemeyen XOR probleminin yanı sıra Hopfield ve Boltzman makinelerinin sınırları da çözüldü ve 1970’den itibaren bitme noktasına gelen YSA’na ilgi yeniden ivme kazanmaya başladı.
- 1985 : ÇKA’lar, “Delta Öğrenme Kuralı” kullanılarak yenileştirildi.
- 1988: Broomhead ve Lowe tarafından, filtreleme probleminde çok başarılı olan, RBF modelini geliştirildi. Daha sonra Specht, ÇKA’ya alternatif olabilen

bu ağı daha gelişmiş şekli olan Probabilistik (Olasılıksal) Sinir Ağlarını (Probabilistic Neural Network, PNN) geliştirdi.

- 1991 : Specht, PNN'nin ardından Genel Regresyon Sinir Ağlarını (General Regression Neural Network, GRNN) geliştirdi.

1991 yılından sonra, birçok bilim dalında ve birçok sektörde, sayısız model ve uygulamaya model olan YSA çalışmalarının literatür taramasını bu yıldan itibaren takip etmek neredeyse imkansızdır.

Günümüzde, ÇKA, LVN, SOM, ART, Hopfield Ağları, İleri Beslemeli Ağ, Neocognitron Ağı, Boltzman Makinesi, PNN, Elman Ağı ve RBN pratik hayatta en çok kullanılan YSA modelleridir¹⁹¹.

2.2.3. Sinir Hücrelerinin Yapısı Ve İşleyişi

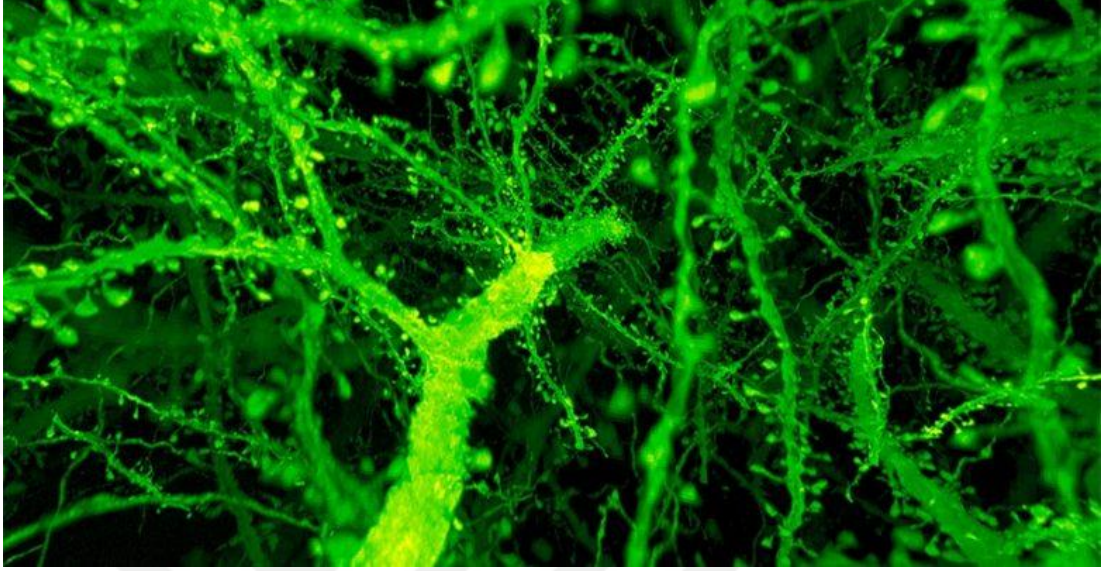
Nöron (beyin hücresi) kelimesi (ve kromozom kelimesi) ilk kez 1890 yılında Alman bilim adamı Heinrich Wilhelm Gottfried von Waldeyer-Hartz tarafından kaleme alınmıştır. Nöronlar, beyinden sinir sinyallerini 200 mil / saat hızla ileten sinir hücreleridir¹⁹².

Nöronlara mikroskop altında bakıldığında, aşağıda Şekil 2.6'daki gibi görünmektedir ve birbirine binlerce dallarla bağlanan bir orman örüntüsünü anımsatmaktadır. Nöron yapılarının ve işlevlerinin anlaşılması için, bu resmin kavramsal bazı alt ünitelere indirgenmiş hali olan Şekil 2.6'ya bakıldığında, nöronun yapısını daha net görmek mümkün olacaktır.

¹⁹¹ Ercan Öztemel, **a.g.e.** s.56.

¹⁹² www.enchantedlearning.com/subjects/anatomy/brain/Neuron.shtml, 30/10/2020.

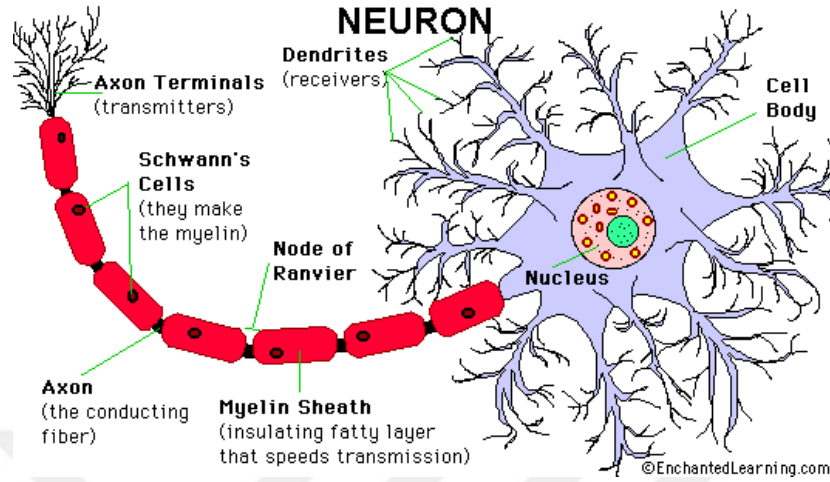
Şekil 0.6: Sinir Hücrelerinin Görünümü



Kaynak: www.sciencenews.org, Neurons, Çevrimiçi, <https://www.sciencenews.org/article/new-ways-image-control-nerve-cells-unlock-brain-mysteries>, 30/10/2019.

Şekil 2.6’da şematize edilmiş sinir hücresine bakıldığında, içerisinde DNA’yı da içeren hücre çekirdeğini barındıran hücre gövdesi ve bu gövdenin etrafını tamamen saran “dendritler”in yani “algılayıcılar”ın, yer aldığı görülmektedir. Gövdeden sonra ise, Schwann hücreleri tarafından örülmüş bir miyelin kılıfın içerisinde geçen ve akson adı verilen sinyal taşıyıcıları ve nihayet bu sinyallerin ulaştırıldığı akson uçları görülmektedir. Hücresinin gövdesini insanın vücuduna, göz, kulak, burun, dil, saç ve deri gibi algılayıcıları ise dendritlere, miyelin kılıfı omurga kemiklerine, aksonları sinir yataklarına ve akson uçlarını da insanın el ve ayaklara benzetebiliriz. Bir insanın aslında fizyolojik anlamda, trilyonlarca mikro-insanın mimarisinden oluştuğunu söylemek, sinir hücresinin anlaşılması bakımından açıklayıcı iyi bir örnek olabilir.

Şekil 0.7: Biyolojik Sinir Hücresi (Nöron)



Kaynak: [www.enchantedlearning.com,](http://www.enchantedlearning.com/subjects/anatomy/brain/Neuron.shtml) Neurons, Çevrimiçi, <https://www.enchantedlearning.com/subjects/anatomy/brain/Neuron.shtml>, 30/10/2019.

Bugünkü nöron bilgisinin temeli, büyük ölçüde, Nobel Fizyoloji ve Tıp Ödülü sahibi, Ramon Cajal'ın (1878) çalışmalarına dayanmaktadır. Ayrıca Cajal'ın Nöron teorisi, 20. yüzyılın en önemli bilimsel keşiflerinden birisidir ve elektron mikroskobu başta olmak üzere, yüz yıldan daha fazla bir süre içerisinde, yeni teknolojiler tarafından onay almaya başlamıştır¹⁹³. Cajal, sinir hücresinde iletilen elektriğin nasıl oluştuğunu da açıklamıştır. Cajal'a göre hücrenin içinde ve dışında iyonlar vardır ve hücre zarının belli bir geçirgenlik özelliği bulunmaktadır. Tıpkı bir oda da oturan insanların, odanın dışındaki çevre hakkında bilgi edinebilmesi için, odaya içerden veya dışardan bazı giriş çıkışların olması gerekir. Odaya birileri girip çıktıkça haberlerin oda içerisinde yayılması gibi, difüzyon yoluyla, hücre içerisine giriş-çıkış yapan iyonlar sayesinde elektrik oluşur ve bu elektrik sinyalleri aksonlar aracılığıyla, akson uçlarına ve oradan da “sinaps” adı verilen bağlantılar aracılığıyla diğer sinir hücrelerine gönderilir¹⁹⁴.

İnsanda bulunan yaklaşık 100 trilyon hücrenin yaklaşık 100 milyarı beyin hücresidir. İnsanı zeki kılan, hücre sayısı değil, bu hücreler arasındaki bağlantı sayılarıdır.

¹⁹³ Francisco López -Muñoz, Boya Jesús , Alamo, Cecilio. “Neuron theory, the cornerstone of neuroscience, on the centenary of the Nobel Prize award to Santiago Ramón y Cajal”, **Brain Research Bulletin**, Cilt: 70, Sayı 4–6, 2006, s.391.

¹⁹⁴ www.scholarpedia.org, Çevrimiçi, http://var.scholarpedia.org/article/Pyramidal_neuron, 05/11/2019.

Örneğin, bir bebek ilk doğduğunda bu nöronların birbirine bağlantı sayıları oldukça azdır. Zamanla, nöronlar görülen, işitilen, koklanan, dokunulan, tadılan uyarılara tepki verirken dendritler yoluyla fiziksel bağlantılar kurarlar ve geliştirirler. Çocuğun aktif yaşantısı, zihinsel çabası ve zengin çevresel uyarıcılar dendritlerin dallanmasını hızlandırır. Böylece çevresel algılama, bir diğer deyişle veri toplama kapasitesi, artarken nöronların arasındaki bağlantılar arttıkça da, veri işleme ve elde edilen veriden kaliteli sonuçlar çıkarma yetisi gelişir. Sonuç olarak bebeğin zekâ yapısı gelişir. Bu, nitekim bir kurumda birimler birbirine ne kadar çok iletişim hattına sahipse, o derece haberleşme ve veri aktarımı sağlanarak, bu da kurumun işleyişinin etkinliğini artırmasına benzetilebilir¹⁹⁵.

YSA'nın, çok katmanlı olması, nöronlara benzeyen düğümler içermesi, açıklayıcı veri girişini dendritler gibi çevresel toplaması ve ileri beslemeli olması gibi özelliklerin dışında nöronlara birebir benzeyen çok fazla özellikleri bulunmamaktadır. Yani yapay sinir hücreleri, biyolojik sinir hücrelerinin çok basit birer taklidi niteliğindedir.

2.2.4. Yapay Sinir Hücreleri

Biyolojik sinir hücrelerini, çeşitli problemleri çözmek için modelleyen basit algılayıcılar, diğer adıyla yapay sinir hücreleri, insan beyinde yer alan bu hücrelerin problem çözmede de kullanılabileceği düşüncesinin 1940'lı yıllarda Walter ve Pitts'in çalışmalarından esinlenilerek geliştirilmeye başlanmıştır. Bunlardan en önemlisi Rosenblatt'ın geliştirdiği “perseptron (perceptron)” modelidir. Perseptronun ardından Widrow ve Hoff tarafından, 1960 yılında geliştirilen ADALINE modeli ve ardından Widrow ve Lehr tarafından açıklanan MADALINE modelleri, günümüzde bilinen basit algılayıcılar olup ilk YSA'nı oluştururlar. Tek katmanlı algılayıcılar bugün, bu şekliyle pratik kullanımdan çok uzaktırlar. Fakat öğrenme yeteneklerinin bulunması ve YSA'nın hatta makine öğrenmesinin temelini oluşturmalarından dolayı oldukça önemlidirler.

¹⁹⁵ Hakan Çuhadar, “Müzik Ve Beyin”, *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, Cilt: 17, Sayı: 2, 2008, s.67-76.

Biyolojik sinir sinir hücresinde gerçekleştirilen faaliyetlerin, yapay sinir hücresinde hangi işlevlere karşılık geldiği Tablo 2.1’de listelenmiştir.

Tablo 0.1: Biyolojik Sinir Ağları İle Yapay Sinir Ağlarının Karşılaştırılması

Biyolojik Sinir Ağı	Yapay Sinir Ağı
Sinir Sistemi	Sinirsel Hesaplama Sistemi
Sinir	Düğüm (Sinir, İşlem elemanı)
Sinaps	Sinirler arası bağlantı ağırlıkları
Dendrit	Toplama işlevi
Hücre gövdesi	Etkinlik İşlevi
Akson	Sinir Çıkışı

Kaynak: Çetin Elmas, **Yapay Zekâ Uygulamaları**, 3. Baskı, Ankara: Seçkin Yayıncılık, 2016, s.90.

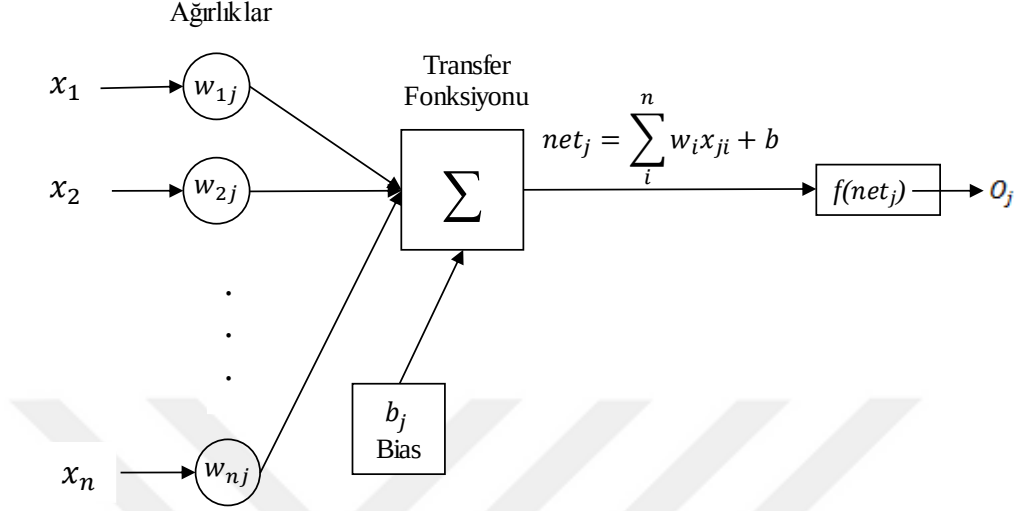
2.2.4.1. Perseptron

Perseptron ilk kez 1958 yılında Frenk Rosenblatt tarafından, birden fazla girdisi ve tek bir çıktısı olan bir sınıflandırma modeli olarak, ileri sürülmüştür¹⁹⁶. Basit algılayıcı mantığında, tüm girdileri sonuçlar üzerindeki etkileri dikkate alarak sıfır ila bir arasındaki değerlerle ağırlıklandırıp toplayarak ve sapmasını ilave ederek, elde edilen sonucun değeri, oluşan fonksiyonun söz konusu girdi değeriyle elde edilen sonuçtan büyükse 1 (bir) değilse 0 (sıfır) olarak sınıflandırma yapılır. Yapay sistemin öğrenmesi, bu ağırlıkların değiştirilerek optimize edilmesi yoluyla gerçekleşir¹⁹⁷. Basit algılayıcı yaklaşımının şematik ve matematiksel gösterimi aşağıdaki gibidir.

¹⁹⁶ Frank Rosenblatt, “The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain”, **Psychonanalytic Review**, Cilt: 65, Sayı: 6, 1958, s.387.

¹⁹⁷ Vasif Nabiye, **Yapay Zekâ**, 5. Baskı, Ankara: Seçkin Yayınevi, 2016, s.593.

Şekil 0.8: Perseptronun Yapısı



Kaynak: Coşkun Hamzaçebi, **Yapay Sinir Ağları Tahmin Amaçlı Kullanımı, MATLAB Ve NEUROSOLUTION Uygulamalı**, 1. Baskı, Bursa: Ekin Yayınevi, 2011, s.37.

$$v_i = \sum_{i=1}^n w_{ij} + b_j \quad (0.16)$$

Sinir çıkış sinyali aşağıdaki eşitlikteki gibidir.

$$y_i = f(w, x) = f(w^T x) = f\left(\sum_{j=1}^n w_{ij} x_i\right) \quad (0.17)$$

Burada $w_i = (w_1, \dots, w_n)^T \in R^n$ ağırlık vektörünü temsil eder. $f(w^T x)$ ise, etkinlik aktarım işlevidir. Ağın değerleri giriş vektörü ile tanımlanır.

$$Ağ = w^T x = w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad (0.18)$$

Burada ilk durum, çıkış sıfır ise aşağıdaki şekildeki gibi hesaplanır.

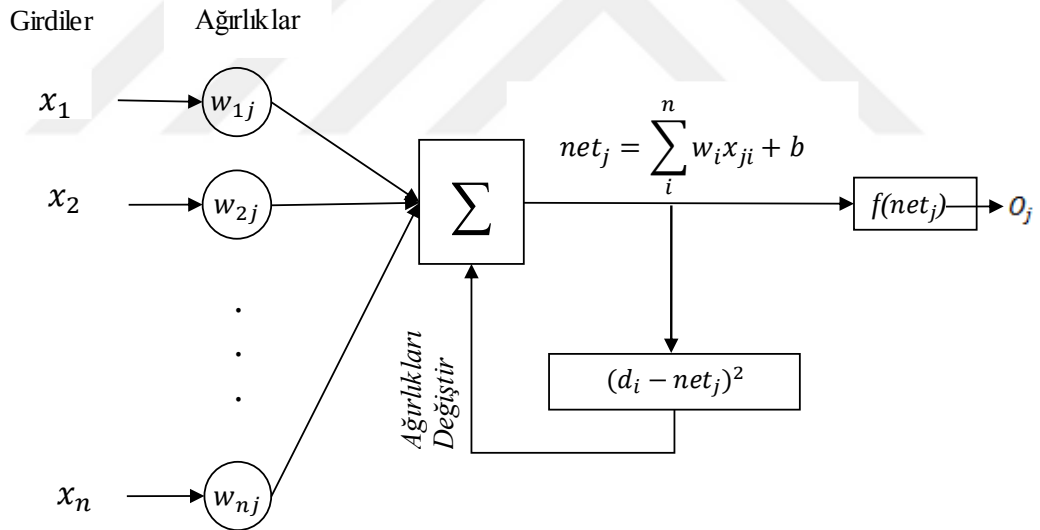
$$o_i = f(ağ) = \begin{cases} 1 & \text{eğer } w^T x, \\ 0 & \text{eğer farklı.} \end{cases} \quad (0.19)$$

Burada b eşik düzeyi olarak adlandırılır ve bu tip düğümlere de, “düzenli eşik düzeyi” denir¹⁹⁸.

2.2.4.2. ADALINE

Adaptive Linear Element (ADALINE), Widrow ve Hoff tarafından 1960 yılından geliştirilmiş bir öğrenme kuralıdır. İlk zamanlar “Adaptive Linear Neuron” olarak bilinen öğrenme modeli daha sonraları “Adaptive Linear Element” olarak anılmaya başlanmıştır. Widrow ve Hoff ADALINE için öğrenme kuralı olarak, “Delta Öğrenme Kuralı” olarak da bilinen “En Küçük Ortalama Kareler (EKOK) - Least Mean Square (LMS)”, yöntemini önermişlerdir. Bu yöntem birden fazla çıkışlı tek katmanlı algılayıcılarda da kullanılabilir. Şekil 2.9’da ADALINE modeli şematize edilmiştir;

Şekil 0.9: ADALINE (Adaptive Linear Element)



Kaynak: Coşkun Hamzaçebi, **Yapay Sinir Ağları Tahmin Amaçlı Kullanımı, MATLAB Ve NEUROSOLUTION Uygulamalı**, 1. Baskı, Bursa: Ekin Yayınevi, 2011, s.37.

Denklem 2.20, ADALINE fonksiyonun çıktısını göstermektedir;

$$O_j = \begin{cases} 1 & \text{eğer } net_j \geq 0, \\ 0 & \text{eğer } \text{değilse}. \end{cases} \quad (0.20)$$

¹⁹⁸ Çetin Elmas, **a.g.e.** s.90.

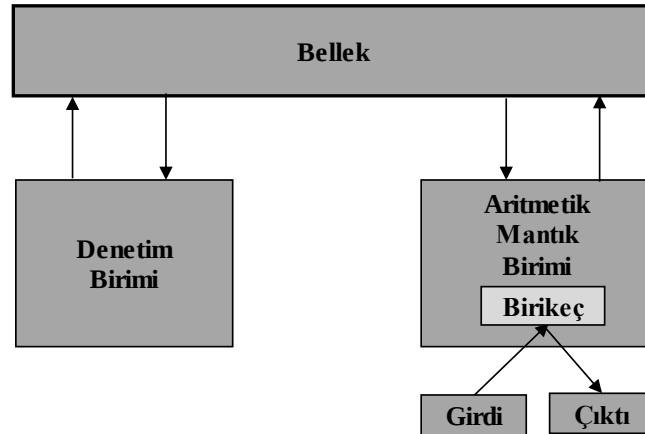
2.2.4.3. MADALINE

Birden çok ADALINE biriminin birbirine bağlanarak oluşturduğu bu modelin ayrıntıları, Widrow ve Lehr tarafından açıklanmıştır. Genel olarak “Multiple Adaptive Linear Neuron (MADALINE)” ağı iki katmandan oluşur ve her katmanda değişik sayıda ADALINE ünitesi bulunur. Ağın çıktı değerleri yine $[1,-1]$ şeklindedir. Öğrenme kuralı ADALINE öğrenme kuralı (Bkz. Denklem 2.20) ile aynıdır¹⁹⁹.

2.2.5. Sayısal Bilgisayarlar Ve Yapay Sinir Ağları

1940’lardan günümüze deksürekli geliştirilen ve işleyişini bildiğimiz sayısal (geleneksel) bilgisayarlar, ilk başta sadece matematiksel hesaplamalar için kullanılmış sonrasında ise resim, metin, sembol ve ses işleme gibi alanlarda da başarılı bir şekilde kullanılmaya başlanmıştır. Aşağıdaki şekilde sayısal bilgisayarların çalışma şekli görülmektedir. Sayısal bilgisayarlar Von-Neumann mimarisi temel alınarak geliştirilmiştir²⁰⁰.

Şekil 0.10: Von Neumann Mimarisinde Bilgisayarın Yapısı



Kaynak: Wikipedia, Çevrimiçi:
https://tr.wikipedia.org/wiki/Neumann_mimarisi#/media/Dosya:VonNeumannMimarisi.PNG,
14/10/2020.

¹⁹⁹ Çetin Elmas. **a.g.e.** s.73.

²⁰⁰ Çetin Elmas, **a.g.e.** s.22-23.

YSA bazı özellikleriyle sayısal bilgisayarlara benzerlerken diğer bazı özellikleri itibariyle de bilgisayarlardan farklılaşmaktadır. Tablo 2.2’de sayısal bilgisayarlar ile YSA arasındaki farklılıklar listelenmektedir.

Tablo 0.2: Sayısal Bilgisayarlar İle Yapay Sinir Ağlarının Karşılaştırılması

Sayısal Bilgisayarlar	Yapay Sinir Ağları
Tümdengelimli üretilen veriden çıkış üretmek için giriş bilgilerine bilinen kurallar uygulanır	Tümevarımlı üretilen veriden çıkış bilgileri (eğitilen örnekler) verilir, kuralları biz koyarız
Hesaplama merkezi, eş zamanlı ve ardışıldır	Hesaplama, toplu, eş zamansız ve öğrenmeden sonra paraleldir
Bellek paketlenmiş, hazır bilgi depolanmış ve yer adreslenebilir	Bellek ayrılmıştır, dâhilidir ve içerik adreslenebilir
Hata toleransı yoktur	Eğer bilgi gürültülü veya kısmi ise kurallar bilinmiyorsa ya da karışıkta hata toleransı uygulanabilir
Hızlıdır	Yavaştır
Bilgiler ve algoritmalar kesindir	Yapay sinir sistemleri deneyimden yararlanır
Bilgi 0 ve 1 serileri ile temsil edilir	Bilgi matematiksel işlev ile temsil edilir
Bilgi belirli bir yerde (bellekte) saklanır	Bilgi tüm ağ boyunca dağıtılmıştır (Dağıtılmış bellek, sonuç çıktıları açısından birleştirilmiş bellek)

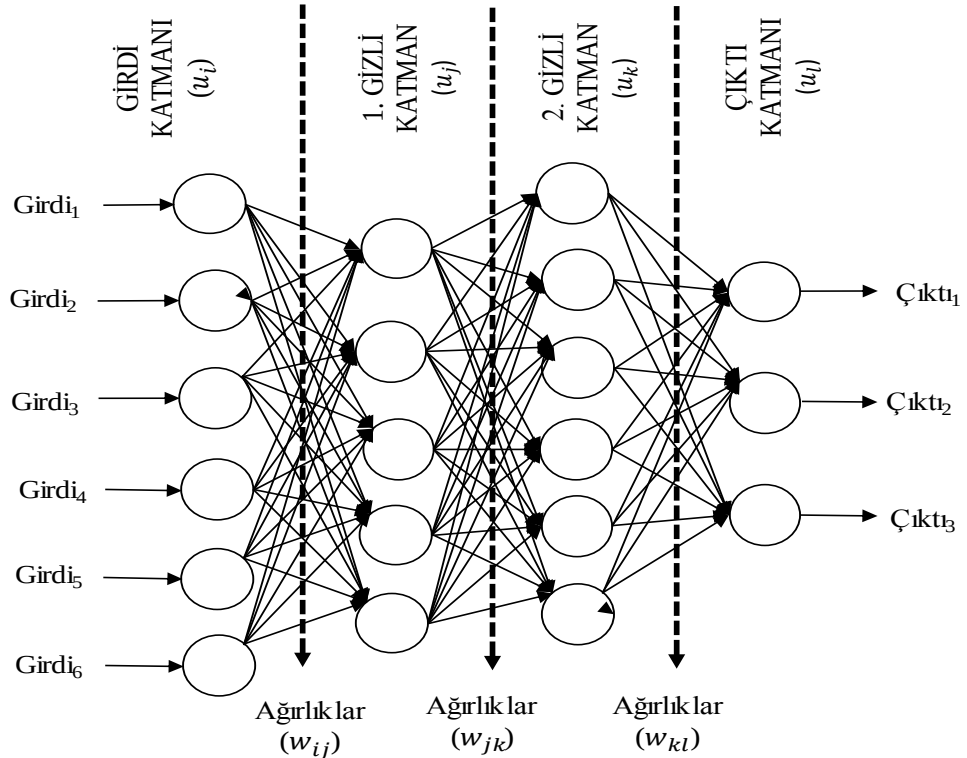
Kaynak: Çetin Elmas, *Yapay Zekâ Uygulamaları*, 3. Baskı, Ankara: Seçkin Yayıncılık, 2016, s. 24.

2.2.6. Çok Katmanlı Ağlar

Perseptronların XOR problemini çözemediğinin ispatlanmasının (1969) ardından, birazda YSA’na dönük abartılı haberlerden dolayı, ölçsüz beklentilerin oluşması sonucu, bir hayal kırıklığı yaşanmış ve sinir ağlarına olan ilgi bitme noktasına gelmiştir. Fakat bazı araştırmacılar, bu algılayıcıların üzerinde çalışmaya devam etmişler ve XOR problemini de çözen ve çok sayıda perseptronun birbirine bağlanmasıyla oluşan, ÇKA adı verilen sinir ağını geliştirdi.

Şekil 2.11’de gördüğümüz yukarıdan aşağı sıralı olan her bir daire dizisi bir katmanı, her bir daire ise bir düğümü temsil etmektedir. Şekil 2.11’de, altı adet girdiden oluşan bir “girdi katmanı”, beş adet düğümünden oluşan birinci “gizli katman”, altı adet düğümünden oluşan ikinci gizli katman ve dört farklı çıktıdan oluşan bir adet “çıkı katmanı” görülmektedir. Girdi ve çıkı katmanları arasında, çeşitli sayılarda, gizli katman bulunabilir. Girdi katmanı, modele girilen veriyi temsil etmektedir. Yani varsayımsal bir katmandır. Girdi katmanında herhangi bir işlem yapılmaz, bundan dolayı hatası da olmaz. Bu katmanda girdiler, sıfır ila bir arasındaki, rasgele atanmış ağırlıklarla çarpılarak ilk gizli katmandaki her nörona (her nörona gönderilmesi ağ mimarisi açısından şart değildir, fakat uygulamaların geneli bu şekildedir) gönderilir. İlk gizli katmandaki nöronlar, kendilerine gelen ağırlıklandırılmış değerleri, bir toplama fonksiyonu yardımıyla, toplar (Bkz. Denklem 2.16) ve elde edilen toplam değeri, yine okların temsil ettiği ağırlıklarla ağırlıklandırılarak, ikinci gizli katmandaki tüm nöronlara iletir. Bu işlemler nöronlara gelen girdiler sonuç sınıflarına atanıncaya kadar devam eder.

Şekil 0.11: Yapay Sinir Ağı (Çok Katmanlı) Modeli



Girdi katmanı, araştırmacının sahip olduğu tüm gözlemleri ifade etmektedir. Çıktı katmanı ise sonuç sınıflarını ifade etmektedir. Kabaca örnek vermek gerekirse, bir araştırmacı, hastanede yatan 100 hastanın astım hastası mı yoksa koah hastası mı olduğunu araştırıyor olsun. Bu durumda, bu araştırmayı yapan kişinin girdi katmanı yüz adet düğümden, çıktı katmanı ise iki adet düğümden oluşacaktır.

Gizli katman, aslında ağın zekâsını temsil etmektedir. Çünkü modeli sonuçlara götüren, matematiksel işlemler burada yapılır. Gizli katman sayısını ve her bir gizli katmanda yer alacak düğüm sayısını araştırmacı, deneme yanılma yoluyla belirleyebilir. En uygun sonucu verecek gizli katman sayısını ve katmanlardaki düğüm sayısını belirlemeye dönük olarak henüz, deneme yanılma yolundan başka bir yol bulunmamaktadır. Bu arada gizli katmanlardaki nöron sayısı çok fazla olursa, aşırı öğrenme yani ezberleme problemiyle karşılaşılabilir. Bu insan beynindeki nöronların bilgi tutmasına benzer. Görüntüyü veya metni, hatasız bir şekilde, zihninde tutan insan, bir bakıma beyninde yer alan çok sayıda nöronu bu iş için harcamış olmasına anlamına gelmektedir. Bir başka açıdan aynı kişi, gördüğü yeni görüntü ya da bilgileri, daha önce tecrübe ettiği görüntü ve metinlerle ilişkilendirme yoluna giderse çok daha az çabayla (nöronla) yenilerinin zihninde tutabilir. Yani yeni bilgileri kodlama yaparak öğrenme yoluna gider ve böylelikle ezberleme hatasından ve zahmetinden kurtulur. “Kodlanarak öğrenilmeyen bir bilgi, ezberlenmiştir”. Ezberlemenin gerçekleştiği durumlarda söz konusu bilgilerde en ufak bir değişiklik meydana geldiğinde, ezber (aşırı öğrenmiş model) işe yaramayacaktır. Gizli katmanlarda yer alan nöron sayısının optimize edilmesi ayrıca bir sorun teşkil edeceği için öncelikle, ağırlık atama işlemi sırasında, düşük sayılarla başlayarak yavaş yavaş yükseltmek, sonuçların negatife döndüğü yerde de durmak tavsiye edilmektedir²⁰¹.

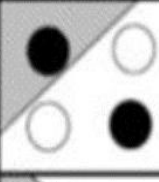
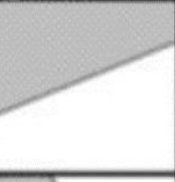
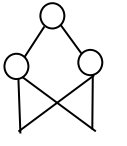
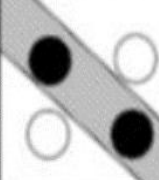
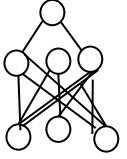
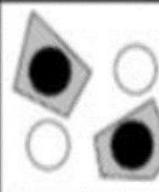
1960’lı yılların sonunda Misnky ve Pappert, çok popüler olan ve bu yüzden ölçüsüz şekilde abartılan perseptronun doğrusal bir problem olmayan, XOR problemini çözemediğini kitaplarında ispatlamalarının ardından sinir ağlarına olan ilgi bitme noktasına gelmiş ve YSA’na olan tüm ilgi, yapay zekânın başka alanlarına yönelmiştir.

²⁰¹ Vasif Nabiyev, **a.g.e.**, s.599.

Bu durum, tâki birden çok perseptronun birbirine yatayda ve dikeyde bağlanmalarıyla elde edilen, ÇKA'nın geliştirmesine (Rumelhart, 1982) ve XOR problemini de çözebildiğinin gösterilmesine (1985, Parker ve Werbos) kadar sürmüş, sonrasında ise sinir ağlarına olan ilgi, yaklaşık on beş yıldan sonra, yeniden canlanmıştır. ÇKA'lar, ileri ve geri beslemeli olarak iki farklı şekilde kullanılabilmekte, geriye yayılma fonksiyonu sayesinde de bugün en çok kullanılan YSA modelleri haline gelmiştir²⁰².

Şekil 2.12'de ÇKA'ların, tek katmanlı olanlara kıyasla, gösterim yeteneği görülmektedir²⁰³:

Şekil 0.12: ÇKA'ların Gösterim Yeteneği

Yapı	Karar Sınırları	XOR Problemi	Sınıflandırma Kesişim Bölgeleri	Genel Alan Şekilleri
 Tek Katman	Hiperdüzlem tarafından sınırlanmış yarım düzlem			
 İki Katman	Keyfi (Karmaşıklık gizli birim sayısı sınırlı)			
 Üç Katman	Keyfi (Karmaşıklık gizli birim sayısı sınırlı)			

Kaynak: Şenol Emir, “Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması: Borsa Endeks Yönünün Tahmini Üzerine Bir Uygulama”, Doktora Tezi, İstanbul Üniversitesi, 2013, s. 87.

²⁰² Talip Torun, **a.g.e.** s.58.

²⁰³ Şenol Emir, **a.g.e.** s.87.

2.2.7. İleri Beslemeli Ve Geri Beslemeli YSA

YSA’nda beslemeden kasıt, bağlantı ağırlıklarını temsil eden okların yönüdür. Bu oklar, soldan sağa doğru olursa ileri beslemeli ağ olarak tanımlanmaktadır. Geri beslemeli ağlarda ise tüm nöronlar birbirlerine oklarla bağlanmıştır.

2.2.7.1. İleri Beslemeli Yapay Sinir Ağları

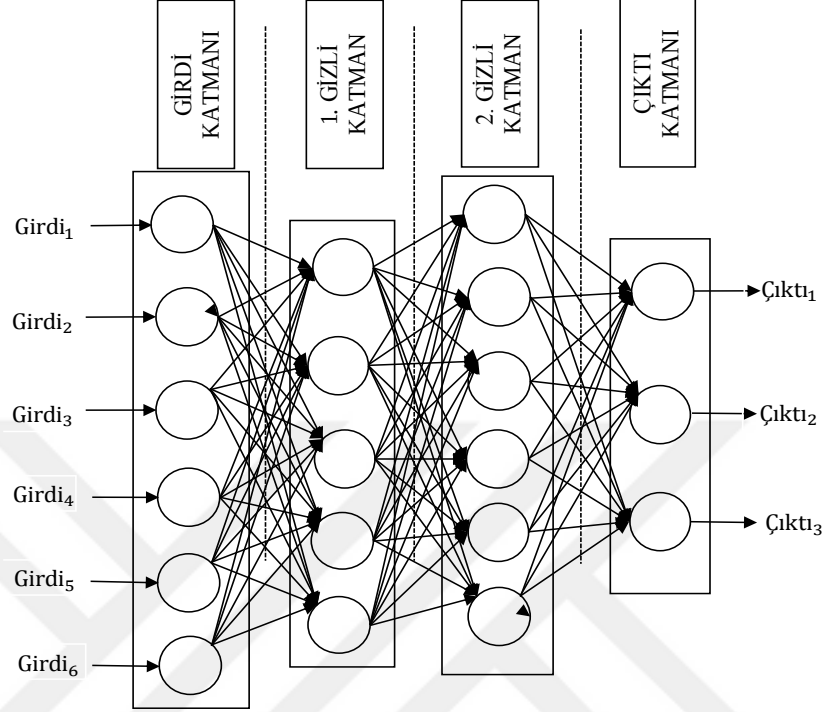
İleri beslemeli (Feed Forward) YSA, aslında en genel anlamıyla bilinen YSA mimarisidir. İleri beslemeli YSA, denetimli ya da yarı denetimli öğrenme gerçekleştirir ve eğitimin söz konusu olduğu mimari, bu mimaridir. Şekil 2.13’te yer alan YSA modelini tekrar göz önüne alalım. i girdisini, j gizli katman düğümüne bağlayan ok, w_{ij} ağırlığının bir işlevi hesaplanarak sonuç bir sonraki nörona iletilir²⁰⁴.

İleri beslemeli YSA mimarisi, fonksiyon kestirme, sınıflandırma ve tahmin problemlerinde kullanılmaktadır²⁰⁵.

²⁰⁴ Tunçhan Cura, “Karar Verme Aracı Olarak Yapay Sinir Ağları Ve Yapay Sinir Ağları İle Portföy Optimizasyonu”, Doktora Tezi, İstanbul Üniversitesi, 2004, s.12.

