

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞLARI YÖNTEMİ İLE KREDİ SKORLAMA

YÜKSEK LİSANS TEZİ

Burak DONEL

İşletme Mühendisliği Anabilim Dalı

İşletme Mühendisliği Programı

HAZİRAN 2012

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞLARI YÖNTEMİ İLE KREDİ SKORLAMA

YÜKSEK LİSANS TEZİ

**Burak DONEL
(507081003)**

İşletme Mühendisliği Anabilim Dalı

İşletme Mühendisliği Programı

Tez Danışmanı: Yrd. Doç. Dr. Cumhuri Ekinci

HAZİRAN 2012

İTÜ, Fen Bilimleri Enstitüsü'nün 507081003 numaralı Yüksek Lisans Öğrencisi **Burak DONEL**, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “**YAPAY SİNİR AĞLARI İLE KREDİ SKORLAMA**” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Yrd. Doç. Dr. Cumhur Ekinci**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Nesrin OKAY**
Boğaziçi Üniversitesi

Prof. Dr. Burç Ülengin
İstanbul Teknik Üniversitesi

Teslim Tarihi : **3 Mayıs 2012**
Savunma Tarihi : **5 Haziran 2012**

ÖNSÖZ

Yüksek lisans öğrenimim sırasında kendisi ile tanışma fırsatı bulduğum, tez çalışmam boyunca bilgi ve tecrübelerini esirgemedi katkıda bulunan, olumlu öneri ve eleştirileri ile beraber bu çalışmayı yöneten çok değerli danışman hocam Sn.Yrd. Doç. Dr. Cumhuri EKİNCİ'ye en derin sevgi, saygı ve teşekkürlerimi sunarım.

Yüksek lisans tez çalışmam için imkân ve olanaklarını sunarak bana destek olan Finansbank Kredi Analitik birimine, Sn. Uğur Aydın'a, Sn. Çağlar Filiz'e, teşekkür ederim

Tez çalışmalarının sıkıntılı zamanlarını, beraber geçirdiğimiz keyifli anlar ve arkadaşlıklarıyla unutturan, bana her konuda destek çok değerli arkadaşlarım; Büke CİHANLI'ya, Özgür ORHON'a ve Erkut ELİUZ'a tüm içtenliğimle teşekkür ederim.

Son olarak, tüm hayatım boyunca her daim yanımda olan, bugünlere gelmemde benden maddi ve manevi desteklerini hiçbir zaman esirgemeyen çok kıymetli aileme en derin duygularıyla teşekkür ederim.

Mayıs 2012

Burak Donel

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	v
İÇİNDEKİLER.....	vii
ŞEKİL LİSTESİ.....	xi
ÇİZELGE LİSTESİ.....	xiii
KISALTMALAR.....	xv
ÖZET.....	xvii
1.GİRİŞ	1
2.KREDİ.....	3
2.1. Kredinin Tanımı	3
2.2. Kredi Türleri	3
2.2.1. Kullanım şekline göre krediler.....	4
2.2.2. Kullanım amacına göre krediler	5
2.2.3. Vadelerine göre krediler	5
2.2.4. Kullanım türlerine göre krediler.....	5
2.2.5. Kredi kaynaklarına göre krediler.....	5
2.2.6. Kredi Teminatlarına göre krediler	6
3. RİSK.....	7
3.1. Riskin Tanımı.....	7
3.2. Bankacılık Sektörü Riskleri	8
3.2.1 Kredi riski	9
3.2.2. Piyasa riski	10
3.2.3. Likidite riski	10
3.2.4. Faiz riski.....	11
3.2.5. Yoğunlaşma riski.....	11
3.2.6. Ülke riski.....	12
3.2.7. Sermaye yeterliliği riski.....	12
3.3. Kredi Riski Yönetimi	13
4.KREDİ SKORLAMA	15
4.1. Kredi Skorlama Tarihçesi.....	15
4.2. Kredi Skorlamada Geleneksel Yaklaşım.....	16
4.3. Kredi Skorlama Metotları.....	17
4.3.1. Lineer regresyon.....	17
4.3.2. Diskriminant analizi	18
4.3.3. Lojistik regresyon.....	19
4.3.4. Karar ağaçları	20
4.3.5. K - En yakın komşu yaklaşımı	21
4.4. Kurulan Modellerin Kontrolü	23
4.4.1. Diverjans istatistiği.....	23
4.4.2. Yanlış sınıflandırma matrisi.....	23
4.4.3. Kullback diverjans ölçüsü.....	24
4.4.4. Kolmogrov-Smirnov istatistiği.....	24

4.4.5. Gini katsayısı	25
4.4.6. Chi-square (χ^2) testi.....	25
4.4.7. Alıcı çalıştırma karakteristiği(ROC)	26
5. YAPAY SİNİR AĞLARI.....	29
5.1. Yapay Zeka.....	29
5.2. Yapay Sinir Ağları Yönteminde Temel Tanımlar	30
5.1.1. Yapay Sinir Ağları Yönteminin Tanımı	30
5.2.1. Biyolojik sinir ağları	31
5.3. Yapay Sinir Ağlarındaki Süreçler.....	32
5.3.1. Girdiler	32
5.3.2. Ağırlıklar	33
5.3.3. Aktivasyon fonksiyonu.....	33
5.3.3.1 Doğrusal aktivasyon fonksiyonu.....	33
5.3.3.2. Adım fonksiyonu.....	34
5.3.3.2. Sigmoid aktivasyon fonksiyonu.....	35
5.3.3.4. Hiperbolik tanjant fonksiyonu	35
5.3.4. Çıktı.....	36
5.4. Yapay Sinir Ağlarının Sınıflandırılması	36
5.4.1. Öğrenme paradigmalarına göre sınıflandırılması	36
5.4.1.1. Danışmanlı öğrenme.....	37
5.4.1.2. Danışmansız öğrenme	37
5.4.1.3. Takviyeli öğrenme.....	38
5.4.1.2. Yapay sinir ağlarında öğrenme kuralları	39
5.4.1.2.1. Hebbian öğrenme kuralı	39
5.4.1.2.2. Delta kuralı.....	39
5.4.1.2.3. Gradyan iniş kuralı	39
5.4.2. Yapay sinir ağlarının yapılarına göre sınıflandırılması.....	39
5.4.2.1. İleri beslemeli ağlar	39
5.4.2.2. Geri beslemeli ağlar	40
5.5. Çok Katmanlı Algılayıcılar	40
5.5. Yapay Sinir Ağlarında Geliştirici Unsurlar.....	41
5.5.1. Öğrenme hızı.....	41
5.5.2. Momentum katsayısı	42
5.5.3. Gizli katman sayısı.....	42
5.5.4. Hata toleransı	43
6. KREDİ SKORLAMA LİTERATÜR ÇALIŞMALARI	45
6.1. Yapay Sinir Ağları Üzerine Yapılan Çalışmalar	46
6.2. İleri Düzey İstatistiksel ve Uzman Sistemler Üzerine Yapılan Çalışmalar	49
6.3. Türkiye’de Kredi Skorlama Üzerine Yapılan Çalışmalar	52
7. LOJİSTİK REGRESYON VE YAPAY SİNİR AĞLARI İLE KREDİ SKORLAMA MODELLERİ	55
7.1. Modelde Kullanılan Değişkenler	57
7.2. Lojistik Regresyon Modeli.....	60
7.2.1. Lojistik regresyon modelinin anlamlılığının test edilmesi	64
7.3. Yapay Sinir Ağları ile Kredi Skorlama Modelinin Kurulması.....	66
7.3.1. Yapay sinir ağlarının yapısı.....	66
7.3.2. Yapay sinir ağlarından elde edilen sonuçlar	68
7.3.3. Yapay sinir ağları öğrenme-test oranları karşılaştırılması.....	70
7.4. Elde Edilen Sonuçların Yorumlanması.....	71
8. SONUÇLAR	75

KAYNAKLAR.....	79
ÖZGEÇMİŞ	83

ŞEKİL LİSTESİ

Sayfa

Şekil 4.1 : Örnek bir karar ağacı yapısının şematik gösterimi.....	21
Şekil 4.2 : ROC eğrisi.....	28
Şekil 5.1 : Biyolojik sinir ağının gösterimi.....	32
Şekil 5.2 : Yapay sinir ağının sistematik gösterimi.....	32
Şekil 5.3 : Doğrusal aktivasyon fonksiyonunun şekilsel gösterimi.....	34
Şekil 5.4 : Adım fonksiyonunun şekilsel gösterimi.....	34
Şekil 5.5 : Hiperbolik tanjant fonksiyonunun şekilsel gösterimi.....	35
Şekil 5.6 : Danışmanlı öğrenme modeli.....	37
Şekil 5.7 : Takviyeli öğrenme modeli.....	38
Şekil 5.8 : Çok katmanlı algılayıcı.....	41