²⁰⁵ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.19.

Şekil 0.13: İleri Beslemeli Yapay Bir Sinir Ağı Modeli



Kaynak: Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008.

İleri beslemeli YSA, geri yayılım algoritmasının çalıştırılabilmesi için bir sonuç değerine ihtiyaç duyar. Dolayısıyla sonuç değerlerinin belli olduğu denetimli öğrenme grubuna girer ve sonuç etiketlerinin önceden belli olduğu sınıflandırma problemlerine uygulanabilir.

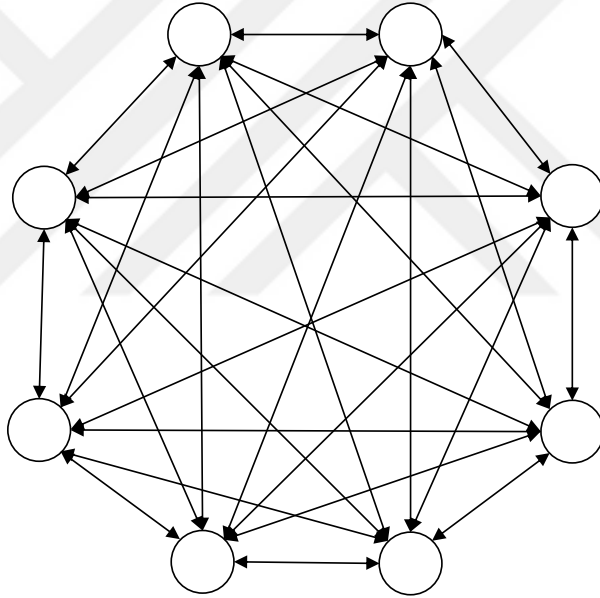
2.2.7.2. Geri Beslemeli Yapay Sinir Ağları

İleri beslemeli YSA’nda, hata terimi belli bir değerin altına düşünce model döngüyü durdurur. Geri beslemeli YSA’nda ise, ileri beslemeli YSA’nın aksine, model durağan hale gelinceye kadar döngüye devam edilir. Geri beslemeli YSA, ilk olarak Hopfield tarafından kullanıldığı için, Hopfield ağları olarak da bilinmektedir. Geri beslemeli ağlar, bazı belirli sınırlamaları sağlayacak şekilde, bilinmeyen parametre değerlerinin

bulunmasını sağlayan problem türü olan, optimizasyon problemlerinin²⁰⁶ çözümünde kullanılır. Geri beslemeli YSA, diğer sezgisel optimizasyon yöntemleriyle kıyaslandığında, kullanımı oldukça nadir görülmektedir. Çünkü diğer yöntemlere kıyasla, başarı oranı düşük ve kullanımı zordur. İleri beslemeli ağlar ise hem diğer yöntemlere de kıyasla daha başarılı, hem de kullanımı kolaydır. Bu yüzden de literatürde en çok tercih edilen yöntemlerden birisidir²⁰⁷.

Şekil 2.14'te görüldüğü gibi geri beslemeli YSA modellerinde tüm nöronlar birbiriyle bağlantılı haldedir.

Şekil 0.14: Geri Beslemeli Yapay Sinir Ağları Mimarisi



Kaynak: Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s. 32.

2.2.8. Geriye Yayılmalı Yapay Sinir Ağları

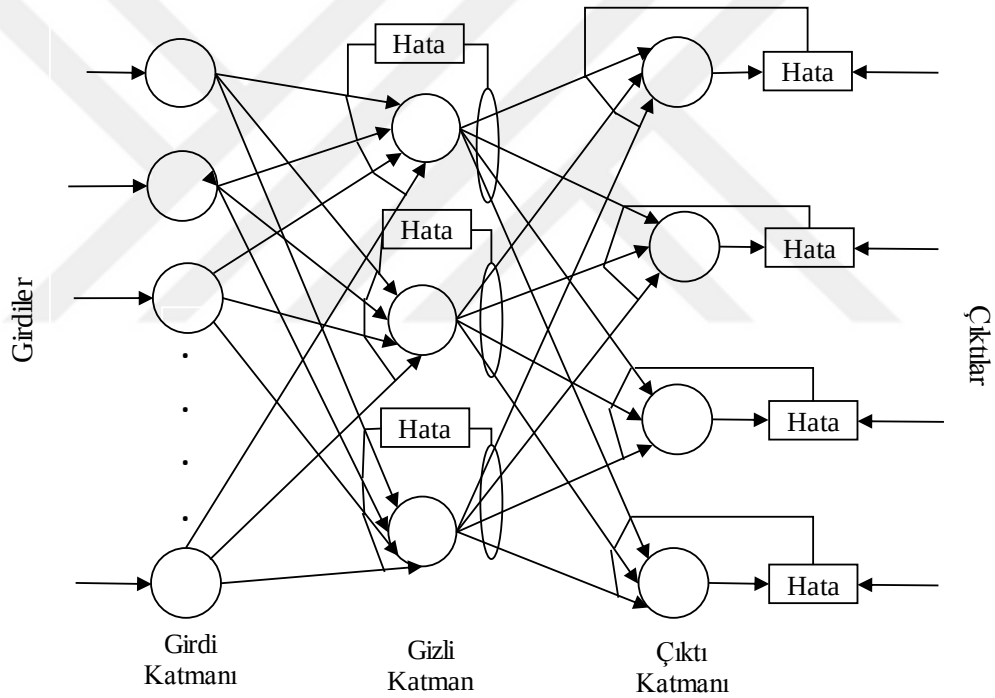
Bir ÇKA modelinde, işlemlerin ve okların yönü soldan sağa doğru olduğu için bu, ileri beslemeli sinir ağı olarak tanımlanır. Sonuç katmanına ulaşıldığında, eğer doğru

²⁰⁶ Derviş Karaboğa, **Yapay Zekâ Optimizasyon Algoritmaları**, 4. Basım, İstanbul: Nobel Akademik Yayıncılık, 2014, s.1.

²⁰⁷ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.31-40.

çıktıya ulaşılmışsa, ağırlıklar değiştirilmez²⁰⁸. Fakat yanlış çıktıya ulaşılmışsa bir hata payı atanır ve bu hata payı geriye doğru tüm ağırlıklara yansıtılır ve ağırlıklar, hata değerleri belli bir değerin altına düşünceye kadar, güncellenir²⁰⁹. Yani bir geriye yayılım (Back Propagation) algoritması söz konusudur. O yüzden bu tür sinir ağlarına “ileri beslemeli geri yayımlı sinir ağı” adı verilmektedir. Böylelikle ağırlıklar geriye döndürülen hata payları sayesinde, alınan tüm türevler yinelenerek ve daha optimize katsayılar dolayısıyla da gerçeğe daha yakın değerler elde edilecektir. Şekil 2.15’te geriye yayımlı bir YSA modeli görülmektedir.

Şekil 0.15: Genel Bir Geriye Yayımlı Ağ Mimarisi



Kaynak: Çetin Elmas, *Yapay Zekâ Uygulamaları*, 3. Baskı, Ankara: Seçkin Yayıncılık, 2016, s. 115.

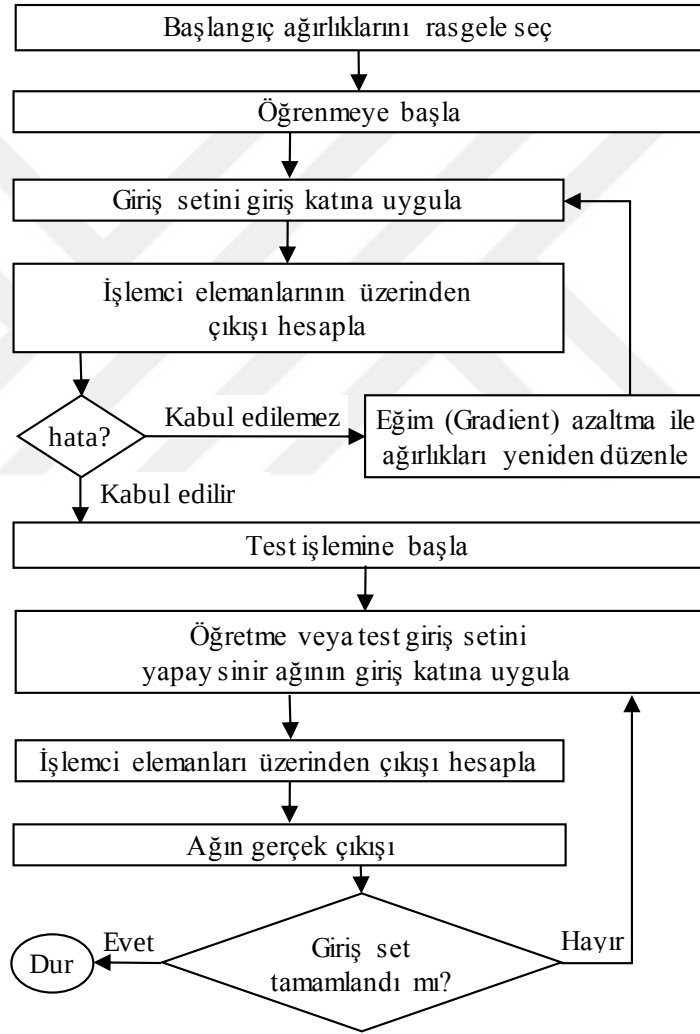
Modelin kendi ağırlıklarını optimize etmesi, modelin öğrenmesi anlamına gelmektedir. Hata payı atayarak geriye döndürülme suretiyle, öğrenme gerçekleşirken her döngüde (iterasyonda) gerçekleştirilen iyileşmeye (hata miktarını azaltmaya), “öğrenme oranı” denilmektedir. Problemin çözümünde en iyi sonuca ulaşabilmek için

²⁰⁸ Talip Torun, a.g.e. s.58.

²⁰⁹ Çetin Elmas, a.g.e. s.88.

öğrenme oranı dikkatli bir şekilde ayarlanmalıdır. Aksi takdirde düşük öğrenme oranıyla yerel minimuma yakalanma ya da çözümün çok fazla zaman alması gibi olumsuzlar yaşanabilecekken, bu oranın yüksek tutulması ise optimum çözümün atlanmasına, dolayısıyla en uygun çözüme hiçbir zaman ulaşamamasına neden olabilir²¹⁰. Şekil 2.16’da geriye yayımlı sinir ağının işleyiş şekli görülmektedir.

Şekil 0.16: Bir Ağda Geri Yayımlı Öğrenme Algoritmasının Uygulanması



Kaynak: Burçin Ataseven, “Yapay Sinir Ağları İle Öngörü Modellemesi”, **Öneri.** Cilt: 10. Sayı: 39, 2013, s.101-115.

²¹⁰ Atınç Yılmaz, Umut Kaya, **Derin Öğrenme**, 1. Baskı, İstanbul: İnkılap Kitabevi, 2019, s.12.

2.2.9. Yapay Sinir Ağlarının İşlevleri

YSA modeli, eldeki girdilerin, çıktılara nasıl dönüştüğüne bakarken açıklayıcı değişkenlerin optimal ağırlıklarını belirlemeye çalışır. Eğitilmiş bir YSA modeline, sonucu belli olmayan ama belli nicelikleri veya nitelikleri bilinen bir vaka sorulduğunda, yüksek olasılıkla makul bir sonuç alınacaktır. Ama bu sonuç ne kadar isabetli olacaktır? Yani model, kendi içinde, ağırlıklarını ne kadar doğru optimize edebilmiştir? Bunu öngörebilmek için araştırmacılar elindeki verinin tamamını eğitim verisi olarak kullanmaz. Onun %80-90 gibi bir oranını eğitim verisi geriye kalanını ise “test verisi” olarak ayırır. Model eğitildikten sonra modelin başarısı, test verileriyle test edilir. Bu yöntemde şu şekilde gerçekleştirilmektedir: araştırmacı elindeki veriyi programa yüklemekte ve bunun örneğin %10-20 sinin test verisi için tesadüfen seçilerek ayrılması için gerekli kodlamaları yapmaktadır. Fakat burada, test verileri tesadüfen seçildiği için, seçilen eğitim verisinin (bir bakıma örneklemin) standart hatası yüksek olabilir. Yani eğer bu seçim, tesadüfen verilerin genelini temsil edebilen bir seçim değilse sonuçlar tatmin edici olmayabilir. Bu yüzden bu işlemin yaklaşık on ila yirmi kez, yine rasgele bir şekilde, tekrarlanıp uygulanıp başarı oranlarının ortalamasının ve standart sapmasının alınması yoluna gidilerek, eldeki veri kümesinden eğitim veri kümesi oluştururken oluşabilecek standart hatayı azaltmak amaçlanır.

YSA tüm bu süreçleri dört temel işlev yardımıyla gerçekleştirir. Bunlar girdi işlevi, ağırlıkların oluşturulması işlevi, toplama işlevi ve son olarak etkinlik işlevidir.

2.2.9.1. Girdi İşlevi

Girdi işlevi, gözlemlerin yani veri kümesinin modele tanıtılmasıdır. Ne kadar çok girdi olursa modelin tecrübesi de o denli güçlü olacaktır. Girdi aşamasında sadece eldeki veriler modele girilirken herhangi bir işlem yapılmaz, yani ağırlıklandırılmaz. Otomatikleştirilmiş YSA sistemleri tıpkı sinir hücrelerinde olduğu gibi etraflarından topladıkları tüm veriyi sisteme alırlar. Düğümler eğitim sırasında büyür ve/veya küçülür. Sinir ağları ile elde edilen sonuçların, her zaman eğitildikleri verilere bağlı olduğu unutulmamalıdır. Sinir ağları, verilerde bulunan kalıpları tanımlama ve

öğrenmede mükemmeldir. Bununla birlikte, bir sinir ağı bir sonucu tahmin etmek için eğitilirse, ağı bu tahmini başarıyla gerçekleştirmeyi öğrenebilmesi için ağı girdiler arasında kestirimci faktörlerin sunulması gerekir. Daha genel bir ifadeyle, sinir ağının (örüntüleri) başarılı bir şekilde öğrenebilmesi için verilerde mevcut örüntüler (korlasyonlar) bulunmalıdır. Veriler herhangi bir model veya öngörücü faktör içermiyorsa, sinir ağı performansı yüksek olamaz. Bu nedenle, sinir ağları yalnızca verilere bağlı değildir, aynı zamanda bu verilerin içerdiği bilgilerle (korelasyonlar) da sınırlıdır²¹¹. Sonuç olarak, örüntü içermeyen verilerle çalışıldığı zaman, sinir ağlarının da yüksek performans göstermesi beklenmemelidir.

2.2.9.2. Ağırlıkların Oluşturulması İşlevi

Araştırmacının deneme ve yanılma yoluyla belirleyeceği gizli katman sayısı ve her bir gizli katmandaki nöron sayısı ve araştırma problemi çerçevesinde çıktı sınıfları belirlenir. Şekil 2.11’den de hatırlanacağı üzere giriş verileri, ilk gizli katmandaki (farklı YSA ağları olmakla birlikte en çok kullanıldığı şekilde) tüm nöronlara oklarla bağlıdır. Giriş verilerinin her biri, herhangi bir işleme tabi tutulmadan, bu nöronların her birine tek tek iletilir. Model, kendi içinde, sıfır ila bir arasında olmak kaydıyla, rasgele ağırlıklar atayarak katmanlar arasındaki iletişimi ve etkileşimi sağlar. Bu ağırlıklar oklarla gösterilir. Veriler ilk gizli katmandan ikinci gizli katmana (ikinci gizli katman yoksa çıktı katmanına) iletilirken, oklarla temsil edilen ağırlıklarla çarpılmak suretiyle, ağırlıklandırılarak ikinci gizli katmana iletilir. Bu aşama ağırlıklar işlevini oluşturur.

2.2.9.3. Toplama İşlevi

Herbir nörona, kendinden önceki katmandaki her bir nöronun gelen ağırlıklandırılmış sonuçların toplanarak söz konusu nöronun net değeri belirlenir. Bu toplama işlemi, yaygın olarak aşağıda gösterildiği (2.21) şekilde yapılır:

²¹¹ Judith E. Dayhoff, James M. DeLeo, “Artificial Neural Networks Opening the Black Box”, **CANCER Supplement**, 15, 2001, Cil: 91, Sayı: 8, s. 1615-1635.

$$n_i^{net} = \sum_{j=1}^{j_i} w_{ji} n_j - \theta_i \quad (0.21)$$

Burada, j_i adet nörondan gelen değerler (n_j), w_{ji} ile yani i . nörondan j . nörona bağlantının atanmış ağırlığı ile çarpılarak toplanır. θ_i ise i nöronunun eşikini (bias) göstermektedir. Eşik aşamayan değerler ise yanmaktadır²¹².

YSA’larda, herbir nöron (diğer bir ifadeyle proses elemanı) diğerlerinden farklı bir toplama fonksiyonuna sahip olabileceği gibi, hepsi aynı toplama fonksiyonuna da sahip olabilir²¹³. Toplama fonksiyonunun dışında kullanılan bazı birleştirme fonksiyonları da bulunmaktadır. Bunlar aşağıda listelenmiştir²¹⁴.

$$\text{Toplam} \quad a_j = \sum_i w_{ji} x_i \quad (0.22)$$

$$\text{Çarpım} \quad a_j = \prod_i w_{ji} x_i \quad (0.23)$$

$$\text{Maksimum} \quad a_j = \text{Maks}(w_{ji} x_i) \quad (0.24)$$

$$\text{Minimum} \quad a_j = \text{Min}(w_{ji} x_i) \quad (0.25)$$

$$\text{Çoğunluk} \quad a_j = \sum_i \text{Sgn}(w_{ji} x_i) \quad (0.26)$$

x_i = i nöronunun çıktısı

w_{ji} = j ve i nöronları arasındaki bağlantının ağırlığı

a_j = j nöronunun net girdisi

²¹² Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.20.

²¹³ Emre Yakut, **a.g.e.** s.54.

²¹⁴ Coşkun Hamzaçebi, “Geleceği Tahminde Yapay Sinir Ağları İçin Sezgisel Öğrenme Algoritması”, Doktora Tezi, Gazi Üniversitesi, 2005.

2.2.9.4. Etkinlik İşlevi

Toplama işlevi de gerçekleştirildikten sonra elde edilen sonuç değeri, iki çıkışlı bir model için, önceden belirlenen bir eşik değerinin üzerinde ise bir sınıfa, altında ise diğer sınıfa atayan bir atama işlevi gerçekleştirilmesi gerekmektedir. Yalnız bu sonuçların sınıflara atanabilmesi bu sonuçlarının değerinin ya $[0,1]$ ya da $[-1,1]$ arasında olması gerekmektedir²¹⁵. Çünkü atama işlevi ancak o zaman gerçekleştirilebilir. Dolayısıyla bu sonuç değerlerini, bir nevi normalize ederek, onları bu iki aralık değerlerinde temsil edilmesini sağlayan etkinlik (aktivasyon) fonksiyonlarının kullanılması gerekmektedir. Etkinlik işlevi Şekil 2.17'den Şekil 2.22'ye kadar gösterilen doğrusal ve eğrisel fonksiyonlardan biri yardımıyla gerçekleştirilmektedir²¹⁶. Etkinlik fonksiyonu, toplama fonksiyonundaki gibi, herbir nöron için aynı olabileceği gibi, farklı olarak da belirlenebilir. Fakat ÇKA, etkinlik fonksiyonunun, türevi alınabilir bir fonksiyon olmasını zorunlu kılar. Hangi etkinlik fonksiyonunun en iyi sonucu vereceğine dair bir formülasyon bulunmayışından dolayı, hangi fonksiyonun kullanılacağına yine deneme yanılma yoluyla karar verilebilmektedir. Tüm nöronlarda, toplama işlevinde olduğu gibi, aynı etkinlik fonksiyonunun kullanılması gerekmez ve her bir nöronda farklı etkinlik fonksiyonu kullanılabilir. En çok kullanılan etkinlik fonksiyonu, değerleri sıfır ila bir arasına taşıyan sigmoid fonksiyondur. Negatif değerleri de alması isteniyorsa hiperbolik tanjant fonksiyonu en çok başvurulan etkinlik fonksiyonudur²¹⁷. Cura, “Modern Sezgisel Teknikler Ve Uygulamaları” isimli kitabında bu fonksiyonları şu şekilde tespit etmektedir²¹⁸:

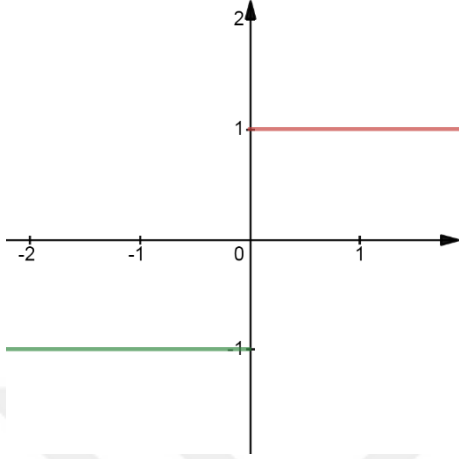
²¹⁵ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.20.

²¹⁶ Elmas Çetin, **a.g.e.** s.24.

²¹⁷ Emre Yakut, **a.g.e.** s.52-53.

²¹⁸ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.21-23 ve Elmas Çetin, **a.g.e.** s.32.

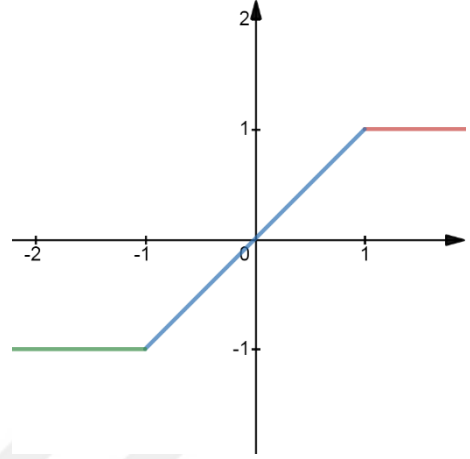
Şekil 0.17: Etkinlik İşlevi İçin Sık Kullanılan Basamak ve Parçalı Doğrusal Fonksiyonlar



$$g(x) = \begin{cases} 1 & \text{eğer } x > 0, \\ 0 & \text{eğer } x = 0, \\ -1 & \text{eğer } x < 0. \end{cases}$$

(0.27)

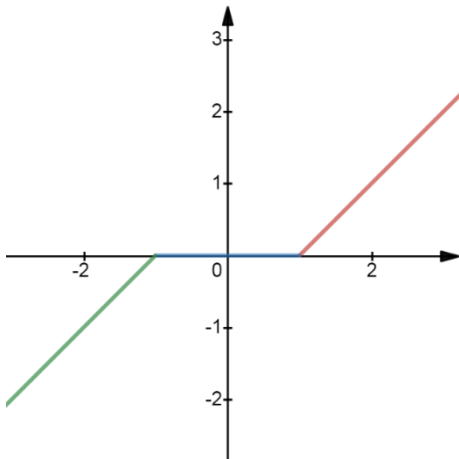
(Basamak [Step] Fonksiyon)



$$g(x) = \begin{cases} 1 & \text{eğer } x > 1, \\ 0 & \text{eğer } -1 < x < 1, \\ -1 & \text{eğer } x < -1. \end{cases}$$

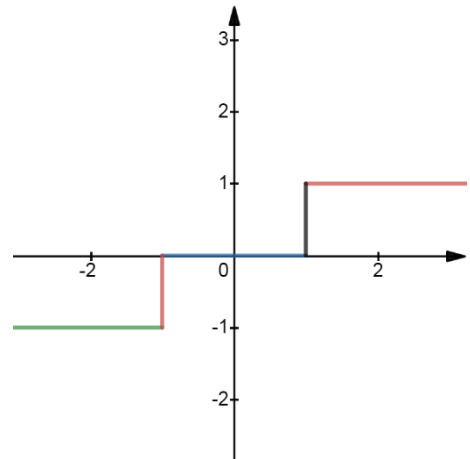
(0.28)

(Parçalı Doğrusal Fonksiyon)



$$g(x) = \begin{cases} x-1 & \text{eğer } x \geq 1, \\ 0 & \text{eğer } -1 < x < 1, \\ x+1 & \text{eğer } x \leq -1. \end{cases}$$

(0.29)

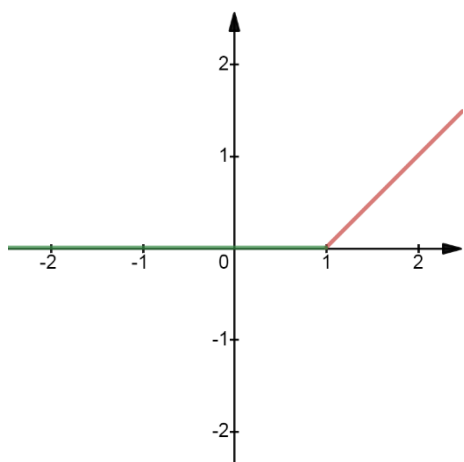


$$g(x) = \begin{cases} 1 & \text{eğer } x > 1, \\ 0 & \text{eğer } -1 < x < 1, \\ -1 & \text{eğer } x \leq -1, \end{cases}$$

$x = 1$ ise $0 < y < 1$,

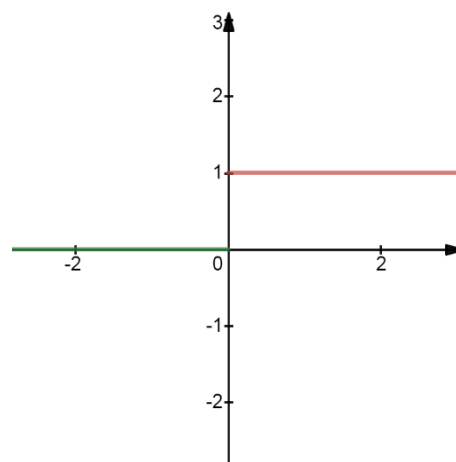
$x = -1$ ise $-1 < y < 0$.

(0.30)



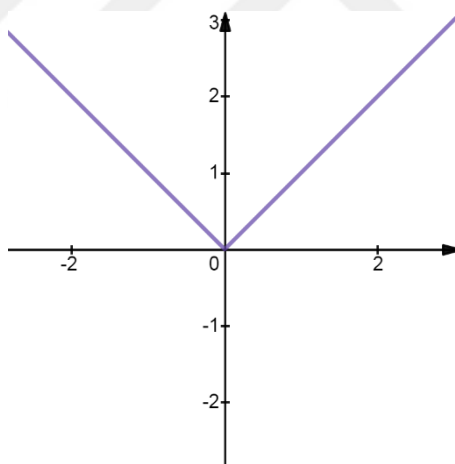
$$g(x) = \begin{cases} x - 1 & \text{eğer } x > 1, \\ 0 & \text{eğer } x \leq 1. \end{cases}$$

(0.31)



$$g(x) = \begin{cases} 1 & \text{eğer } x > 0, \\ 0 & \text{eğer } x \leq 0. \end{cases}$$

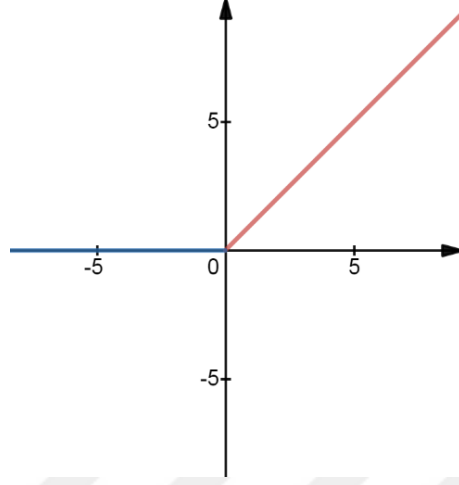
(0.32)



$$g(x) = |x|$$

(0.33)

Şekil 0.18: ReLU Etkinlik Fonksiyonu



$$g(x) = \begin{cases} x & \text{eğer } x > 0, \\ 0 & \text{eğer } x \leq 0. \end{cases} \quad (0.34)$$

Not: ReLU etkinlik işlevinin en büyük faydası, sıfırdan büyük tüm girdiler için ağıın daha hızlı eğitilmesini sağlayan, sabit türev değerine sahip olmasıdır²¹⁹.

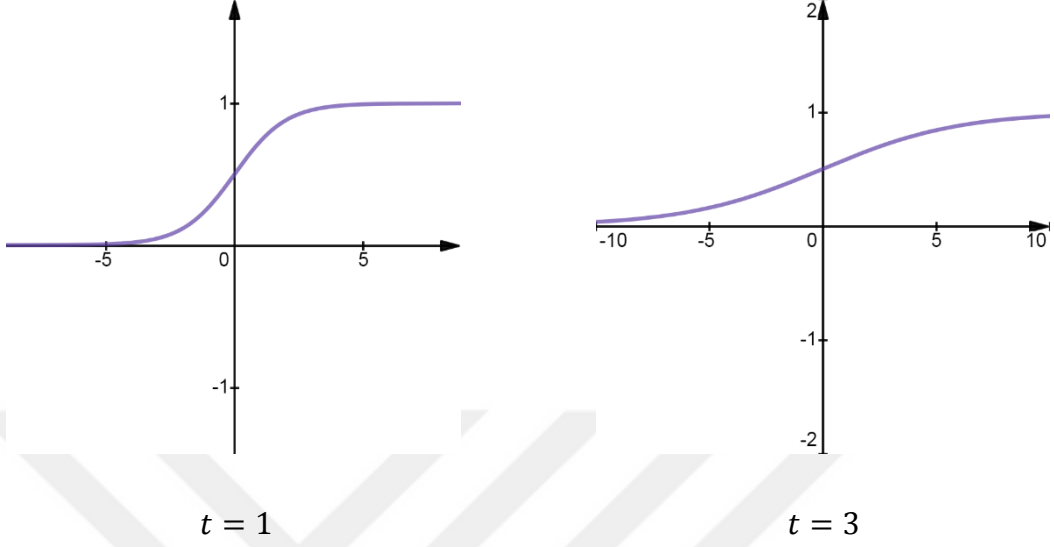
Lojistik Dönüştürme Fonksiyonu: Lojistik dönüştürme fonksiyonu bir nörondaki değerin sıfır ila bir arasında kalmasını sağlamaktadır.

Not: Aşağıdaki grafiklerde kullanılan t parametresi bir tür hassaslık derecesi olarak kullanılır.

$$g(x) = \frac{1}{1 + e^{-\frac{x}{t}}} \quad (0.35)$$

²¹⁹ Atınç Yılmaz, Umut Kaya, **a.g.e.** s.9.

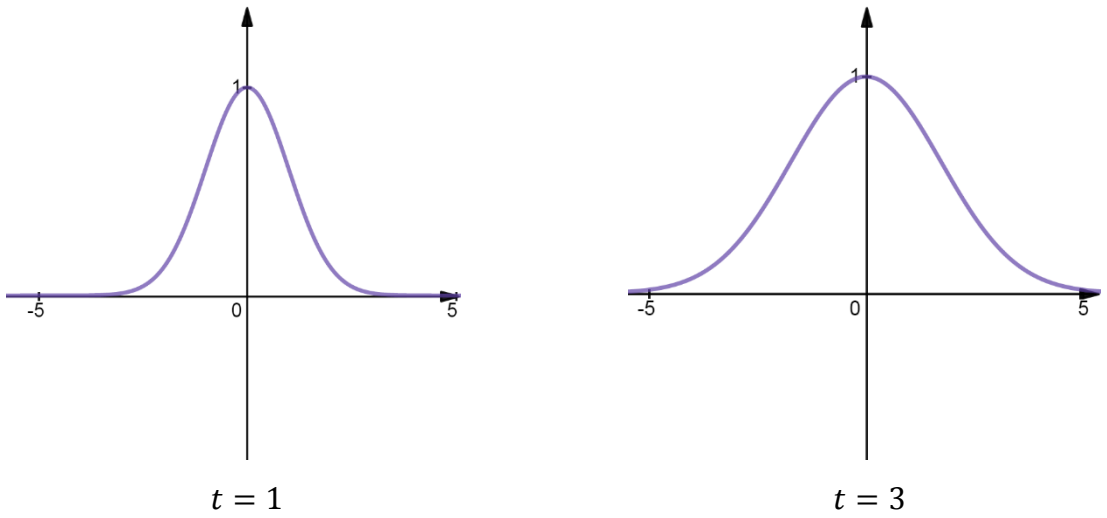
Şekil 0.19: Lojistik Dönüştürme Fonksiyonu



Gauss Dönüştürme Fonksiyonu

$$g(x) = e^{-\frac{x^2}{2t}} \quad (0.36)$$

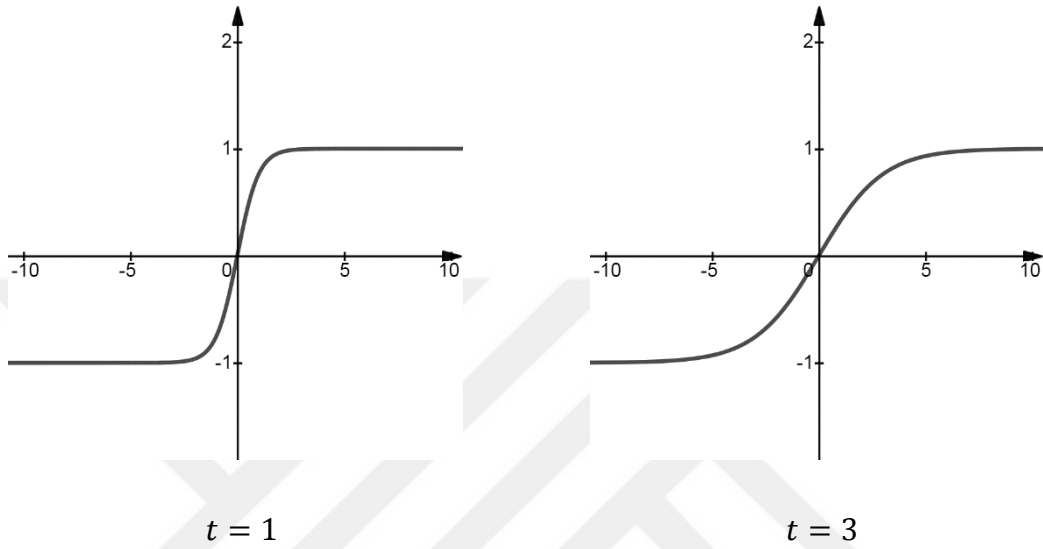
Şekil 0.20: Gauss Dönüştürme Fonksiyonları



Hiperbolik Tanjant Dönüştürme Fonksiyonu: Hiperbolik tanjant dönüştürme fonksiyonu bir nörondaki değerin $[-1,1]$ kalmasını sağlar.

$$g(x) = \frac{e^{\frac{x}{t}} - e^{-\frac{x}{t}}}{e^{\frac{x}{t}} + e^{-\frac{x}{t}}} \quad (0.37)$$

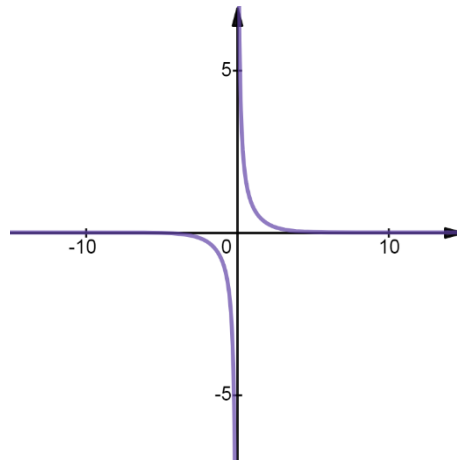
Şekil 0.21: Hiperbolik Tanjant Dönüştürme Fonksiyonları



Hiperbolik Sekant Dönüştürme Fonksiyonu

$$g(x) = \begin{cases} 1 & \text{eğer } x = 0, \\ \frac{2}{e^x - e^{-x}} & \text{değilse.} \end{cases} \quad (0.38)$$

Şekil 0.22: Hiperbolik Sekant Dönüştürme Fonksiyonu



Kaynak: Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.21-23 ve Elmas Çetin, **a.g.e.** s.32.

Tüm bu işlevler yerine getirildikten sonra elde edilen, ya $[0,1]$ ya da $[-1,1]$ arasında bu sonuçlarının, araştırmacı tarafından belirlenen bir atama kuralı yardımıyla, sınıflara atanmasıyla model çalıştırılmış olur.

2.2.10. XOR Problemi

Tablo 2.3'te AND, OR ve XOR (exclusive OR) problemlerinde sınıf değerleri ve isimleri görülmektedir. Şekil 2.23a ve Şekil 2.23b'de AND ve OR probleminin, doğrusal içerikli problemler olması neticesinde, sınıflar düz bir çizgiyle (doğrusal bir sınıflandırıcıyla) ayrılabilir. Şekil 2.24a'ya bakıldığında ise XOR problemini, doğrusal bir sınıflandırıcı kullanarak ayırmanın mümkün olmadığı, fakat Şekil 2.24b'de ise, birden fazla sınıflandırıcı (tek katmanlı değil çok katmanlı bir algılayıcı) kullanarak, çözülebildiği görülmektedir²²⁰.

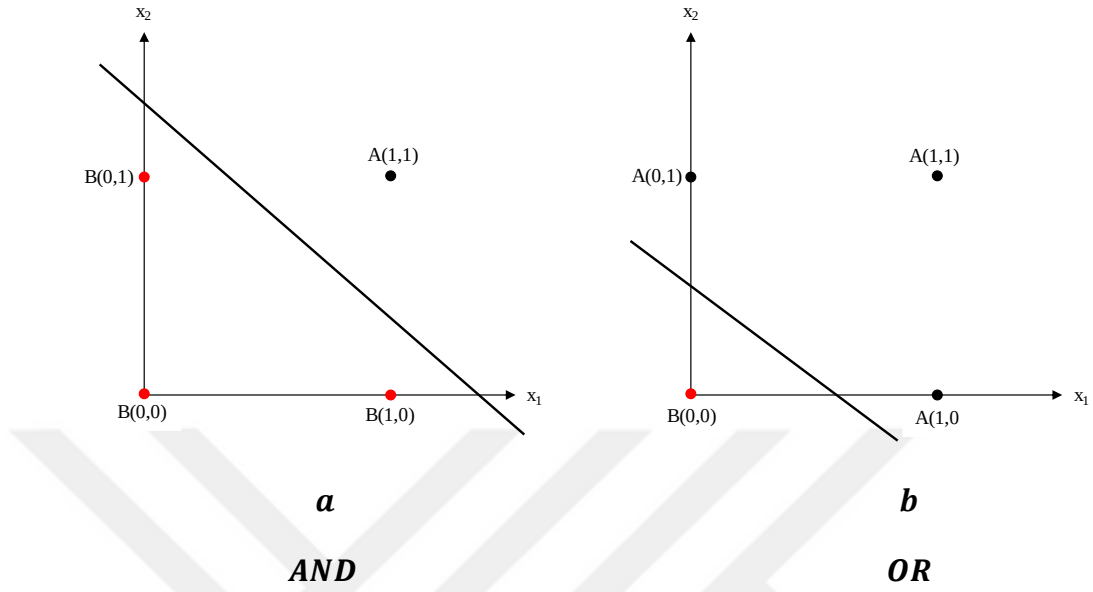
Tablo 0.3: OR, AND ve XOR Problemlerinde x ve y Değerleri

AND OR VE XOR PROBLEMLERİNDE x VE y DEĞERLERİ							
x ₁	x ₂	VE (AND)		VEYA (OR)		ÖZEL VEYA (XOR) (Exclusive OR)	
		y	SINIF	y	SINIF	y	SINIF
0	0	0	B	0	B	0	B
0	1	0	B	1	A	1	A
1	0	0	B	1	A	1	A
1	1	1	A	1	A	0	B

Kaynak: Sergios Theodoridis, Konstantinos Koutroumbas, **Pattern Recognition**, 4. Baskı, San Diego: Elsevier Inc., 2009, s.119.

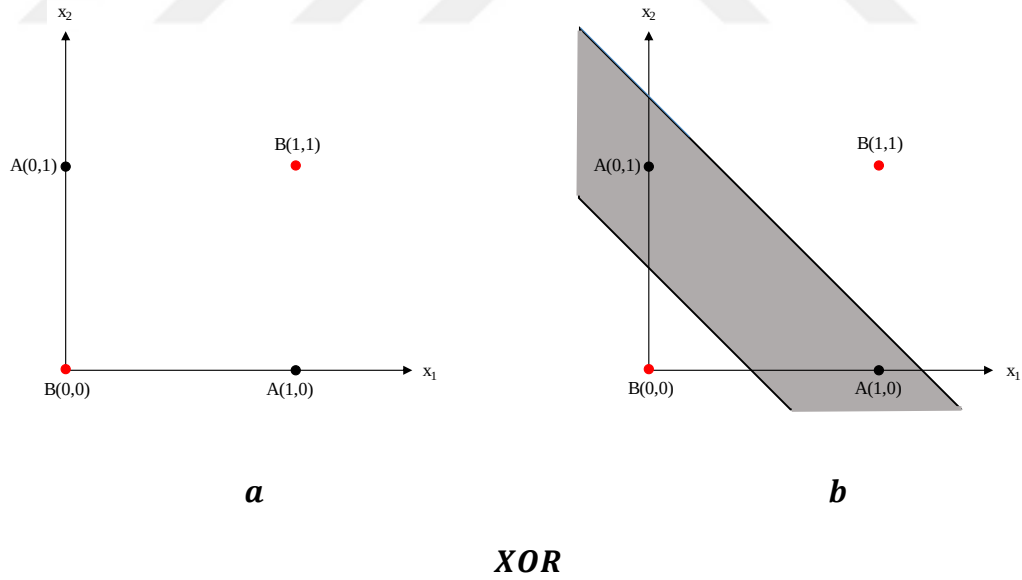
²²⁰ Sergios Theodoridis, Konstantinos Koutroumbas, **Pattern Recognition**, 4. Baskı, San Diego: Elsevier Inc., 2009, s.119.

Şekil 0.23: AND ve OR Probleminde Sınıflandırma



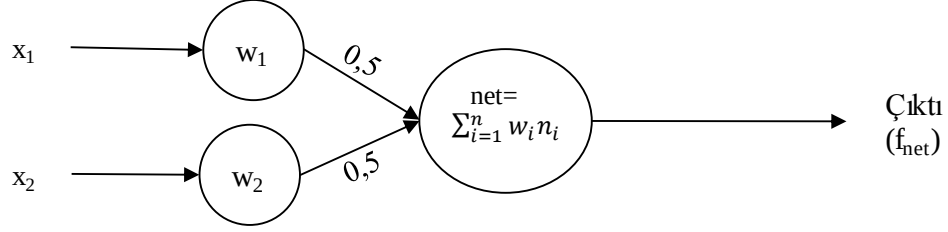
Kaynak: Sergios Theodoridis, Konstantinos Koutroumbas, **Pattern Recognition**, 4. Baskı, San Diego: Elsevier Inc., 2009, s.119.

Şekil 0.24: XOR Probleminde Sınıflandırma



Kaynak: Sergios Theodoridis, Konstantinos Koutroumbas, **Pattern Recognition**, 4. Baskı, San Diego: Elsevier Inc., 2009, s.119.

Şekil 0.25: AND ve OR Problemlerinin Tek Katmanlı YSA İle Çözümü



AND problemi için çıktının değerini vermesi için kullanılan sınıflandırma fonksiyonu:

$$f(x) = \begin{cases} 1 & \text{eğer } x = 1, \\ 0 & \text{eğer } x < 1. \end{cases}$$

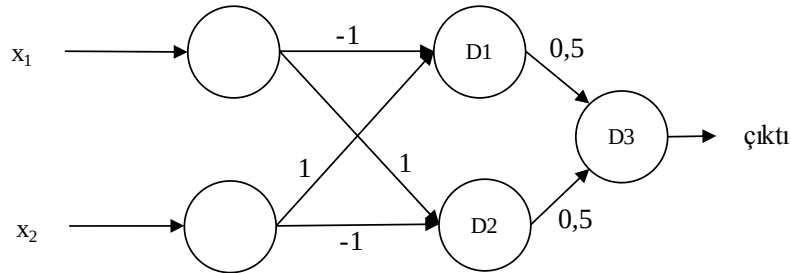
OR problemi için çıktının değerini vermesi için kullanılan sınıflandırma fonksiyonu:

$$f(x) = \begin{cases} 1 & \text{eğer } x > 0,5, \\ 0 & \text{eğer } x \leq 0,5. \end{cases}$$

XOR Probleminin Çok Katmanlı YSA Kullanılarak Çözülmesi

Perseptron da dâhil olmak üzere hiçbir doğrusal model Şekil. 2.24'den de anlaşılacağı gibi XOR problemini çözememektedir. Bunun için XOR problemi perseptron modeline bir gizli katman ilavesi yapılarak çözülebilmektedir. Şekil 2.26'daki tek gizli katmana sahip ÇKA ağı üzerindeki ağırlıklarla XOR problemi doğru bir şekilde çözülebilmektedir. Çözüm işlemlerinin ayrıntısına Çetin Elmas'ın kitabından ulaşılabilir²²¹.

Şekil 0.26: XOR Probleminin Çok Katmanlı YSA İle Çözümü



²²¹ **Kaynak:** Çetin Elmas, *Yapay Zekâ Uygulamaları*, 3. Baskı, Ankara: Seçkin Yayıncılık, 2016.

D1, D2 ve D3 düğümleri için kullanılan sınıflandırma fonksiyonu:

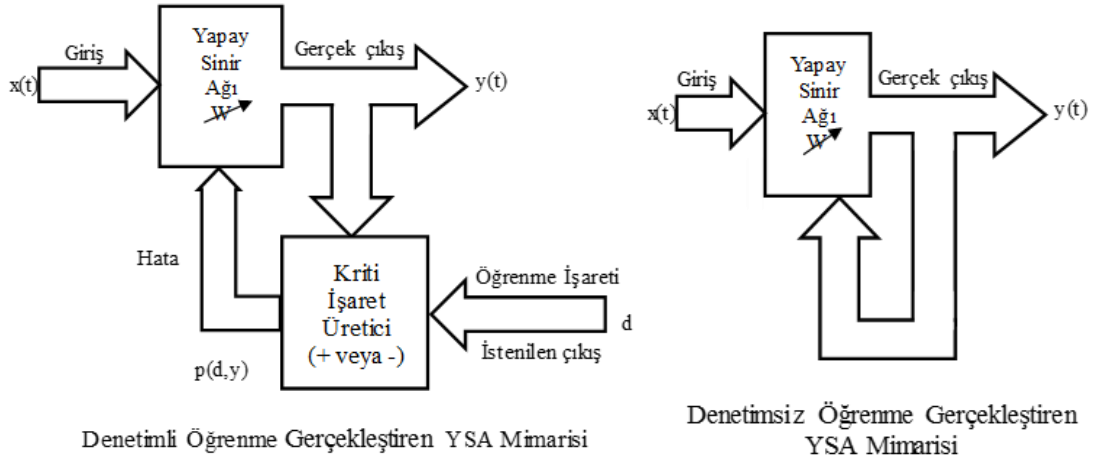
$$f(x) = \begin{cases} 0 & \text{eğer } x < 0,5, \\ 1 & \text{eğer } x \geq 0,5. \end{cases}$$

Delta kuralı kullanılarak ağırlıklar, iteratif bir şekilde, belirlenebilmektedir.

2.2.11. Yapay Sinir Ağlarının Eğitilmesi

YSA’nda öğrenme; yeni bağlantılar oluşturma, var olan bağlantıları silme, bağlantı ağırlıklarını değiştirme, birimlerin eşik değerlerini değiştirme, etkinlik fonksiyonunu değiştirme, yeni birimler oluşturma, var olan birimleri veya bağlantıları kaldırma vb. birçok yöntemle gerçekleştirilebilir. Fakat bunların içinde en çok kullanılanı, bağlantı ağırlıklarını değiştirme ya da optimize etme diyebileceğimiz yöntemdir²²². Sonuç değerlerinin modele tanıtılıp tanıtılmaması, eğitimin denetimli mi yoksa denetimsiz mi olacağını belirlemektedir ve bu iki farklı sistematik, şekil 2.27’de gösterilmektedir.

Şekil 0.27: Yapay Sinir Ağlarında Denetimli Ve Denetimsiz Öğrenme Mimarileri



Kaynak: Burçin Ataseven, “Yapay Sinir Ağları İle Öngörü Modellemesi”, **Öneri**. Cilt: 10. Sayı: 39, 2013, s.103.

Eğitimin denetimli mi yoksa denetimsiz gerçekleştirileceği, elbette elde sonuç değerlerinin elde bulunup bulunmamasına bağlıdır ve bulunması durumunda elbette

²²² Şenol Emir, **a.g.e.**, s.72.

en iyi modelin oluşturulabilmesi için sonuç değerleri kullanılır, yani denetimli öğrenme gerçekleştirilir. Denetimli ya da denetimsiz olsun, bağlantı ağırlıklarının, geriye döndürülme yoluyla, iteratif bir şekilde, güncellenmesi sırasında yaklaşımsal olarak bazı öğrenme kuralları kullanılır. Bunlar; Hebb, Hopfield, Delta, Eğimli İniş ve Kohonen öğrenme kurallarıdır.

Öğrenme kurallarının bazıları çevrimiçi (on-line) öğrenebilmekte ve gerçek zamanlı çalışabilmektedirler. Yani işlevlerini yerine getirirken diğer yandan öğrenmeye (veri girişine) devam etmektedirler. ART ağının öğrenme kuralı ve Kohonen öğrenme kuralı bu sınıfa dâhil edilebilir. Delta öğrenme kuralı gibi diğer bazı çevrimdışı (off-line) öğrenme yöntemleri ise belli bir eğitim verisi ile eğitimi tamamlandıktan sonra çevrime kapatılır, eğitim durdurulur ve model kullanılır. Tekrar eğitilmek istendiğinde kullanımı durdurulur yeniden eğitilir ve tekrar eğitimi durdurularak modelin kullanımına devam edilir²²³.

2.2.11.1. Hebb Öğrenme Kuralı

Birçok sinir ağı için öğrenme, sinaptik ağırlıkların iki yönlü modifikasyonu ile gerçekleşir. Bu öğrenme türlerinden biri de, Hebbian (Hebb ve Anti-Hebb) kurallara benzeyen, basit aktiviteye bağlı kurallara göre gerçekleşen öğrenme yöntemidir. Bu tür ağlar birden fazla hafızayı depolayabilir, giriş modelleri için seçicilik geliştirebilir, optimum çözümler bulabilir ve topografik haritaları düzenleyebilirler²²⁴.

Doğal olarak akla gelen ilk soru, tam olarak öğrenme veya hatırlama sürecinin bir parçası olarak düşünülmesi gereken şey, nöronlar arasındaki ağırlıkların neye göre artırılacağı ya da azaltılacağıdır. Bu, özellikle yeni YSA mekanizmaları tasarlanırken, zorlu bir süreci temsil etmektedir. Sinir ağlarındaki sinapslar, denetimli veya denetimsiz bir ağırlık değişim süreci ile güçlendirilebilir. Fakat ilerleyen çalışmalarla birlikte Sinir Fizyoloğu Donald Hebb zaman içinde bir ağırlık değişimi kuralı önerdi. Bu kuralın uygulanması, YSA'nın tasarımında oldukça verimli olmuştur. Temel kural,

²²³ Ercan Öztemel, **a.g.e.** s.26.

²²⁴ John Lisman, "A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory", **Proc. Natl. Acad. Sci. USA**, Cilt: 86, 1989, s.9574-9578.

sinaptik verimlilikteki değişikliklerin, karşılık gelen nöronların yerel özelliklerine ve aktivitesine dayandırılmasıdır. Yani, insan beyninin biyolojik sinir ağlarıyla yaptığı gibi, özellikleri ortak olan obje ve olaylar arasındaki çağrışımların kuvvetlendirilmesi gibi, Hebb kuralıyla öğrenen YSA, birbirleriyle bağlantıları olan nöronların çıktı değerleri, pozitiflik ya da negatiflik noktasında özdeşse (matematiksel olarak işaretleri aynıysa) bu nöronlar arasındaki ağırlıklar artırılır, değilse azaltılır²²⁵.

Hebb öğrenme kuralında bağlantılar arası ağırlık değişimi aşamasında korelasyon kullanılmaktadır. Hebb öğrenme kuralı, girdi verisi birbirinden bağımsız ve dikey yapılı ise, yüksek verim sunabilmektedir. Bu yöntemde, bir hücre aktif ise kendine bağlı olan hücreleri de aktif, pasif ise kendine bağlı olan hücreleri de pasif hale getirmeye çalışmaktadır. Bu öğrenme kuralında sıfıra yakın değerler ile başlanılır. Matematiksel olarak, artı (+) ya da eksi (-) aynı işaretli hücreler arasındaki ağırlıkları artıran bir yöntemdir. Hebb öğrenme kuralına göre giriş ve çıkış değerlerinin çarpımı pozitif ise w_{ij} ağırlığında bir artış, negatif ise bir azalış gerçekleşir. Ağırlık vektöründeki artış ise aşağıdaki şekilde gerçekleştirilir²²⁶:

Öğrenme sinyali:

$$f(w_i^t x) \quad (0.39)$$

Ağırlık vektöründeki artış:

$$\Delta w_i = c f(w_i^t x) x \quad (0.40)$$

veya,

$$\Delta w_{ij} = c o_i x_j \quad (0.41)$$

şeklinde olur. 1949 yılında geliştirilen Hebb öğrenme kuralı, bilinen en eski öğrenme kuralı olup, diğer tüm yöntemlerin temelini teşkil etmektedir.

²²⁵ Eduard Kuriscak, Petr Marsalek, Julius Stroffek, Peter G. Toth, "Biological Context Of Hebb Learning In Artificial Neural Networks, A Review", **Neurocomputing**, Cilt: 152, 2015, s.32.

²²⁶ Emre Yakut, **a.g.e.** s.67-68.

2.2.11.2. Hopfield Öğrenme Kuralı

Hopfield ağı fark denklemi tarafından yönetilen bir sinir ağıdır. Fark denklemi (2.42):

$$x_i(t + 1) = \text{sgn} \left(\sum_{j \neq i} w_{ij} x_j(t) \right) \quad (0.42)$$

Burada $x_i(t)$, i . nöronun $\{+1, -1\}$ durumunu, w_{ij} simetrik ağırlık matrisini ve $x_i(n) = \pm 1$ ise i . nöronun eş zamansız olarak gerçekleştirilen n . güncellemesinin durumunu göstermektedir. Ağ başlangıç durumunda kurulur, sonrasında serbest bırakılır ve dinamikler ağı en yakın sabit noktaya doğru götürür. Bir Hopfield ağının öğrenme kuralının görevi, gerekli modelleri ağ dinamiklerinin sabit noktaları olarak saklayan, bazı w_{ij} ağırlık matrisini bulmaktır. Böylece desenler gürültülü veya eksik de olsa, ilk girişlerden geri çağırılabilir²²⁷.

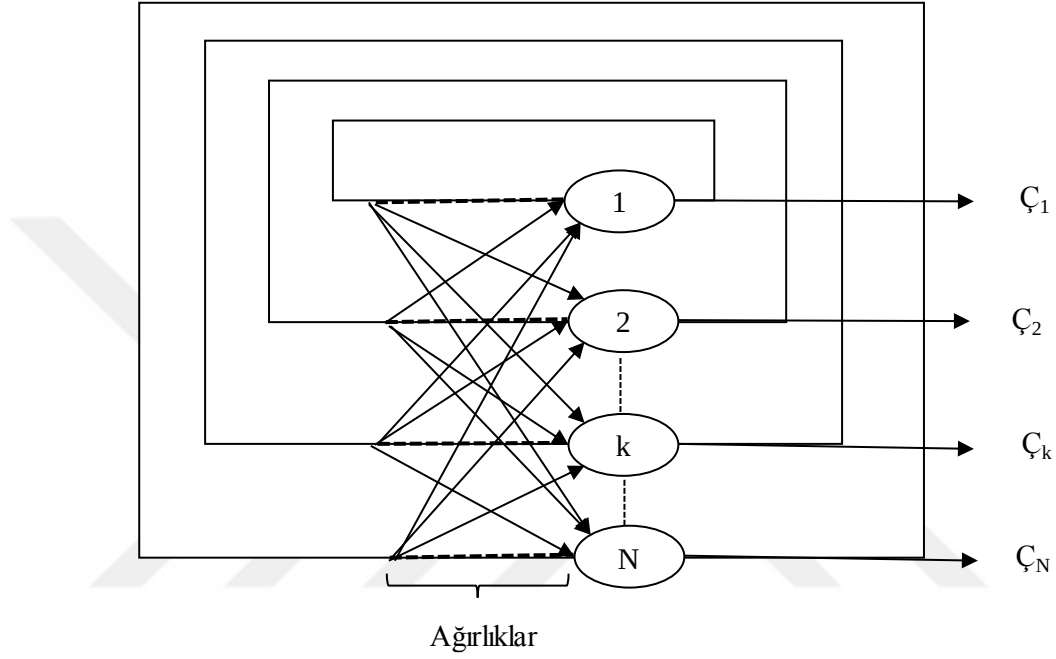
Hopfield öğrenme kuralı, Hebb öğrenme kuralı ile aynı yaklaşıma (her iki hücrede aktif ya da pasif ise aradaki bağlantının güçlendirilmesi) sahip olup, farkı bu işlemi sıfır ila bir arasındaki, kullanıcı tarafından belirlenen bir sabit ve pozitif bir değer olan “öğreme katsayısı” yardımı ile gerçekleştiriyor olmasıdır. Hopfield, tek katmanlı (girdi katmanı, gizli katman ve çıktı katmanı şeklinde ayrılmış katmanları olmayan) ve geri dönüşümlü bir ağ olup, proses elemanlarının hepsinin hem girdi hem de çıktı olarak kullanıldığı bir ağıdır. Hopfield ağı, çağrışımlı bellek (associative memory) olarak kullanılan kesikli (discrete) ve kombinatorial optimizasyon problemlerinde kullanılmakta ve sürekli ya da kesikli şeklinde ikiye ayrılmaktadır. Bu iki yaklaşım arasındaki temel fark ise, kesikli ağlarda aktivasyon adına signum fonksiyonu kullanılırken, sürekli ağlarda sigmoid fonksiyon kullanılır. Hopfield ağlarının en önemli uygulama alanı, geleneksel optimizasyon yöntemleri ile çözülemeyen, “gezgin satıcı” problemi olmuştur²²⁸.

²²⁷ A.J. Storkey, R. Valabregue, “The Basins Of Attraction Of A New Hopfield Learning Rule”, **Neural Networks**, Cilt: 12, 1999, s.870.

²²⁸ Ercan Öztemel, **a.g.e.** s.26.

Hopfield ağlarında hafıza kararlıdır ve geri çağrılabilir. Belirsizlikler istatistiksel olarak çözülür. Ağın genelleme kapasitesi vardır ve hafızanın zaman sıralaması da kodlanabilmektedir²²⁹. Şekil 2.28’de Hopfield ağının yapısı görülmektedir.

Şekil 0.28: Hopfield Ağının Yapısı



Kaynak: Emre Yakut, “Veri Madenciliği Tekniklerinden C5.0 Algoritması Ve Destek Vektör Makineleri İle Yapay Sinir Ağlarının Sınıflandırma Başarılarının Karşılaştırılması: İmalat Sektöründe Bir Uygulama”, Doktora Tezi, Atatürk Üniversitesi, 2012, s.92.

2.2.11.3. Delta Öğrenme Kuralı

Widrow ve Hoff ADALINE için öğrenme kuralı olarak, “Delta Öğrenme Kuralı” olarak da bilinen “En Küçük Ortalama Kareler (EKOK) - Least Mean Square (LMS)”, yöntemini önermişlerdir²³⁰. Regresyondaki en küçük hata karelerine kısmen benzeyen bu yöntem, yine hata kareleri toplamının minimize etme mantığıyla, çıkış nöronundaki hataları, giriş katsayılarını optimize ederek giderme yaklaşımıyla çalışan bir öğrenme kuralıdır. Kabaca, hatalı ürün üreten bir fabrikanın, ürünün neden hatalı üretildiğini

²²⁹ J. J. Hopfield, “Neural Networks And Physical Systems With Emergent Collective Computational Abilities”, **Proc. NatL Acad. Sci. USA**, Cilt: 79, 1982, s.2558.

²³⁰ Coşkun Hamzaçebi, **Yapay Sinir Ağları Tahmin Amaçlı Kullanımı, MATLAB Ve NEUROSOLUTION Uygulamalı**, 1. Baskı, Bursa: Ekin Yayınevi, 2011, s.37.

geriye doğru soruşturan ve hataların düzeltilmesini sağlayan bir kalite kontrol sürecine benzetilebilir. Sonuç olarak bu yöntem giriş-çıkış verileri arasındaki farkı gidermek için, belli bir öğrenme oranı yardımıyla, ağırlıkların yavaş yavaş değiştirilmesi mantığına dayanmaktadır. Formülasyon ise aşağıdaki gibidir²³¹:

$$E_m = B_m - \zeta_m \quad (0.43)$$

Burada E hatayı, B beklenen çıktıyı ζ ise gerçekleşen çıktıyı ve son olarak m ise m . proses elemanını göstermektedir. Yani bu formül m . proses elemanının hatasını göstermektedir. Çıktı katmanı için oluşan toplam hatayı (TH) bulmak için ise tüm hataları toplamak gerekir. Pozitif ve negatif değerli hataların birbirini yok etmesini engellemek için hata kareleri alınır ve bu durumda toplama hata formülü Denklem 2.44'deki gibi olur:

$$TH = \frac{1}{2} \sum_m E_m^2 \quad (0.44)$$

Ayrıntılı hesaplama bilgisi için, Ercan Öztemel'in kitabına başvurulabilir.

Bilindiği üzere, $f(x)$ şeklindeki bir fonksiyonun türevi $f'(x)$ olur ve bir fonksiyonun birinci türevi, o fonksiyonun x noktasındaki eğimine eşittir. Parabolun tepe noktaları bulunurken kullanıldığı gibi, bir fonksiyonun, yerel olsun evrensel olsun, maksimum ya da minimum olsun, tepe noktalarında birinci türevi sıfıra eşit olur. Bu durumda bir fonksiyonun grafiği üzerinde, eğer fonksiyonun birinci türevi olan $f'(x)$ negatifse (ki burada fonksiyonun değişkeni olan x 'de negatif olacaktır) sağa doğru; yok eğer fonksiyonun birinci türevi $f'(x)$ pozitif ise (ki burada fonksiyonun değişkeni olan x 'de pozitif olacaktır) sola doğru, fonksiyonun birinci türevi $f'(x)$ sıfıra (0) eşit oluncaya dek ilerlenir. Sonra fonksiyonun birinci türevi $f'(x) = 0$ olduğu nokta bize fonksiyonun minimum değerini verecektir. Bu durum Şekil 2.29'da daha net görülmektedir.

²³¹ Ercan Öztemel, **a.g.e.** s.77-80.

YSA’da eğitim süreci, optimum ağırlıkların belirlenme sürecidir. Fakat optimum sonuç için, türevlerin sıfıra eşitlenerek bulunduğu minimum noktanın tek olması gerekliliği vardır. Eğrisel bir konveks fonksiyonun herhangi bir değişkenin belirli bir noktadaki eğimine, yani kısmi türevine, ters yönde hareket edilmesiyle hatanın en az olduğu noktaya, yani eğimin sıfır olduğu noktaya ulaşılır. Bu yöntemle gradyan azaltma (gradient descent) yöntemi denilmektedir²³². Gradyan azaltma yöntemi, eğimli iniş algoritması olarak da bilinmekle birlikte, bir benzetme yapacak olursak bir tepeden aşağı inilmek istendiğinde, eşit mesafelerde en az eğimi veren yol haritasını takip ederek minimum yükseklikteki noktayı bulmayı sağlayan bir yöntemdir. Matematiksel açıdan bir fonksiyonunun yerel minimumunu bulmak için, ilgili fonksiyonun negatif eğimine göre ilerlenmelidir²³³. Gradyan azaltma yöntemi, fonksiyonun en küçük olduğu noktayı ararken, Şekil 2.29’da görüldüğü gibi kısmi türev;

$$\Delta x'_i = f'(x_i^{t-1})\varphi \quad (0.45)$$

$$x'_i = x_i^{t-1} + \Delta x'_i \quad (0.46)$$

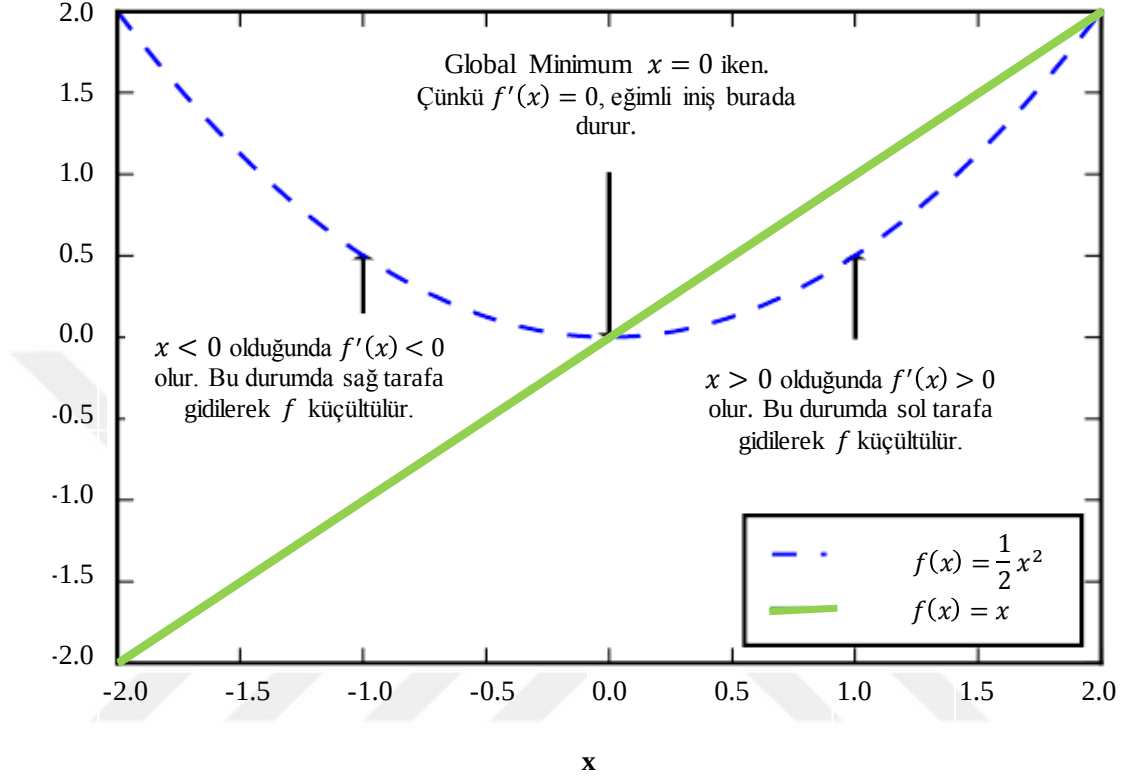
kullanır. Bunun yanı sıra nöron hataları kısmi türev yoluyla geriye doğru döndürülür²³⁴.

²³² Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.26.

²³³ Atınç Yılmaz, Umut Kaya, **a.g.e.** s.12.

²³⁴ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.26-27.

Şekil 0.29: Delta Öğrenme Kuralı ve Gradyan Azaltma

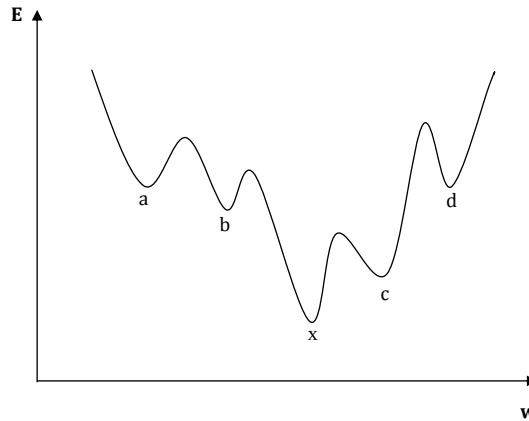


Kaynak: Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **Deep Learning**, 1. Baskı, Cambridge: The MIT Press, 2017s. 81.

Delta kuralında hata düzeltme, belli bir öğrenme oranı yardımıyla, katsayılar yavaş yavaş yani iteratif olarak değiştirilme yoluyla yapılır. Bu hata düzeltme, Şekil 2.29'da gösterildiği gibi, gradyan azalma (gradient descent) yöntemi kullanılarak yapılır. Böylece, ağırlıkları artırma ya da azaltma yoluyla, en küçük hata karelerini temel alarak, en kısa yoldan evrensel (global) minimuma ulaşılmış olur. Fakat bu çözüm, modelin evrensel minimum çözümünü bize verir. Bir hipotez uzayında, birçok yerel minimum nokta bulunmasından ve modelin bu noktalara takılıp kalma ihtimalinden dolayı, en küçük ortalama kareler yöntemi, gerçek evrensel çözümü malesef garanti edememektedir. Bulunan çözümün gerçek evrensel çözüm olması da ihtimal dâhilindedir. Bu durum Şekil 2.30'da daha net olarak görülmektedir.

YSA başlangıçta atanan rasgele ağırlıklarla başlar. Sonra delta öğrenme kuralını kullanarak, en küçük hata kareleri toplamını dikkate alarak, eğimli iniş yapar ve bulduğu ilk minimum noktadaki ağırlık değerlerini saptar ve modelini bu ağırlıklarla oluşturur. Minimum noktaya gelince yöntem durur. Fakat böyle problemlerde, grafiksel bakış açısıyla (Bkz. Şekil 2.30) daha net anlaşılacağı üzere, göz önüne alınması gereken en önemli sorunlardan biri, birden fazla minimum nokta bulunabileceği ve yöntemin bunlardan birine takılıp kalması ihtimalidir. Yani yöntemin, yerel minimuma yakalanmadan, evrensel minimuma ulaşması her zaman mümkün olamayabilir. Hatta bulunan minimumun evrensel mi yerel mi olduğu da meçhuldür. Şekil 2.30'a bakıldığında a, b, c ve d noktaları yerel minimum noktaları x noktası ise evrensel minimum noktasını göstermektedir. Grafikten de anlaşılacağı üzere bulunan ilk nokta a, b, c ve d noktalarından birisi olursa elde edilen sonuç, hatanın minimizasyonunun tam anlamıyla gerçekleştiği, evrensel minimum nokta olmayacaktır. Bu da hatanın tam olarak minimize edilememesi sonucunu doğurmaktadır. Grafikteki yatay eksen (w) ağırlıkları, dikey eksen (E) ise hata kareleri toplamını göstermektedir. Literatürdeki bazı çalışmalarda YSA'nın, yerel minimuma yakalanma konusunda, aktivasyon fonksiyonu kullanması dolayısıyla, doğrusal olmayan bir yapı kazanması sayesinde, K en yakın komşu gibi diğer doğrusal yöntemlere göre, daha avantajlı olduğu fakat sezgisel yöntemlerle kıyaslandığında ise, daha geri kaldığına dair tespitler bulunmaktadır²³⁵.

Şekil 0.30: YSA'nda Yerel Ve Evrensel Minimum Noktaları

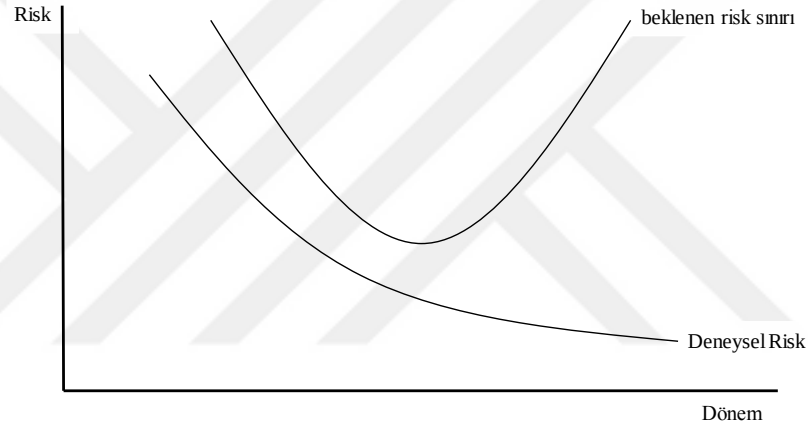


²³⁵ Ercan Öztemel, **a.g.e.** s.27.

Delta öğrenme kuralı, diğer öğrenme yöntemlerinde olduğu gibi Hebb öğrenme kuralı mantığıyla (matematiksel olarak aynı işaretli olan hücre bağlantılarının güçlendirilmesi) çalışmakta, farklı olarak da bu işi, iterasyon yoluyla, sürekli bir hale getirmesi ve hangi çözümün en iyi olduğuna ise hata kareleri ortalamasını referans alarak karar vermesidir²³⁶.

Delta öğrenme kuralı kullanarak, eğitim hatasının, deneysel riskin minimize edilmesi yoluyla, minimize edilmesi uygulaması Şekil 2.31'deki grafikte görülmektedir.

Şekil 0.31: Deneysel Risk ve Beklenen Risk İlişkisi



Kaynak: Şenol Emir, “Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması: Borsa Endeks Yönünün Tahmini Üzerine Bir Uygulama”, Doktora Tezi, İstanbul Üniversitesi, 2013, s.101.

Çoğu derin öğrenme algoritması, stokastik gradyan inişi adı verilen bir optimizasyon algoritmasına dayanır²³⁷.

2.2.11.4. Kohonen Öğrenme Kuralı

Kendi Kendini Düzenleyen Haritalar (Self-Organizing Maps-SOM) veya diğer bir deyimle, Kohonen ağları 1980’lerde geliştirilen en önemli ağ yapılarından birisidir. Kendi kendini düzenleyen haritaların ana fonksiyonu, n-boyutlu uzaydaki girdi

²³⁶ Ercan Öztemel, **a.g.e.** s.27.

²³⁷ Ian Goodfellow, Yoshua Bengio ve Aaron Courville, **a.g.e.** s. 98.

verilerini, orijinal topolojik ilişkilerini koruyarak daha düşük boyutlu (genellikle bir veya iki) çizimle haritalaştırmasıdır²³⁸.