ÇİZELGE LİSTESİ

Sayfa

Çizelge 4.1 : Sınıflandırma tablosu	27
Çizelge 7.1 : Bağımlı ve bağımsız değişkenler ve kodları	57
Çizelge 7.2 : Çocuk sayısı frekans tablosu	58
Çizelge 7.3 : Meslek grupları ve kodlanmaları	58
Çizelge 7.4 : Meslek frekans tablosu	59
Çizelge 7.5 : Müşteri ilişki süresi frekans tablosu	59
Çizelge 7.6 : Cinsiyet frekans tablosu	59
Çizelge 7.7 : Eğitim durumu frekans tablosu	60
Çizelge 7.8 : Eğitim durumu frekans tablosu	60
Çizelge 7.9 : Veriseti test ve öğrenme adet ve yüzdeleri	60
Çizelge 7.10 : Kategorik değişkenlerin kodlanması	61
Çizelge 7.11 : Modele dahil edilmeyen değişkenler	62
Çizelge 7.12 : Modele dahil edilen değişkenler	63
Çizelge 7.13 : Hosmer ve Lemeshow Testi	64
Çizelge 7.14 : Sınıflandırma Tablosu	65
Çizelge 7.15 : 16 gizli katmanlı algılayıcın yapay sinir ağı momentum katsayıları ve öğrenme katsayıları karşılaştırılması	68
Çizelge 7.16 : 21 gizli katmanlı algılayıcın yapay sinir ağı momentum katsayıları ve öğrenme katsayılar karşılaştırılması	68
Çizelge 7.17 : Yapay sinir ağları sınıflandırma tabloları	69
Çizelge 7.18 : Kredi skorlama modellerinin karşılaştırılması	70
Çizelge 7.19 : Yapay sinir ağları öğrenme- test oranları karşılaştırılması	70
Çizelge 7.20 : Yapay sinir ağları öğrenme- test oranları karşılaştırılması	71

KISALTMALAR

ANN	: Artificial Neural Network
CART	: Classification and regression tree
İE	: İşlemci elemanlar
KS	: Kolmogrow – Smirnov
LVQ	: Lineer vector quantization
MARS	: Multiple adaptive regression splines
MLP	: Multi layer perceptron
SOM	: Self organizing map
YSA	: Yapay sinir ağları
5C	: Karakter, kapasite, sermaye, teminat, dış faktörler

YAPAY SİNİR AĞLARI İLE KREDİ SKORLAMA

ÖZET

Kredi riski geçmiş dönem ve günümüzde finans sektörünün karşılaştığı en önemli risklerden biri olmuştur. Kredi riskinin makul seviyelerde tutulamamasından dolayı firma, ülkesel ve küresel bazda krizler gerçekleşmiştir. Finans sektörünün en büyük kuruluşları olan bankalar açısından kredi riskinin yönetilmesi önem kazanmıştır. Bankalar açısından kredi riskinin iyi yönetilmesi müşterilerin temerrüde düşme olasılığını tespit etmeyi ve müşterileri temerrüt düşme olasılıklarına göre sınıflandırmayı gerektirmektedir. Bankalar tarafından yapılan bu sınıflandırma çeşitli yöntemlerle elde edilen skorkartlar vasıtasıyla sağlanmaktadır. Skorkart kavramı finansal kuruluşlar tarafından birçok uygulamada kullanılmaktadır. Yeni müşterilerin kabul edilmesinde, mevcut müşterilerin davranışlarının belirlenmesinde, tahsilat skorlarının belirlenmesinde, müşteri skoru hesaplanmasında skorkartlar kullanılmaktadır.

Çalışmada, literatürde mevcut kredi skorlama modelleri açıklanmış ve ticari bir bankaya ait veriseti üzerinden skorkartlar oluşturulmaya çalışılmıştır. Veriseti 2008 yılına ait tüketici kredisi başvurusu yapmış müşterilerin bilgilerini içermektedir. Literatürde farklı yöntemler kullanılarak çeşitli skorkartlar üretilmiştir. Bu çalışmada, skorkartların modellenmesi amacıyla lojistik regresyon yöntemi ve yapay sinir ağları yöntemi kullanılmıştır. Öncelikle, literatürde kabul görmüş olan lojistik regresyon yöntemiyle skorkart elde edilmiştir. Çok katmanlı algılayıcı yapay sinir ağı yapısı formülleştirme ve deneme -yanılma yöntemi ile girdi-ara-çıkı katmanı, momentum katsayısı ve öğrenme katsayısı ve iki adet çok katmanlı yapay sinir ağı elde edilmiştir. Skorkartlardan elde edilen değerler birinci tip hata ve ikinci tip hata değeri üzerinden karşılaştırılmıştır. Birinci tip hata (yani iyi kredilerin kötü kredi olarak belirtilmesi) tespitinde, lojistik regresyon modelinin daha iyi bir performans gösterdiği görülmüştür. Bankalar ve finans kuruluşları açısından daha maliyetli olan 2.tip hatanın (yani kötü kredilerin iyi kredi olarak belirtilmesi) tespitinde, yapay sinir ağları daha başarılı olmuştur. Toplam hata yüzdesinde yapay sinir ağları daha düşük hatayla kredi skorlama modeli oluşturmuştur.

Elde edilen yapay sinir ağı yapılarının performansını arttırabilmek amacıyla, öğrenme ve sına veriseti çeşitli oranlarda değiştirilerek ağ yapısının performansı birinci ve ikinci tip hata cinsinden ve alıcı çalıştırma karakteristiği (ROC) değerleri üzerinden karşılaştırılmıştır. Sonuç olarak, yapay sinir ağlarının yapılarındaki farklılıklar itibarıyla öğrenme-sınama veriseti yüzdeleri değiştirilerek performanslarının iyileştirilebildiği görülmüştür.

CREDIT SCORING WITH ARTIFICIAL NEURAL NETWORK

SUMMARY

Credit risk has been one of the most important risks that financial sector ever encounters. Because it could not be capped in resonable exposure, domestic and international financial crisis occured. Management of this risk becomes really important for banks that are largest institutions in financial sector. The management of the credit risk for banks involves investigation about default risks and classification of the customers according to the risk that has been investigated. This type of classification has been made by the score cards that have been calculated by various methods.

Under a competitive perspective, screening borrowers and differentiating the prices accordingly, given the borrowers' standing and their contributions to the bank's portfolio risk–return profile, are key issues. Not doing so results in adverse economics for banks. Banks who do not differentiate risks lend to borrowers rejected by banks who better screen and differentiate risks. By overpricing good risks, they discourage good borrowers. By underpricing risks to risky customers, they attract them. By discouraging the relatively good ones and attracting the relatively bad ones, the less advanced banks face the risk of becoming riskier and poorer than banks adopting sound risk-based practices at an earlier stage. Those banking institutions that actively manage their risks have a competitive advantage. They take risks more consciously, they anticipate adverse changes, they protect themselves from unexpected events and they gain the expertise to price risks. The competitors who lack such abilities may gain business in the short-term. Nevertheless, they will lose ground with time, when those risks materialize into losses.

As regards credit risk, and any other situation involving trust (which includes all economic activities involving contracts or liabilities), the contracting parties must be aware of the possibility that things may not be as they seem. Where trust is low, lenders will increase their charges to cover the risks. Trust can however, be enhanced through collateral, other security or more information. In ages past, credit was often only extended against collateral, but the cost of realising its value is high. The modern information age allows lenders to enhance trust, by using data about borrowers' financial and other circumstances, whether at time of application or ongoing thereafter.

Scoring refers to the use of a numerical tool to rank order cases (people, companies, fruit, countries) according to some real or perceived quality (performance, desirability, saleability, risk) in order to discriminate between them, and ensure objective and consistent decisions (select, discard, export, sell). Available data is integrated into a single value that implies some quality, usually related to desirability or suitability. Scores are usually presented as numbers that represent a single quality, while grades may be presented as letters (A, B, C, etc.) or labels (export quality, investment grade) to represent one or more qualities.

Credit scores provide the greatest value, when they are used to guide decisions that affect the customer. In decision processes, lenders define different scenarios using scores and policy, and then the action to be taken in each case—like accept/reject, maximum loan value or repayment, interest rate, loan term, etc. Alternatively, underwriters may consider scores as one of several inputs into a credit decision. The cost benefits of decision automation are placing incredible pressure upon organizations to limit the use of underwriters, to cases where their specialist knowledge is absolutely essential, especially where there is significant information that cannot be captured within the scoring process, and potential profits are high.

In 1950's with the invention of the credit-cards, there were too many applications for having it. Furthermore the classical method where the credit experts had not the capacity to analyses all of the applications and the credit scoring applications are developed. Firstly it was used simple methods, they are statistical methods; linear regression and discriminant analyses. In the credit scoring applications there are two types of the methods; parametric methods and non-parametric methods. Parametric methods have some assumptions about the data. But they have too many assumptions about the data and with the structure about the problem, they were not fit well. Logistic regression was used to develop a scorecard. With the nature of the logistic regression, there are not any assumptions about the data and the results of the logistic regression are very appropriate to choose which credits are good and which credits are bad. With the study it was supposed that the logistic regression models are performing better than linear regression models and discriminant analyses to develop scorecards.

With the extensive progress in the computers, it was developed to much intelligent learning paradigms to resolve complex problems which were not resolved with classical statistical methods. In real world, the natures of problems are not linear so the interest in the intelligent learning paradigms were grew. In the credit scoring problems there are too much data about the problem and all of the data are not linear with outputs of the models. In the end of the 1990's the researcher were interested with different methods to resolve the credit scoring problems. The decision trees were non-parametric methods was used in credit-scoring applications. With the help of the computers, in the big dataset the decision trees were performed well in the credit scoring problems and with the demonstration of decision tree, the analyses of model are very easy. But in the small dataset the credibility of the decision trees are uncertain, also with the growth of the dataset, the demonstration of the models is not easy to understand.

Artificial neural networks are the intelligent machine systems which were developed to simulate humans nerve systems. Their working schedule starts with the given of the input data to the system. There are three layer, input, interval and output layer. Input layers are connected with interval layers and interval layer are connected with output layers. The system work like nerve system, with the input data the weights of all the interval layers change and all of interval layers have a value. Finally, the weights between interval and output layers change, it given a value about the problem.

In the artificial neural networks there are no assumptions about the dataset and it works well with the problems non-linear. They are performed to the binary solutions problems like credit scoring problems where there are two solutions, good credit or bad credit. They are too many artificial neural network systems, they are classified

about the paradigm of learning of the network and the direction of the networks. Multi-layer perceptron is one of the artificial neural network systems which are generally used for the credit scoring. There are two difficulties to use the artificial neural networks; firstly there is no information about the solutions of the models because of it the scientist name “black box” to the artificial neural networks, secondly in the artificial neural networks architecture there is no a true architecture for all of the problems. In the artificial neural networks, architecture mean the number of interval layer, learning rate and momentum rate which are used in the neural networks. All of these numbers affect directly to the performance of the neural networks.

In this thesis, credit scoring methods in the literature are explained and scorecards are calculated according to the dataset of a commercial bank. Dataset includes information about the consumers who applied for a loan in 2008. In dataset there is ten independent variables and one dependent variable. Dependant variable is; the credit will default or not default.

In the implementation of the models firstly, scorecards were calculated according to the logistic regression model. The regression model was established and was tested by receiving operating characteristic and Hosmer-Lemeshow Test. Then, multi-layer perceptron artificial network system formulation, input interval outcome layer with trial and error approach, momentum coefficient, learning coefficient and two multilayer perceptron artificial network systems were acquired. The aim of the preparing two different artificial neural networks is to compare the formulation and trial and error approach. If the multi-layer perceptron artificial network system formulation had better performance than trial and error approach, we could not spend much time to derive best architecture structure. Values that are acquired from score cards were compared with Type one and Type two error values. In the evaluation of Type one error, which means admitting good credits as bad, logistic regression model performs superior than other models. In the evaluation of Type two error, which means admitting bad credits as good, artificial network system showed superior performance. Artificial network system formed credit scoring model with lower error in the total error percentage.

In literature research, it was seemed that generally it worked with higher learning rate of data and lower trial rate of data, approximately around %70 percent of the data were divided to learning dataset and around %30 percent of the data were trial dataset. With lower learning dataset, it suggested that in Type two error the correct percentage was increased but in the Type one error the correct percentage was decreased. For the financial institutions, Type two error is more important than Type one error because higher percent of bad credit acceptance is higher costly than higher percent of good credit reject. In comparing trial and error approach and formulation approach, trial and error approach performs better than formulation approach to detect good and bad credits.

In the final section of our implementation, to increase the performance of artificial network systems, learning and trial datasets were replaced with various proportions and values obtained from Type one, Type two error and Receiver Operating Characteristic (ROC) were compared. As a result, by changing learning and trial datasets, performance of the artificial nerve system was increased.

1.GİRİŞ

Kriz beklenmeyen, öngörülemeyen bir durumun ortaya çıkması olarak ifade edilmektedir. Dünya ve Türkiye’de meydana gelen krizler sonucunda finansal kuruluşlar için çeşitli uygulamalar geliştirilmiş ve bu kurumları denetleyecek kurumlar oluşturulmuştur. Finansal kurumların en önemlilerini bankalar oluşturmaktadır. Bankaların temel görevi elde ettikleri kaynakları etkili ve etkin bir biçimde kullandırarak ülkenin kalkınmasına katkı vermektir.

Türkiye’de bankacılık sektörü, 1980’li yıllardan 2000’li yıllara kadar geçen yüksek enflasyonlu süreçte kredi vermekten çok finansal piyasalarda işlemler yaparak kârlılıklarını artırma yoluna gitmişlerdir. 2000’li yıllar itibarıyla düşük enflasyon ve düşük faiz dönemi ile kredi vermek en önemli kâr kaynaklarını oluşturmaya başlamıştır. Esas bankacılığın kredi vermek olarak adlandırıldığını düşünürsek bu dönem itibarıyla esas bankacılık yapılmaya başlanmıştır.

Bankacılık sektörünün büyüklüğü ve önemi itibarıyla taşıdığı riskler, hükümetler ve denetleyici kurumlar tarafından yakından izlenmektedir. Bankacılık sektörünü etkileyen en önemli risk olarak kredi riski kavramı öne çıkmaktadır. Kredi riski kullandırılan kredinin ödenmemesi durumunda ortaya çıkan risk olarak tarif edilebilir. Bankalar kredi verdikleri müşterileri iyi tanımaya ve verdikleri kredilerin geri ödenebilmesi konusunda sorun yaşamamaya çalışırlar. Fakat günümüzde bankalardan kredi kullanan kişi sayısının milyonlara ulaşması sonucunda güvene ve müşteriye dayalı kredi verme süreci işlemesi oldukça güçtür.

Kredi skortlama teknikleri olarak adlandırılan ve müşteri bilgilerini kullanarak müşterilere kredi verilip, verilmeyeceğini ölçen teknikler günümüzde bankalar arasında oldukça yaygın olarak kullanılmaktadır. Kredi skortlama, müşteri bilgilerinin kullanılarak kredi başvurusunda bulunan müşterilere kredi verip vermeme kararını veren tekniklerdir. Kredi skortlama yöntemleri ile kredi verildiğinde geri ödemede sıkıntı yaşamayacak kredi müşterilerinin seçiminde daha yüksek oranda başarı sağlamak ve verilen kredilerde geri ödememe durumuyla karşılaşılacak kredi müşterisi yüzdesini azaltmak sağlanmalıdır. Bankaların ellerinde bulundurdıkları

müşteri bilgileri doğrultusunda istatistiksel ve istatistiksel olmayan yöntemler kullanılarak müşteriler kredi skorlamadan geçirilir ve bir skorları oluşur.

Kredi skorlama yöntemleri olarak en yaygın kullanılan ve kabul görmüş yöntem lojistik regresyondur. Lineer regresyon ve lineer diskriminant analizine göre daha üstün gelen lojistik regresyon skorlama çalışmalarında oldukça yaygın kullanılmıştır. Günümüzde bilgisayar teknolojilerindeki gelişmeler doğrultusunda uzman sistemler olarak adlandırılan yapay sinir ağları, karar ağaçları, genetik algoritmalar, çoklu vektör makineleri gibi yöntemler de kredi skorlama yöntemi olarak kullanılmaktadır.