Kohonen öğrenme kuralında nöronlar, ağırlıklarını değiştirmek için birbiri ile yarışır. Yarışın kazananı, en büyük çıktıyı üreten hücredir ve bağlantı ağırlıkları değiştirilir. Bu işlem o hücrenin, çevredeki hücrelerden daha kuvvetli olması demektir. Dolayısıyla sistematik olarak, kazanan hücrelerin ve diğer çevresel hücrelerin ağırlıklarının değiştirilmesine izin verilir²³⁹. Kazanan hücre doğru sınıfa düşerse, ilgili hücrenin ağırlıkları girdi vektörüne yaklaştırılır, aksi takdirde uzaklaştırılır. Her iki durumda da β öğrenme katsayısı, A ağırlık vektörü, X girdi vektörü olmak üzere ağırlıkların değiştirilmesi formülü şu şekilde ifade edilir²⁴⁰:

$$A_{yeni} = A_{eski} + \beta (X - A_{eski}) \quad (0.47)$$

2.2.11.5. Grossberg Öğrenme Kuralı

İnsan beyninin biyolojik fonksiyonlarına yoğunlaşan Grossberg 1976 yılında Adaptif Rezonans Teorisini (Adaptive Resonance Theory, ART) geliştirmiştir. Grossberg modelini, insan beyninde yer alan üç farklı biyolojik fonksiyondan esinlenerek geliştirmiştir. Bunlar, çevrede meydana gelen büyük değişimlere uyum sağlama anlamına gelen “normalizasyon”, yine çevrede meydana gelen değişimi ayrıntılı olarak fark edebilme anlamına gelen “ayırıştırma” ve son olarak da algılanan değişimin saklandığı “hafıza” fonksiyonlarıdır. Hafıza fonksiyonu, anlık kararlar için kısa dönemli hafızayı, normal olarak verilen kararlar içinse uzun dönemli hafızayı kullandığı belirtilmektedir. ART ağları, girdi özelliklerinin gösterildiği F1 katmanından ve ayrıştırılmış sınıfların gösterildiği F2 katmanından oluşmaktadır. Girdiler F1 katmanından alınır, sınıflandırma ise F2 katmanında yapılır. Sonrasında yapılan sınıflandırma F1 katmanına gönderilerek bir karşılaştırma yapılır ve uymuyorsa yeni bir kategori oluşturulur. ART ağları aşağıdan yukarıya ve yukarıdan

²³⁸ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.21-23.

²³⁹ Ercan, Öztemel, **a.g.e.** s.27.

²⁴⁰ Emre Yakut, **a.g.e.** s.67.

aşağıya olmak üzere iki yönlü olarak çalışır. ART ağları, birçok ağın tersine çevrimiçi çalışabilir, yani gerçek zamanlı öğrenebilir. Bu ağlar denetimsiz olarak öğrenebilmektedir. Değişik örnek yapılarına karşı kararlıdırlar. Öğrenme sürekli olduğu için ağdaki ağırlıklar sürekli güncellenir. Bu ağda girdi değerleri otomatik olarak normalize edilir. Bir deneme ile hızlı ve birçok deneme ile yavaş öğrenme gerçekleştirebilirler²⁴¹.

2.2.12. Yapay Sinir Ağları İle İlgili Diğer Kavramlar

2.2.12.1. Yığınlar

YSA’nda eğitim verileri girdi olarak modele girilirken aynı veri kümesi, ya tamamen ya da yığınlar halinde, modele dâhil edilebilir. Yığınlar rastgele, eşit büyüklükte ve birbirinden bağımsız yığınlara bölünerek ağda eğitilmektedir. Verilerin yığınlar halinde modele girilmesi yaklaşımı, verilerin tamamının tek seferde modele girilmesi yöntemine göre daha yaygındır²⁴².

2.2.12.2. Yineleme Sayısı

Araştırmacı kurduğu modeli tekrarlı döngüye tutarak minimum hatayı (minimum maliyet olarak da adlandırılmaktadır) verecek değere yakınsamaya çalışırken bu yinelemelerin (iterasyonların) sayısı yüksek belirlenirse, model daha iyi bir sonuca yaklaşabilir. Fakat bu durumun iki dezavantajı vardır. Birincisi zaman maliyeti, ikincisi ise model veri kümesini ezberleyebilir, yani aşırı uyum sağlayabilir. Yineleme sayısı çok düşük tutulursa, bu kez eksik uyum gerçekleşebilir. Bu yüzden araştırmacı, bu durumu dikkate alarak yineleme sayısını, gerekirse deneme yanılma yoluyla, dengeli bir şekilde belirlemelidir.

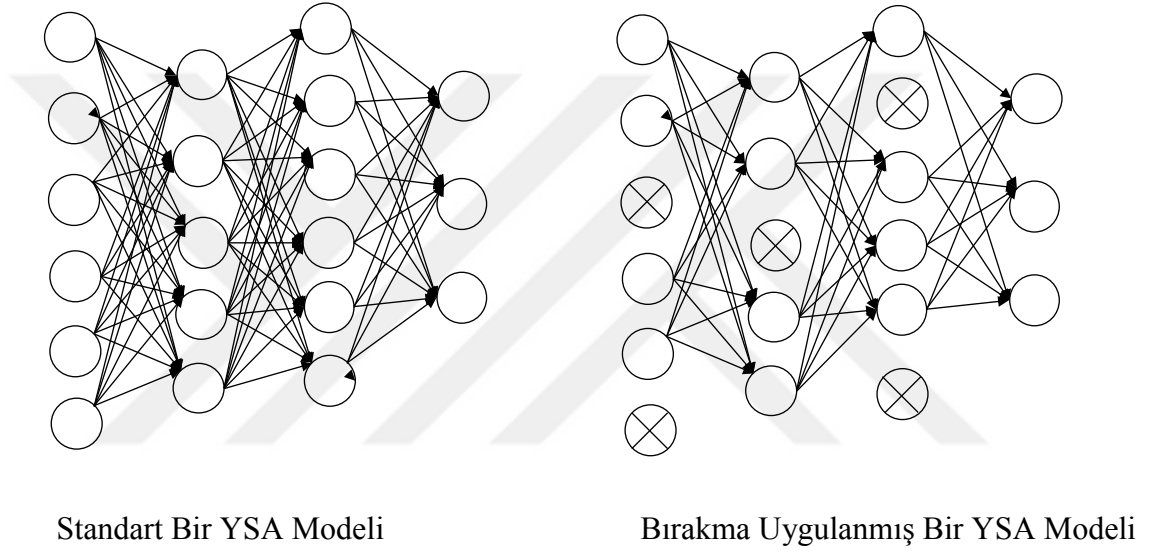
²⁴¹ Emre Yakut, **a.g.e.** s.89-92.

²⁴² Atınç Yılmaz, Umut Kaya, **a.g.e.** s.14.

2.2.12.3. Bırakma

Modelin eğitim verisine aşırı uyumuna engel olmak için, ağda her nöronun bir sonraki her nörona değil de bazı nöronlara bağlantı sağlanarak model kurma, “bırakma” olarak bilinmektedir. Şekil 2.32 bırakmanın öncesinde ve sonrasında ağın durumunu göstermektedir.

Şekil 0.32: Yapay Sinir Ağlarında Bırakma Uygulaması Öncesi Ve Sonrası



Kaynak: Atıncı Yılmaz, Umut Kaya, **Derin Öğrenme**, 1. Baskı, İstanbul: İnkılap Kitabevi, 2019, s.9.

2.2.12.4. Normalizasyon

Modele girilen eğitim kümesi çok farklı sayısal değerleri olan açıklayıcı veriler içeriyor olabilir. Hatta bu duruma birçok uygulamada rastlanır. Örneğin insanların gelir durumunu demografik bir değişken olarak modele ilave eden bir araştırmacı, birbirinden çok farklı sayısal veriler kullanabilir. Ülkelerin işsizlik, enflasyon, yabancı yatırım miktarı vs değişkenlerini araştıran bir kişi, sıfır ila bir arasında finansal oranlara bir veri kalemi olarak sahipken, aynı araştırmacı bu ülkelerin bu kez yüz milyonlara hatta milyarlarca ulaşan gayri safi hasıla vb. sayısal değerleri, aynı modelin içinde kullanmak istiyor olabilir. YSA, veri kümesi üzerinde oluşturduğu eğriler üzerindeki minimum noktayı arayan bir yöntem olması itibarıyla, bu kadar inişli ve çıkışlı veriler modeli yorar ve sonuç performansını olumsuz etkiler. Bu durumun

önüne geçmek için veri kümesi çeşitli normalizasyon yöntemleriyle normalize edilir. Böylelikle model birbirine yakın değerler üzerinden daha rahat ve doğru sonuçlar elde edebilir.

Birbirinden farklı, Z-Skoru, min-maks, median, sigmoid, ortalama ve standart sapma, istatistiksel sütun vb. şekillerde normalizasyon metotları bulunmaktadır. Bazı deneysel sonuçlar, sinir ağları performansının seçilen normalleştirme yöntemine de bağlı olduğunu göstermiştir. Sonuç olarak, normalizasyon, eğitilmiş ağın güvenilirliğini artırmaktadır²⁴³. Bu tez çalışmasının uygulama aşamasında, birçok çalışmada iyi sonuçlar veren, min-maks normalizasyonu kullanılmıştır.

2.2.12.5. Model Geçerleme

Eğitim verisi üzerinde çok sayıda model geliştirilebilir. Sonra bu modeller test verisi üzerinde test edilebilir. Bu durumda modeller test verisini de ezberleyebilir. Bu durum her yıl aynı sınav sorularını soran bir öğretmenin, bir öğrencisinin geçen yılın sorularına ulaşması neticesinde yapılan sınavın bir anlamı kalmaması durumuna benzetilebilir. Bu yüzden veri eğitim ve test verisi şeklinde iki gruba değil de, eğitim geçerleme ve test verisi şeklinde üç gruba ayrılabilir ve geçerleme verisi üzerinde modeller denendikten sonra gerçek performanslarına test verisi üzerinde karar verilebilir²⁴⁴.

2.2.13. Yapay Sinir Ağlarının Kullanım Alanları

Yapılan bir doğal dil tercümesi (machine translation), arka planda YSA çalıştıran çeviri yazılımları yardımıyla, hataları ayıklanabilir ya da kelimelerin en uygun olan karşılıkları kullanılarak ilgili metin daha akıcı hale getirilebilir.

Çeşitli kan değerleri yardımıyla, kişinin diyabet hastası olup olmadığını tahmin edilebilir. Diyabet hastası olduğu ve olmadığı bilinen kişilerin kan değerleri üzerinden

²⁴³ T.Jayalakshmi, Dr. A. Santhakumaran, “Statistical Normalization And Back Propagation For Classification”, **International Journal Of Computer Theory And Engineering**, Cilt: 3, Sayı 1, 2011, s.92.

²⁴⁴ Şenol Emir, **a.g.e.** s.41-42.

model eğitilerek bu tarz tahminler YSA ile rahatlıkla yapılabilir. Ayrıca nefes darlığı çeken bir hastanın astım hastası mı yoksa koah hastası mı olduğu yine mevcut veriler üzerinden YSA kullanılarak sınıflandırılabilir.

YSA genel anlamda:

- Probabilistik fonksiyon kestirimleri,
- Sınıflandırma (İlişkilendirme veya örüntü tanıma),
- Zaman serileri analizleri,
- Veri sıkıştırma,
- Doğrusal olmayan sistem modelleme,
- Optimizasyon,
- Doğrusal olmayan kontrol,

problemlerinde kullanılmaktadır. Özelde ise YSA birçok amaç için kullanılmaktadır. Bunlar²⁴⁵:

- Veri madenciliği kapsamında; müşteri ve pazar profilleri oluşturma, mal değeri ve satış tahminleri üretme, üretim ihtiyacını belirleme, pazar performansını ve çevresel ekonomik riskleri tahmin etme,
- Bankalarda, kredi müracaatlarını değerlendirme,
- Kredi kartı sahtekârlıklarını saptama, borçlanma ve risk değerlendirmeleri,
- Banka başarısızlıklarının tahmini, iflas tahmini, finansal sıkıntı tespiti,
- Tahvil dereceleme, hisse senedi fiyat kalıplarının tespiti ,
- Araçlar ve robotlar için optimum güzergah tespiti,
- Hava tahminleri, jüri panelleri,
- At yarışı, futbol vb. müsabaka sonuçlarının tahmin edilmesi,
- Güvenlik sistemlerinde ve soruşturmalarda, kart, göz, yüz, avuç içi, parmak, imza ve yazı tanıma, çek okuma, şekil tanıma, spektrum tanımlaması,
- Robot hareketlerinin kontrol edilmesi,
- Mekanik parçaların ömürlerinin ve kırılmalarının tahmin edilmesi,

²⁴⁵ Ercan Öztemel, **a.g.e.** s.36 ve Çetin Elmas, **a.g.e.** s.88 ve Talip Torun, **a.g.e.** s. 47-48.

- Meydana gelebilecek, arıza, patlama, yıkılma vb tahmin edilmesi,
- Kalite kontrol süreçleri, işlem kontrolü, sistem kontrolü, kimyasal yapılar, dinamik sistemler, işaret karşılaştırma, plastik kalıpcılık, kaynak kontrolü,
- İş çizelgelerini ve iş sıralamasını belirlenmesi,
- İletişim kanallarındaki geçersiz ekoların filtrelenmesi,
- İletişim kanallarındaki trafik yoğunluğunu kontrol edilmesi ve anahtarlanması,
- Radar ve sonar sinyallerinin sınıflandırılması,
- Üretim planlaması ve çizelgelemesi,
- Kan hücreleri reaksiyonlarının tahmin edilmesi, hücre tiplerinin tanımlanması ve kan analizlerinin sınıflandırılması,
- Kanseri ve diyabet gibi dâhil değişik tıbbi teşhislerin yapılması ve kalp krizinin tedavi edilmesi
- Beynin modellenmesi çalışmaları.

YSA'nın kullanım alanlarına dair, istatistiksel bakış açısı çerçevesinde bir fikir edinmek de mümkündür. Tablo 2.4'te hangi istatistiksel tekniklerin, hangi YSA modeliyle örtüştüğüne dair bir liste, Tablo 2.5'te ise hangi problemler için hangi YSA modelinin daha kullanışlı olduğuna dair bir sınıflandırma verilmiştir.

Tablo 0.4: Bazı İstatistiksel Testlere Karşılık Gelen YSA Modelleri

İstatistiksel Modeller	YSA Modelleri
Doğrusal Regresyon	Tek Katmanlı Doğrusal Algılayıcı
Çoklu Doğrusal Regresyon	Basit Doğrusal Algılayıcı
Lojistik Regresyon	Basit Doğrusal Olmayan Algılayıcı
Doğrusal Diskriminant Fonksiyonu	ADALINE
Basit Doğrusal Olmayan Regresyon Ve Çoklu Doğrusal Olmayan Regresyon	ÇKA

Kaynak: Vildan Gülpınar, “Yapay Sinir Ağları İşleyişinin Sosyal Ağ Analizi Yardımı İle Çözülmesi”, Doktora Tezi, Marmara Üniversitesi, 2013, s.16.

Tablo 2.5'te ise işlevlerine göre YSA modelleri tasnif edilmiştir.

Tablo 0.5: Yapay Sinir Ağı Modelleri Seçim Tablosu

Öğrenme Probleminin Türü	Problemin Çözümü İçin Kullanılan YSA Modeli	Modelin İşlevi
Tahmin	➤ Geriye Yayılım (Back Propagation, BP) ➤ Delta Bar Delta (Delta Bar Delta) ➤ Genişletilmiş Delta Bar Delta (Extended Delta Bar Delta) ➤ Yönlendirilmiş Rastgele Arama (Directed Random Search) ➤ Yüksek Dereceli YSA (Higher Order Neural Networks) ➤ Geriye Yayılımlı Kendisini Düzenleyen Model (Self Organizing Map into Back- Propagation)	Bazı çıktıları tahmin etmek için girdi değerlerini kullanır (örneğin, borsadaki en iyi hisse senetlerini seçme, hava durumunu tahmin etme, kanser riski olan kişileri belirleme)
Sınıflandırma	➤ LVN ➤ Sayaçlı Yayılım (Counter-Propagation) ➤ PNN	Sınıflandırmayı belirlemek için girdi değerlerini kullanır (örneğin, girdi A harfi mi, video verilerinin bloğu bir düzlem ve ne tür bir düzlem olduğunu bulma)
Veri Birleştirme	➤ Hopfield Ağı (Hopfield Network) ➤ Boltzman Makinesi (Boltzman Machine) ➤ Hamming Ağı (Hamming Network) ➤ Çift Yönlü Çağrışimli Bellek (Bidirectional Associative Memory, BAM) ➤ Uzaysal - Mekansal Geçici Örüntü Tanıma (Spatio-Temporal Pattern Recognition)	Sınıflandırma gibi, ancak aynı zamanda hata içeren verileri de tanır (örneğin, yalnızca taranan karakterleri değil, aynı zamanda tarayıcının düzgün çalışmadığı zamanları da tanımlama)
Veri Kavramsallaştırma	➤ ART ➤ SOM	Gruplama ilişkilerinin çıkarılabilmesi için girdileri analiz eder (örneğin, belirli bir ürünü satın alma olasılığı en yüksek olanların adlarını bir veri tabanından çıkarma)
Veri Filtreleme	➤ Resirkülasyon (Recirculation)	Bir giriş sinyalini yumuşatır (örneğin, bir telefon sinyalindeki gürültüyü çıkarma)

Kaynak: Dave Anderson and George McNeill, “Artificial Neural Networks Technology”, A DACS State-of-the-Art Report, 1992, s.31.

YSA’nın kullanım alanlarını, genel başlıklar halinde ifade edecek olursak, Tablo 2.5’ten de görülebileceği gibi, sınıflandırma (örüntü tanıma), hata azaltma

(optimizasyon), tahmin ve modelleme (fonksiyon kestirme) şeklinde de özetlenebilir²⁴⁶.

Sonuç olarak problemlerin büyük çoğunluğunun sınıflandırma problemi şeklinde olduğu görülmektedir. Günümüzde yapılan bazı düzenlemelerle, problemlerin büyük bir çoğunluğu, önce bir sınıflandırma problemine dönüştürülmekte, sonrasında ise çözüme gidilmektedir. YSA ise iyi bir sınıflandırma yöntemidir. Fakat her problemin doğası, istenen sonuçlar ve veri kümesi diğerlerinden farklı olacağı için problemin genel çerçevesi göz önünde bulundurulmalı ve mümkün olan en iyi modelle çözüme gidilmelidir²⁴⁷.

2.2.14. Yapay Sinir Ağlarının Avantajları

Makine öğrenmesi yöntemleri ile yapılan uygulamaların başarısı, literatürde bugüne kadar yapılan araştırmalardan da anlaşılacağı üzere, diğer deterministik yöntemlerde olduğu gibi, hiçbir zaman %100 olmaz. Olması durumu, verinin dengesizliğinden (imbalance) yani kurulan modelin yakınsamayışından (not convergency) kaynaklanıyor olabilir. Bu durumda da uygulama sonuçlarına güvenilmez. Bu salt başarının bu tür modellerde yakalanamayışının temel sebebi, bilinmek istenen herhangi bir değişkeni etkileyen tüm açıklayıcı değişkenlerin tespit edilerek modele dâhil edilmesinin gerçek hayatta maalesef mümkün olmamasıdır. Çünkü gerçek hayatta, her şey birbiriyle etkileşim halindedir ve ilgili açıklayıcı değişkenleri, sayısal formda tespit ederek modele ilave edebilmek kısmen mümkün olabilmektedir. Zaten salt başarı durumunda, mevcut problem deterministik bir forma dönüşür ve simpleks gibi geleneksel sayısal yöntemlerle de, tam doğru olarak, çözülebilir. Fakat bu tarz problemlere, genellikle pozitif bilimlerde rastlanmaktadır. Ancak sosyal bilimlerde bağımlı değişkenin tahmini, gerçek sonuçlardan, açıklanamayan veya “belirlenemeyen

²⁴⁶ Tunçhan Cura, “Karar Verme Aracı Olarak Yapay Sinir Ağları Ve Yapay Sinir Ağları İle Portföy Optimizasyonu”, Doktora Tezi, İstanbul Üniversitesi, 2004, s.9. Ercan Öztemel, **a.g.e.** s.36 ve Çetin Elmas, **a.g.e.** s.88.

²⁴⁷ Ercan Öztemel, **a.g.e.** s.36.

tesadüfilik fenomeni” (random phenomena) çerçevesinde muhakkak bir miktar sapma gösterecektir²⁴⁸.

YSA vb. doğrusal olmayan yaklaşımların doğrusal yaklaşımlardan genel olarak daha üstün olmasının birkaç nedeni bulunmaktadır. İlk olarak finansal oranlar ile temerrüt tahmini arasındaki ilişkilerde doygunluk etkileri olduğu söylenebilir. Örneğin, kazançlar/toplam varlıklar 0,1'de iken 0,2 miktarında değişiyorsa, bu oran 1,0'dan 1,2'ye değiştiğinden çok daha büyük bir etkiye sahip olacaktır. Sinir ağları, doğrusal olmayan yapısıyla genel olarak mevcuttaki doğrusal yöntemlerden daha iyi performans göstermektedir²⁴⁹.

YSA'nın başarı performansı, yine literatüre bakıldığında, geleneksel istatistikî yöntemlerle yapılan testlerden daha düşük değildir ve çoğu testte de daha yüksek olduğu görülmektedir. Çünkü doğrusal olmadığı gibi, parametrik bir yöntem de değildir. Yani herhangi bir ön varsayıma dayanmaz. Bu yönüyle de çok pratiktir²⁵⁰.

YSA, matematiksel olarak modellenmesi mümkün olmayan, gerçek hayat problemlerini tüm zorluk ve karmaşıklıklarına rağmen başarılı bir şekilde çözebilmektedir. Bu zorluk ve karmaşıklık aslında gerçek hayattaki olayların arkasında yatan ilişkilerin gücünün tamamen saptanamamasına dayanmaktadır. Geleneksel modellerde bu ilişkilerin tamamının tespit edilip modele ilave edilmesi ya da yok sayılması gerekmektedir. Ama YSA tüm bu ilişkileri, arasındaki ilişkinin gücünü bize kantitatif olarak vermese de, örneklerden öğrenerek, tüm örüntüyü dikkate alır ve oldukça yüksek doğruluk oranlarıyla, iş süreçlerinin verimliliğini ciddi ölçüde artırır. YSA'nın paralel çalışabilmeleri ise gerçek zamanlı kullanımına imkân tanır²⁵¹.

ART ve SOM modelleri, eğitim verileriyle eğitildikten sonra, her yeni vakada, oluşturduğu tüm ağırlıkları günceller. Artık yeni edindiği gözlemler de model için birer

²⁴⁸ Ahmet Mete Çilingirtürk, **İstatiksel Karar Almada Veri Analizi**, 1. Baskı, Ankara: Seçkin Yayınevi, 2011, s.150.

²⁴⁹ Amir F. Atiya, **a.g.e.** s.930.

²⁵⁰ Tian-Shyug Lee, I-Fei Chen, **a.g.e.** s.747.

²⁵¹ Ercan Öztemel, **a.g.e.** s.208.

eğitim verisine dönüşür. Böylelikle dinamik yani sürekli öğrenen bir programalamaya sahiptir. Dolayısıyla sahip olduğu eğitim düzeyi sürekli artarken, hiç karşılaşılmamış olaylarla karşılaştığında tecrübeli insanların yaptığı gibi en uygun kararı verebilen bir başarı performansına erişir. Bu sürekli öğrenen özelliği de YSA'nı, geleneksel yöntemlere üstün kılmaktadır.

YSA, istatistiki modeller gibi varsayımlara bağlı değildir ve aşırı uç değerlerden etkilenmez. Ayrıca çok çeşitli istatistiki testlerden hangisinin en uygun test olduğuna karar vermek gibi bir zorluk taşımaz.

Uzman karar süreçlerini modellemek için bir sinir ağı mimarisi geliştirmenin önemli avantajları, ağıın bilgisinin doğrudan çıkarılmaya çalışılmasının maliyetli, zaman alıcı ve hataya meyilli olmasıdır. Hatta bu nedenle uzmanlar, olumsuz neticelerin sorumlusu olmamak için bu karar verme yükünden kaçınmaktadırlar. Dolayısıyla karar örneklerinden öğrenen ve çözüm üreten bir ağ hiçbir karar yükünden geri durmayacaktır. Sinir ağlarının ikincil avantajları ise ağıın, belirsizlik koşulları altında karar vermede tipik olarak bulunan gürültülü, eksik ve çarpık verileri işleme kabiliyetidir²⁵².

Eksik veri probleminden de etkilenmeyen YSA, temel olarak doğrusal olmama, yüksek paralellik, hata ve gürültü toleransı, öğrenebilme, genelleme, esneyebilme ve uyarlanabilme yetenekleriyle ilgili dikkat çekici bilgi işlem özellikleri ön plana çıkmaktadır²⁵³.

Genelleme, “eğitim ya da öğrenme sürecinde karşılaşılmayan girişler için de ağıın uygun tepkileri üretmesi” şeklinde tanımlanmaktadır²⁵⁴. YSA gizli katmanları sayesinde ve birden çok nöron (proses elemanı) kullanması ve aynı zamanda iteratif bir yöntem olması sayesinde, yüksek genelleme yeteneğine sahip olan ve veri yapısındaki dejenerasyondan geç etkilenen, oldukça verimli bir yöntemdir. Her ağıın, kendi mimarisine özgü olarak, diğer yöntemlere kıyasla, iyi olduğu veya iyi olmadığı

²⁵² Louis W. Glorfeld, **a.g.e.** s.37.

²⁵³ I.A. Basheer, M. Hajmeer, **a.g.e.** s.3–31.

²⁵⁴ Emre Yakut, **a.g.e.** s.107.

alanlar mutlaka vardır. Bu alanları bilerek uygulama yapmak sonuçların başarısı açısından elbette önemlidir. Bazı araştırmacıların yanlış ağ seçiminden kaynaklanan ve bu yüzden uzun süreler boyunca ağını iyi sonuçlar verecek şekilde eğitemeyen bu yüzden de YSA'nın yetersiz olduğunu ya da uygun bir yöntem olmadığını ileri sürenler olmuştur. Fakat doğru bir YSA mimarisinin, düzgün bir örnek seti ve iyi belirlenmiş bir öğrenme algoritması eşliğinde çözemeyeceği problem sayısı gerçekte çok azdır²⁵⁵.

2.2.15. Yapay Sinir Ağlarının Dezavantajları

Ne yazık ki, nöral ağ tabanlı modellerin önemli bir kısıtı, çıkarım sürecinin opaklığı, yani şeffaf olmayan hesaplama sürecidir. Geleneksel uzman sistemler karar destek araçlarının aksine, karar vericiler genellikle sinir ağı kararlarının temelini anlayamazlar. Dolayısıyla, örnek vermek gerekirse, YSA'nın bir kredi müşterisinin kredi talebini neden kabul ettiği ya da neden reddettiği açık değildir. Bu sorun genellikle bu sistemleri karar destek uygulamaları için, istenmeyen bir hale getirmektedir²⁵⁶. Bu durum hesap verilebilirlik ve yönetimde şeffaflık gibi unsurlar açısından olumsuz sonuçlar doğurmaktadır. Örneğin bir müşteri temerrüde düştüğünde, bu müşteriye neden kredi verildiği sorusunun somut bir cevabı olmayacaktır. Çünkü sinir ağı kullanan model, gerçek hayat verilerinden elde ettiği bağlamları değerlendirerek, başvuruyu iyi kredi olarak sınıflandırmıştır.

Kara kutular konusunu biraz irdeleyelim. Daha iyi anlaşılması için başka bir model olan regresyonla karşılaştıralım. Bir regresyon analizi sonucunda oluşan regresyon formülüne (Denklem 2.2) bakıldığında modelin sabitini, hatasını, her bir açıklayıcı değişkenin bağımlı değişkeni ne ölçüde etkilediğini, bir açıklayıcı değişkende meydana gelecek bir birimlik bir değişimin bağımlı değişkende ne kadarlık bir etki yaratacağını görebiliriz. Ayrıca, determinasyon katsayısı yardımıyla, açıklayıcı değişkenlerin bağımlı değişkenin ne kadarını açıkladığını da görebiliriz. YSA'nda,

²⁵⁵ Ercan Öztemel, **a.g.e.** s.207.

²⁵⁶ Louis W. Glorfeld, **a.g.e.** s.37.

modelin kendisi bunu bilir, ama aşırı karmaşık olduğu için bir regresyon formülü gibi bakıp fikir edinmek mümkün değildir²⁵⁷.

Geleneksel yöntemler, belirli sayıda işlem içerirler. Dolayısıyla bilgisayarlar yardımıyla çok kısa bir sürede, hemen hemen tüm testler gerçekleştirilebilmektedir. Fakat genelde yapay zekâ modelleri, özelde ise YSA, arka planda o kadar çok işlem den geçer ki, bilgisayar işlemcilerini çok yüksek düzeyde kullanan bu modeller, işlemci kapasitesi zayıf olan bilgisayarlarda yapılamamaktadır. Problemin büyüklüğüne göre, bazı YSA'nın çalıştırılıp sonuç alınması günlerce sürebilmekte hatta sonuç alınamayıp sezgisel yöntemlere başvurmak zorunda kalınabilmektedir.

Geleneksel istatistiki yöntemlerin 1930'lu yıllarda, giderek karmaşıklaşan çok fazla işlemin artık elle kâğıt üzerinde yapılamaması nedeni ile tıkanıp bilinmektedir. 1960'lı yıllardan itibaren, bilgisayarların yayılması ile istatistiğin yeniden gelişim göstermeye başlaması beklenen bir durumdu. Yapay zekâ da çok yeni değildir. Fakat Alan Turing'den bu yana az çok bilinen ve sürekli gelişen yapay zekâ uygulamaları, bilgisayar işlemcilerinin yeterince gelişmemiş olmasından dolayı ciddi anlamda gecikme yaşamıştır. Sonuç olarak, bilgisayarların gelişmişlik düzeyi yapay zekâ çalışmalarını kısıtlamaktadır. Bunun önemli bir sebebi yapay zekâ yöntemlerinin, geleneksel yöntemlere kıyasla, çok daha fazla sayıda işlem gerektirmeleri dolayısıyla işlemci kapasitesini çok fazla kullanmalarıdır.

YSA'nın bir dezavantajı da en uygun çözümü verecek olan gizli katman sayısının ve bu katmanlarda yer alacak düğüm sayısının önceden bilinmemesidir. Hatta hangi nöronda hangi aktivasyon fonksiyonunun kullanılacağı, öğrenme oranı, momentum katsayısı, durdurma kriteri, verinin alfanümerik hale getirilmesinde kullanılacak kategoriler (örneğin çıkış sınıflarını 0,1,2,3 şeklinde kodlamakla, 1,2,3,4 şeklinde kodlamak) sonuçların performansını etkileyebilir²⁵⁸. Dolayısıyla araştırmacı uygun mimariye karar verebilmek için çok sayıda denemeler yapmak zorunda kalmaktadır. Bu durum, yani YSA mimarisi oluşturulurken genel bir kurallar manzumesinin

²⁵⁷ Eliana Angelini, Giacomo Di Tollo Ve Andrea Roli, **a.g.e.** S.739.

²⁵⁸ Ercan Öztemel, **a.g.e.** s.90-105.

bulunmayışı, YSA'nın en çok eleştirildiği ve uygulayıcılarını en çok yoran konuların başında gelmektedir. İlgili parametrelerden bazılarının optimaliteden uzaklaşması halinde ağıın ne yönde etkileneceği Tablo2.6'da gösterilmektedir.

Tablo 0.6: Tasarım Parametrelerinin Uç Değerlerinin Eğitim Yakınsaması Ve Ağ Genellemesi Üzerindeki Etkisi

Tasarım parametresi	Çok yüksek veya çok büyük	Çok düşük veya çok küçük
Gizli katman sayısı	Aşırı uyum (genelleme hatası yüksek)	Yetersiz uyum (eğitim hatası yüksek: ysa, verilere gömülü temel kuralları elde edemiyor)
Öğrenme oranı (η)	Optimal çözüm hakkında kararsız salınımlar yapan YSA (ağırlıklar)	Yavaş eğitim, yerel minimuma yakalanma
Momentum katsayısı (μ)	Yerel minimum riskini azaltır, eğitimi hızlandırır ve çözümü atlama riskini artırır	Yerel minimumda tuzaklanma riskinin potansiyel olarak artmasına yol açar ve eğitimi yavaşlatır
İterasyon sayısı	İyi hatırlama YSA (yani, aşırı uyum, YSA'nın verileri ezberlemesi) ve eğitimsiz verilere kötü genelleme	Verileri temsil edemeyen bir YSA modeli üretir
Eğitim alt kümesinin boyutu	İyi hatırlama ve genelleme yapan YSA	YSA sorunu tam olarak açıklayamıyor ve sınırlı veya kötü genelleme yapıyor
Test alt kümesinin boyutu	YSA'nın genelleme başarımını test etme yeteneği artırıyor	YSA'nın genelleme başarımını test etme yeteneği azalıyor

Kaynak: I.A. Basheer ve M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, Journal of Microbiological Methods, Cilt: 43, 2000, s. 24.

Olasılık dağılımının bilinmediği durumlarda açıklayıcı tekniklerin kullanılmasına ihtiyaç duyulur. Güncel uygulamaların çoğundaki yegâne bilgi, sonlu veri kümesidir ki, o bilgi de çok boyutlu ve oldukça sınırlı sayıdadır. Bu yüzden öğrenme makinesinin bu koşullarda dahi işlem yapabilmesi istenir. Maalesef, sinir ağları ve bulanık sistemler gibi geleneksel metotlar bu yeteneğe sahip değildir ve bu tip durumlarda güvenilir

olmayan sonuçlar üretirler. Ayrıca, yüksek boyutlu uzay ise “boyut sıkıntısı (curse of dimensionality)” problemine sebep olur²⁵⁹.

Yerel minimum problemini aşmak için, her birinin farklı stratejiler sunduğu, çeşitli sezgisel yöntemler geliştirilmiştir. En yaygın kullanılanları benzetilmiş tavlama ve GA olmak üzere; parçacık sürü optimizasyonu, tabu arama, karınca kolonileri, arı kolonileri vb. birçok alternatif yöntem bulunmakta olup, bu yöntemlerle de bu ağırlıklar oluşturulabilmekte ve böylelikle evrensel minimuma ulaşma ihtimali daha da artmaktadır²⁶⁰.

YSA, ilişkili bellek özelliği ve genelleme yeteneği ile kredi riski belirlemeye dönük modellerde önemli bir alternatif haline gelmiştir. Bununla birlikte, potansiyel girdi değişkenlerinin önemi, ağırlık topolojisinin kararı ve eğitim süresinin oldukça uzun olması, sıklıkla eleştirilmektedir ve bu nedenle kredi riskini belirleme problemini ele almadaki uygulamasını da maalesef sınırlandırmaktadır²⁶¹.

Algoritmaların çoğu, probleme özgü algoritmalar olup, üzerinde geliştirildiği problemlerin dışındaki problemlere uygulandığında iyi sonuçlar vermemektedir. YSA’nda da her problemin kendine özgü bir mimari olgunluk yakalamak ise zaman almaktadır. Bu zorluğu giderebilmek için, tüm problemlere uygulanabilecek türden algoritma çalışmaları son yıllarda sürekli artmaktadır²⁶².

2.3. Destek Vektör Makineleri

İlk olarak 1960 yılında, Vladimir Vapnik ve Alexey Chervonenkis tarafından istatistiksel öğrenme teorisiyle temeli atılan yöntem bir araştırma problemine, ilk kez yine Rus Matematikçi Vladimir Vapnik tarafından, 1995 yılında uygulanmıştır. Bu tarihten sonra popülerlik kazanan DVM’nin teorik yetkinliğe erişmesi ve YSA gibi yöntemlere alternatif olarak kabul edilmesi süreci, yaklaşık olarak otuz beş (35) yılı bulmuştur.

²⁵⁹ Emre Çomak, **a.g.e.** s.19.

²⁶⁰ Tunçhan Cura, **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008, s.28.

²⁶¹ Tian-Shyug Lee, Chih-Chou Chiu, Chi-Jie Lu, I-Fei Chen, “Credit Scoring Using The Hybrid Neural Discriminant Technique”, **Expert Systems With Applications**, Cilt: 23, 2002, S.245–254.

²⁶² Derviş Karaboğa, **a.g.e.** s.18.

Diğer bir ifadeyle DVM'nin gelişim süreci, 1960'lı yıllara ve büyük çoğunluğu, Vladimir Vapnik'in çalışmalarına dayanmaktadır.

Vapnik – Chervonenkis (VC) teorisi olarak da bilinen, İstatiksel Öğrenme Teorisi (Statistical Learning Theory, SLT), 1960'ların sonlarında tanıtılmıştır. 1990'lara kadar tamamen veri toplama ve verilerden fonksiyon tahmin etme probleminin teorik bir analizi olarak kalmıştır. Fakat teori, ilk kez 1995 yılında, metodolojinin kuramsal ve matematiksel mimarı Vladimir Vapnik tarafından, yeni geliştirilen bir tür öğrenme algoritmasını (DVM) kullanarak, bir araştırma problemine uygulanmış ve başarılı bir sonuç elde edilmiştir. Bu tarihten sonra literatürde ve uygulamada VC teorisini kullanarak çok iyi sonuçlar alan bir algoritma olan DVM yöntemi, sınıflandırma ve regresyon tabanlı problemlerde, çok yaygın bir şekilde kullanılmaya başlanmıştır²⁶³. 1960'lardan 1995'lere kadar teorinin yaşamış olduğu bu sessizlik, tıpkı 1969'dan itibaren YSA'nın temeli perseptronun, doğrusal bir model olmasından dolayı, XOR problemini çözemediğinin ipatlnmasının ardından, çözüldüğü 1985 yılına kadarki yaşadığı ilgisizlik dönemine benzetilebilir.

DVM, “genelleme teorisinin kavrayışını benimseyen ve optimizasyon teorisini kullanan, doğrusal öğrenme makinelerini, çekirdek kaynaklı özellik uzaylarında, verimli bir şekilde eğiten bir sistem olarak tanımlanmaktadır”²⁶⁴.

DVM, ile ilgili en kapsayıcı tanım, “istatistiksel öğrenme teorisine dayalı bir evrensel yapısal öğrenme süreci”dir. DVM'nin evrensel olmasını sağlayan yönü, yöntemin sinirsel ağlar (sigmoid aktivasyon), RBF, splinelar, polinom tahmin ediciler vb. gibi çeşitli gösterimleri öğrenmek için kullanılabilen bir yöntem olmasıdır. Bununla birlikte DVM, tümevarımsal olmayan öğrenme modellerine de uygulanabilir. Örneğin, destek vektör parametreleştirme, lineer denklemler, bilgisayarlı tomografi, sinyal sıkıştırma, görüntü sıkıştırma ve benzeri büyük problemlerin çözümünde kullanılabilir. Aizerman, Braverman ve Rozoner, 1964 yılında, çekirdek kullanma

²⁶³ Vladimir Vapnik, “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.988.