Yapay sinir ağları özellikleri itibarıyla ikili sonuçlara ulaşılmasının isteneceği durumlarda etkili bir yöntem olarak düşünülmüş ve kredi skorlama çalışmalarında bir yöntem olarak kullanılmıştır. Yapay sinir ağlarında en büyük zorluk kullanılan sinir ağının ağ yapısının belirlenmesidir.

Çalışmada ilk olarak kredi türleri ve finansal kuruluşların maruz kaldıkları riskler anlatılmıştır. Kredi skorlamanın ne olduğu, yöntemleri, yapay sinir ağları ile ilgili ayrıntılı bilgi verilmiştir. Literatür taramasında yapay sinir ağları ve diğer uzman sistemlerle yapılan çalışmalar irdelenmeye çalışılmış ve Türkiye’de kredi skorlama üzerine yapılan çalışmalar gözden geçirilmiştir. Yapay sinir ağlarının oluşturulmasında kurulacak olan ağ yapısı mimarisi üzerine çeşitli çalışmalar yapılmış olsa da, kesin kabul görmüş bir ağ mimarisi yapısının mevcut olmaması en büyük sorunu oluşturmaktadır. Bu açıdan uygulama kısmında lojistik regresyon ve yapay sinir ağları ile kredi skorlama modelleri oluşturulmuştur. Elde edilen sonuçlar karşılaştırılmış ve kullanılan veri setleri farklı öğrenme-sınama oranlarında değiştirilerek yapay sinir ağları performanslarının nasıl arttırabileceği irdelenmiştir.

2.KREDİ

2.1. Kredinin Tanımı

Gelişmiş bir ekonomide finansal sistem, fon fazlası olan kişiler ile fon gereksinimi bulunan kişileri buluşturarak fon alışverişinin düzenli ve güvenli bir biçimde gerçekleşmesini sağlar. Finansal sistemde fon ihtiyacı olanlarla fon fazlası olanları buluşturan başlıca kurumlardan olan bankalar, fon aktarma işlevini ağırlıklı olarak krediler aracılığı ile yerine getirirler.

Kredi, bankanın gerçek ya da tüzel olan bir kişiye ödünç para vermesi veya bu kişi lehine bir işin yapılması ya da bir borcun ödenmesi konusunda taahhütte bulunması veya garanti vermesidir.

Ülkemizde kredi işlemi Bankalar Kanunu’nun 48. maddesi, 1. fıkrasında şu şekilde tanımlanmıştır; “Bankalarca verilen nakdi krediler ile teminat mektupları, kontr-garantiler, kefaletler, aval, ciro, kabul gibi gayri nakdi krediler ve bu niteliği haiz taahhütler, satın alınan tahvil ve benzeri sermaye piyasası araçları, tevdiatta bulunmak suretiyle ya da herhangi bir şekil ve surette verilen ödünçler, varlıkların vadeli satışından doğan alacaklar, vadesi geçmiş nakdi krediler, tahakkuk etmekle birlikte tahsil edilmemiş faizler, gayri nakdi kredilerin nakde tahvil olan bedelleri, ters repo işlemlerinden alacaklar, vadeli işlem ve opsiyon sözleşmeleri ile diğer sözleşmeler nedeniyle üstlenilen riskler, ortaklık payları ve Kurulca kredi olarak kabul edilen işlemler izlendikleri hesaba bakılmaksızın bu Kanun uygulamasında kredi sayılır.” (Bankalar Kanunu, 1999).

2.2. Kredi Türleri

Bankacılık sisteminde krediler genel anlamıyla vadelerine veya teminat yapılarına göre sınıflandırılırlar. Daha geniş olarak kredi türlerinin ayrımı şu şekilde yapılabilir;

1. Kullanırım şekline göre krediler;

- nakdi krediler,
- gayri nakdi kredilerdir.

2. Kullanım amacına göre krediler;

- yatırım kredileri,
- işletme kredileridir.

3. Vadelerine göre krediler;

- kısa vadeli krediler,
- orta vadeli krediler,
- uzun vadeli kredilerdir.

4. Kullanım türlerine göre krediler;

- ticari krediler,
- proje kredileri,
- sektörel krediler,
- ihracat kredileridir

5. Kredi kaynaklarına göre krediler;

- yurtiçi kaynaklı krediler,
- yurtdışı kaynaklı kredilerdir.

6. Kredi teminatlarına göre krediler;

- teminatsız krediler,
- teminatlı kredilerdir.

2.2.1. Kullandırım şekline göre krediler

Kullandırım şekline göre, nakdi krediler belirli bir gelir getirecek şekilde faiz ve/veya komisyon karşılığında nakit para şeklinde kullandırılan kredilerdir. Bu tür krediler borçlu cari, iskonto, iştirak veya avans gibi farklı şekillerde kullandırılabilir. Gayri nakdi kredilerde bankalar müşteri veya firmaya direkt olarak para vermezler. Bu tür kredilerde firmaların üçüncü kişilere olan borçlarına çeşitli yollardan ödeme garantisi verirler. Banka, borçluya bir tür kefil olmaktadır. Gayri nakdi kredilerin bankalar tarafından en yaygın olarak kullandırılanları; akreditifler, teminat mektupları, harici garantiler ve kabul aval kredileridir (Tekindağ, 2010).

2.2.2. Kullanım amacına göre krediler

Kullanım amacına göre krediler, yatırım kredileri ve işletme kredileri olarak iki grupta incelenebilir. Yatırım kredileri bina, makine, teçhizat gibi sabit kıymetlerin modernizasyonun yapılması, yenilenmesi, satın alınması gibi durumlarda 3-10 yıl gibi süreçler için kullandırılan kredilerdir. İşletme kredileri ise işletmenin çeşitli ihtiyaçlarını (depolama, hammadde satın alma vs.) karşılamak amacıyla kullandırılan kredilerdir, genellikle orta vadeli krediler olarak kullandırılır ve geri ödeme süreleri 1 ile 5 yıl arasında değişir.

2.2.3. Vadelerine göre krediler

Kısa vadeli krediler, kredinin kullandırım tarihinden itibaren geri ödendiği tarih ile arasındaki süresinin 365 günden kısa olduğu kredilerdir. Orta vadeli krediler 1 ile 5 yıl arasında kullandırılan krediler, uzun vadeli krediler ise 5 yıldan fazla süre için kullandırılan kredilerdir. Uzun vadeli krediler genel olarak yatırım finansmanı amacıyla kullandırılmaktadır.

2.2.4. Kullanım türlerine göre krediler

Ticari krediler, ticari işletmelere kullandırılan vadesi 1 yıldan az olan, işletmenin sermaye ihtiyacını karşılamak amacıyla kullandırılan kredilerdir. Proje kredileri bir proje kapsamında ve proje konusu kapsamında (yeni ürün üretimi, yeni proses geliştirme vs.) verilen kredilerdir. Genellikle orta ve uzun vadede kullandırılan proje kredileri, proje kapsamında projenin nakit akışına bağlı olarak bankalar tarafından verilen kredilerdir. Sektörel krediler, belirli sektörlerle destek olmak için verilen kredilerdir. Bu tür kredilerde devlet tarafından bankalar teşvik edilmektedir. Bu tür kredilere tarım sektörüne destek vermek amacıyla kullandırılan tarım kredileri örnek olarak verilebilir. İhracat kredileri, ihracatın finansmanı amacıyla kısa veya orta vadeli kullandırılan kredilerdir.

2.2.5. Kredi kaynaklarına göre krediler

Yurtiçi kaynaklı krediler bankaların kendi özkaynakları, sermaye, mevduat ve sendikasyon kredileri ile elde ettikleri kaynakları kullanarak müşterilerine kullandırdıkları kredilerdir. Bu kaynaklar bankaların bilançolarında yer almasından dolayı yurtiçi kaynaklı krediler olarak adlandırılmaktadır. Yurtdışı kaynaklı krediler,

bankaların yurtdışı kaynaklarından temin ederek belli bir sektör, müşteri grubu veya projeye kullandıkları kredilerdir (Ataçoğlu, 2006).

2.2.6. Kredi Teminatlarına göre krediler

Teminatsız krediler herhangi bir teminata bağlı olmadan kullanılan kredilerdir. Literatürde bu tür kredilere ‘açık’ krediler de denilmekte olup en yüksek risk derecesine sahip kredilerdir. Bu tür yüksek riske sahip krediler bankalar tarafından çok yüksek derecede kredibilitesi olan müşterilere kullanılmaktadır. Teminatlı krediler müşterilerden kredi kullandırımı karşılığında maddi bir teminat olarak kullanılan kredilere denir. Alınan teminatlar çeşitli çok farklı çeşitlerde olabilmektedir. Bunlar;

- kefalet karşılığı kredi,
- senet karşılığı kredi,
- emtia karşılığı kredi,
- ipotek karşılığı kredi,
- temlik karşılığı kredi,
- menkul kıymet karşılığı kredi,
- nakit karşılıklı kredidir.

3. RISK

3.1. Riskin Tanımı

Finansal kuruluşlar ve bankalar finansal piyasalarda kararlar alırken bazı ölçütlere göre hareket etmektedirler. Bu ölçütlerden en önemlisi risk kavramıdır. Finansal yatırımlarda dikkate alınması gereken üç durum söz konusudur; belirlilik durumu, riskli durum ve belirsizlik durumudur. Belirsizlik ve riskli durumda gelecekteki olayların kesin olarak bilinemeyeceği söz konusudur. Bu iki kavram birbirinden ayrı kavramlar olarak ele alınmalarına rağmen her iki durumunda gelecekte kesinlikle bilinemeyeceğini savunanlar, riski, beklenen bir sonucun belirsizlik derecesi olarak kabul etmektedirler.

Literatürde risk kavramına karşılık gelen birçok açıklama mevcuttur:

- Risk, gelecekte bir anda gerçekleşecek getirinin, beklenen getiriden sapması durumu olarak tanımlanabilir.
- Risk, bir isleme ilişkin, bir parasal kaybın ortaya çıkması veya bir giderin ya da zararın oluşması nedeniyle, ekonomik faydanın azalması ihtimalidir.
- Teknik anlamda risk, getirilere ilişkin olasılık değerlerinin ortalama değer etrafındaki dağılımı ile ifade edilebilir.

Bu tanımlardan da anlaşılabileceği üzere bazı riskler kontrol edilebilmekte bazı riskler kontrol edilememektedir. Literatürde de riskler sistematik olan ve sistematik olmayan riskler olarak ayrılmıştır. Sistematik riskler mevcut piyasa koşullarında ortaya çıkan riskler olarak algılanırken, sistematik olmayan riskler, dış faktörler nedeniyle ortaya çıkan riskler olarak algılanmaktadır. Bankacılık ve finansal sistemlerde sistematik riskler bazı tekniklerle kontrol edilip etkileri minimize edilmeye çalışılırken, sistematik olmayan risklere karşı kuruluşlar önlem alma çabası içindedir.

3.2. Bankacılık Sektörü Riskleri

Bankacılık sektörü içerisinde, riskler genel olarak içsel ve dışsal riskler olarak sınıflandırılmaktadır. Sektörün kendi yapısından kaynaklı risklere içsel, sektörün dışından kaynaklı risklere dışsal riskler denilmektedir. Bankacılık sektöründe karşılaşılan risk grupları içerisinde en önemlisi ise, mali riskler denilen, bankaların ve sektörün kendi yapı ve operasyonlarından kaynaklanan risk grubudur.

Bankacılık sektöründe karşılaşılan risklerin gruplandırılması ve çeşitlendirilmesi çeşitli şekillerde yapılmıştır. Dalkılıç (2010) tarafından yapılan çalışmada 4 ana gruba bölünen riskler;

- piyasa riski,
- kredi riski,
- operasyonel risk,
- diğer riskler,

olarak adlandırılmış ve ana grupların altında alt-grup riskler incelenmiştir.

Öker (2007) tarafından yapılan çalışmada riskler 8 grupta incelenmiş ve bu gruplar:

- Likidite riski,
- kredi riski,
- faaliyet riski,
- operasyonel risk,
- faiz oranı riski,
- döviz kuru riski,
- piyasa riski,
- ülke riskidir.

Bankacılık Denetleme ve Düzenleme Kurumu (BDDK)'nın 3 ayda bir yayınlamakta olduğu Finansal Piyasalar Raporunda ele alınan riskler bu çalışmada Bankacılık sektörü tarafından maruz kalınan en önemli riskler olarak ele alınacak olup bu risklerin detaylı anlatımı yapılacaktır. Bu riskler;

- kredi riski,

- piyasa riski,
- yapısal faiz oranı riski,
- likidite riski,
- yoğunlaşma riski,
- sermaye yeterliliği riski,
- ülke riskidir.

3.2.1 Kredi riski

Kredi riski kullanılan kredinin geri ödenmeme durumunda ortaya çıkan bir risktir. Buradaki riskte, müşteriler ön planda olmakta ve geri ödemelerinde sıkıntıya düşebilecekleri ifade edilmektedir.

Kredi riskinin şu şekilde tanımlanması da mümkündür; kredi riski, banka müşterisinin yapılan sözleşme gereklerine uymayarak, yükümlülüğünü kısmen ya da tamamen zamanında yerine getirememesinden dolayı bankanın karşılaştığı durumu ifade eder. Kredi riskinin yönetilmesinin amacı; uygun parametreler içinde bankanın maruz kalabileceği riskleri yöneterek, bankanın risk ayarlı getirisini maksimize etmektir.

Kredi riski literatürde 2 boyutlu olarak incelenmektedir;

- risk miktarı,
- riskin kalite derecesidir.

Risk miktarı müşteriye kullanılan kredi miktarı olarak ifade edilmekte, riskin kalite derecesi ise temerrüde düşme ihtimali olan riskin temerrüde düşme ihtimalinin sayısal olarak ifade edilmesi sürecidir. Bankalar çeşitli teknikler kullanarak elde ettikleri bilgiler doğrultusunda risk kalitesini ölçmeye çalışırlar. Bankaların ellerindeki bu bilgiler; derece, sınıflama, müşterilerin eski bilgileri, müşterilerin finansal ve demografik bilgileridir. Geri ödememe durumunun irdelenmesinde aşağıdaki üç grup model olarak ele alınabilir, bu modeller (Yıldırım, 2007);

- kalitatif modeller,
- kredi skorlaması modeli,

- Newer Modelleridir.

3.2.2. Piyasa riski

Bu risk, bankaların sahip oldukları bir ya da birden fazla ticari varlığın işleme tabi tutulabileceği süre dâhilinde, piyasada meydana gelen beklenmeyen olumsuz dalgalanmaların sebep olduğu kayıp veya beklenenden düşük seviyedeki kar halini ifade eder. Piyasa riski herhangi bir zaman zarfında meydana gelebilir.

Piyasa riskini en aza indirmek piyasa disiplinin sağlanmasıyla mümkündür. Piyasa disiplini, piyasadaki kurumlarla ilgili bilgilerin zamanında doğru ve şeffaf şekilde alınmasını içerir. Bankacılık sektöründe piyasa disiplinin sağlanmasıyla birlikte piyasadaki ilgili birimler, çok daha sağlıklı değerlendirmeler yapabilecekler ve böylece istenmeyen riskler en aza indirilir.

3.2.3. Likidite riski

Bankaların nakit akışlarını iyi planlayamamaları nedeniyle nakit giriş ve çıkışlarında karşılaştıkları risktir. Bankalar, günlük faaliyetleri kaydi para şeklinde işlemler yaparak ilerletmektedirler. Kaydi paranın tanımı şu şekilde yapılabilir: Bankalar aracılığıyla efektif kullanmaksızın tedavül eden satın alma gücünü ifade eder. Hesaptan hesaba nakil yapmak, keşide etmek ve takas usulünden faydalanmak suretiyle gerçekleştirilen ödemelerin satın alma gücü kaydi paradır.” (BDDK,2011). Kaydi para ile yapılan işlemlerin yanı sıra bankalar ellerinde bir miktar nakit para bulundurmak zorundadırlar. Bunun amacı çeşitli müşterilerin günlük ihtiyaçlarını karşılamaktır. Bankalar, ellerindeki nakit paranın miktarını iyi bir şekilde yönetemezlerse müşterilerin ihtiyaçlarına cevap veremez duruma düşebilirler. Bu tür durumlarda oluşan riske likidite riski denilmektedir.