²⁶⁴ C. Stoean and R. Stoean, **Support Vector Machines and Evolutionary Algorithms for Classification**, İsviçre: Springer International Publishing, 2014, s.7.

fikrini ortaya atarken, Mangasarian 1965'te matematiksel programlamada, sınıflandırma için DVM benzeri bir doğrusal optimizasyon modeli ileri sürmüştür. Ancak bu öneriler, istatistiksel öğrenme kuramının sağlamış olduğu teorilerden yoksun olduğu için pratik öğrenme algoritmaları arasında yer edinememiştir. DVM'nin kendisi ise, doksanlı yıllardan itibaren, birçok sınıflandırma problemine başarıyla uygulanabilmiş ve uygulanmaya da devam etmektedir²⁶⁵.

Doksanlı yılların ortasında, istatistiksel öğrenme teorisine dayalı, DVM olarak adlandırılan, yeni öğrenme algoritması türleri önerilmiştir. Bu durum, istatistiksel öğrenme teorisini sadece teorik analiz için bir yardımcı değil, aynı zamanda çok boyutlu fonksiyonların tahmini için kullanışlı algoritmalar oluşturabilen bir araç haline getirmiştir²⁶⁶.

DVM yönteminin temeli, istatistiksel öğrenme teorisini, diğer adıyla, VC katsayısını birlikte bulan ve bu katsayıyı kullanarak, ampirik risk minimizasyonuna farklı bir kriter ekleyerek yeni bir risk minimizasyonu yaklaşımı (Yapısal Risk Minimizasyonu, YRM) geliştiren Vapnik ve Chervonenkis'in soyisimlerinin başharflerini içeren "VC teorisini" adı da verilen bu matematiksel yaklaşıma dayanmaktadır. İstatistiksel öğrenme teorisinin dayandığı temel fikir olan Yapısal Risk Minimizasyonu (Structural Risk Minimization, SLT) ise, eğitim hatasını minimize eden ampirik risk minimizasyonunun, genelleme hatasını da eğitim hatası ile birlikte minimize edebilmesi için bir nevi modifikasyona tabi tutulmuş şeklidir²⁶⁷.

2.3.1. İstatistiksel Öğrenme Teorisi

Vapnik – Chervonenkis (VC) teorisini, "sonlu sayıdaki örneklerle", istatistiksel tahmin yapmak için mevcuttaki en iyi teori olarak bilinmektedir. İstatistiksel öğrenme teorisini, şimdiye kadar ki hemen her istatistiksel tahmin kavramını titizlikle tanımlamakta ve öğrenme problemlerinin kapsamlı bir çerçevesini sunmaktadır. YSA, cezalandırma ve

²⁶⁵ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.408.

²⁶⁶ Vladimir Vapnik, "An Overview of Statistical Learning Theory", **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.988.

²⁶⁷ J. Shen, "Fusing Support Vector Machines and Soft Computing for Pattern Recognition and Regression", Doktora Tezi, Kansas State University, 2005.

Bayesçi çıkarım gibi diğer metotların aksine, istatistiksel öğrenme teorisi sonlu örneklerle tahmin amaçlı öğrenme için matematiksel kanıtlar sunmaktadır. İstatistiksel öğrenmenin dayandığı temel felsefe, sistemin arkaplanında yer alan tüm matematiksel bağlantıları, analizini yapmak suretiyle, ispatlamaktır. Yapay zekâ yöntemleri ise bundan farklı olarak, mevcut sistemi, taklit etme anlayışını benimser ve bu yönüyle klasik istatistik yaklaşımından tamamen farklılaşır. Yani yapay zekâ yöntemlerinin istatistiki yaklaşımdan farkı “analitik değil, pratik olması”dır. Örnek vermek gerekirse Doğal dil tercümesi gerçekleştiren yapay zekâ yazılımları, tercüme metinlerini karşılıklı olarak değerlendirip, metnin hangi kelimelerle diğer dilde ifade edileceğini belirlemektedir. Yani bu yazılım tek bir gramer kuralını dahi bilmemektedir. Burada da yine ilham alınan, insandır. Yürümeyi öğrenen bir bebek mekaniğin kurallarından habersiz olduğu gibi, araba kullanan bir sürücü de gerekli fizik kurallarından habersizdir ve yine de bu eylemlerinde son derece başarılıdırlar. Modelleme söz konusu olduğunda ise artık ilkelere dayanan klasik modellemekten, verilerden model geliştirmeye doğru evrilen bir paradigma değişimi görülmektedir²⁶⁸.

Şimdiye kadar, veriden öğrenmenin gerçekleştirilebilmesi adına, model geliştirebilmek için birçok farklı yöntem kullanılmıştır. Söz konusu veriden öğrenme modelleri, evet oldukça başarılı sonuçlar verdiği görülmüştür. Fakat eğitim veri kümesi üzerinde geliştirilen bu modeller, eğitim veri kümesinin dışında, geliştirilen modelin tamamen yabancı olduğu yeni bir veri kümesi üzerinde uygulandığında, modellerin başarısının belirgin ölçüde düştüğü gözlemlenmiştir. Bunun sebebi ise ilgili modellerin sadece eğitim hatasını minimize etmeye odaklanmaları ve genelleme hatasını göz önünde bulundurmamalarıdır. Bu da, veriyi ezberleme anlamına gelen, aşırı öğrenme (over fitting) problemini gündeme getirmekte ve modellerin, yeni veri setleri üzerindeki, doğruluk oranlarına olumsuz bir şekilde yansımaktadır (Bkz. Kısım, 2.1.8.). Bu da modellerin karmaşıklık düzeylerinin gerekenden çok daha fazla olmasından kaynaklanmaktadır. Modellerin karmaşıklık seviyesinin (model complexity) optimize edilmesi (control of model complexity) böylelikle genelleme

²⁶⁸ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.100.

hatasını da indirgemeleri, doğru sonuçlar için çok önemli olmakla birlikte, modellerin birbirine kıyasla başarı sıralamasını da belirlemektedir.

İstatiksel öğrenme teorisi (kısaca VC teorisi), temel yaklaşımı dikkate alındığında, “amacı bilinmeyen sistemin doğru tanımlanmasından ziyade pratik amacı iyi bir genelleme olan tüm uygulamaların tamamı için uygun bir yöntemdir”. Zaten sistemin tamamen analiz edilmesi, boyut laneti (curse of dimensionality) göz önünde bulundurulduğunda, gerçekçi bir hedef olma imkânını da yitirmektedir²⁶⁹. Dolayısıyla sistemi tamamen analiz etme düşüncesiyle eğitim hatasını minimize etme yaklaşımında, sistemin tamamen analizi mümkün olmayacağından ve aşırı öğrenme problemine (aşırı karmaşık bir model oluşmasına) sebep olacağından dolayı, modelin yeni veriler üzerindeki başarısı olumsuz etkilenmektedir.

Model karmaşıklığının optimizasyonu konusunda, birçok model Ampirik Risk Minimizasyonu (ARM)(Empirical Risk Minimization, ERM) yaklaşımını temel almaktadır. Çünkü, istatiksel olarak, olasılık dağılımı bilinmeyen bir veride, eğitim hatasını minimize edecek, en genel ve en bilinen risk minimizasyonu yaklaşımı ARM’dur. Esasında VC teorisi de ARM’nun tutarlılık sınırlarını belirleme amacıyla çalışmalar yapılırken geliştirilmiş, teorik bir yaklaşımdır²⁷⁰.

2.3.1.1. Ampirik Risk Minimizasyonu

Regresyon analizinde risk (tahmin edilen sonuçla, beklenen sonucun farklı çıkması durumu) minimizasyonu için kullanılan En Küçük Kareler (EKK)(Least Squares, LS) yöntemi ve yoğunluk tahmini problemlerinde kullanılan Maksimum Olabilirlik (MO) (Maximum Likelihood, ML) yöntemleri, aslında bu problemlerin kayıp fonksiyonları için problemin türüne göre özelleştirilmiş, birer ARM uygulamalarıdır. Dolayısıyla ARM, risk minimizasyonu prensibini temel alan birçok yöntem için, oldukça genel bir uygulamadır. Sonuç olarak ARM için geçerli olan teorilerin tamamı, bu risk

²⁶⁹ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.100-101.

²⁷⁰ Vladimir Vapnik, “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.990.

minimizasyonu ilkesini kullanan tüm özelleştirilmiş yöntemler için de geçerli olmaktadır²⁷¹.

Şimdi bir problem üzerinde ARM'nin kullanımını ele alalım. Bağımsız x ve y girdileri ve bu doğrultuda oluşan bağımlı z çıktıları bulunsun. Burada $z = f(x, y)$ çıktılarının oluşturacağı olasılık yoğunluk fonksiyonu ise $p(x)$ olsun. Burada $p(x)$ fonksiyonu, evet açıklayıcı ve özdeş fakat dağılımı bilinmeyen (independent identically distributed, i.i.d.) bir fonksiyondur. Yani $p(x)$, olasılık yoğunluk fonksiyonu bilinmeyen, $Z_n = \{z_1, z_2, z_3, \dots, z_n\}$ açıklayıcı ve özdeş dağılmış n adet örneklem verilerinden oluşmaktadır. Bu problemin kayıp fonksiyonu ise $Q(z, \omega)$, $\omega \in \Omega$ olsun. Burada öğrenme probleminin amacı, aşağıdaki risk fonksiyonunu (2.48) minimize edecek bir $Q(z, \omega_0)$ fonksiyonunun bulunmasıdır²⁷².

$$R(\omega) = \int Q(z, \omega)F(z) \text{ veya } R(\omega) = \int Q(z, \omega)P(z)d(z) \quad (0.48)$$

Burada $Q(z, \omega) = L(y, f(x, \omega))$, sınıflandırma ve regresyon gibi spesifik öğrenme problemlerinin, kayıp fonksiyonlarının bir kümesini gösterir. Mesela bu regresyonda;

$$Q(z, \omega) = (y - f(x, \omega))^2 \quad (0.49)$$

veya,

$$Q(z, \omega) = |y - f(x, \omega)| \quad (0.50)$$

şeklindedir. Sonrasında, fonksiyon uzayından, $Q(z, \omega)$ fonksiyonlarının içerisinde yukarıda formülleri verilen kayıp fonksiyonlarından değeri minimum olan bir $Q(z, \omega_0)$ fonksiyon çekilerek, mümkün olan en iyi model, bir başka deyişle, minimum kayıpla hesaplama yapan fonksiyon elde edilmiş olur. Öğrenme problemlerinde, beklenen risk fonksiyonu, yani her uygulama probleminde gerçekleşecek kayıp

²⁷¹ Vladimir Vapnik, "An Overview of Statistical Learning Theory", **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.989.

²⁷² Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.102.

önceden kesin olarak bilinemeyeceği için, bunun yerine bilinen ARM (2.51) ilkesi kullanılarak, bir kayıp fonksiyonu hesaplanmaya çalışılır²⁷³.

$$R_{amp}(\omega) = \sum_{k=1}^n Q(z_i, \omega) \quad (0.51)$$

Burada öğrenme makinesini zorlayan tek konu, eğitim verisinin $F(z)$ Kümülatif Yoğunluk Fonksiyonunun (Cumulative Density Function, CDF) bilinmemesidir. $F(z)$ 'nin dağılımı ile alakalı tek geçerli bilgi, eğitim verisinin kendisi, $Z(n)$ 'dir. Bu darboğazı klasik istatistik yaklaşımı, bilinen veri $Z(n)$ üzerinden mümkün olan en iyi $Q(z, \omega_0)$ tahminleyicisini bularak aşmaktadır. Bir diğer yaklaşım olan ARM ise, gerçek kayıp fonksiyonu bilinemeyeceği için, eldeki ampirik kayıp fonksiyonunu minimize edecek olan tahminleyici fonksiyonu bulmak için kullanır. ARM, öngörülü öğrenmede yaygın olarak kullanılmaktadır. Sonlu örneklerde ARM yaklaşımının, yoğunluk tahmininde de tercih edilebilir olduğu da ileri sürülmüştür²⁷⁴.

Herhangi bir tümevarımsal prensip için gerekli olan, genel bir özellik, asimtotik tutarlılıktır; bu, ARM tarafından sağlanan tahminlerin, eğitim örneklerinin sayısı arttıkça, gerçek değerlere veya mümkün olan en iyi değerlere uyması gerekir. Yani örneklem miktarı arttıkça, tahmin kalitesinin artmasını gerektiren tutarlılık prensibi, burada da geçerlidir. Bu çerçevede düşünüldüğünde, denklem 2.52 ve 2.53'te gösterildiği gibi, hem klasik istatistikteki kayıp fonksiyonu hem de ARM yaklaşımı gözlem sayısı arttıkça, asimptotik tutarlılık çerçevesinde, gerçek değere yakınsayacaktır. Gerçek riski minimize edecek tahmin fonksiyonuna $Q(z, \omega_0)$ demiştik. Ampirik riski minimize edecek tahmin fonksiyonuna ise $Q(z, \omega^*)$ diyelim. O halde bu fonksiyonlara ait risk fonksiyonlarını sırasıyla $R(\omega_0)$ ve $R(\omega_n^*)$ şeklinde gösterilirse, bu durumda, gözlem sayısı arttıkça ampirik beklenen riske yaklaşacaktır. Bu durum Denklem 2.52 ve 2.53'te görülmektedir²⁷⁵.

²⁷³ Vladimir Vapnik, "An Overview of Statistical Learning Theory", **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.989.

²⁷⁴ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.102.

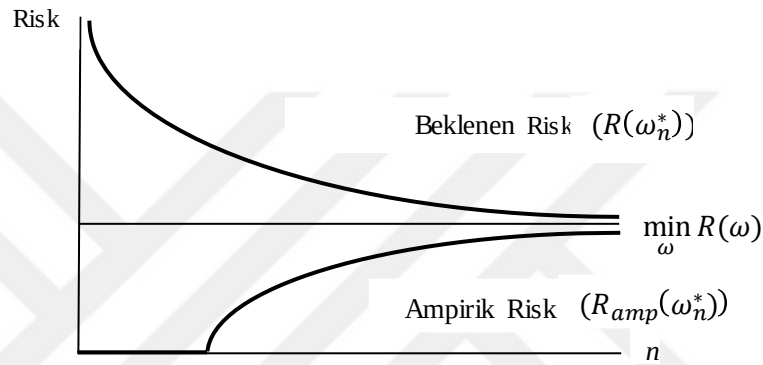
²⁷⁵ Vladimir Cherkassky, Filip Mulier, **a.g.e.** s.103.

$$R(\omega_n^*) \rightarrow R(\omega_0) \Rightarrow n \rightarrow \infty \quad (0.52)$$

$$R_{amp}(\omega_n^*) \rightarrow R(\omega_0) \Rightarrow n \rightarrow \infty \quad (0.53)$$

Bu asimptotik yaklaşım Şekil 2.33’te, grafiksel olarak görülmektedir.

Şekil 0.33: Ampirik Riskin, Eğitim Verisi Arttıkça, Beklenen Riske Asimptotik Olarak Yakınsaması



Kaynak: Vladimir Cherkassky, Filip Mulier, **Lerning From Data: Concepts, Theory and Methods**, 2. Baskı, New Jersey, John Wiley & Sons, Inc., 2007, s.103.

VC teorisinin, ARM’nun tutarlılık sınırlarını belirleme çalışmaları sırasında geliştirildiğini belirtmiştik. Asimptotik tutarlılık, istatistikteki “Merkezi Limit Teoremi” çerçevesinde, öğrenme teorisinin sahip olması gereken bir tutarlılıktır. Bunun, sadece bir eğitim veri kümesi için değil, tüm veri setleri için geçerliliğini test eden, ama her seferinde uygulanması zor olan anahtar teoremi takip etmektedir. Sonrasında model karmaşıklığının kontrolü önemli bir konudur. ARM’nin tutarlılığı için bir diğer koşul büyüme fonksiyonunun tespiti²⁷⁶.

Büyüme fonksiyonu, herhangi bir $n > n_0$ için,

$$P\left((R(\omega) - R(\omega^*)) < \varepsilon\right) = e^{-c n \varepsilon^2} \quad c, \text{pozitif bir sabit} \quad (0.54)$$

sınırı sağlanıyorsa, asimptotik yakınsama oranı hızlıdır denilir. VC teorisine göre, dağılımdan açıklayıcı olarak, ARM ilkesinin tutarlılığı ve hızlı yakınsaması için hem gerek hem yeter şart olmasıdır, bu bağlamda:

²⁷⁶ Şenol Emir, **a.g.e.** s. 150.

$$\lim_{n \rightarrow \infty} \frac{G(n)}{n} = 0 \quad (0.55)$$

büyüme fonksiyonu bir olasılık dağılımına bağlı olmadığı için, denklem 2.55’te belirtilen koşul da dağılımdan açıklayıcı olup hızlı yakınsamayı da garanti etmektedir. Sonuç olarak, akla “amaç küçük (sınırlı) örneklem için algoritmalar oluşturmak ise neden bir tutarlılık teorisine ihtiyaç duyulur?” sorusunun cevabı; VC teorisinin ARM ilkesinin yakınsaması için sadece yeter şartların değil, gerek şartların da neler olduğunu belirtmesidir²⁷⁷.

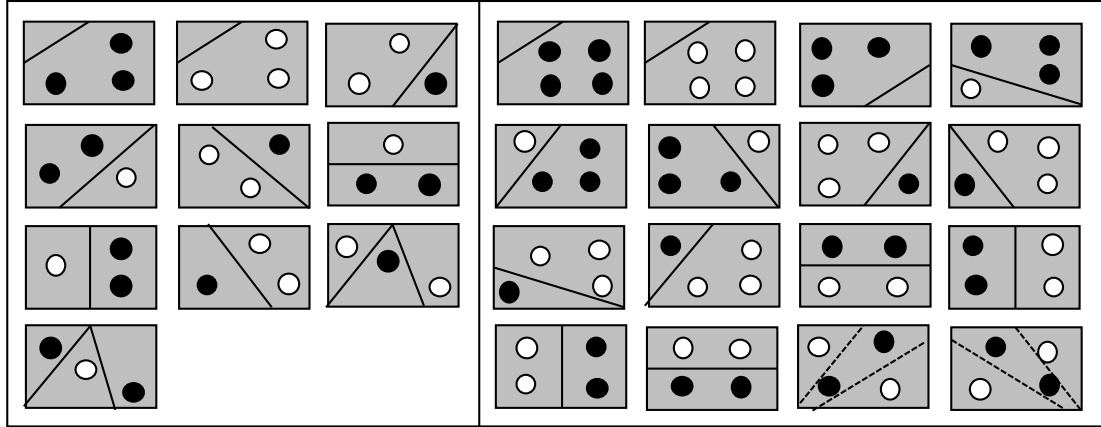
2.3.1.2. VC Boyutu

$Q(z, \alpha), \alpha \in \Lambda$ şeklinde gösterilen gösterge fonksiyonlarının bir kümesinin VC boyutu (h ile gösterilir), bu işlevler tarafından parçalanmış bu kümenin, tüm 2^h muhtemel yolla kullanılabilen fonksiyonlarının, ayrılabilen $z_1, z_2, \dots, z_h$ vektörlerin maksimum h sayısıdır²⁷⁸. Daha basit bir tanımla, VC boyutunu şu şekilde de tanımlayabiliriz: Eğer n adet örneklem temsilci fonksiyonların tamamı tarafından olası 2^n durumun tamamında ayrılabiliyorsa, o zaman bu örneklem kümesinin bu fonksiyon kümesi ile ayrıştırılabildiği söylenir ve VC boyutu n olur. Örnek vermek gerekirse üç elemanlı ($n = 3$) bir örneklem düşünelim. Muhtemel sekiz adet ($2^3 = 8$) yolla ayrışabilir. Fakat burada uzayın boyutu da önemlidir. Söz konusu örnek için kâğıt üzerinde, yani iki boyutlu uzayda, Şekil 2.34a’da gösterildiği gibi, üç gözlemden oluşan örneklem muhtemel tüm yollarla ayrıştığı görülmektedir. Fakat gözlem sayısı dörde çıktığında, iki boyutlu uzayda tüm muhtemel durumların ($2^4 = 16$) ilgili fonksiyon kümesi tarafından ayrıştırılamadığı (son iki olasılıkta XOR problemiyle karşılaşıldığı), şekil 2.34b’de görülmektedir. Bu durumda iki boyutlu uzay için maksimum ayrılabilir gözlem sayısı üç ($h = 3$) olmaktadır.

²⁷⁷ Şenol Emir, **a.g.e.** s.115.

²⁷⁸ Vladimir Vapnik, “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.992.

Şekil 0.34: İki Boyutlu Uzayda, Üç Ve Dört Gözlemlilik İçin VC Boyutu



a

$n = 3$, muhtemel olasılık $2^3 = 8$

Ayrılabilir

b

$n = 4$, muhtemel olasılık $2^4 = 16$

Ayrılmıyor

İki boyutlu uzayda VC boyutu, $h=3$.

Kaynak: Emre Çomak, “Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar”, Doktora Tezi, Selçuk Üniversitesi, 2008, s.30.

Örnekte de görüldüğü gibi, uzayın boyutu iki (2) iken VC boyutu üç (3) olmaktadır. Buradan, şu şekilde bir genelleme yapmak mümkündür. Doğrusal problemlerin birçoğunda, VC boyutu uzay boyutundan bir fazla yani, $h = n + 1$ olmaktadır. Fakat ayırıcı hiperdüzlemlerin oluşturulması için kullanılan kenar (margin) sayısı arttıkça VC boyutu $n + 1$ ’ in altına düşmektedir. Şekil 2.34’te gösterilen şematik ve hesap VC boyutu hesaplamasının en temel boyutunu temsil etmektedir. Gerçek hayat problemlerinde, uzay boyutları da çok değiştiği için, VC boyutunu hesaplamak yukarıdaki gibi kolay olmamaktadır. Nitekim herhangi bir n gözlem sayısı için, $Q(z, \alpha), \alpha \in \Lambda$ kümesi tarafından parçalanabilen bir n vektör kümesi varsa bu durumda VC boyutu sonsuza eşit olmaktadır²⁷⁹.

Fonksiyonlar grubunun VC boyutunun sonlu olması, ARM yönteminin gerekli ve yeterli öğrenme koşullarını tanımlamakta, yani yöntemin tutarlılığını sağlamaktadır.

²⁷⁹ Vladimir Vapnik, “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.992-993.

Bu, VC boyutunun sadece veri kümesinin elemanlarının miktarını değil, aynı zamanda fonksiyonlar grubunun çeşitliliğinin ölçüsünü de karakterize ettiği anlamına gelmektedir. Dolayısıyla fonksiyonlar kümesi çok çeşitli olmamalıdır. Sonlu VC boyutuna sahip işlevlerin iç içe geçmiş alt kümeleri üzerindeki yapıyı kullanan (sonsuz VC boyutuna sahip olabilen işlev kümelerinde de tanımlanan) Yapısal Risk Minimizasyonu (YRM) ilkesi, riskin bu yapı için, mümkün olan en iyi riske yakınsamasını garanti eder²⁸⁰.

Ayrıca Vapnik ve Chervonenkis, büyüme fonksiyonunun doğrusal olduğunu ya da n sayısının logaritmik bir fonksiyon ile sınırlandığını, Şekil 2.35'te de görülebileceği gibi, kanıtladı. Büyümenin yavaşlamaya başladığı $n \frac{1}{4} h$ noktasına “VC boyutu” denir. VC boyutu sonlu ise, büyüme fonksiyonu yeterince büyük örnekler için doğrusal olarak büyümez ve;

$$G(n) \leq h \left(1 + \ln \frac{n}{h} \right) \quad (0.56)$$

şeklinde bir logaritmik fonksiyonla sınırlıdır. VC boyutu, bir dizi fonksiyonun bir özelliğidir. Bu boyutun sonlu olması, hızlı yakınsama ve ARM öğrenmesinin dağıtımdan açıklayıcı tutarlılığı için gerekli ve yeterli koşulları sağlamaz. Aksine, sınır herhangi bir n için doğrusal kalırsa;

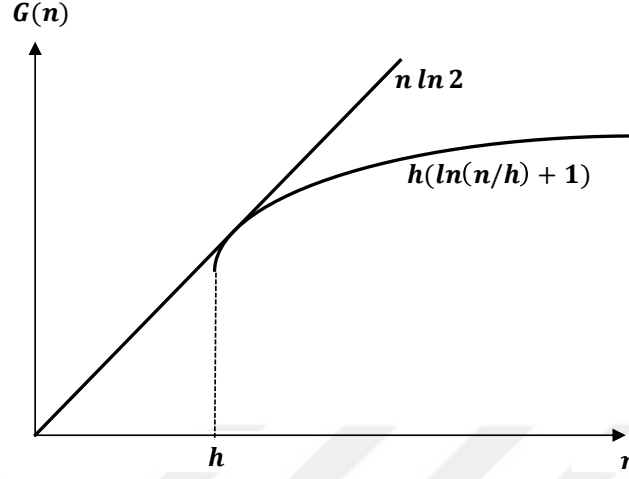
$$G(n) = n \ln 2 \quad (0.57)$$

şeklini alır²⁸¹.

²⁸⁰ Vladimir Vapnik, Rauf Izmailov, “Knowledge Transfer In SVM And Neural Networks”, **Ann Math Artif Intell**, 2017, Cilt: 81: s.8.

²⁸¹ Vladimir Cherkassky ve Filip Mulier, **a.g.e.** s.107.

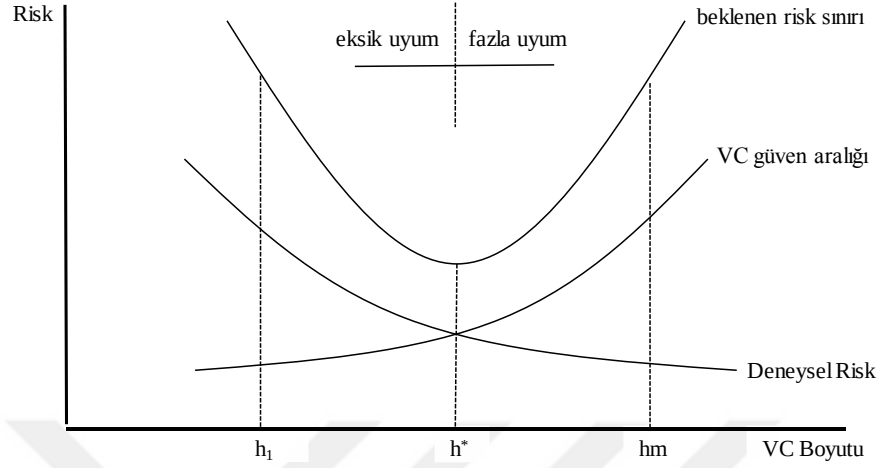
Şekil 0.35: Büyüme Fonksiyonunun Yapısı



Kaynak: Vladimir Cherkassky ve Filip Mulier, **Lerning From Data: Concepts, Theory and Methods**, 2. Baskı, New Jersey, John Wiley & Sons, Inc., 2007, s.107.

Hiyerarşik fonksiyonlar sınıfını $H_1 \in H_2 \in \dots \in H^* \in \dots \in H_m$ şeklinde sıralayacak olursak, her bir hipotez uzayı kendinden daha karmaşık olan hipotez uzayının bir alt kümesi olacaktır. Burada m uzayı en karmaşık, 1. Uzay ise en az karmaşık olan uzaydır. Bu hipotez uzaylarının VC boyutlarının ise sırasıyla $h_1 \in h_2 \in \dots \in h^* \in \dots \in h_m$ olduğunu varsayalım. Birinci (1.) uzayda, VC boyutu (h_1) en düşük, eğitim hatası en yüksek, m . uzay ise VC boyutunun (h_m) en yüksek fakat eğitim hatasının minimum olduğu hipotez uzayıdır. Fakat eğitim hatasının minimum tutulması, genelleme hatasının da minimum olmasını sağlamayacaktır. Eğitim hatasının minimize edilmesi, aşırı karmaşıklığa, o da beklenen riskin minimum noktadan uzaklaşmasına dolayısıyla da yüksek genelleme hatasına neden olacaktır. Sonuç olarak VC boyutunu, dolayısıyla hipotez uzayının büyüklüğünü, yani modelin karmaşıklığını, ampirik risk ile birlikte, minimize etmek beklenen riskin minimize edilmesini sağlayacaktır. Bu durum Şekil 2.36'da görülmektedir.

Şekil 0.36: VC Boyutu Kullanılarak Beklenen Riskin Minimize Edilmesi



Kaynak: Vladimir Cherkassky, Filip Mulier, **Lerning From Data: Concepts, Theory and Methods**, 2. Baskı, New Jersey, John Wiley & Sons, Inc., 2007, s.124.

Şekil 2.31'e tekrar dönüp bakılacak olursa, yukarıdaki Şekil 2.36'yla olan benzerliği ve farkları görülecek ve YSA ile DVM arasındaki fark daha da belirginleşecektir. YSA deneysel riski, en küçük hata karelerini minimize etme yoluyla, minimize etmektedir. Dolayısıyla YSA'nın bu yöntemle ulaştığı nokta, Şekil 2.36'da h_m noktasına denk gelen, kesikli çizgi olacaktır. Burası ise beklenen riskin minimize edildiği h^* noktasından uzak olup, modelin aşırı öğrendiği bölgeye denk gelmektedir. DVM ise VC boyutunu kullanarak, modelini h^* noktasını izleyen parametre değerleriyle oluşturacağı için aşırı ya da eksik öğrenme hatasına düşmeyip beklenen riski minimize edecektir.

Son olarak, hipotez uzayı konusunu diğer makine öğrenmesi yöntemlerine de genelleleyebiliriz. Şöyle ki, m . uzayı, m . dereceden bir fonksiyon, ya da m tane bulanık kuralın kullanıldığı bir bulanık mantık veya m tane gizli katman kullanan bir YSA olarak da düşünebiliriz²⁸².

²⁸² Emre Çomak, **a.g.e.** s.28.

2.3.1.3. Yapısal Risk Minimizasyonu

YRM, karmaşıklıkları açısından sıralanan, bir dizi model minimizasyonunu kullanır. Karmaşıklık genellikle serbest parametre sayısı ile belirlenir. VC boyutu, model karmaşıklığının başka bir ölçüsüdür. YRM tarafından model seçimi, modelin düzen açısından en basit ve verilerdeki ampirik hata açısından ise en iyisinin bulunmasına karşılık gelir²⁸³.

ARM prensibi büyük bir örneklem büyüklüğü ile başa çıkmak için geliştirilmiş ve verimli bir şekilde, aynı zamanda yaygın olarak kullanılmaktadır. Örneğin bir veri kümesindeki gözlem sayısına “ l ” diyelim. Eğer l/h büyükse, gerçek risk ampirik riskin değerine yakındır. Bu durumda, ampirik riskin küçük bir değeri, evet beklenen riskin küçük bir değerine karşılık gelir. Bununla birlikte, l/h küçükse, küçük bir $R_{amp}(\alpha_1)$ bile küçük bir risk değerini garanti etmez. Bu durumda, R_α minimizasyonu, her iki terimin, yani hem beklenen hem ampirik riskin, eşzamanlı minimizasyonuna dayanan yeni bir prensip gerektirir; bunlardan biri ampirik riskin değerine bağlı, ikincisi ise veri kümesinin VC boyutuna bağlıdır. Bu durumda riski en aza indirebilmek için; ampirik riskin değerini en aza indirmenin yanı sıra, öğrenme makinesinin VC boyutunu da kontrol eden bir yöntem bulmak gerekmektedir. YRM olarak adlandırılan aşağıdaki ilkenin, fonksiyon grubunun hem ampirik risk hem de VC boyutuna göre fonksiyonel riski en aza indirmesi amaçlanmaktadır²⁸⁴. Üstelik YRM ilkesi, aslında yaklaşıklığın kalitesi ile yaklaşık işlevin karmaşıklığı arasında bir denge önermektedir. YRM teorisinin temel sonuçları şöyledir; herhangi bir dağıtım fonksiyonu için YRM yöntemi, muhtemel en iyi çözüme yakınsaklık sağlar. Başka bir deyişle, YRM yöntemi, evrensel olarak, oldukça tutarlıdır²⁸⁵.

²⁸³ Ethem Alpaydın, **Introduction To Machine Learning**, 2. Baskı, Londra: The MIT Press, 2010, s.82.

²⁸⁴ Bkz. Bu konunun matematiksel arka planına, Vladimir Vapnik’in “An Overview of Statistical Learning Theory” isimli makalesinde özet olarak, daha ayrıntılı olarak 1995 yılındaki çalışmasına ve matematiksel ispatları da içeren ve 1998 yılında basılan “Statistical Learning Theory” isimli kitabından ulaşılabilir.

²⁸⁵ Vladimir Vapnik, “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.992-993.

Genelleme sınırını minimize etmek için kullanılabilecek, iki farklı YRM stratejisi bulunmaktadır. Bunlar:

1. “Model karmaşıklığını (VC boyutunu) sabit tutup, ampirik riski minimize etmek”.
2. “Ampirik riski sabit (küçük) tutup, VC boyutunu minimize etmek”.

YSA’da dâhil olmak üzere, birçok istatistiksel yöntem birinci stratejiyi uygularken, DVM ise ikinci stratejiyi uygulamaktadır²⁸⁶.

$$R(\alpha) = R_{emp}(\alpha) + \sqrt{\frac{h \left(\log \left(\frac{2n}{h} \right) + 1 \right) - \log \left(\frac{\eta}{4} \right)}{n}} \quad (0.58)$$

İstatiksel öğrenme teorisi, öğrenme makinelerinin genelleme yeteneğini, denklem 2.58 deki risk fonksiyonunu minimum tutarak, kontrol eder. Burada $R(\alpha)$ gerçek riski, $R_{emp}(\alpha)$ ampirik riski, h VC boyutu değerini, n gözlem sayısını ve η , $[0,1]$ aralığında bir değeri temsil etmektedir²⁸⁷.

YRM prensibine göre, bir öğrenme problemini sınırlı veri ile çözebilmek için, yaklaşan fonksiyonlar kümesi üzerinde öncül bir yapı oluşturulması gerekir. Bunun için kayıp fonksiyonları kümeleri birbirinin altkümeleri olacak şekilde bir yapı (2.59) oluşturulur.

$$S_1 \subset S_2 \subset \dots S_m \subset \dots \quad (0.59)$$

Her bir alt kümenin, VC boyutu sonludur ve h_i ile gösterilir (2.60).

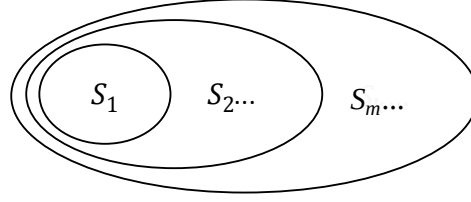
Tanım gereği, oluşturulan yapı, altkümelerin kapasitelerine (VC boyutuna) göre sıralı olmasını sağlar.

$$h_1 \leq h_2 \leq \dots \leq h_m \quad (0.60)$$

²⁸⁶ Wang, Xiangying, Zhong, Yixin: “Statistical Learning Theory and State of the Art in SVM”, **Proceedings of the Second IEEE International Conference on Cognitive Informatics**, 2003, s.57.

²⁸⁷ Emre Çomak, **a.g.e.** s.27.

Şekil 0.37: Kapasitelerine Göre Belirlenmiş Fonksiyon Kümeleri



Kaynak: Vladimir Cherkassky, Filip Mulier, **Lerning From Data: Concepts, Theory and Methods**, 2. Baskı, New Jersey, John Wiley & Sons, Inc., 2007, s.122.

Bu yapı belirlendikten sonra optimal olan modelin seçimi:

- Yapının optimal kapasiteye sahip elemanını seçmek
- Bu eleman üzerinden modeli tahmin etmek

adımlarını içerir.

YRM ilkesinin uygulanabilmesi için

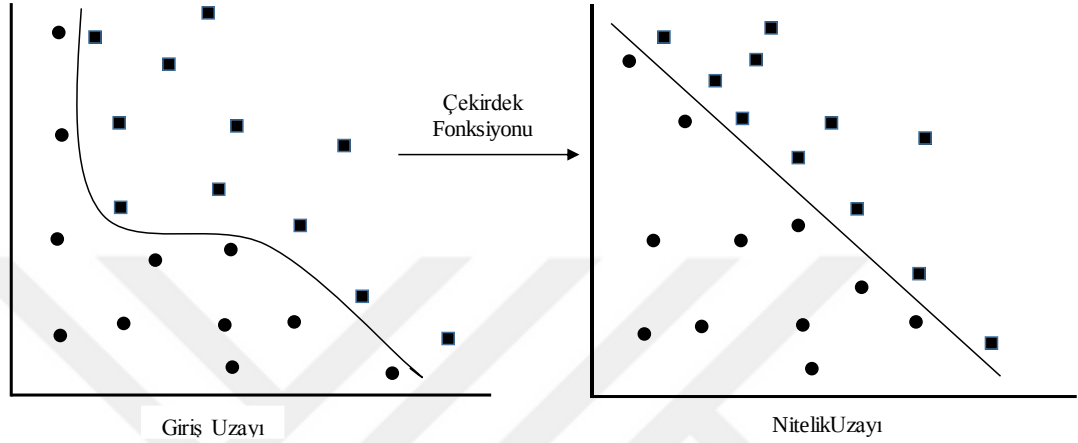
- Yapıdaki her bir S_m alt kümesinin VC boyutu hesaplanır.
- Her bir alt küme için ampirik risk minimize edilir²⁸⁸.