Likidite riski genelde varlık ve yükümlükler arasındaki vade farkı uyumsuzluğundan kaynaklanır. Bu durumla karşılaşan bankalar, ihtiyaç duydukları nakidi farklı kaynaklardan temin ederler. Ancak bu durumda, kaynaklardan uygun maliyetle temin edilip edilemediği konusu da önemlidir. Likidite riski ile karşılaşmak istemeyen bankaların yapması gereken, mevcut risklerini taşıyabilecek kadar kuvvetli bir sermaye yapısına sahip olup olmadıklarını iyi hesaplamaktır (Yıldırım, 2007).

3.2.4. Faiz riski

Faiz riski, risk oranlarında oluşan ters yönlü hareketlerin bir bankanın mali yapısında yarattığı etki olarak tanımlanmaktadır. Faiz riski, faiz oranlarındaki hareketler nedeniyle bankanın pozisyon durumuna bağlı olarak maruz kalabileceği zarar ihtimalinin ifade edilmesidir. Faiz riski, bankanın gelirleri ile ilişkilendirilip faiz oranı riskinin tanımı, faiz oranlarındaki hareketler nedeniyle gelirlerdeki azalma riski olarak da ifade edilmektedir.

Bankaların bulundurdıkları faiz riski, bir miktar arttırılarak karlılıkları arttırılabilir, fakat aşırı derecedeki faiz riskinin faiz oranlarındaki değişimler doğrultusunda bankaların faize duyarlı gelir ve gider kalemlerini etkileyebileceği unutulmamalıdır. Faiz oranlarının değişmesi banka varlıklarını, yükümlülüklerini ve bilanço dışı işlemlerini etkiler.

Faiz riskinin nedenleri çok farklı olabilir. Faiz riski; vade farklılıkları nedeniyle riskin yeniden fiyatlandırılması sonucu bilanço kalemlerinin fiyatlandırılması şeklinde olabileceği gibi, verim eğrisindeki değişmelerin ya da temel faiz riski aynı fiyatlandırma özelliklerine sahip farklı enstrümanlara uygulanan faiz oranlarının ayarlanmasında yapılan hatalı korelasyonlardan da kaynaklanabilir (Ataçoğlu, 2006).

3.2.5. Yoğunlaşma riski

Risk yoğunlaşması (temerküzü), bir risk tutarının veya grup halindeki risk tutarlarının, bankanın esas faaliyetlerini yürütebilme kabiliyetini tehdit edebilecek derecede yüksek (bankanın sermayesi, toplam aktifleri veya bütüncül risk seviyesi ile orantılı olarak) zararlara neden olabilme potansiyelidir.

Risk yoğunlaşmaları, bankaların aktiflerinde, pasiflerinde ya da bilanço dışı kalemlerinde, işlemlerin yürütülmesi veya gerçekleştirilmesi (ürün ya da hizmet) esnasında veya bu geniş kategorilerdeki risk tutarlarının farklı kombinasyonları şeklinde ortaya çıkabilmektedir. Kredilendirme faaliyeti, birçok bankanın en temel faaliyeti olduğundan kredi riski yoğunlaşmaları, banka içerisindeki en önemli risk yoğunlaşmasını oluşturmaktadır.

3.2.6. Ülke riski

Ülke riski, uluslararası kredi işlemlerinde, kredi alan kişi ya da kuruluşun faaliyette bulunduğu ülkenin ekonomik, sosyal ve politik yapısı nedeniyle yükümlülüğünü tamamen ya da kısmen yerine getirememesi anlamına gelmektedir.

Ülke riskinin ortaya çıkma nedenleri üç genel başlıkta sıralanabilir;

- ülkenin siyasi koşulları (savaş, işgal, isyan, yönetim sorunları),
- sosyal koşullar (sosyal dengesizlik, dinsel huzursuzluklar),
- ekonomik koşullardır (enflasyon, durgunluk, ödemeler dengesi açıkları).

Ülke riski, çeşitli yabancı bankalar ve özel derecelendirme kuruluşları tarafından ölçülmektedir. Ülke analizlerinde; siyasal rejim, yasal düzenlemeler, ülkenin üye olduğu uluslararası örgütlenmeler ilişkiler, büyüme, döviz rezervi, ödemeler dengesi, enflasyon oranı, ithalat-ihracat dengesi, dış borçlar, borç yükü, nüfus, istihdam, gelir dağılımı gibi değişkenler kullanılmaktadır. Ülke riski ölçümü, hem kantitatif hem de kalitatif yöntemler kullanılarak yapılmaktadır.

3.2.7. Sermaye yeterliliği riski

Sermaye yeterliliği riski bankaların ödeme yapma yeteneğini gösteren en önemli göstergelerden biridir. Basel Komitesi tarafından belirlenen bu risk oranı, bankalara kanuni zorunluluk getirilerek sermayelerini belirli bir seviyenin üzerine çıkarmalarını gerektirmektedir. 1988 yılında Uluslararası Ödemeler Bankası tarafından belirlenen minimum sermaye yeterlilik oranı % 8 olarak belirlenmiştir. Ülkemiz Bankacılık Denetleme ve Düzenleme Kurulu (BDDK), bankalar üzerindeki yaptırımları doğrultusunda bankaların sermaye yeterlilik oranını minimum % 12 olmasını istemektedir.

Yeterli miktarda sermayenin, ani likidite sıkışıklıklarında yedek bir güvence olarak bulunmasının, bankaların sorunlu dönemleri daha rahat aşmasına neden olacağı iddiası, sermaye miktarında bir standardizasyona gidilmesini savunanların temel dayanaklarından birini oluşturmaktadır. Ayrıca sermaye miktarı arttıkça, bankaların kaybedecekleri aratacağından, riskli yatırımlardan uzak durmaları ihtimali kuvvetlenir (Ünsal, 2007).

3.3. Kredi Riski Yönetimi

Bankalar bir önceki bölümde anlatılan risk kavramlarına maruz kalmaktadırlar. Bankaların esas amaçları kar etmektir. Bu doğrultuda almış oldukları ve alacakları kararları irdelerken, içsel ve dışsal çeşitli faktörleri dikkate alarak stratejiler geliştirmelilerdir. Diğer riskler dışında, kredi riski, bankalar açısından yönetilmesi en zor olan risktir. Kredi riski, kredi kullandırımıyla başlayıp kredinin kapatılması anına kadar geçen süre zarfında, banka tarafından her an karşılaşılabilecek bir risktir. Bu riskin tamamen bertaraf edilmesi gibi bir durum söz konusu değildir (Öker, 2007).

Kredi riski verilen kredinin geri ödenmeme durumu olarak açılırken, bu riskin doğmasını önlemeye yönelik politikalar geliştirilebilir. Bunlar;

- risk tutarının küçültülmesi yoluyla çeşitlendirilmesi,
- riskin sınırlandırılması: kredi limitlerinin belirlenmesi,
- riskin farklılaştırılması: sektör, müşteri ve kredi türü bağımlılığının azaltılması,
- risk primi uygulaması,
- riskin yansıtılması (Teminatlandırma),
- kredi değerlendirmesinde derecelendirme/skorlama yöntemlerinin uygulanması,
- erken uyarı sisteminin kullanılmasıdır.

Risk tutarının küçültülmesi yoluyla çeşitlendirilmesi: Bankalar toplam kredilerini daha küçük miktarlarda kullanarak daha fazla müşteriye kullandırması bu politikanın amacıdır. Burada hedeflenen, bankaların müşteri portföylerini çeşitlendirerek batık oranını düşürmeye çalışmalarıdır.

Riskin sınırlandırılması-kredi limitlerinin belirlenmesi: Bu politika sadece kredi kullanıcısı üzerine veya tek bir kredi türüne ilişkin risklerin önlenmesine yönelik bir politika olmayıp, temel olarak bankanın vereceği toplam kredilere üst sınırlar getirerek kontrolü sağlamaktır. Bu konuda, Bankalar Kanunu, yasal çerçeveyi belirleyen ana unsurdur. Kanunda, kullanıcıya ve kredinin türüne bağlı olarak banka özkaynakları ile ilişkilendirilmiş limitleri geçmemek kaydıyla kredi kullandırımı yapılabilmektedir.

Riskin farklılaştırılması–sektör, müşteri ve kredi türü bağımlılığının azaltılması: Risklerin kredi kullanım miktarları olarak değil de, nitelik olarak farklılaşmasını ve kredilerin farklı gruplara yayılması ve farklılaştırılmasını esas alan politikalardır. Örneğin; müşteri grupları, tüketiciler, küçük ticari müşteriler, büyük ticari müşteriler ve yatırım projelerinin finansmanı vb. müşteri grupları oluşturulur ve belirlenen bu alanlara ilişkin kredi limitleri belirlenirse hem kredi risk sınırlaması hem de alanlar/müşteriler anlamında risk farklılaştırması sağlanmış olacaktır. Risk farklılaştırmasının arkasındaki düşünce, Markowitz'in Menkul Kıymet Portföyü için geliştirdiği Portföy Teorisine dayanmaktadır. Bu teoriye göre; portföyün toplam riski, ancak portföyün birbirinden farklı, genel ekonomik gelişmeler karşısında aynı yönde ve derecede etkilenmeyen yatırımlardan oluşturulması durumunda kısmen azaltılabilir.

Risk primi uygulaması: Risk primi sisteminin iki temel amacı vardır. Bunlardan ilki, müşteri riskliliğine göre kredinin kullandırılıp kullandırılmaması kararının verilmesi, kredi kullandırılacaksa riskliliğin fiyata yansıtılmasıdır. İkincisi ise, bankanın plasmanlarının takibe düşmesini ve takibe düşen alacaklarının bankanın mali bünyesini etkilemesini önlemektir. Fiyatlamada temel ölçüt, kredi başvurusu yapan firmanın derecelendirme/skorlama notudur. Kullandırılacak olan kredi için, derecelendirme notu ve risk primlerini kapsayan bir matristen faydalanarak belirlenecek risk primi oranı kredi faizine eklenir.

Kredi değerlendirmesinde derecelendirme/skorlama yöntemlerinin uygulanması: Bankaların veya kredi kurumlarının özellikle kredi risk yönetimi politikalarına bağlı olarak kredi analiz sürecini olabildiğince objektif unsurlarla gerçekleştirmesi ve kredi karar sürecine destek sağlayıcı derecelendirme/skorlama modellerinin kullanımının sağlanması gerekmektedir. Çünkü kredi talebinde bulunan firmanın, krediyi geri ödememe riskinin belirlenmesi, kredi risk politikasının en önemli aracıdır.

Riskin yansıtılması: Kredi riskinin müşteriden alınarak teminat niteliğinde başka bir unsura alınması, yani kredilendirme ilkeleri içerisinde yer alan güvenlik ilkesinin uygulanmasına dayanır. Kredinin anapara veya faizinin ödenmemesi durumunda bankanın zararını tazmin edecek bir teminatın alınması ve böyle bir durumda bu teminata başvurarak zararın tamamen veya kısmen telafisi sağlanmaktadır.

4.KREDİ SKORLAMA

Kredi skarlama kredi iin bařvuru yapan kiřinin veya kuruluřun kredi alma yeterlilięini len bir puanlama sistemi yntemidir.

4.1. Kredi Skarlama Tarihesi

Latince kkenli kredi szcę, gvenmek anlamına gelen ‘credo’ kelimesinden tremiřtir. Gnmzde ise kredi, ‘bugn al daha sonra de’ anlamını tařımaktadır. Skarlama, gerek veya algılanan kalite doęrultusunda sayısal olarak durumların sıralanmasını iermektedir. Bu durumlar arasından, istenen dereceye ve hedeflere varılmasını iermektedir. Skarlalar, sayılar olabileceęi gibi harf veya semboller olabilmektedir.

Kredi skarlama, ilk defa 1958’li yıllarda Amerika’da kredi kart kullanıcıları ve mortgage kredileri bařvuruları iin kullanılmaya bařlanmıřtır. Bařlangıta istatistiksel modeller ile kredi vermenin zor olduęunu dřnen kuruluřlar, kredi skarlama modelleri ile karřılıklı olarak fayda saęlandıęını grnce fikirlerini deęiřtirmiřlerdir. Kredi kartı rn olarak yeni olması ve kredi kartı verme deneyimine sahip insanların az olmasından dolayı kredi skarlama ile giderlerin azaltıldıęı grlmř ve mřteri kredi kartı taleplerine cevap vermede daha hızlı davranıldıęı grlmřtir.

Kredi skarlama, istatistiksel yntemler kullanılarak mevcut verinin sayısal leklere indirgenip kredi kararının verilmesi olarak aıklanabilir. Kredi skarlama yntemleri, 1960 yıllarda kullanılmaya bařlamıř olup gnmze kadar geen srede hem kullanılan yntemler hem de uygulama alanları aısından geliřimini srdrmřtir. Klasik anlamda kredi skarlama riskin, gelirin, tepkinin ve kaparonun llmesinden oluřmaktadır (Anderson,2007).

4.2. Kredi Skorlamada Geleneksel Yaklaşım

Geleneksel yaklaşımda; kredi kullandıranlar kredi kullanan kişinin 5C'sine göre karar vermektedirler. Bunlar; karakter, kapasite, sermaye, teminat ve dış faktörlerdir. Bu yaklaşımda, kredi kullandıracak kişinin tecrübesi, geçmiş bilgilerinden yararlanması ve kredi kullanacak kişinin gelecekteki durumuna ilişkin görüşleri önemli bir yer tutmaktadır. Kredi skorlama metotları ile kredi kullandıracak kişi veya firmanın riskleri belirli modellerle minimize edilmektedir.

Kredi skorlama yöntemlerinin en büyük faydası müşterileri etkileyecek bir karar verme durumunda ortaya çıkmaktadır. Karar verme sürecinde kredi verecek kişi veya kurum değişik senaryo ve politikalara göre kredi verme eyleminde bulunarak; kabul etme/etmeme, kredi faizi, kredi süresi gibi konularda karar vermektedir. Kredi skorlama yaklaşımı, skorlama için yeterli derecede veriseti bulunmadığı durumlarda ve olası karın çok yüksek olduğu durumlarda, yerini kredi uzmanları tarafından geleneksel yöntemlere bırakmaktadır.

Kredi skorları kullanım alanlarına göre farklı isimler almaktadır. Bunlar;

- **Başvuru Skoru:** Yeni çalışılmaya başlanan müşteriler için yapılan bu skorlama tekniğinde, kullanılan veriler müşterilerin geçmiş anlaşmaları ve kredi bürolarından elde edilen veriler dâhilinde skorkart hazırlanmaktadır.
- **Davranış Skoru:** Mevcut müşterilerin hesaplarındaki hareketlere göre müşterinin davranışı (tutumu) belirlenip limit koyma, izin verme gibi eylemlerde bulunmak için kullanılmaktadır.
- **Tahsilat Skoru:** Tahsilat süreci için kullanılan bir skorlama yöntemidir.
- **Müşteri Skoru:** Birçok hesaptaki müşteri davranışlarının incelenerek müşterilerin hesaplarının yönetimi ve çapraz satış yapılabilmesi için kullanılmaktadır.
- **Büro Skoru:** Kredi büroları tarafından hazırlanan büro skoru, gecikme ve iflasların tahmini için kullanılan bir skordur..

Skorlama çalışmaları, amaçları doğrultusunda farklı isimlendirilmelerine rağmen çeşitli ortak özelliklere sahiptir.

1. Veri olarak müşteri, içsel ve dışsal veriler kullanılmaktadır.

2. Müşteri davranışlarının 4 şekli vardır. Bunlar; risk, gelir, tepki ve kaparodur.
3. Tüm bu skorlar, pazarlamada, yeni iş süreçlerinde, tahsilâtlarda, reklam kampanyalarında kullanılabilir (Anderson, 2007).

4.3. Kredi Skorlama Metotları

Kredi skorlama yöntemleri, tahmin edici metotlar(algoritmalar) aracılığıyla, geçmiş deneyimlerini kullanarak durumların gelecekte iyi mi kötü mü olacağını tahmin ederler. Kredi skorlama yaklaşımlarında farklı algoritma ve metotlar kullanılmasına karşın en kabul gören yaklaşım istatistiksel bir metot olan regresyondur. Kredi skorlama yöntemlerinin geliştirilmeye başlanmasından skorartların geliştirilmesine kadar farklı metotlar kullanılmıştır. Bu metotlar, günümüzde parametrik olan ve parametrik olmayan olmak üzere ikiye ayrılmaktadır. Parametrik metotlar, kullanılan veri üzerinde bazı varsayımları kabul ederken, parametrik olmayan yöntemlerde kullanılan veriler üzerinde herhangi bir varsayım bulunmamaktadır (Anderson, 2007).