2.3.2. Çekirdek Fonksiyonları

Çekirdek (kernel) fonksiyonları, doğrusal olarak ayıramayan giriş uzayı verilerini doğrusal olarak ayrılabilen yüksek boyutlu nitelik uzayına dönüştürmek (Şekil 2.38) ve böylelikle boyut sıkıntısını gidermek için kullanılan fonksiyonlardır.

²⁸⁸ Vladimir Cherkassky ve Filip Mulier, **a.g.e.** s.123.

Şekil 0.38: Girdilerin Doğrusal Olarak Ayrılamadığı Giriş Uzayından, Çekirdek Fonksiyonu Yardımı İle, Doğrusal Olarak Ayrılabilirdiği Yüksek Boyutlu Nitelik Uzayına Dönüştürülmesi



Kaynak: Emre Çomak, “Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar”, Doktora Tezi, Selçuk Üniversitesi, 2008, s.32.

Mercer Teoremi, DVM’de önceden atanmış bir çekirdek işlevi, örtük olarak bir eşleme işlevini (giriş alanından özellik uzayına) ve özellik uzayını, yani gerek ve yeter koşulları tanımlar. Bir başka deyişle bir fonksiyonun çekirdek fonksiyon olabilmesi için Mercer koşullarını sağlaması gerekir. Pozitif tanımlı ve bulanık tabanlı veya üyelik fonksiyonları ile bulanık bir sistem oluşturmak, özünde orijinal girdi uzayında doğrusal olmayan bir karar veya yaklaşım fonksiyonlarına karşılık gelen çekirdek fonksiyonları, küçültülmüş yüksek boyutlu özellik uzayında bir hiper düzlem bulmakla aynıdır²⁸⁹.

Tablo 2.7’de literatürde en çok rastlanılan çekirdek fonksiyonları listelenmektedir (γ, r, τ ve d çekirdek parametreleridir).

²⁸⁹ J. Shen, **a.g.e.** s.204-213.

Tablo 0.7: Çekirdek Fonksiyonları

Çekirdek Fonksiyonunun Adı	Fonksiyon
Doğrusal	$x_i^T x_j$ (0.61)
Radyal tabanlı fonksiyon	$\exp(-\gamma \ x_i - x_j\ ^2), \quad \gamma > 0$ (0.62)
Polinom	$(\gamma x_i^T x_j + r)^d, \quad \gamma > 0$ (0.63)
Sigmoid	$\tanh(\gamma x_i^T x_j + r)^d, \quad \gamma > 0$ (0.64)
Çoklu Kuadratik	$(\ x_i - x_j\ + \tau)^{1/4}$ (0.65)
Ters Çoklu Kuadratik	$(\ x_i - x_j\ + \tau)^{-1/4}$ (0.66)

Kaynak: Selçuk Alp ve Ersoy Öz, **Makine Öğrenmesinde Sınıflandırma Yöntemleri Ve R Uygulamaları**, Nobel Kitabevi, Ankara, 2019, s.175.

Kullanılan veri kümesi doğrusal olarak ayrılabilirse, verileri yüksek boyutlu uzaya taşıma gereği duyulmadan, doğrusal fonksiyonla, hızlı bir şekilde çözülebilir. Polinom fonksiyon ise, doğrusal fonksiyonun daha genelleştirilmiş hali olup, d parametresinin 1 olması durumunda yine doğrusal bir fonksiyona dönüşür. Radyal tabanlı fonksiyon, en çok kullanılan çekirdek fonksiyonudur. Radyal tabanlı fonksiyonda γ yarıçapın kontrolünü sağlar²⁹⁰.

DVM uygulaması sırasında çekirdek fonksiyonlarının seçimi, YSA'daki gizli katman ve nöron sayısı seçimi gibi önemli bir konudur. Çünkü çekirdek fonksiyonunun türü de uygulama sonuçlarının performansına etki etmektedir.

2.3.3. Optimal Ayırıcı Hiperdüzlem

DVM, gözlem verilerini temsil eden uzayda yer alan gözlem vektörlerini, temel olarak, iki ayrı sınıfa bir platformla (hiperdüzlem), ayırma yaklaşımı içermektedir. DVM'nin teorik arka planı; gözlemleri, yani gözlem uzayındaki bu vektörleri, sınıflandırırken ayırıcı hiperdüzlemin (hyperplane) her iki sınıfın, kendisine en yakın gözlemlerine

²⁹⁰ Emre Yakut, **a.g.e.** s.51.

olan mesafesini maksimize edecek (maksimum marj) şekilde, sınırlarının belirlenmesi amacına dayanmaktadır. Başlangıçta gözlemleri iki farklı sınıfa ayırma mantığıyla temelleri atılan DVM, günümüzde genişletilmiş teorik çerçevesi sayesinde, çok boyutlu ve doğrusal olmayan birçok sınıflandırma problemine de rahatlıkla uygulanabilmektedir. Yöntem ele alınırken de bu sıranın takip edilmesi anlaşılması bakımından önemlidir.

DVM, maksimum marj sınıflandırıcısı olarak da bilinir. Boyutu n olan bir uzayda, eğitim örnekleri şu formülle (2.67) birbirinden ayrılabilir:

$$f(x) = w \cdot x + b = 0 \quad (0.67)$$

Eğitim verisinin ayrılabilirdiğinin varsayıldığı durumlarda, tüm örnekler;

$$|w \cdot x_i + b| = 1, i = 1, 2, \dots, N, \quad N: \text{Örneklem Sayısı} \quad (0.68)$$

Şeklindeki bir formda (2.68) temsil edilir. Bunlardan bazıları;

$$f(x) = w \cdot x + b = 1 \quad (0.69)$$

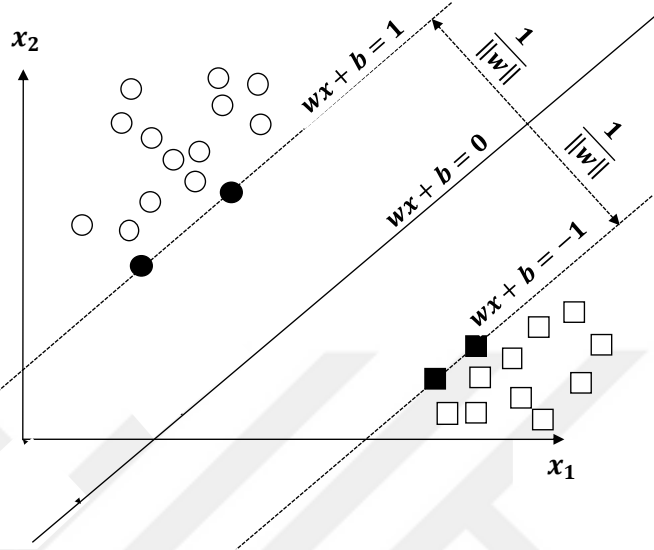
şeklinde bazıları ise;

$$f(x) = w \cdot x + b = -1 \quad (0.70)$$

şeklindedir. Bu ayırıcı fonksiyonun oluşturduğu sınırın toplam mesafesi, Şekil 2.39'dan da anlaşılacağı üzere, $2/\|w\|$ olur. Ayırıcı fonksiyonlar kümesinin VC boyutu, ağırlık vektörü (w) ile olan sınırı ayarlayarak kontrol edilir. Böylece en yakın gözlem noktalarının ayırıcı sınıra, yani optimal ayırıcı hiperdüzleme, mesafesi $1/\|w\|$ olacaktır. Matematiksel açıdan düşünüldüğünde, ağırlık vektörü (w) değerinin minimum tutulması, örnekleri temsil eden noktaların optimal ayırıcı hiperdüzlem ile olan mesafesinin maksimum, VC boyutunun ise minimum olmasını sağlayacaktır²⁹¹.

²⁹¹ Emre Çomak, **a.g.e.** s.36.

Şekil 0.39: Optimal Ayırıcı Hiperdüzlem



Kaynak: Michal Haltuf, Support Vector Machines for Credit Scoring, Yüksek Lisans Tezi, University of Economics in Prague, 2014, s.15.

2.3.3.1. Doğrusal Olarak Ayrılabilir Veri Kümeleri için Doğrusal Maksimal Marj Sınıflandırıcı

Literatürde birçok doğrusal sınıflandırıcı modeli bulunmaktadır (Bkz. Tablo 3.1). DVM’nde, farklı olarak, iki sınıfı birbirinden ayırma fonksiyonunun yanı sıra, bu sınıflar arasındaki mesafeyi maksimize etme işlevi de bulunmaktadır. Bu işlev sayesinde DVM, yeni veri kümelerine uygulanması açısından, daha genellenebilir bir forma sahip olmaktadır²⁹².

Bir X eğitim setinin özellik vektörleri x_i , ve $i = 1, 2, \dots, N$ olsun. Bu vektörlerde, doğrusal olarak ayrılabilirdiği varsayılan, iki ayrı sınıfa (ω_1, ω_2) ait olsun. Burada amaç bir hiperdüzlem oluşturmaktır. Marj sınıflandırma konusunda, iki boyutlu

²⁹² Sergios Theodoridis, Konstantinos Koutroumbas, **Pattern Recognition**, 4. Baskı, San Diego: Elsevier Inc., 2009, s.119.

gösteriminin mümkün olabilmesi için, iki boyutlu bir girdi uzayını ele alalım ve girdi verilerini de Denklem 2.71’deki gibi gösterelim²⁹³;

$$(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l), x \in R^n, y \in \{+1, -1\} \quad (0.71)$$

Bu veri kümesinin doğrusal olarak ayrılabilirliğini düşünelim. O halde bu veriler, onlarca farklı ayırıcı düzlemlerle (hiperdüzlem), Şekil 2.40’ta gösterildiği gibi, birbirinden ayrılabilir. Bu ayırma çizgisi bir başka deyişle “karar çizgisi”;

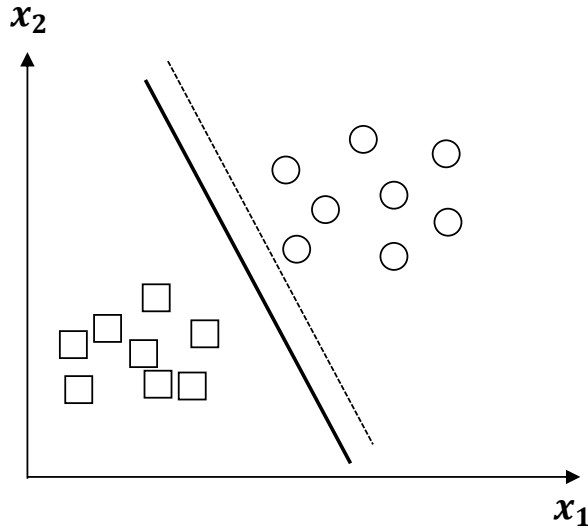
$$w_1x_1 + w_2x_2 + b = 0 \quad (0.72)$$

özetle,

$$g(x) = w^T x + b = 0 \quad (0.73)$$

formülüyle bulunur. Bu şekilde oluşturulan bir sınıflandırıcı, iki muhtemel sınıflandırıcıdan biri gibi, vektörleri sınıflara doğru bir şekilde ayıracaktır.

Şekil 0.40: Mümkün İki Farklı Sınıflandırıcı İle Doğrusal Olarak Ayrılabilen İki Sınıflı Bir Örnek



Şekil 2.40’daki hem kesikli çizgi ile gösterilen ayırıcı hiperdüzlem hem de düz çizgi ile gösterilen hiperdüzlem, söz konusu veri kümesini sınıflara doğru bir şekilde

²⁹³ Vojislav Kecman, “Support Vector Machines – An Introduction”, **StudFuzz**, Cilt: 177, 2005 s.12.

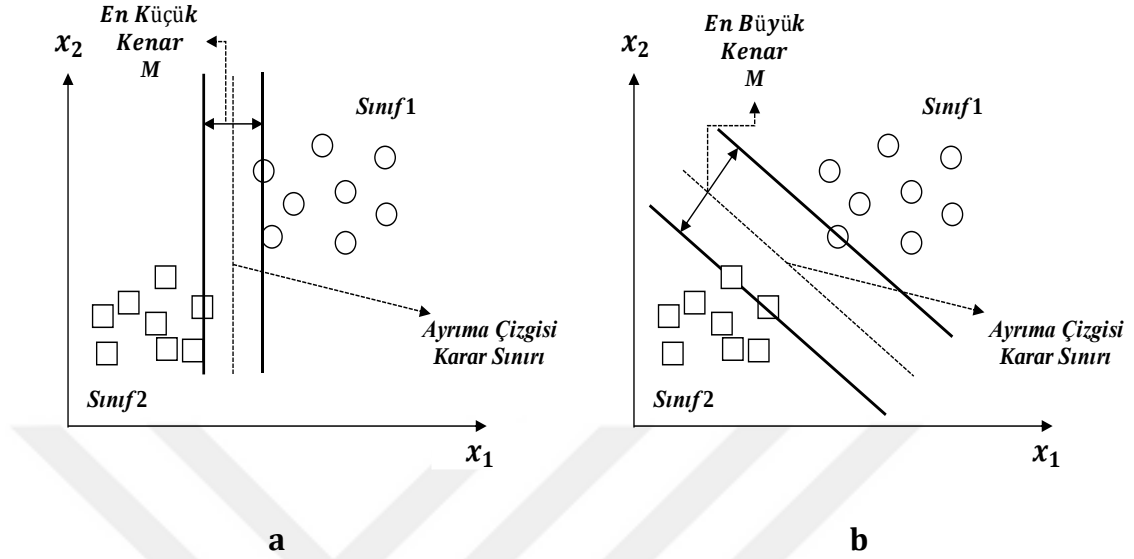
ayırabildiği görülmektedir. Fakat bu ayırma çizgilerinden hangisi, eğitim verisinin dışındaki, farklı veri kümeleri için daha duyarlı bir ayırıcı hiperdüzlem olacaktır? Elbette düz çizgi ile gösterilen sınıflandırıcı, daha güvenilir ve genelleme kabiliyeti açısından daha etkin bir sınıflandırıcı olacaktır. Çünkü bu hiperdüzlem, her iki tarafında muhtemel yeni gözlemleri için daha çok yer sağlamaktadır. Ayrıca mevcuttaki gözlemlerin değerinde bir değişme olursa bu, herhangi bir hataya sebep olmaksızın rahat hareket edebilecekleri anlamına gelmektedir. Bu da ayırma işlemini yaparken, “sınırlandırıcının performansını genelleme (generalization performance of the classifier)” olarak bilinen bir işlevin de yerine getirilmesini sağladığı gibi, eğitim verisi yardımıyla oluşturulan sınıflandırıcının performansını da artıracaktır²⁹⁴. Bu aşamada sorulması gereken soru; “bu sınıflandırma performansının maksimizasyonunun nasıl sağlanacağı?” sorusudur.

Sınıflandırıcının performansını maksimize etmek için, sınırlı sayıda ve dağılımı bilinmeyen bir eğitim verisi için, en iyi optimal ayırıcı hiperdüzlemi bulmak gerekmektedir. Ampirik hatayı minimuma, sınıflar arasındaki mesafeyi (marjı) ise maksimuma ulaştıracak ve bilinmeyen veriler için de optimal bir genelleme sağlayacak bir hiperdüzlemin oluşturulması esastır. Şekil 2.41a’daki hiperdüzlem sınıfları birbirinden doğru bir şekilde ayırabilmektedir. Fakat Şekil 2.41b’deki hiperdüzlem, farklı sınıfların birbirine en yakın gözlemlerine olan mesafeyi maksimize ederek, maksimal ayırıcı hiperdüzlemi oluşturma mantığıyla çizilmiştir. Dolayısıyla Şekil 2.41b’deki hiperdüzlemin beklenen riski daha küçük tutması beklenir²⁹⁵.

²⁹⁴Sergios Theodoridis ve Konstantinos Koutroumbas, **a.g.e.** s.119-120.

²⁹⁵ Vojislav Kecman, **a.g.e.** s.12.

Şekil 0.41: Birbirinden Farklı Hiperdüzlemlerle Ayrılmış İki Boyutlu Girdi Uzayı

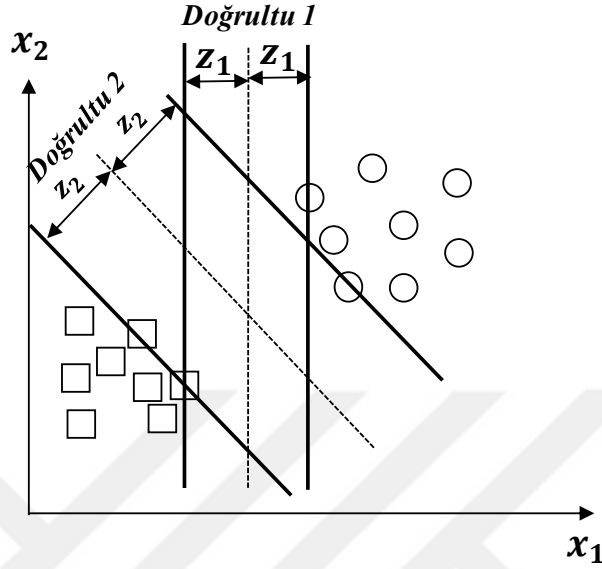


Kaynak: Vojislav Kecman, “Support Vector Machines – An Introduction”, **StudFuzz**, Cilt: 177, 2005 s.13.

İki sınıfın arasında, binlerce sayıda farklı hiperdüzlem çizilebilir. Çizilebilecek bu hiperdüzlemlerden hangisinin “doğrusal olarak ayrılabilir veriler için optimal olan, doğrusal maksimal marj sınıflandırıcı” olduğunu belirlemek gerekmektedir. Bunun için, her bir hiperdüzlem doğrultusuyla (w) ve uzaydaki tam pozisyonuna olan mesafesi, yani bias sapması, b ile temsil edilsin. Oluşturulmak istenen hiperdüzlem, her iki sınıfın birbirine en yakın gözlemlerine eşit uzaklıkta olması gerekir. Bu durum Şekil 2.42’de görülmektedir. Burada Doğrultu 1 için marjin $2z_1$ iken, Doğrultu 2 için marjin $2z_2$ dir²⁹⁶.

²⁹⁶ Sergios Theodoridis ve Konstantinos Koutroumbas, **a.g.e.** s.120.

Şekil 0.42: Mümkün İki Farklı Sınıflandırıcı İle Doğrusal Olarak Ayrılabilen İki Sınıflı Bir Örnek



Her bir hiperdüzlem bir ölçek faktör ile belirlenir. Uygun ölçeklendirme ile tüm aday hiperdüzlemler dikkate alınabilir. Bir gözlemin bir hiperdüzleme uzaklığı;

$$z = \frac{|g(x)|}{\|w\|} \quad (0.74)$$

formülüyle hesaplanır. Burada $g(x)$ fonksiyonu, Şekil 2.43'daki düz çizgi ile gösterilen fonksiyonu, yani doğrusal olarak ayrılabilir veriler için doğrusal maksimal marj sınıflandırıcıyı temsil etmektedir. z hiperdüzlemin her iki sınıfın en yakın vektörlerine olan eşit uzaklığı göstermektedir. $g(x)$ Sınıflardan biri için 1 değerine, diğeri içinse -1 değerine karşılık gelir. Bu durumda;

$$z = \frac{1}{\|w\|} \quad (0.75)$$

olur. Sınıflar arası maksimum marj ise;

$$2z = \frac{1}{\|w\|} + \frac{1}{\|w\|} = \frac{2}{\|w\|} \quad (0.76)$$

olur. Dolayısıyla,

$$w^T x + b \geq 1, \text{ ise } \forall x \in \omega_1$$

$$w^T x + b \leq 1, \text{ ise } \forall x \in \omega_2 \quad (0.77)$$

şeklinde vektörler, sınıflara atanır.

Denklem 2.76'dan da görüleceği üzere, dikkat edilmesi gereken, marjini maksimize etmek için, w 'yi minimize etmek gerekliliğidir. Bunun için;

$$J(w, b) \equiv \frac{1}{2} \|w\|^2 \quad (0.78)$$

fonksiyonunu minimize etmek gerekmektedir. Tabii;

$$y_i(w^T x_i + b) \geq 1, \quad i = 1, 2, \dots, N \quad (0.79)$$

olur²⁹⁷.

Burası ise problemin, artık doğrusallıktan çıkıp, doğrusal olmayan (nonlinear) optimizasyon problemine dönüştüğü bir noktadır ve Karush-Kuhn-Tucker koşulları geçerlidir.

$$\frac{\partial}{\partial w} L(w, b, \lambda) = 0 \quad (0.80)$$

$$\frac{\partial}{\partial b} L(w, b, \lambda) = 0 \quad (0.81)$$

$$\lambda_i \geq 0, \quad i = 1, 2, \dots, N \quad (0.82)$$

$$\lambda_i [y_i(w^T x_i + b) - 1] = 0, \quad i = 1, 2, \dots, N \quad (0.83)$$

λ , Lagrange çarpanlarının (λ_i) bir vektörüdür ve $L(w, b, \lambda)$ bir Lagrange fonksiyonudur ve;

²⁹⁷ Sergios Theodoridis ve Konstantinos Koutroumbas, **a.g.e.** s.120.

$$L(w, b, \lambda) = \frac{1}{2} w^T w - \sum_{i=1}^N \lambda_i [y_i (w^T x_i + b) - 1] \quad (0.84)$$

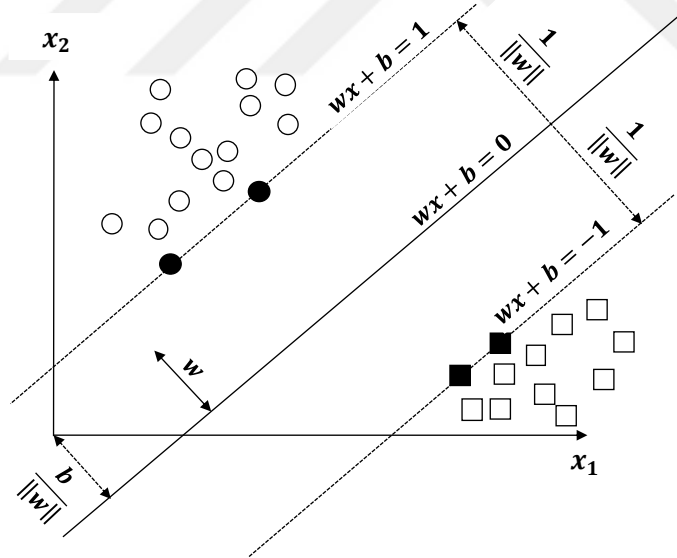
Sonuç olarak, bu denklemler birleştirilerek aşağıdaki denklemler elde edilir²⁹⁸;

$$w = \sum_{i=1}^N \lambda_i y_i x_i \quad (0.85)$$

$$w = \sum_{i=1}^N \lambda_i y_i = 0 \quad (0.86)$$

Sonuçta gözlemlerin sınıflara ayrımı, Şekil 2.43'teki gibi olur.

Şekil 0.43: Doğrusal Olarak Ayrılabilir Veriler İçin Doğrusal Maksimal Marj Sınıflandırıcının Matematiksel Gösterimi



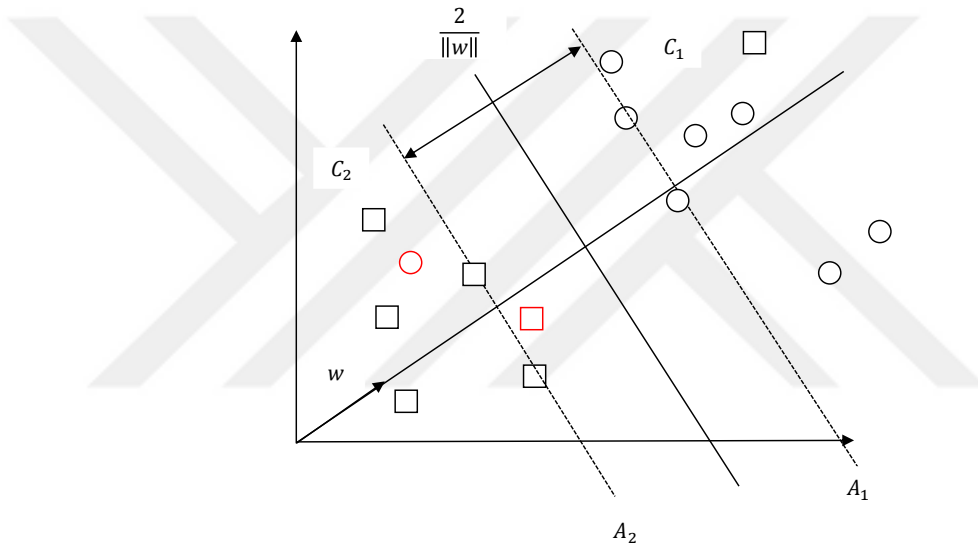
Kaynak: Michal Haltuf, Support Vector Machines for Credit Scoring, Yüksek Lisans Tezi, University of Economics in Prague, 2014, s.15.

²⁹⁸ Sergios Theodoridis ve Konstantinos Koutroumbas, **a.g.e.** s.120.

2.3.3.2. Doğrusal Olarak Ayrılamayan Veri Kümeleri için Doğrusal Maksimal Marj Sınıflandırıcı

Bazı veri grupları doğrusal olarak ayıramayabilir. Bu durumda, çizilecek herhangi bir doğrusal hiperdüzlem ile iki sınıfı birbirinden, her iki sınıfın tüm gözlemlerini doğru sınıfa atayacak şekilde, ayırmak mümkün olamayacaktır. Bu durum Şekil 2.44'te görülmektedir.

Şekil 0.44: Doğrusal Olarak Ayrılamayan İki Sınıflı Bir Örnek



Kaynak: Emre Yakut, “Veri Madenciliği Tekniklerinden C5.0 Algoritması Ve Destek Vektör Makineleri İle Yapay Sinir Ağlarının Sınıflandırma Başarılarının Karşılaştırılması: İmalat Sektöründe Bir Uygulama”, Doktora Tezi, Atatürk Üniversitesi, 2012, s.47.

Vektörlerin doğrusal olarak ayıramadığı durumlarda, gözlemleri üç farklı kategoride incelemek gerekmektedir ve bu kategorilerin formülasyonunun ise şu şekilde olması gerekmektedir²⁹⁹:

- I. Hiperdüzlemin dışına düşen ve tamamen doğru sınıflandırılan vektörler;

$$y_i(w^T x_i + b) \geq 1, \quad i = 1, 2, \dots, N \quad (0.87)$$

²⁹⁹ Sergios Theodoridis ve Konstantinos Koutroumbas, **a.g.e.** s.125.

II. Marjin bölgesine düşen ama yine doğru sınıflandırılan vektörler;

$$0 \leq y_i(w^T x_i + b) < 1, \quad (0.88)$$

III. Marjin bölgesine düşen ve doğru sınıflandırılmayan vektörler;

$$y_i(w^T x_i + b) < 0 \quad (0.89)$$

Tüm bu formülleri, yeni bir değişken seti, yani aylak değişken (ξ_i) tanımlayarak, tek bir çatı altında toplayabiliriz³⁰⁰;

$$y_i = +1 \text{ için } y_i(w^T x_i + b) \geq 1 - \xi_i \quad (0.90)$$

$$y_i = -1 \text{ için } y_i(w^T x_i + b) \geq 1 + \xi_i \quad (0.91)$$

$$\xi_i \geq 0, \forall_i \quad (0.92)$$

Burada birinci veri sınıfı $\xi_i = 0$ eşitliğine (x_i örneği doğru sınıflandırılmış), ikinci sınıf $0 \leq \xi_i \leq 1$ aralığına (x_i örneği doğru sınıflandırılmış, ancak A_1 ve A_2 hiperdüzlemleri arasında yer alıyor) ve üçüncü sınıf için ise $\xi_i > 1$ aralığına (x_i örneği yanlış sınıflandırılmış) karşılık gelmektedir. Burada ξ_i değişkenleri gevşek ya da aylak değişkenler (slack variables) olarak bilinir. Burada optimizasyon hedefi daha karmaşık bir hale gelir. Yapılması gereken, gevşek değişkenin mümkün olduğu kadar sıfır değerini aşmamasına dikkat ederek, marjini maksimize etmektir. Bu ise Denklem 2.93'te yera alan J fonksiyonunu minimize etmekle mümkün olur;

$$J(w, b, \xi) \equiv \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N I(\xi_i) \quad (0.93)$$

ξ , burada ξ_i parametrelerinin bir vektörüdür ve

$$I(\xi_i) = \begin{cases} 1 & \text{eğer } \xi_i > 0 \\ 0 & \text{eğer } \xi_i = 0 \end{cases} \quad (0.94)$$

³⁰⁰ Selçuk Alp ve Ersoy Öz, **a.g.e.** s.173.

olur³⁰¹.

2.3.4. Destek Vektör Makinelerinin Avantajları

DVM, güçlü bir algoritmaya ve ispatlanmış bir matematiksel altyapıya sahiptir. DVM, DVS (Destek Vektör Sınıflandırma) olarak bilinen sınıflandırma işlemlerinde ve DVR (Destek Vektör Regresyon) olarak bilinen tahmin işlemlerinde oldukça başarılıdır. DVM, yüksek boyutlu, az sayıda veri içeren veri kümelerinde de başarılı olabilmesiyle, diğer yöntemler arasında, ön plana çıkmaktadır. Çekirdek fonksiyonlarının kullanımı ile birlikte, DVM, yüksek boyutlu nitelik uzayında doğrusal bir modele dönüşür³⁰².

DVM'nin popülaritesi, onun “kernel hilesi” olarak da bilinen, kernel fonksiyonlarını kullanabilmesinden kaynaklanmaktadır. Bu güçlü özellik onu yüksek boyutlu uzayda, bu uzayın boyutlarından etkilenmeyen, doğrusal bir sınıflandırıcıya dönüştürmektedir.³⁰³

DVM dışındaki yöntemler sadece eğitim hatasını dikkate almaktadır. Genelleştirme hatasını, yani öğrenme yönteminin genelleştirme kapasitesini dikkate almamaktadırlar. Bu da, girdi verileri yüksek bir boyuta sahipse veya arkasında yer alan işlev çok karmaşıksa, öğrenme için çok daha fazla veri gerektirmektedir. Ek olarak, “boyut laneti” de (yani, YSA’nda çok sayıda gizli nöron veya bulanık modellerde çok fazla bulanık kural) yüksek boyutlu giriş boşluğunun yol açtığı sonuç olacaktır. Bu iki sorunu çözmek için yeni öğrenme makinesi algoritmalarına ihtiyaç duyulmaktadır. DVM, istatistiksel öğrenme teorisi çerçevesinde yer alan, yüksek boyutlu girdi alanından ve küçük veri boyutundan öğrenmeye uygun, yeni nesil bir öğrenme makinesidir³⁰⁴.

DVM, sınıflandırma konusunda, özellikle yüksek boyutlu ve sınırlı sayıda gözlem olması nedeniyle kötü davranışlı olan (ill-behaved) dağılımlarda, farklı bir esneklik ve

³⁰¹ Sergios Theodoridis ve Konsantinos Koutroumbas, **a.g.e.** s.125.

³⁰² Emre Çomak, **a.g.e.** s.5.

³⁰³ Sergios Theodoridis ve Konstantinos Koutroumbas, **a.g.e.** s.203.

³⁰⁴ J. Shen, **a.g.e.** s.1.

genelleme kapasitesi sunar. Gerçi bunun aksi de olabilir. Kolay durumlarda, yetersiz sınıflandırma performansı verebilir. Öklit uzaklık sınıflandırıcısını kullanan LDA, bu tip zor durumlarda DVM'den daha iyi sonuçlar verebilir, fakat bunun için çok fazla eğitim verisine ihtiyaç duyarlar. DVM'nin esnek sınıflandırmasının arkasında, kullandığı uygun çekirdek fonksiyonları aracılığıyla sağladığı, ileri eğrisellik (doğrusallıktan uzak olması) yatmaktadır. DVM ayrıca tüm veriyi eğitim verisi olarak kullanabilme ve ek test verisine ihtiyaç duymama ayrıcalığına da sahiptirler. Böylelikle daha eğitilmiş bir çerçeve sunarlar. Yöntemin esneklikleri, eğitim verileri üzerinde, otomatik olarak ve ilkeli bir şekilde kontrol edilebilmektedir³⁰⁵.

DVM'nin kullandığı yöntem, yani sınıf sınırları arasındaki mesafenin maksimizasyonu ile modellerin eğitimi, örneğin hata kareleri ortalaması gibi diğer risk fonksiyonlarını minimize eden yöntemlere bir alternatiftir. Bu yöntem maksimum kaybı minimize etme prensibine dayanmakta ve bazı özellikler içermektedir. Bunlar sınıflandırma fonksiyonunun otomatik kapasite ayarı ve eğitim verilerinden az sayıda destek modelinin çıkarılması ve çözümün benzersizliğidir. Yüksek dereceli polinom veya RBF sınıflandırıcılar gibi çok büyük kapasiteli lineer sınıflandırıcılar için etkili bir öğrenme algoritmasından yararlanırlar. Burada anahtar, karar fonksiyonunun iki boyutlu uzayda, çok daha düşük boyutlara sahip bir çift alanda temsil edilmesidir. Eğitim algoritması, çözüm uzayının boyutları üstel veya sınırsız olduğu durumlarda bile, eğitim modellerinin sayısı için bir polinomdur. Sonuç olarak da model, son derece basite indirildiği için, çok kısa bir zamanda sonuç verebilmektedir³⁰⁶.

Makine öğrenmesi ya da veri madenciliği veya istatistiksel veri analizi şeklindeki başlıklar altında toplayabileceğimiz birçok modelleme yönteminin amacı, veriden (örneklerden) öğrenmedir. YSA'da dâhil olmak üzere birçok öğrenme makinesi, regresyon, örüntü tanıma, sistem kontrolü ve modellemesi, tahmin vb. alanlarda

³⁰⁵ A.I. Belousov, S.A. Verzhakov, Ve J. Von Frese, "A Flexible Classification Approach With Optimal Generalisation Performance: Support Vector Machines", **Chemometrics And Intelligent Laboratory Systems**, Cilt: 64, 2002, s.24.

³⁰⁶ Bernhard E. Boser, Isabelle M. Guyon, Vladimir Vapnik, "A Training Algorithm For Optimal Margin Classifiers", Çevrimiçi, <file:///C:/Users/gokhanno/Desktop/Downloads/download.pdf>, 26/04/2020.

yoğunlukla kullanılmaktadır. Fakat bu modellerin çoğu; eksik öğrenme, aşırı öğrenme, düşük yakınsama, yerel minimuma yakalanma ve çok fazla veriye ihtiyaç duyma gibi handikapları mevcuttur. Bu problemlerin önemli bir kısmı, bu yöntemlerin ARM ilkesini temel almasından kaynaklanmaktadır. DVM ise, ARM'den farklı olarak YRM prensibini temel alan, özgün bir yöntemdir. DVM, kullandığı modelin karmaşıklığını da kontrol eden YRM prensibini kullanarak, söz konusu problemleri büyük ölçüde aşabilmektedir. DVM, bu özelliği sayesinde, iyi bir genelleme yeteneği sunar ve az miktarda veri ile de öğrenebilir ve genellikle evrensel optimum noktaya ulaşabilir³⁰⁷.

2.3.5. Destek Vektör Makinelerinin Dezavantajları

DVM hakkındaki bazı tartışmalar ise devam etmektedir. Bunlar, uygulama sırasında uygun çekirdek fonksiyonu seçimine dair bir yöntemin olmaması ve henüz çekirdeklerin kullanımına dair bir yönergenin oluşmamış olmasıdır. Ayrıca ilk başta iki sonuçlu sınıflandırma problemleri için geliştirilen DVM, evet çok sınıflı problemlere de uyarlanmıştır. Fakat çok sınıflı DVM'nin verimliliği, iki sınıflı olana kıyasla düşük kalmaktadır³⁰⁸.

Büyük veri kümelerindeki hesaplama maliyetinin yüksekliği ve bunun yol açtığı eğitim zamanının uzaması DVM uygulamalarında sorun olmaktadır. Ayırıcı düzlemin, sadece az sayıda eğitim örneği (destek vektör) referans alınarak oluşturulmasından dolayı aykırı veri noktalarına karşı oluşan hassasiyet, yöntemin başarımına olumsuz tesir etmektedir. Öğrenme parametrelerinin, sınıflandırma başarımını doğrudan etkilemesine karşın bunların, optimum değerinin bulunmasını sağlayacak bir yaklaşımın, orijinal DVM metodu içinde, yer almaması yöntemin diğer bir dezavantajıdır. Ayrıca dengeli olmayan veri kümelerinde sınıflandırma başarısının azalması ve orijinal DVM sınıflandırıcısının sadece +1, -1 çıkışlarını üretebilmesi yine yöntemin handikapları arasındadır³⁰⁹.

³⁰⁷ Emre Çomak, **a.g.e.** s.35.

³⁰⁸ A.I. Belousov, S.A. Verzakov, Ve J. Von Frese, **a.g.e.** s.24.

³⁰⁹ Emre Çomak, **a.g.e.** s.17.

ÜÇÜNCÜ BÖLÜM

KREDİ RİSKİNİN BELİRLENMESİNDE YAPAY ZEKÂ YAKLAŞIMLARI VE BİR UYGULAMA

3.1. Uygulamanın Amacı

Gelişen yapay zekâ modellerinin, finansal başarısızlık ve kredi riskinin belirlenmesi vb. daha öncesinde istatistiksel yöntemler kullanılarak analizi yapılan problemlere ne gibi bir katkısının olabileceği merak konusudur. İlk zamanlar YSA'nın vb. bazı problemleri çözememesi, ya da aşırı genelleme, yerel minimuma yakalanma benzeri handikaplarının olması neticesinde yaklaşık on altı (16) yıl boyunca (1969-1985) ilgi görmemiş, fakat daha sonra bu modellerde kaydedilen teorik gelişim sayesinde artık YSA ve DVM gibi yöntemler, şimdilerde hemen her disiplin ve sektörde yaygınlaşmış ve başarıyla kullanılmaktadır. Dolayısıyla kredi riskini belirleme gibi kadim problemlere, oldukça gelişmiş olan bu yöntemlerin katkısını ölçebilmek için, önde gelen iki sınıflandırma yöntemi olan YSA ve DVM, geleneksel yöntemlerle de kıyaslayabilmek adına LR yöntemi kullanılmıştır.

Klasik istatistik yöntemleri, özellikle konumuz açısından LDA ve LR, çok boyutlu problemlere çözüm arama konusunda eldeki ham veriyi değerleyecek daha iyi yöntemler yokken oldukça kullanışlı ve yaygındı. Bugün de hala birçok problemde etkin olarak kullanılabilmekte ve bazı problemlerde, özellikle LR, yapay zekâ yöntemlerinden daha iyi sonuçlar verebilmektedir. Fakat, geçtiğimiz birkaç on yılda, daha öncesinin aksine, yapay zekâ yöntemleri oldukça gelişti, çeşitlendi ve genelleme sınırı oluşturma gibi bir çok optimizasyon yöntemi sayesinde, daha da zenginleşti. Ayrıca LDA gibi istatistikî yöntemler doğrusal yöntemlerdir. Problemler için doğrusal modeller belirleyip daha sonraki verilerde kullanılmaktadır. Fakat gerçek hayat verileri çoğunlukla doğrusal olmamaktadır. Yapay zekâ yöntemleri, doğrusal olmayan modeller kurma yetenekleriyle, gerçek hayat verileri içerisinde yer alan örüntüleri daha iyi deşifre edebilmekte ve daha iyi sonuçlar verebilmektedirler. Tabii veri içeriğinin doğrusal olduğu nadir durumlarda doğrusal içerikli istatistikî yöntemler

daha başarılı olabilmektedir. O yüzden bu çalışmada, yapay zekâ yöntemlerinden YSA ve DVM'nin yanı sıra, istatistiki bir yöntem olan ve kategorik sınıflandırma yapan ve yine doğrusal olmayan LR yöntemi de kullanılmış ve sonuçları diğer yöntemlerle mukayese edilmiştir.

Probleme dönecek olursak, kredi yönetiminin etkinliği, özellikle Türkiye gibi, gelişmekte olan, sektör ya da politika kaynaklı olarak ekonomik dalgalanmaların çok yaşandığı piyasalarda, iş süreçlerinin sağlıklı yürütülebilmesi için, son derece önemlidir. Kredi yönetiminin etkinliği ise büyük ölçüde kredi riskinin doğru belirlenmesi ile alakalıdır. Bu yüzden kredi riskini doğru belirlemeye yarayan tüm yaklaşımlar, ana faaliyet konusu (yaklaşık üçte ikisi, $2/3^{310}$) kredi vermek olan ve ekonomilerin dayandığı temel aracı kuruluşlar olan ve finansman ve güvence kaynağı bankaların güvenebileceği önemli araçlardır.

Kredi risk derecelendirme konusunda yapılmış çok çeşitli çalışmalar literatürde bulunmaktadır. Biz bu çalışmada, bir Türk katılım bankasından alınan bireysel ve kurumsal kredi müşterilerinin verileriyle, yani gerçek hayat verileriyle, ilgili yöntemlerin performanslarını değerleyerek, bu yöntemlerin pratik hayatta finansal kurumlar tarafından kredi riski belirleme aşamasındaki kullanım imkânlarını araştırmaktır.

3.2. Uygulamada Kullanılan Yöntemler

Sınıflandırma problemleri için kullanılan doğrusal, doğrusal olmayan, metrik, metrik olmayan vs. niteliklere sahip yapay zekâ ve istatistik yöntemlerinin tamamını, karakteristik özellikleriyle, Lu ve Weng³¹¹ aşağıdaki şekilde sınıflandırmışlardır.

³¹⁰ Birgül Şakar, **a.g.e.** s.1-2.

³¹¹ D. Lu & Q. Weng, "A survey of image classification methods and techniques for improving classification performance", **International Journal of Remote Sensing**, Cilt: 28, No, 5, 2006, s.830.

Tablo 0.1: Sınıflandırma Problemlerinde Kullanılan Yöntemler Ve Özellikleri

Özelliği	Türü	Karakteristik Özellikleri	Sınıflandırıcı Örnekleri
Eğitim örneklerinin kullanılıp kullanılmamasına göre	Denetimli sınıflandırma yaklaşımı	Yeterli referans verisi mevcuttur ve eğitim örneği olarak kullanılır. Eğitim örneklerinden üretilen modeller daha sonra sınıflandırıcıyı spektral verileri tematik bir harita olarak sınıflandırmak üzere eğitmek için kullanılır.	_Maksimum olabilirlik, _Minimum mesafe, _YSA, _Karar ağacı
	Denetimsiz sınıflandırma yaklaşımı	Kümeleme tabanlı algoritmalar. Sınıfların önceki tanımları kullanılmaz. Analist, sınıfları anlamlı sınıflara etiketlemekten ve birleştirmekten sorumludur.	_ISODATA, _K-ortalamar kümeleme algoritması
Ortalama vektör ve kovaryans matrisi gibi parametrelerin kullanılıp kullanılmadığına göre	Metrik (parametrik)	Gauss dağılımı olduğu varsayılır. Parametreler (örn., Ortalama vektör ve kovaryans matrisi) genellikle eğitim örneklerinden üretilir. Model karmaşık olduğunda, parametrik sınıflandırıcılar genellikle "gürültülü" sonuçlar üretir. Bir diğer önemli dezavantaj, yan verileri, mekânsal ve bağlamsal özellikleri ve istatistiksel olmayan bilgileri bir sınıflandırma prosedürüne entegre etmenin zor olmasıdır.	_Maksimum olabilirlik (MO) _LDA
	Metrik olmayan (non-parametrik)	Veriler hakkında herhangi bir varsayım gerekli değildir. Parametrik olmayan sınıflandırıcılar, sınıf ayrımını hesaplamak için istatistiksel parametreler kullanmaz.	_YSA, _Karar ağacı, _Delil akıl yürütme, _DVM, _Uzman sistem.

Kaynak: D. Lu & Q. Weng, "A survey of image classification methods and techniques for improving classification performance", **International Journal of Remote Sensing**, Cilt: 28, No, 5, 2006, s.830.

Tabloda yer almayan ve sınıflandırma problemlerinde kullanılan, LR ve EKK-DVM gibi yöntemler de tabloya ilave edilebilir. Uygulama probleminde, sonuç verileri de bulunduğu için, bir sınıflandırma problemi olarak tasarlanmış ve denetimli öğrenme gerçekleştirileceği için, yöntembilim açısından uygun yapay zekâ yöntemleri olarak, YSA ve DVM, istatistiki olarak da LR kullanılmıştır.

3.2.1. Yapay Zekâ Yöntemleri

Kredi riskini belirleme problemini dikkate alacak olursak yapay zekâ yöntemlerinden YSA ve DVM, literatürde bu problemin çözümünde öne çıkan iki temel yöntemdir. Bu tez çalışmasında da bu iki yapay zekâ yöntemi kullanılmıştır. Ayrıca K-ortalamalar algoritması da söz konusu problemde denenmiş, fakat YSA ve DVM'ye göre doğruluk oranı iyi çıkmayınca, söz konusu yöntemle ilgili çıktılar uygulama sonuçlarına dâhil edilmemiştir.

3.2.1.1. Yapay Sinir Ağları

YSA, uygulama kolaylığı ve yüksek doğruluk oranları sayesinde, literatürde ve reel sektörde çok yaygın bir şekilde kullanılmakta ve hatta bankalar başta olmak üzere birçok işletme, YSA vasıtasıyla, yapay zekânın karar verme yeteneklerini kendi kurumsal sistemlerine entegre etmeye çalışmaktadırlar.

YSA, kredi riskinin belirlenmesi probleminde uzun yıllardır kullanılmaktadır. Hatta bunun için birçok banka, artık kendine özgü YSA mimarisini oluşturmuş ve patentini de almıştır. Bankacılık kredi faaliyetleri açısından, artık monotonlaşmış başvuru işlemlerini cevaplandırma süreçleri başta olmak üzere, kredi kararları, risklerin ve sahtekârlıkların tespiti gibi kritik karar süreçlerindeki kararların birçoğunu, YSA ve YSA'yı temel alan derin öğrenme (deep learning) yöntemleriyle bu süreçlerini robotize etmektedirler. Böylelikle birçok iş süreci için insan faktörünü aracı olmaktan çıkarmakta ve bu sayede bankalar, ilgili personelinin, iş geliştirme, ürün geliştirme, ARGE vb. uzman eksenli, daha efektif işlerde istihdam etme olanağı yakalamaktadırlar.

3.2.1.2. Destek Vektör Makineleri

DVM, son yıllarda, yüksek doğruluk oranları ve yer yer YSA'dan daha iyi sonuç vermeleri (Bkz. Bölüm 1.4), kernel fonksiyonlarını kullanmaları, matematiksel altyapısı ve çok boyutlu (çok fazla açıklayıcı değişkene sahip) fakat az gözlemlili veri kümelerinde başarılı olmaları gibi avantajlarıyla ön plana çıkmakta ve YSA gibi diğer

modellere alternatif teşkil edebilmektedirler. Sonuçta en karmaşık olan model değil, probleme özgü olan model daha başarılı olacaktır³¹². Modellerin başarı sıralamasının problemden probleme ve bir veri kümesinden diğer veri kümesine değişebileceği ihtimaliyle, bir problem söz konusu olduğunda uygun modellerin bir kıyasının yapılması gerekliliği gündeme gelmektedir. İkinci bir yapay zekâ modeli olan DVM'nin, ilgili problem için, gerçek hayat verileriyle, uygulama sonuçlarının bu çalışmaya dâhil edilmiş olması, çalışmanın hem daha güncel hem de daha kapsamlı olması açısından önemlidir.

3.2.2. İstatistiki Yöntemler

Kredi riski probleminin çözümünde, daha önceki yıllarda LDA de uzun yıllar kullanılmış fakat istatistik özelinde LR'nun ve yapay zekâ özelinde ise YSA ve DVM gibi yöntemlerin tartışmasız daha iyi sonuçlar vermesinden dolayı, son zamanlarda yapılan çalışmalarda artık kullanılmaz olmuştur (Bkz. Bölüm 1.4). MARS yöntemi gibi bazı alternatifler de bulunmaktadır. Fakat bu tezin esas amacı, yapay zekâ yöntemlerinin bu kadim problemin çözüm kalitesine getireceği marjinal faydanın tartışılması ve bunun kategorik sınıflandırmada en çok kullanılan, geleneksel, istatistiki bir yöntemle de tescil edilmesidir.

3.2.2.1. Lojistik Regresyon

Logit bir fonksiyon yardımıyla gözlemleri sınıflara taşıyan bu yüzden logit regresyon ismi de verilen LR, kategorik sınıflandırma problemlerinde oldukça başarılı sonuçlar veren bir yöntemdir. 1930'lu yıllardan itibaren Fisher'in öncülüğünde gelişen ve sınıflandırma problemleri için kullanılan LDA'nin yerini alan LR analizi LDA'nin varsayımlarını gerektirmemektedir. Ayrıca kolaylıkla uygulanabilmesi, yapay zekâ yöntemlerine kıyasla zamandan oldukça tasarruf sağlaması, kolay ve hızlı bir yöntem olması gibi nedenlerle sadece kredi risk değerlendirme problemlerinde değil birçok sınıflandırma probleminde başvurulmaktadır (Bkz. Bölüm 1.4). Dolayısıyla bu tez

³¹² David Corfield, Bernhard Schölkopf, Vladimir Vapnik, **a.g.e.** s.1.

çalışmasında, kredi riski belirleme konusunda öne çıkan istatistiksel yöntem olan LR kullanılmıştır.

3.3. Uygulamada Kullanılan Yazılımlar

Çalışmada verilerin alfa-nümerik hale getirilmesi, başlıklandırılması ve kategorize edilmesi aşamasında Excel'in 2016 sürümü kullanılmıştır. Yine tabloların hazırlanması ve grafiklerin çizilmesi aşamasında aynı Excel sürümü kullanılmıştır.

Analizlerin yapılması aşamasında ise Python 3.8.1 sürümü ve bunun için PyCharm Programı kullanılmıştır.

Formüllerin ve tezin yazımı aşamasında da Microsoft Word'ün yine 2016 sürümü kullanılmıştır.

Grafiklerin çizilmesi aşamasında www.desmos.com sitesinden yararlanılmıştır.

3.4. Uygulamada Kullanılan Veri Kümeleri

Veri toplama aşamasında, finansal kuruluşlara ait kredi müşterisi verisi gerektiği için, çeşitli finansal kuruluşlardan kredi verisi istenmiş ve bireysel ve kurumsal olmak üzere iki tür veri kümesi edinilmiştir.

3.4.1. Bireysel Veri Kümesi

Yapılan araştırma neticesinde toplam 34343 adet bireysel müşteri kredi verisine ulaşılmıştır. Fakat yapılan incelemede, bu müşterilerden yalnızca 150 tanesinin (yaklaşık %0,00436'sı) kusurlu müşteri verisi olduğu tespit edilmiştir. Dolayısıyla analiz aşamasında yakınsama (convergency) problemi yaşanmış, daha sonra veri riskli ve risksiz müşteri sayıları daha dengeli (62 kusurlu ve 345 kusursuz müşteri, %15,23) hale getirilerek yakınsama probleminin aşılabildiği, yeni bir "*Bireysel Veri Kümesi*" oluşturulmuştur.

3.4.2. Kurumsal Veri Kümesi

Yapılan arama neticesinde, toplam 12921 adet kurumsal müşteri verisine ulaşılmıştır. “*Kurumsal Veri Kümesi*” olarak isimlendirdiğimiz veri kümesinde müşterilerden 1590 tanesinin (yaklaşık % 12,3’ü) kusurlu müşteri verisi olduğu tespit edilmiştir. Kurumsal veri kümesi uygulamalarında, bireysel veri kümesinin aksine, yakınsama problemi yaşanmamıştır.

3.4.3. Dengeli Kurumsal Veri Kümesi

Kurumsal veride toplam 1590 adet kusurlu müşteri olduğu için; 1590 tanesi kusurlu, 1590 tanesi ise kusursuz müşterilerden oluşan üçüncü bir “*Dengeli Kurumsal Veri Kümesi*” oluşturularak, yine modellerin performansı bu kez kusurlu ve kursursuz müşteri verisinin eşit olduğu bu “dengeli” veri kümesine uygulanmıştır. Tahmin edileceği üzere, bu veri kümesiyle yapılan uygulamada yakınsama problemi yaşanmamıştır.

3.5. Uygulamada Kullanılan Değişkenler

Uygulamada, YSA ve DVM gibi (geleneksel yöntemlerin performansı ile da kıyaslayabilmek için LR kullanılmıştır) nispeten yeni yöntemlerinin, gerçek hayat verileri üzerindeki performansını kıyaslayabilmek için bankalardan, bireysel ve kurumsal olarak ayrı ayrı olmak kaydıyla, kredi müşterilerinin verileri istenmiştir. Edinilen müşteri verilerinde herhangi bir kısıtlamaya gidilmemiş yani bankanın, müşteriye dair olan bilgilerinin tamamı istenmiş ve aşağıda başlıklar halinde sıralanan bilgilere (değişkenlere) ulaşılabilmektedir.

3.5.1. Bireysel Veri Kümesi Değişkenleri

Araştırmada, elde edilen bireysel kredi müşterilerine ait, açıklayıcı değişken isimleri şu şekildedir;

1. Otomatik Reddetme Tahsisatı Var mı?
2. Bölümü Reddet

3. İptal edildi
4. Avukat Takibi
5. Son Tahsis Edilen Kredi Tarihi
6. Güncel Gecikmeli Gün Sayısı
7. Son Altı Ay Kredi Maksimum Gecikmeli Gün Sayısı
8. Son On iki Ay Kredi Maksimum Gecikmeli Gün Sayısı
9. Son Altı Ay Gecikmeli Taksit Sayısı
10. ***Son On iki Ay Gecikmeli Taksit Sayısı***
11. Teverruk var mı?
12. Hesap Açılış Tarihi
13. Gri Listede
14. Katılma Hesabı Bakiyesi
15. Son Altı Ay Toplam Verimlilik Geliri
16. Son On İki Ay Toplam Verimlilik Geliri
17. Son On iki Ay Ortalama Cari Bakiye
18. Son Üç Ay Ortalama Cari Bakiye
19. Son Altı Ay Ortalama Cari Bakiye
20. Son On iki Ay Ortalama Valor Bakiye
21. Geçmiş Aylar Ortalama Valor Bakiyesi
22. Son Üç Ay TCMB Protesto Çek Kontrol Kaydı
23. Son Altı Ay TCMB Protesto Çek Kontrol Kaydı
24. Son On iki Ay TCMB Protesto Çek Kontrol Kaydı
25. Son Ağaç Ayları TCMB Çekilmiş Çek Kontrol Kaydı
26. Son Altı Ay TCMB Çekilmiş Çek Kontrol Kaydı
27. Son On iki Ay TCMB Çekilmiş Çek Kontrol Kaydı
28. Hukuki Takipsizliği Var mı? Kapalı Kredi
29. Cari Risk Miktarı
30. Müşteri segmenti
31. Yer Tipi
32. Personel mi?
33. Maksimum Firma Yaşı
34. Meslek

35. Riskli Yaş mı?

3.5.1.1. Bireysel Veri Kümesi Bağımlı Değişkeni

Bireysel müşteri verisine ait olan yukarıdaki değişkenlerden, 10. sırada yer alan, “*Son On iki Ay Kredi Maksimum Gecikmeli Taksit Sayısı*” değişkeni, bağımlı değişken olarak atanmış ve “kusursuz”, “az kusurlu”, “orta seviyede kusurlu” ve “çok kusurlu” şeklinde müşteriler kategorize edilmiştir.

Kategorize edilen müşteri sınıflarının özellikleri, sınıf etiketleri ve her sınıfa düşen müşteri sayısı Tablo 3.2’de gösterilmektedir.

Bu bireysel veri kümesi çok sınırlı olan gözlem sayısı (407) ve çok boyutluluğu ile (34 adet açıklayıcı değişken kullanılmıştır) boyut laneti problemine maruz kalma olasılığı çok yüksek ve modelleri oldukça zorlayacak bir veri kümesidir. Çünkü Tablo 3.2’den de anlaşılacağı gibi, ikinci kategoriye düşen gözlem sayısı toplam gözlem sayısının sadece %1.96’sına denk gelmektedir. Böyle bir veri yapısı ise modellerin etkin çalışması için oldukça zorlayıcı, sınırlı bir veri kümesi anlamına gelmektedir.

Tablo 0.2: Bireysel Veri Sonuç Sınıfları Ve Sınıf Etiketleri

Son 12 Ay Kredi Maksimum Gecikmeli Taksit Sayısı	Sınıfı	Sınıf Adı	Müşteri Sayısı	Verideki Miktarı (%)
0	0	Kusursuz Müşteri	345	84,76
1	1	Az Kusurlu Müşteri	38	9,33
2-4	2	Orta Seviye Kusurlu Müşteri	8	1,96
5 ve üstü	3	Çok Kusurlu Müşteri	16	3,93
Toplam Bireysel Kusursuz Müşteri			345	84,76
Toplam Bireysel Kusurlu Müşteri			62	15,23
Toplam Bireysel Müşteri			407	100,00

3.5.1.2. Bireysel Veri Kümesi Açıklayıcı değişkenleri

Yukarıda listelenmiş olan, otuz beş (35) adet açıklayıcı değişkeni bulunan, bireysel müşteri verileri içinde onuncu (10.) sırada yer alan, “*Son On iki Ay Kredi Maksimum Gecikmeli Taksit Sayısı*” değişkeninin dışında yer alan 34 adet değişken ise modele açıklayıcı değişken olarak ilave edilmiştir.

Açıklayıcı değişkenlerde herhangi bir kategorilendirme yapılmamıştır. Fakat modelin stabilizasyonunun sağlanması için, veri kümesi, Min-Max Normalleştirilmesi kullanılarak, normalize edilmiştir.

3.5.2. Kurumsal Veri Kümesi Değişkenleri

Araştırmada edinilen kurumsal veri kümesinde yer alan değişkenler ise şu şekildedir:

1. TL ile Katılım Bakiyesi
2. Son On İki Ay TCMB Protesto Çek Kaydı
3. Son Altı Ay TCMB Protesto Çek Kaydı
4. Son Üç Ay TCMB Protesto Çek Kaydı
5. Son Üç Ay TCMB Çekilmiş Çek Kaydı
6. Son Altı Ay TCMB Çekilmiş Çek Kaydı
7. Son On İki Ay TCMB Çekilmiş Çek Kaydı
8. Son Dönem Toplamı
9. Son Üç Dönem Toplamı
10. Son Altı Dönem Toplamı Varsayılanı
11. Teverruk var mı?
12. Son Altı Ay Gecikmeli Taksit Sayısı
13. *Son On İki Ay Gecikmeli Taksit Sayısı*
14. Son Altı Ay Kredi Maksimum Gecikmeli Gün Sayısı
15. Son On İki Ay Kredi Maksimum Gecikmeli Gün Sayısı
16. Grup Aktif Sınırı
17. Geçerli Gecikme Günü Sayısı

3.5.2.1. Kurumsal Veri Kümesi Bağımlı Değişkeni

Kurumsal veri kümesinde, on üçüncü (13.) sırada yer alan; “*Son 12 Ayda Geciktirilen Taksit Sayısı (Last Twelve Months Delayed Installment Number)*” değişkeni bağımlı değişken olarak atanmıştır.

Oluşturulan müşteri sınıflarının özellikleri, sınıf etiketleri ve her sınıfa düşen müşteri sayısı Tablo 3.3’te görülmektedir.

Tablo 0.3: Kurumsal Veri Sonuç Sınıfları Ve Sınıf Etiketleri

Son 12 Ayda Geciktirilen Taksit Sayısı	Sınıfı	Sınıf Adı	Müşteri Sayısı	Verideki Miktarı (%)
0	0	Kusursuz Müşteri	11331	87,69
1	1	Az Kusurlu Müşteri	753	5,83
2-4	2	Orta Seviye Kusurlu Müşteri	272	2,11
5 ve üstü	3	Çok Kusurlu Müşteri	565	4,37
Toplam Kusursuz Kurumsal Müşteri			11331	87,69
Toplam Kusurlu Kurumsal Müşteri			1590	12,31
Toplam Kurumsal Müşteri			12921	100,00

Modellerin dengeli veri üzerindeki doğruluk oranını da test etmek için kurumsal veri kümesi ayrıca kusurlu ve kusursuz müşteri veri kümesinin eşit olduğu, dengeli, ikinci bir kurumsal veri kümesi oluşturulmuştur. Oluşturulan bu “Dengeli Kurumsal Veri Kümesi” nin kategorik dağılımı ise Tablo 3.4’teki gibidir.

Tablo 0.4: Dengeli Kurumsal Veri Sonuç Sınıfları Ve Sınıf Etiketleri

Son 12 Ayda Geciktirilen Taksit Sayısı	Sınıfı	Sınıf Adı	Müşteri Sayısı	Verideki Miktarı (%)
0	0	Kusursuz Müşteri	1590	50,00
1	1	Az Kusurlu Müşteri	753	23,68
2-4	2	Orta Seviye Kusurlu Müşteri	272	8,55
5 ve üstü	3	Çok Kusurlu Müşteri	565	17,77
Toplam Kusursuz Kurumsal Müşteri			1590	50,00
Toplam Kusurlu Kurumsal Müşteri			1590	50,00
Toplam Kurumsal Müşteri			3180	100,00

3.5.2.2. Kurumsal Veri Kümesi Açıklayıcı değişkenleri

Kurumsal veri kümesinde, bağımlı değişken olarak atanan, “*Son 12 Ayda Geciktirilen Taksit Sayısı (Last Twelve Months Delayed Installment Number)*” değişkeninin dışında yer alan diğer on altı (16) adet değişken ise açıklayıcı değişken olarak araştırmada kullanılan modellere tanıtılmıştır.

Bireysel veri kümesinde olduğu gibi, kurumsal veri kümesinde de açıklayıcı değişkenlerde herhangi bir kategorilendirme yapılmamış, yine modelin stabilizasyonu için, “normalizasyon” işlemi yapılmıştır. Normalizasyon işleminde, yine “Min-Max Normalleştirilmesi” (Bkz. 3.1) kullanılmıştır.

3.6. Normalizasyon

Açıklayıcı değişkenlerin değerini sıfır ila bir [0,1] arasındaki değerlere dönüştüren ve özellik ölçeklendirmesi (feature scaling) ismi de verilen, “normalizasyon” işlemi tüm verilere uygulanmıştır. Normalizasyon işleminde, aşağıda formülü de (3.1) yer alan “Min-Max Normalleştirilmesi” kullanılmıştır;

$$x' = \frac{x - x_{min}}{x_{maks} - x_{min}} \quad (0.1)$$

3.7. Uygulama Sonuçları

Uygulamada, YSA ve DVM gibi yöntemlerin, gerçek hayat verileri üzerindeki performansını kıyaslayabilmek için, bankalardan, bireysel ve kurumsal olmak kaydıyla, kredi müşterilerinin verileri istenmiştir. İstenen verilerde herhangi bir kısıtlamaya gidilmemiştir. Yani bankada, müşteriye dair açıklayıcı her ne özellik (açıklayıcı değişken) varsa hepsi istenmiş ve yukarıda sıralanan açıklayıcı değişkenlere ulaşılabilmiş ve elde edilen tüm değişkenler uygulamaya dâhil edilmiştir.

3.7.1. Bireysel Müşteriler İçin Uygulama Sonuçları

Bireysel veri kümesi üzerinde, bu tez çalışmasında kullanılan, YSA, DVM ve LR modelleri uygulanmıştır ve sonuçları ise Tablo 3.5'te görülmektedir.

Tablo 0.5: Bireysel Müşteriler Veri Kümesi Uygulama Sonuçları

Bireysel Veri Kümesi Uygulama Sonuçları				
Uygulamada Kullanılan Veri Kümeleri	Doğruluk Oranları ve Standart Sapma	Uygulamada Kullanılan Yöntemler		
		YSA	DVM	LR
Bireysel Veri Kümesi ile Yapılan Uygulama	Doğruluk Oranı (%)	79	85	84
	Standart Sapma (%)	8	2	3

Yapılan uygulamalar, örneklemin çok iyi ya da çok kötü seçilmiş olması ihtimaline karşın, 10 farklı rasgele örneklem seçimiyle yinelenmiş ve bunların doğruluk oranlarının ortalamalarına ve standart sapmalarına sonuç tablolarında yer verilmiştir. YSA modelinde iterasyon sayısı on bin (10000)'dir. Yapılan denemeler sonucu, doğruluk oranı en yüksek olan sonucu veren ve on beş (15) nörona sahip tek gizli katmanlı ve on bin (10000) iterasyona sahip YSA mimarisinin sonuçları yukarıdaki tabloya (Tablo 3.5) işlenmiştir. Bireysel müşteri verisi içerisinde kusurlu müşteri sayısı oldukça az olduğu için kusurlu ve kusursuz müşterileri de içeren, yakınsama probleminin yaşanmadığı, sınırlı (407 gözlem) bir veri kümesi oluşturulmuştur. Kurumsal müşteri verisinde ise hem tüm veriden oluşan kurumsal müşteri veri kümesi

ve bu kümenin riskli ve risksiz müşteri sayılarının eşitlendiği, dengeli kurumsal veri kümesi uygulamaları ayrı ayrı yapılmıştır.

3.7.2. Kurumsal Müşteriler İçin Uygulama Sonuçları

Kurumsal müşteri verisi, dengeli dağılmış (orantısız) bir veri değildir. Yani veri sınıflara dağıldığında yukarıdaki Tablo 3.3'den de görülebileceği sınıflara homojen bir şekilde, yani eşit sayılarda dağılmamaktadır. Bu durumun uygulamada meydana getirebileceği herhangi bir tutarsızlığın boyutlarını görebilmek adına analizde orijinal verinin yanı sıra eşit sayıda kusurlu ve kusursuz müşteriden oluşan dengeli veri kümesiyle de sonuçlar ayrıca incelenmiştir.

Tablo 0.6: Kurumsal Veri Kümesi Uygulama Sonuçları

Kurumsal Veri Kümesi Uygulama Sonuçları				
Uygulamada Kullanılan Veri Kümeleri	Doğruluk Oranları ve Standart Sapma	Uygulamada Kullanılan Yöntemler		
		YSA	DVM	LR
Kurumsal Veri Kümesi İle Yapılan Uygulama*	Doğruluk Oranı (%)	90	95	87
	Standart Sapma (%)	1	1	2
Dengeli Kurumsal Veri Kümesi İle Yapılan Uygulama**	Doğruluk Oranı (%)	68	96	57
	Standart Sapma (%)	8	3	4

*: 11331'i kusursuz, 1590'ı kusurlu olmak üzere, toplam 12921 gözlemde oluşan veri kümesiyle yapılan uygulama.

** : 1590'ı kusursuz, 1590 kusurlu olmak üzere, toplam 3180 gözlemde oluşan veri kümesiyle yapılan uygulama

Kullanılan modeller, örneklemin çok iyi ya da çok kötü seçilmiş olması ihtimaline karşın, uygulamalar 10 farklı rasgele örneklem seçimiyle yinelenmiş ve bunların doğruluk ortalamalarına ve standart sapmalarına yukarıdaki Tablo 3.6'da yer verilmiştir. Yapılan denemeler sonucu, en yüksek doğruluk oranını veren, beş (5) nörona sahip tek gizli katmanlı ve on bin (10000) iterasyona sahip YSA mimarisinin

sonuçları Tablo 3.6'ya işlenmiştir. Dengeli kurumsal veri kümesi üzerinde yapılan çeşitli YSA mimarilerinden elde edilen sonuçlar ise Tablo 3.8'de listelenmiştir.

Bireysel veri, kurumsal veri ve dengeli kurumsal veri üzerinde yapılan uygulamaların tamamı Tablo 3.7'de toplu olarak görülebilmektedir.

Tablo 0.7: Tüm Uygulamaların Sonuçları

Tüm Uygulamaların Sonuçları				
Uygulamada Kullanılan Veri Kümeleri	Doğruluk Oranları & Standart Sapma	Uygulamada Kullanılan Yöntemler		
		YSA	DVM	LR
Bireysel Veri Kümesi İle Yapılan Uygulama (407, %15,23)	Doğruluk Oranı (%)	79	85	84
	Standart Sapma (%)	8	2	3
Kurumsal Veri Kümesi İle Yapılan Uygulama (12921, %12,31)	Doğruluk Oranı (%)	90	95	87
	Standart Sapma (%)	1	1	2
Dengeli Kurumsal Veri Kümesi İle Yapılan Uygulama (3180, %50,00)	Doğruluk Oranı (%)	68	96	57
	Standart Sapma (%)	8	3	4

Tablo 0.8: Dengeli Kurumsal Veri Kümesi YSA Uygulamasında, Farklı Sayıda Nöron Ve Döğüm Sayısı Kullanıldığında, Ulaşılan Doğruluk Oranları Ve Standart Sapmaları (Maksimum İterasyon: 10000, cv=10, Veri: Dengeli Kurumsal Veri)

Deneme Sayısı	1. Gizli Katman Nöron Sayısı	2. Gizli Katman Nöron Sayısı	3. Gizli Katman Nöron Sayısı	Doğruluk Oranı (%)	Standart Sapma (%)
1	1	-	-	55	5
2	2	-	-	59	4
3	3	-	-	64	8
4	4	-	-	67	8
5	5	-	-	68	8
6	10	-	-	66	12
7	15	-	-	66	6
8	5	5	-	58	5
9	5	10	-	61	3
10	5	15	-	62	3
11	10	5	-	59	5
12	10	10	-	63	7
13	10	15	-	63	4
14	15	15	-	61	3
15	15	10	-	61	4
16	15	5	-	58	5
17	5	5	5	60	5
18	5	5	10	59	5
19	5	5	15	58	5
20	5	10	5	61	3
21	5	10	10	60	7
22	5	10	15	60	5
23	5	15	5	60	3
24	5	15	10	60	4
25	5	15	15	63	6
26	10	5	5	59	4
27	10	5	10	58	5
28	10	5	15	60	3
29	10	10	5	59	4
30	10	10	10	61	6
31	10	10	15	59	3

SONUÇLAR

Bu araştırma, bankaya bireysel ya da kurumsal kredi başvurusu yapan müşterilerin kredi riskinin belirlenmesine dönük olarak, YSA ve DVM özelinde bir risk değerlendirmesi yapabilecek modeller oluşturma ve oluşturulan bu modellerin birbirlerine kıyasla performanslarını değerlendirme amacıyla yapılmıştır.

Gerçek hayat verisi üzerinde yapılan uygulamalarda, bireysel veri kümesine bakıldığında, yalnızca 407 adet gözleme karşılık tam 34 farklı açıklayıcı değişken bulunmaktadır. Boyut sayısı çok fazla ve gözlem sayısı ise çok az olan bu veri kümesinde, boyut laneti (the curse of dimensionality) probleminin yaşanması kaçınılmazdır. Çünkü boyut sayısının artması, gözlem sayısının üstel olarak artması ile telafi edilebilmektedir. Sonuç olarak da bu boyut sıkıntısından en az etkilenen yöntemin DVM (%85) olduğu sonrasında LR (%84) ve son olarak da YSA (%79) olduğu görülmüştür.

DVM, yapılan bireysel veri uygulamalarında %85, kurumsal veri üzerinde %95, degelenmiş kurumsal veri üzerinde ise %96 doğruluk oranı ile, her üç veri seti üzerinde de, en yüksek doğruluk oranlarını veren yöntem olmuştur. Kurumsal veri kümesinden (12921 gözlem), kusurlu ve kusursuz müşteri sayısının eşitlendiği, dengeli kurumsal veri kümesine (3180 gözlem) geçildiğinde, eğitim veri kümesinin söz konusu çıktı sınıflarına daha düzgün dağılmış yani homojenize edilmiş bir yapıya kavuşması nedeniyle, doğal olarak daha yüksek doğruluk oranı beklenmekteydi. Fakat veri sayısındaki yaklaşık %75,39'luk azalış, DVM yöntemini olumsuz etkilemezken (doğruluk oranı %1 artmıştır) YSA'nın doğruluk oranını %22, LR'un doğruluk oranını ise %30 oranında düşürmüştür. Dolayısıyla DVM'nin, diğer yöntemlere göre, az sayıda gözlemle daha iyi sonuçlar veren yöntem olduğu tespitinin doğruluğu, bu tez çalışmasında da, görülmüştür. DVM uygulamasının bir başka sonucu ise, en az standart sapmaya sahip sonuçları vermiş olmasıdır. Bu da yöntemin yüksek güvenilirlikli ya da stabil bir yöntem olduğunun bir göstergesi olarak görülebilir.

Çalışmanın bir başka sonucu, gözlem sayısının az olduğu bireysel veri kümesinde, LR'nun YSA'dan daha başarılı sonuçlar vermiş olmasıdır. Fakat gözlem sayısının

olukça yeterli olduđu diğeri veri kümelerinde ise YSA'nın, LR'na bariz üstün olduđu görölmektedir. Kurumsal veri kümesinden dengeli kurumsal veri kümesine geçildiğinde her iki yöntemin de başarı oranı çok ciddi oranlarda düşüş göstermiştir. Fakat YSA yönteminde, hatırlanacak olursa, uygulayıcının optimal olarak karar vermesi gereken yaklaşık yedi parametresi bulunmakta ve bu parametrelerin optimal değerinin nasıl hesaplanacağı ise bilinmemektedir. Dolayısıyla YSA mimarisinin, eğer mümkünse, daha iyi bir şekilde dizayn edilmesi neticesinde, yöntemin performansında az da olsa artış sağlanabilir. YSA evet güvenilir, iyi sonuçlar veren, ama probleme özgü mimari geliştirildiğinde daha çok verim alınabilen bir yöntem olduđu için, bankaların zaman için de, YSA uygulamalarına dayalı olarak, kendilerine ve kredi portföylerine uygun bir mimari oluşturmaları ve kullanmaları çok yüksek başarı oranı sağlayacaktır. Ayrıca YSA'nın, veri sayısı yüksek olduğunda iyi sonuçlar veren bir yöntem olduđu görüşü, bireysel veri kümesi uygulamasında LR'dan geri kalması ve yeterli veri sayısı olan diğeri veri kümesine uygulandığında LR'dan daha iyi sonuç vermiş olmasından dolayı bu tez çalışmasında da doğrulanmıştır.

Edinilen verilerde, bireysel müşteriler için, findeks puanı gibi bireylerin fatura ve kredi ödeme alışkanlıklarını göz önünde bulundurularak onların kredi puanını hesaplayan bir değişken bu çalışmada yoktur. Findeks gibi açıklayıcı bir değişkenin ya da kurumsal veya bireysel müşterilere ait herhangi bir kredi notunun (Moody's, Sandart and Poors vs.) bulunmaması yine modellerin etkinliğini yine olumsuz etkileyen bir durumdur. Yine edinilen veri kümelerinde kurumsal müşteriler için, işletme büyüklüğü, kuruluş sermayesi, şirket türü, kredi notu, çalışan sayısı, borsaya kote olup olmadığı, cirosu, sektörü, findeks puanı ve diğeri bilanço verileri gibi önemli verilere ulaşamamıştır. Sonuç olarak, bu çok önemli bilgilerin yer almamasına rağmen, yüksek doğruluk oranları elde edilebilmiştir. Kredi borcunu geciktirme bağımlı değişkeni üzerinde etkili olabilecek bu çok önemli bilgi kalemlerine rahatlıkla ulaşabilecek bankaların ise, bu modellerin yardımıyla, çok yüksek doğruluk oranlarıyla, kredi müşterilerini doğru sınıflandırabilmeleri oldukça mümkün görünmektedir.

İçsel değerlendirme açısından çok önemli olan bu yapay zekâ yöntemleri, maliyetlerin sürekli düşürülmeye, niteliği zayıf olan ve bilgisayarlarla hızla sonuçlandırılmaya çalışan tipik kredi uygulamalarını titizlikle yapmak zorunda olan günümüz bankaları için önemli bir işlevi yerine getirmektedir. Bununla birlikte karlılığını ve rekabetçiliğini korumaya çalışan bu yüzden işlerini sürekli hatasız yapmaya mecbur olan bankalar için, yapay zekâ uygulamaları rekabetçiliği ve sürekliliği sağlamanın vazgeçilmez bir argümanı olarak görmek mümkündür. Çünkü daha öncesinde günler hatta haftalar süren kredi başvurularını sonuçlandırma işlemi, artık yapay zekâ yöntemleriyle süreçlerini robotize eden yani insan faktörünü neredeyse tamamen aradan çıkaran bankarda, saniyeler içerisinde sonuçlandırılabilir. Bu aynı zamanda ilgili personelin ve diğer bankacılık kaynaklarının daha verimli, daha çok gelecek ve kar vaadeden işlere kanalize edilmesinin de bir yolu olarak görülebilir.

KAYNAKÇA

- Alp Selçuk,
Ersoy Öz : **Makine Öğrenmesinde Sınıflandırma Yöntemleri Ve R Uygulamaları**, Nobel Kitabevi, Ankara, 2019.
- Alp, Selçuk,
Savaş Özdemir,
Arzu Kilitci : **Veri Tabanı Yönetim Sistemleri**, İstanbul: Türkmen Kitabevi, 2011.
- Alpaydın, Ethem : **Introduction To Machine Learning**, 2. Baskı, Londra: The MIT Press, 2010.
- Altan, Mikail : **“Fonksiyonlar ve İşlemler Açısından Bankacılık**, Beta Basım Yayın, 2001.
- Altay, Erdinç : **Bankacılıkta Risk, Piyasa Riski Kredi Riski Ve Operasyonel Riskin Ölçümü Ve Yönetimi**, İstanbul: Derin Yayınları, 2015.
- Altman, Edward : “Financial Ratios, Discriminant Analysis And The Prediction Of Corporate Bankruptcy”, **Journal Of Finance**, Cilt: 13, 1968, s.589-609.
- Anderson, Dave,
George McNeill, : “Artificial Neural Networks Technology”, A DACS State-of-the-Art Report, 1992.

- Anderson, Dave,
George McNeill : “Artificial Neural Networks Technology”, A
DACs State-of-the-Art Report, Contract Number
F30602-89-C-0082, 1992.
- Angelini, Eliana,
Giacomo Di Tollo,
Andrea Roli : “A Neural Network Approach For Credit Risk
Evaluation”, **The Quarterly Review of
Economics and Finance**, Cilt: 48, 2008, s.733-
755.
- Arminger, G.,
D. Enache,
T. Bonne : “Analyzing Credit Risk Data: A Comparison Of
Logistic Discrimination Classification Tree
Analysis And Feedforward Networks”,
Computational Statistics, Cilt: 12, 1997, s.293–
310.
- Arslan, İbrahim : “Basel Kriterleri Ve Türk Bankacılık Sektörüne
Etkileri”, **Selçuk Üniversitesi Sosyal Bilimler
Dergisi**, 2007, s.49-66.
- Ataseven, Burçin : “Yapay Sinir Ağları İle Öngörü Modellemesi”,
Öneri, Cilt: 10, Sayı:39, 2013, s.101-115.
- Atiya, Amir F. : “Bankruptcy Prediction for Credit Risk Using
Neural Networks: A Survey and New Results”,
IEEE Transactions On Neural Networks,
Cilt:12, Sayı: 4, 2001, s.929-935.

- Aven, T. : “On risk defined as an event where the outcome is uncertain”, **Journal of Risk Research**, Cilt: 12, O. Renn Sayı: 1, 2009, s.1-11.
- Baesens, Bart, : “Benchmarking State-Of-The-Art Classification Algorithms For Credit Scoring”, **Journal of the Operational Research Society**, Cilt: 54, 2003, T. Van Gestel, s.627–635. S.Viaene, M. Stepanova, J. Suykens, J. Vanthienen
- Baesens, Bart, : “Using Neural Network Rule Extraction And Rudy Setiono, Decision Tables For Credit-Risk Evaluation”, Christophe Mues, **Management Science**, Cilt: 49, Sayı: 3, 2003, Jan Vanthienen s.312-329.
- Bank of International : About BIS Overview, Çevrimiçi, Settlements (BIS) <https://www.bis.org/about/index.htm>, 28/10/2019.
- Bank of International : “Principles for the Management of Credit Risk”, Settlements (BIS) 2000, Çevrimiçi, <http://www.bis.org/publ/bcbs75.pdf>, 29/06/2011, s.4.
- Bankacılık Düzenleme Ve : “10 Soruda Yeni Basel Sermaye Uzlaşısı (Basel-II)”, 2005, s.2, Çevrimiçi, Denetleme Kurumu https://www.bddk.org.tr/ContentBddk/dokuman/duyuru_basel_0001_38.pdf, 07.09.2020

- Bankacılık Düzenleme Ve Denetleme Kurumu : “Sorularla Basel III”, Risk Yönetimi Dairesi, Aralık, 2010, s.3-4.
- Bardos, Mireille : “Detecting The Risk Of Company Failure At The Banque De France”, **Journal Of Banking & Finance**, Cilt: 22, Sayı: 10–11, 1998, s.1405-1419.
- Barney, Douglas K., O. Finley Graves, John D. Johnson : “The farmers home administration and farm debt failure prediction”, **Journal of Accounting and Public Policy**, Cilt: 18, 1999, s.99-139.
- Basheer, I.A., M. Hajmeer : “Artificial Neural Networks: Fundamentals, Computing, Design, And Application”, **Journal of Microbiological Methods**, Cilt: 43, 2000, s.3-31.
- Basu, Jayanta Kumar, Debnath Bhattacharyya, Tai-hoon Kim : “Use of Artificial Neural Network in Pattern Recognition”, **International Journal of Software Engineering and Its Applications**, Cilt: 4, Sayı: 2, 2010, s.23-34.
- Belousov, A.I., S.A. Verzhakov, J. von Frese : “A Flexible Classification Approach With Optimal Generalisation Performance: Support Vector Machines”, **Chemometrics And Intelligent Laboratory Systems**, Cilt: 64, 2002, s.15-25.

- Boser, Bernhard E.,
Isabelle M. Guyon,
Vladimir Vapnik,
: “A Training Algorithm For Optimal Margin Classifiers”, Çevrimiçi, <file:///C:/Users/gokhanno/Desktop/Downloads/download.pdf>, 26/04/2020.
- Burges, Christopher J.C.
: “A Tutorial On Support Vector Machines For Pattern Recognition”, **Data Mining and Knowledge Discovery**, Cilt: 2, Sayı:2, 1998, s.121-167.
- Candan Hasan,
Alper Özün
: **Bankalarda Risk Yönetimi Ve Basel II**, İstanbul: Türkiye İş Bankası Kültür Yayınları, 2006.
- Chen, Mu-Chen,
Shih-Hsien Huang
: “Credit Scoring And Rejected Instances Reassigning Through Evolutionary Computation Techniques”, **Expert Systems with Applications** Cilt: 24, 2003, s.433–441.
- Cheng, Bing,
D. M. Titterington
: “Neural Networks: A Review from a Statistical Perspective”, **Statistical Science**, 1994, Cilt: 9, Sayı: 1, s.2-54.
- Cherkassky, Vladimir,
Xuhui Shao,
Filip Mulier,
Vladimir Vapnik
: “Model Complexity Control for Regression Using VC Generalization Bounds”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.1075-1089.

Cherkassky, Vladimir,
Filip Mulier

: **Lerning From Data: Consepts, Theory and Methods**, 2. Baskı, New Jersey, John Wiley & Sons, Inc., 2007.

Chuang, Chun-Ling,
Rong-Ho Lin,

: “Constructing a reassigning credit scoring model”, **Expert Systems with Applications**, Cilt: 36, Sayı: 2, 2009, s.1685-1694.

Coffman, J. Y.,

: “The Proper Role Of Tree Analysis In Forecasting The Risk Behaviour Of Borrowers, **Management Decision Systems**, Atlanta, MDS Reports 3,4,7 ve 9, 1986.

Cohn, David,
Les Atlas,
Richard Ladner

: “Improving Generalization with Active Learning”, **Machine Learning**, Cilt: 15, 1994, s.201-221.

Corfield, David,
Bernhard Schölkopf,
Vladimir Vapnik

: “Popper, Falsification And The VC-Dimension”, **Max Planck Institute for Biological Cybernetics**, Teknik Rapor, Sayı: 145, 2005, s.1-4.

Corfield, David,
Bernhard Schölkopf,
Vladimir Vapnik

: “Falsificationism and Statistical Learning Theory: Comparing the Popper and Vapnik-Chervonenkis Dimensions”, **J Gen Philos Sci**, Cilt: 40, 2009, s.51–58.

Cura, Tunçhan : **Modern Sezgisel Teknikler ve Uygulamaları**, 1. Basım, İstanbul: Papatya Yayıncılık, 2008.

Cura, Tunçhan : “Karar Verme Aracı Olarak Yapay Sinir Ağları Ve Yapay Sinir Ağları İle Portföy Optimizasyonu”, Doktora Tezi, İstanbul Üniversitesi, 2004.

Çilingirtürk, Ahmet Mete : **İstatistiksel Karar Almada Veri Analizi**, Birinci Baskı, Ankara: Seçkin Yayınevi, 2011.

Çomak, Emre : “Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar”, Doktora Tezi, Selçuk Üniversitesi, 2008.

Çuhadar, Hakan : “Müzik Ve Beyin”, **Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi**, Cilt: 17, Sayı: 2, 2008, s.67-76.

Davis, R. H.,
D. B. Edelman,
A. J. Gammerman, : “Machine Learning Algorithms For Credit-Card Applications”, **Journal Of Mathematics Applied In Business And Industry**, Cilt: 4, 1992, s.43-51.

Dayhoff, Judith E.,
James M. DeLeo, : “Artificial Neural Networks Opening the Black Box”, **CANCER Supplement**, 15, 2001, Cilt: 91, Sayı: 8, s.1615-1635.

De Gooijer, Jan G.,
K. Ray Bonnie,
Horst Krager : “Forecasting exchange rates using TSMARS”, **Journal of International Money and Finance**, Cilt: 17, 1998, s.513-534.

Demirkol, Ömer Faruk : “Basel II Sermaye Yeterliliği Uzlaşısı Süreci
Emel ABA İçinde Gelen Ek Düzenlemeler Seti: Basel III
Kriterleri”, **Niğde Üniversitesi İİBF Dergisi**,
2012,Cilt: 5, Sayı: 2, s.253-264.

Desai, V. S., : “A Comparison Of Neural Networks And Linear
J. N. Crook, Scoring Models In The Credit Union
Jr. G. A.Overstreet Environment” **European Journal Of
Operational Research**, Cilt: 95, 1996, s.24–37.

Desai, Vijay S., : “A Comparison Of Neural Networks And Linear
N. Crook Jonathan Scoring Models In The Credit Union
A. George Overstreet Environment”, **European Journal Of
Operational Research**, Cilt: 95, 1996, s.24-37.

Duda, Richard O., : **Pattern Classification**, 2. Baskı, Kanada: John
Peter E. Hart, Wiley & Sons, INC, 2000.
David G.Stork

Elmas, Çetin : **Yapay Sinir Ağları**, 3. Baskı, Ankara: Seçkin
Yayıncılık, 2016.

Emir, Şenol : “Yapay Sinir Ağları ve Destek Vektör Makineleri
Yöntemlerinin Sınıflandırma Performanslarının
Karşılaştırılması: Borsa Endeks Yönünün Tahmini
Üzerine Bir Uygulama”, Doktora Tezi, İstanbul
Üniversitesi, 2013.

Eroğlu, Ergün,
Timur Keskintürk

: “Warehouse Location Problem With Genetic Algorithm”, **35th 656 International Conference On Computers And Industrial Engineering**, 2005, s.655-660.

Ersoy, Esen,
Ömer Karal

: “Yapay Sinir Ağları Ve İnsan Beyni”, İnsan ve Toplum Bilimleri Araştırmaları Dergisi, Cilt:1, Sayı:2, 2012.

Findeks

: Findeks Kredi Notu, Çevrimiçi, <https://www.findeks.com/urunler/findeks-kredi-notu>, 28/10/2019.

Fogarty, Terence C.,
Neil S.Ireson,

: “Evolving Bayesian Classifiers For Credit Control—A Comparison With Other Machine-Learning Methods”, **IMA Journal Of Mathematics Applied In Business & Industry**, Cilt: 5, Sayı: 4, 1993, s.63-75.

Friedman, Jerome H,
C.B. Roosen

: “An Introduction To Multivariate Adaptive Regression Splines”, **Statistical Methods in Medical Research**, Cilt: 4, 1995, s.197-217.

Friedman, Jerome H.

: “Multivariate Adaptive Regression Splines”, **Annals Of Statistics**, Cilt: 19, Sayı: 1, 1991, s.1-141.

Glorfeld, L. W.,
B. C. Hardgrave

: “An Improved Method For Developing Neural Networks: The Case Of Evaluating Commercial Loan Creditworthiness”, **Computers & Operations Research**, Cilt: 23, Sayı: 10, 1996, s.933–944.

Glorfeld, Louis W.

: “A Methodology for Simplification and Interpretation of Backpropagation-Based Neural Network Models”, **Expert Systems With Applications**, Cilt: 10, Sayı: 1, 1996, s.37-54.

Goodfellow, Ian,
Yoshua Bengio,
Aaron Courville,

: **Deep Learning**, 1. Baskı, Cambridge: The MIT Press, 2017.

Griffin, W.L.,
N.I. Fisher,
J.H. Friedman,
C.G. Ryan

: “Statistical Techniques For The Classification Of Chromites In Diamond Exploration Samples”, **Journal of Geochemical Exploration**, Cilt: 59, 1997, s.233-249.

Gülpınar, Vildan

: “Yapay Sinir Ağları İşleyişinin Sosyal Ağ Analizi Yardımı İle Çözümlemesi”, Doktora Tezi, Marmara Üniversitesi, 2013.

Gürel, Eymen,
Esra Burcu Bulgurcu Gürel

: “Basel III Kriterleri”, **BSAD Bankacılık ve Sigortacılık Araştırmaları Dergisi**, Cilt: 1, Sayı:3-4, 2012, s.16-28.

- Gürsoy, U. T. Şimşek : **Veri Madenciliği Ve Bilgi Keşfi**, 1. Baskı, Ankara: Pegem Yayınevi, 2009.
- Haltuf, Michal : “Support Vector Machines for Credit Scoring”, Yüksek Lisans Tezi, University of Economics in Prague Faculty of Finance Department of Banking and Insurance, Financial Engineering, 2014.
- Hamzaçebi, Coşkun : **Yapay Sinir Ağları Tahmin Amaçlı Kullanımı, MATLAB Ve NEUROSOLUTION Uygulamalı**, 1. Baskı, Bursa: Ekin Yayınevi, 2011.
- Hamzaçebi, Coşkun : “Geleceği Tahminde Yapay Sinir Ağları İçin Sezgisel Öğrenme Algoritması”, Doktora Tezi, Gazi Üniversitesi, 2005.
- Harrell, Frank, Jr. Kerry L. Lee : “A Comparison Of The Discrimination Of Discriminant Analysis And Logistic Regresyon Under Multivarite Normality”, **S.K. Se (Ed.), Biostatistics: Statistics in biomedical, public health, and environmental sciences**, Cilt: 1985, s.333-343.
- Henley, W. E., D. J. Hand : “A K-Nearest Neighbor Classifier For Assessing Consumer Credit Risk”, **Statistician**, Cilt: 44, Sayı: 1, 1996, s.77–95.

Hopfield, J. J. : “Neural Networks And Physical Systems With Emergent Collective Computational Abilities”, **Proc. NatL Acad. Sci. USA**, Cilt: 79, 1982, s.2554-2558.

Hsieh, Nan Chen : “Hybrid Mining Approach In The Design Of Credit Scoring Models”. **Expert Systems with Applications**, Cilt: 28, Sayı: 4, 2005, s.655–665.

Huang, Cheng-Lung,
Mu-Chen Chen : “Credit Scoring And Rejected Instances Reassigning Through Evolutionary Computation Techniques”, **Expert Systems With Applications**, Cilt: 24, Sayı: 4, 2003, s.433–441.

Huang, Cheng-Lung,
Mu-Chen Chen,
Chieh-Jen Wang : “Credit Scoring With A Data Mining Approach Based On Support Vector Machines”, **Expert Systems with Applications**, Cilt: 33, 2007, s.847–856.

Huang, Zan,
Hsinchun Chen,
Chia-Jung Hsu,
Wun-Hwa Chen,
Soushan Wu : “Credit Rating Analysis With Support Vector Machines And Neural Networks: A Market Comparative Study”, **Decision Support Systems**, Cilt: 37, 2004, s.543– 558.

Jayalakshmi T.,
Dr.A. Santhakumaran : “Statistical Normalization And Back Propagation For Classification”, **International Journal Of Computer Theory And Engineering**, Cilt: 3, Sayı: 1, 2011, s.89-93.

- Karaboğa, Derviş : **Yapay Zekâ Optimizasyon Algoritmaları**, 4. Baskı, İstanbul: Nobel Akademik Yayıncılık, 2014.
- Kartal, Mustafa Tevfik : **Bankaların Finans Sektöründeki Önemi**, Finansal İktisat, Orion Yayıncılık, 2018.
- Kecman, Vojislav : “Support Vector Machines – An Introduction”, **StudFuzz**, Cilt: 177, 200, s.1-47.
- Kecman, Vojislav : “Support Vector Machines – An Introduction”, **StudFuzz**, Cilt: 177, 2005.
- Kelleher, John D.,
Brian Mac Namee,
Aoife D’arcey : **Fundamentals Of Machine Learning For Predictive Data Analytics, Algorithms, Worked Examples, And Case Studies**, 1. Baskı, USA: MIT Press, 2015.
- Kocamaz, Kaan : “Hastane Bilgi Yönetim Sistemlerinde Veri Madenciliği Ve Konya Meram Tıp Fakültesi Hastanesindeki Hastane Bilgi Yönetim Sistemi Uygulamasının İncelenmesi”, Yüksek Lisans Tezi, Selçuk Üniversitesi, 2007.
- Kuhnert, Petra M.,
Kim-Anh Do,
Rod McClure : “Combining Non-Parametric Models With Logistic Regression: An Application To Motor Vehicle Injury Data”, **Computational Statistics & Data Analysis**, Cilt: 34, 2000, s.371-386.

Kuriscak, Eduard,
Petr Marsalek,
Julius Stroffek,
Peter G.Toth
: “Biological context of Hebb learning in artificial neural networks, a review”, **Neurocomputing**, Cilt: 152, 2015, s.27-35.

Laitinen, Erkki K.,
Teija Laitinen
: “Bankruptcy prediction Application of the Taylor's expansion in logistic regression”, **International Review of Financial Analysis**, Cilt: 9, 2000, s.327 – 349.

Lee, Tian Shyug,
I-Fei Chen
: “A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks And Multivariate Adaptive Regression Splines”, **Expert Systems With Applications: An International Journal**, Cilt: 28, Sayı: 4, 2005, s.589-609.

Lee, Tian-Shyug
Chih-Chou Chiu,
Chi-Jie Lu,
I-Fei Chen,
: “Credit Scoring Using The Hybrid Neural Discriminant Technique”, **Expert Systems With Applications**, Cilt: 23, 2002, s.245–254.

Lee, Tian-Shyug,
Chih-Chou Chiu,
Yu-Chao Chou,
Chi-JieLu
: “Mining The Customer Credit Using Classification And Regression Tree And Multivariate Adaptive Regression Splines”, **Computational Statistics & Data Analysis**, Cilt: 50, Sayı: 4, 2006, s. 1113-1130.

Lee, Tian-Shyug,
I-Fei Chen : “A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks And Multivariate Adaptive Regression Splines”, **Expert Systems With Applications**, Cilt: 28, 2005.

Leng, S.,
M. Hutter : “A Collections Of Definitions Of Intelligence”, **Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms**.

Lisman, John : “A Mechanism For The Hebb And The Anti-Hebb Processes Underlying Learning And Memory”, **Proc. Natl. Acad. Sci. USA**, Cilt: 86, 1989, s.9574-9578.

López Muñoz Francisco,
Jesús Boya,
Cecilio Alamo : “Neuron Theory, The Cornerstone Of Neuroscience, On The Centenary Of The Nobel Prize Award To Santiago Ramón Y Cajal”, **Brain Research Bulletin**, Cilt: 70, Sayı: 4–6, 2006, s.391-405.

López -Muñoz, Francisco,
Boya Jesús ,
Alamo, Cecilio : “Neuron theory, the cornerstone of neuroscience, on the centenary of the Nobel Prize award to Santiago Ramón y Cajal”, **Brain Research Bulletin**, Cilt: 70, Sayı 4–, 2006, s.391.

Lu, D.,
Q. Weng : “A survey of image classification methods and techniques for improving classification

performance”, **International Journal of Remote Sensing**, Cilt: 28, Sayı: 5, 2006, s.823-870.

Lu, Huimin,
Yujie Li,
Min Chen,
Hyoungseop Kim,
Seiichi Serikawa,

: “Brain Intelligence: Go beyond Artificial Intelligence”, **Mobile Netw Appl**, Cilt: 23, 2018, s.368–375.

Makowski, S.

: “Credit Scoring Branches Out”, **The Credit World**, Cilt: 75, 1985, s.30–37.

Malhotra, R.,
D. K., Malhotra

: “Differentiating Between Good Credits And Bad Credits Using Neuro-Fuzzy Systems”, **European Journal Of Operational Research**, Cilt: 136, Sayı: 1, 2002, s.190-211.

Martell, T.F.,
R.L. Fitts

: “A Quadratic Discriminant Analysis Of Bank Credit Card User Characteristics”, **Journal Of Economics And Business**, Cilt: 33, 1981, s.153-159.

Mevzuat Dergisi

: “Kredi Sözleşmelerinin Hukuki Niteliği Ve Türleri”, Çevrimiçi,
<https://www.mevzuatdergisi.com/2002/01a/01.htm>, 11/03/2020.

- Mevzuat Dergisi : “5411 Sayılı Bankacılık Kanunu”, Çevrimiçi, <https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5411.pdf>, 05.09.2020.
- Muzır, Erol : “Basel II Düzenlemeleri Doğrultusunda Kredi Riski Analizi Ve Ölçümü: Geleneksel Ekonometrik Modellerin Yapay Sinir Ağı Ve Mars Modelleriyle Karşılaştırılmasına Yönelik Ampirik Bir Çalışma”, Doktora Tezi, İstanbul Üniversitesi, 2011.
- Nabiyev, Vasif V. : **Yapay Zekâ**, 5. Baskı, Ankara: Seçkin Yayınevi, 2016.
- Nguyen-Cong, V,
G Van Dang,
BM Rode : “Using Multivariate Adaptive Regression Splines To QSAR Studies Of Dihydroartemisinin Derivatives”, **Eur J Med Chem**, Cilt: 3, Sayı: 1, 1996, s.797-803.
- Nilson, Nils J. : **Yapay Zekâ Geçmişi ve Geleceği**, 1. Baskı, İstanbul: Boğaziçi Üniversitesi Yayınevi, 2019.
- Ohmann, C.,
V. Moustakis,
Q. Yang,
K. Lang : “Evaluation Of Automatic Knowledge Acquisition Techniques In The Diagnosis Of Acute Abdominal Pain”, **Artificial Intelligence in Medicine**, Cilt: 8, 1996, s.23-36.

Ong, Chorng-Shyong,
Jih-Jeng Huang,
Gwo-Hshiung Tzeng

: “Building Credit Scoring Models Using Genetic Programming”, **Expert Systems with Applications**, Cilt: 29, 2005, s.41-47.

Orhunbilge, Neyran

: **Uygulamalı Regresyon Ve Korelasyon Analizi**, 2. Baskı, İstanbul: İşletme Fakültesi Yayınları, 2002.

Orhunbilge, Neyran

: **Uygulamalı Regresyon Ve Korelasyon Analizi**, 2. Baskı, İstanbul: İşletme Fakültesi Yayınları, 2002.

Overstreet, G.A.,
Jr. E.L. Bradley,
Jr. R.S. Kemp

: “The Flat-Maximum Effect And Generic Linear Scoring Model: A Test”, **IMA Journal Of Mathematics Applied In Business And Industry**, Cilt: 4, 1992, s.97-109.

Özekes, Mustafa

: “Peabody Resim Kelime Testi 3.01-3.12 Yaş Aralığı İzmir Bölgesi Standardizasyonu Çalışması”, **Ege Eğitim Dergisi**, 2013, Cilt: 14, No: 1, s.90-107.

Öztemel, Ercan

: **Yapay Sinir Ağları**, İstanbul, Papatya Yayıncılık, 2012.

Öztürk, Kezban

: **Kredi Politikası Ve Değerlendirmesi**, Ankara: Siyasal Kitabevi, 2015.

- Pang S.,
Y. Wang,
Y. Bai
: “Credit Scoring Model Based On Neural Network”, **In Proceedings Of The First International Conference On Machine Learning And Cybernetics**, 2002.
- Piramuthu, S.,
M. J. Shaw,
J. A. Gentry
: “A Classification Approach Using Multi-Layered Neural Networks”, **Decision Support Systems**, Cilt: 11, 1994, s.509–525.
- Piramuthu, Selwyn
: “Financial credit-risk evaluation with neural and neurofuzzy systems”, **European Journal of Operational Research**, Cilt: 112, 1999, s.310-321.
- Pirim, Harun
: “Yapay Zekâ”, **Journal Of Yasar University**, Cilt: 1, Sayı: 1, s. 81-93.
- Pohar, Maja,
Mateja Blas,
Sandra Turk
: “Comparison of Logistic Regression and Linear Discriminant Analysis: A Simulation Study”, **Metodološki zvezki**, Cilt: 1, 2004, s.143-161.
- Quinlan, J. R.
: “**C4.5: Programs For Machine Learning Discriminanting**”, Morgan Kaufman, San Mateo, CA, 1993.
- Rong-Zhou, L.,
S.Su-Lin,
X. Jian-Min
: “Neural network credit-risk evaluation model based on back-propagation algorithm”, **In Proceedings of the first international**

conference on machine learning and cybernetics, 2002, s.1702-1706.

Rosenblatt, Frank : “The Perceptron: A Porobabilistic Model For Information Storage And Organization In The Brain”, **Psychonanalytic Review**, Cilt: 65, s.386-408.

Safavian, S.R.,
D. A. Landgrebe : “Survey Of Decision Tree Classifier Methodology”. **IEEE Transactions On Systems, Man And Cybernetics**, Cilt: 21, 1991, s.660-674.

Shen, J. : “Fusing Support Vector Machines and Soft Computing for Pattern Recognition and Regression”, Kansas State University Manhattan, Kansas, Doktora Tezi, 2005.

Shyug Lee, Tian,
I-Fei Chen, : “A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks And Multivariate Adaptive Regression Splines”, **Expert Systems With Applications: An International Journal**, Cilt: 28, Sayı: 4, 2005, s.743-752.

Smola, Alex,
S.V.N. Vishwanathan : **Introduction To Machine Learning**, Cambridge: Cambridge University Press, First Edition, 2008.

Stoean, C.,
R. Stoean : **Support Vector Machines and Evolutionary Algorithms for Classification**, Switzerland: Springer International Publishing, 2014.

Storkey, A.J.,
R.Valabregue : “The Basins Of Attraction Of A New Hopfield Learning Rule”, **Neural Networks**, Cilt: 12, 1999, s.869-876.

Şakar, Birgül : **Banka Kredileri Ve Yönetimi**, İstanbul: Beta Yayınları, 6. Baskı, 2015.

Theodoridis, Sergios,
Konstantinos Koutroumbas : **Pattern Recognition**, 4. Baskı, San Diego: Elsevier Inc., 2009.

Thomas, Lyn. C. : “A Survey Of Credit And Behavioural Scoring: Forecasting Financial Risk Of Lending To Consumers”, **International Journal Of Forecasting**, Cilt: 16, Sayı: 2, 2000, s.149-172.

Tolun, Seda : “Destek Vektör Makineleri: Banka Başarısızlığının Tahmini Üzerine Bir Uygulama”, Doktora Tezi, İstanbul Üniversitesi, 2008.

Torsun, I. S. : “A Neural Network For A Loan Application Scoring System” **The New Review Of Applied Expert Systems**, Cilt: 2, 1996, s.47–62.

Torun, Talip : “Finansal Başarısızlık Tahmininde Geleneksel İstatistiki Yöntemlerle Yapay Sinir Ağlarının

Karşılaştırılması Ve Sanayi İşletmeleri Üzerinde Uygulama”, Doktora Tezi, Erciyes Üniversitesi, 2007.

Tosh, Colin R.
Graeme Di Ruxton

: **Modelling Perception With Artificial Neural Networks**, 1. Baskı, Cambridge: Cambridge University Press, 2010.

Tunay, K. Batu

: “Türkiye’de paranın gelir dolaşım hızlarının MARS yöntemiyle tahmini”, **ODTÜ Gelişme Dergisi**, Cilt: 28, Sayı: 3-4, 2001, s.175-196.

Turgut, Ahmet

: “Bankalar Ve Kobi’ler İçin Dönüm Noktası: Basel II Standartları”, **Karamanoğlu Mehmetbey Üniversitesi Sosyal Ve Ekonomik Araştırmalar Dergisi**, Cilt: 2007, Sayı 2, s.29-44.

Türk Dil Kurumu

: **Genel Türkçe Sözlük**, Çevrimiçi, <https://sozluk.gov.tr/?kelime=risk>, 28/10/2019.

Türkiye Bankalar Birliği

: **Kredi Riskinin Yönetimine İlişkin İlkeler**, Çevrimiçi, https://www.tbb.org.tr/Dosyalar/Arastirma_ve_Raporlar/risk_yonetim.doc, 28/10/2019.

Vapnik, Vladimir N.

: “An Overview of Statistical Learning Theory”, **IEEE Transactions On Neural Networks**, Cilt: 10, Sayı: 5, 1999, s.988-999.

- Vapnik, Vladimir,
Rauf Izmailov : “Knowledge Transfer In SVM And Neural Networks”, **Ann Math Artif Intell**, Cilt: 81, 2017, s.3-19.
- Velioglu, Yilmaz : “Memzuç Nedir? Memzuç Kayıtlarının İçeriği ve Finans Dünyasındaki Yeri”, Çevrimiçi, http://www.muhasabetr.com/yazarlarimiz/yilmaz_velioglu/011/, 11/03/2020.
- Vellido, A.,
S.J.G. Lisboa,
J. Vaughan : “Neural networks in business: a survey of applications (1992–1998)”, **Expert Systems with Applications**, Cilt: 17, 1999, s.51–70.
- Wang, Xiangying,
Yixin Zhong : “Statistical Learning Theory and State of the Art in SVM”, **Proceedings of the Second IEEE International Conference on Cognitive Informatics**, 2003, s.55-59.
- West, David : “Neural Network Credit Scoring Models”, **Computers & Operations Research**, Cilt: 27, 2000, s.1131-1152.
- Westgaard, Sjur,
Nico Van Der Wijst : “Default Probabilities In A Corporate Bank Portfolio: A Logistic Model Approach”, **European Journal Of Operational Research**, Cilt: 135, 2001, s.338-349.

Wiginton, John C. : “A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behavior”, Cilt: 15, Sayı: 3, 1980, s.757-770.

Wikipedia, : Çevrimiçi,
https://tr.wikipedia.org/wiki/Ockhaml%C4%B1_William, 1004/2020.

Wikipedia : Çevrimiçi:
https://tr.wikipedia.org/wiki/Neumann_mimarisi#/media/Dosya:VonNeumannMimarisi.PNG, 14/10/2020.

Wikipedia, : Çevrimiçi,
https://tr.wikipedia.org/wiki/Mant%C4%B1ksal_pozitivizm, 07/04/2020.

Williamson, A. G. : “Refining, A Neural Network Credit Application Vetting System With A Genetic Algorithm” **Journal Of Microcomputer Applications**, Cilt: 18, 1995, s.261–277.

Wu, C.,
X. M. Wang : “A Neural Network Approach For Analyzing Small Business Lending Decisions”, **Review Of Quantitative Finance And Accounting**, Cilt: 15, 2000, s.397-408.

www.enchantedlearning.com : Çevrimiçi,
<https://www.enchantedlearning.com/subjects/anatomy/brain/Neuron.shtml>, 19/12/2019.

www.enchantedlearning.com : Çevrimiçi,
<https://www.enchantedlearning.com/subjects/anatomy/brain/Neuron.shtml>, 30/10/2020.

www.enchantedlearning.com, : Çevrimiçi,
<https://www.enchantedlearning.com/subjects/anatomy/brain/Neuron.shtml>, 30/10/2020.

www.scholarpedia.org : Çevrimiçi,
http://var.scholarpedia.org/article/Pyramidal_neuron

www.scholarpedia.org, : Çevrimiçi,
http://var.scholarpedia.org/article/Pyramidal_neuron, 05/11/2019

www.sciencenews.org : Çevrimiçi,
<https://www.sciencenews.org/article/new-ways-image-control-nerve-cells-unlock-brain-mysteries>

www.sciencenews.org, : Çevrimiçi,
<https://www.sciencenews.org/article/new-ways-image-control-nerve-cells-unlock-brain-mysteries>, 30/10/2020.

www.sciencewal.com : Çevrimiçi,
[https://tr.sciencewal.com/43028demystifying-
optimizations-for-machine-learning-
c6c6405d3eea-30](https://tr.sciencewal.com/43028demystifying-optimizations-for-machine-learning-c6c6405d3eea-30)

Yakut, Emre : “Veri Madenciliği Tekniklerinden C5.0 Algoritması Ve Destek Vektör Makineleri İle Yapay Sinir Ağlarının Sınıflandırma Başarılarının Karşılaştırılması: İmalat Sektöründe Bir Uygulama”, Doktora Tezi, Atatürk Üniversitesi, 2012.

Yao, Xin : “Evolving Artificial Neural Networks”, **Proceedings Of The IEEE**, Cilt: 87, Sayı: 9, 1999, s.1423-1447.

Yılmaz, Atınç : **Yapay Zekâ**, 2. Baskı, İstanbul: Kodlab Yayınevi, 2017.

Yılmaz, Atınç, Umut Kaya : **Derin Öğrenme**, 1. Baskı, İstanbul: İnkılap Kitabevi, 2019.

Yılmaz, Ender Aykut : “Basel II Kapsamında Kredi Riski Ve Temerrüt Olasılığı Hesaplama Yolu İle Tespiti”, Gazi Üniversitesi, Doktora Tezi, 2010.

Yılmaz, Şebnem Koltan, Said Patir, : “Kümeleme Analizi Ve Pazarlamada Kullanımı Kümeleme Analizi Ve Pazarlamada Kullanımı”,

Akademik Yaklaşımlar Dergisi, Cilt: 2, Sayı: 1, 2011, s.91-113.

Yılmaz, Şebnem Koltan,
Said Patir

: “Kümeleme Analizi Ve Pazarlamada Kullanımı”,
Akademik Yaklaşımlar Dergisi, Cilt: 2, Sayı: 1, 2011, s. 91-113.

Zhang, Guoqiang,
Michael Hu,
B. Eddy Patuwo,
Daniel C. Indro

: “Artificial Neural Networks In Bankruptcy Prediction: General Framework And Cross-Validation Analysis,” **European J. Operational Research**, Cilt: 116, 1999, s.16-32.

ÖZGEÇMİŞ

1982 yılında, Kayseri’de doğan Gökhan Korkmaz, ilköğreniminin ilk dört yılını Kayseri’de, beşinci yılını İzmir’de, orta ve lise öğrenimini ise yine Kayseri’de tamamlamıştır. 2002 yılında, Erciyes Üniversitesi, İktisadi Ve İdari Bilimler Fakültesi, İşletme Bölümüne yerleşmiştir. Bir yıllık İngilizce hazırlık eğitiminin ardından, 2007 yılında bu fakülteden mezun olmuş, kısa bir süre sonra ise askerlik hizmetini yerine getirmiştir. Askerlik hizmetinden önce ve sonra olma üzere, kısa süreli olarak, özel sektörde tecrübe edinmiş, ardından 2010 yılında Şırnak Üniversitesi, İktisadi Ve İdari Bilimler Fakültesi, İşletme Bölümü’nde, Araştırma Görevlisi olmuştur.

Yüksek lisansını, Çukurova Üniversitesi, İktisadi Ve İdari Bilimler Fakültesi, İşletme Bölümü, Yönetim Ve Organizasyon Anabilim Dalı’nda, 2015 yılında tamamlayan Korkmaz, doktora Sayısal Yönelimler alanını tercih etmiş ve 2016 yılında İstanbul Üniversitesi, İşletme Fakültesi, Sayısal Yöntemler Anabilim Dalı’nda doktora eğitimine başlamıştır. Yüksek lisans ve doktora aşamalarında, sırasıyla Çukurova ve İstanbul Üniversitelerinde, geçici görevlendirmeyeyle, sırasıyla iki ve üç yıl görev yapan Korkmaz, evli ve bir çocuk babasıdır ve Şırnak Üniversitesi’ndeki görevine devam etmektedir.