Parametrik metotlar;

1. lineer regresyon,
2. diskriminant analizi,
3. lojistik regresyondur.

Parametrik olmayan metotlar;

1. karar ağaçları,
2. yapay sinir ağları,
3. K en yakın komşu yaklaşımıdır.

4.3.1. Lineer regresyon

En basit ilişkilerden bir tanesi lineer ilişkidir; bir değer artarken diğeri bir kat sayısı değerinde değişir ve değeri bilinmektedir. Fakat finansal piyasalarda lineer ilişkiye sahip değişkenlerin sayısı oldukça azdır. Buna karşın lineer ilişki bulunan değişkenlerde modelleme için lineer regresyon kullanılabilir.

Basit bir lineer regresyon formülü şu şekilde ifade edilebilir;

$$y_i = \beta_0 + \beta_1 x_{i1} + e_i \quad (4.1)$$

Lineer regresyon kullanılmasıdaki en büyük problem birçok varsayımı bulunmasıdır. Bunlar;

1. Lineerlik,
2. hedef değişkenin sürekli ve normal dağılıma sahip olması,
3. bağımsız hata terimi,
4. uygun değişkenlerin kullanılmasıdır.

Kredi skorlama modellerinde sonuç ikili bir çıktıdır. Bu doğrultuda lineer olasılık modeli şu şekilde formülleştirilebilir;

$$P(\text{Good})_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + e_i \quad (4.2)$$

Her i kayıtlı olasılığı, sabitlerin, değişkenin ağırlık ile çarpılması sonucunda elde edilen sonuç ile toplanması ile bulunur. Burada iki çeşit sonuç elde edilmektedir; 0 ve 1. Bu modellerde sonuç 1 ve 0 den daha büyük veya daha küçük olmasından dolayı lineer olasılık modelleri kredi skorlama yöntemi olarak eleştiri almaktadır. Regresyon modelinin ne kadar doğru çalıştığını anlamak için genel olarak 2 yöntem kullanılmaktadır. Bunlar standart hata ve R^2 'dir.

$$s_e = \sqrt{\sum (\hat{Y}_i - Y_i) / (n-1-k)} \quad (4.3)$$

Burada s_e =standart hata,

$$E = \hat{Y} - Y = \text{hata},$$

n= değerlendirilen örnek sayısı

k = kullanılan açıklayıcı değişken sayısıdır.

$$R^2 = 1 - (\sum (\hat{Y}_i - Y_i)^2 / \sum (\tilde{Y}_i - Y_i)^2) \quad (4.4)$$

R^2 değişkenlik katsayısı, model tarafından ne kadar hatanın açıklanıp açıklanmadığını göstermektedir. 0 ile 100% arasında değer almaktadır.

4.3.2. Diskriminant analizi

Diskriminant analizi iki veya daha fazla grup olması durumunda, sonuçları gruplara bölmek için kullanılan istatistiksel bir tekniktir. Diskriminant analizi, sınıflandırma

araçlarını kullanarak bir grup içerisindeki örnekler arasındaki mesafeyi kısaltmaya ve farklı gruplar arasındaki farkı maksimum seviyeye getirmeye çalışır. Diskriminant analizinde süreçlere aşağıda yer verilmiştir. Bunlar;

1. grupların tanımlanması,
2. model formunun tanımlanması (genellikle bazı regresyon modelleri kullanılmaktadır.),
3. seçilen istatistiksel teknik ile modelin türetilmesi,
4. test örneklerin kullanılması,
5. modelin kullanılmasıdır.

Kredi skorlama için kullanılan diskriminant analizleri iki gruplu yapıyı kabul etmektedirler, diğer gruplar yapıda kabul edilmemektedir. Örneklerin hangi gruplara dâhil olacakları, aldıkları skorlar doğrultusunda diskriminant modelinde belirlenmektedir.

Genelde kullanılan lineer diskriminant analizi, lineer olasılık modellerini kullanmaktadır. Bu yapı, küçük grupların kullanıldığı modellerde sınıflandırma yanlışları yapmaktadır. Bu nedenle her grup için eşit derecede örnek kullanılması gerekmektedir.

4.3.3. Lojistik regresyon

Lojistik regresyon kredi skorlama modellerinin gelişimi döneminde oldukça yaygın olarak kullanılan modeldir. Bunun nedeni kredi skorlama modellerindeki hedef değişkenlerin ikili (binary) olmasıdır. Lojistik regresyon maksimum olasılık tahmin sürecini kullanmaktadır (Anderson, 2007). Bu süreç;

- (1) bağımlı değişkenlerin logaritmik bir fonksiyona dönüştürülmesi,
- (2) hangi katsayılar olması gerektiği,
- (3) katsayı değişikliklerinin karar verilmesi ve logaritmik olasılığın maksimize edilmesine bağlıdır.

$$\ln(p(\text{Good})/ 1-p (\text{Good})) = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + e \quad (4.5)$$

Lojistik regresyon kullanılabilmesi için gerekli olan varsayımlar;

1. kategorik hedef değişken,

2. logaritmik odds fonksiyonunda lineer ilişki,
3. bağımsız hata terimi,
4. ilişkisiz tahmin ediciler,
5. uygun değişkenlerdir.

Günümüzde kredi skorlama modelleri geliştirilmesinde Lojistik regresyon en önemli yöntem olarak kabul edilmektedir. Bunun nedenleri;

1. ikili çıktıları sonuçlandırma için dizayn edilmesi,
2. sonuç olasılıklarının 0 ile 1 arasında kalması,
3. verilen bilgiler ile oldukça başarılı olasılık tahminleri yapabilmesidir.

Lojistik regresyon, diskriminant analizi ve lineer regresyon yöntemleri ile karşılaştırıldığında şu şeklide avantajları bulunmaktadır;

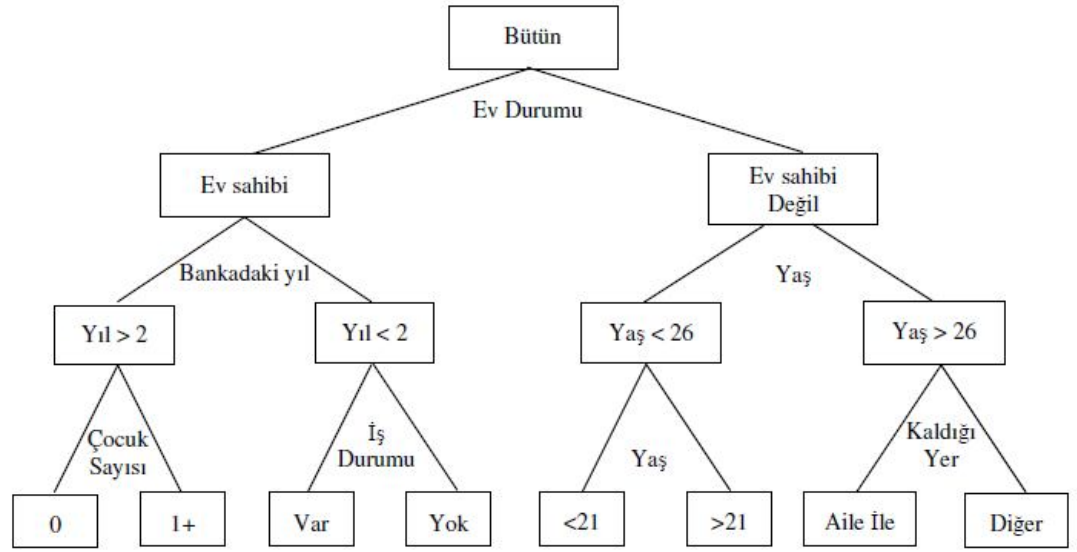
1. Lojistik regresyon kullanılması için bağımsız değişkenlerin normal dağılması gibi bir varsayıma ihtiyaç yoktur.
2. Grup büyüklükleri arasında büyük farklılıklar bulunması durumunda lojistik regresyon iyi çalışabilmektedir.
3. Birçok kişi tarafından sunulan modellerin oldukça anlaşılır olmasıdır.

4.3.4. Karar ağaçları

Karar ağacı, karar vericilere olayların olası sonuçlarını gösteren grafiksel bir araçtır. Karar ağaçları sınıflandırma problemlerinde ve tahmin problemlerinde de kullanılmaktadır. Gelişmiş yöntemlerin biri de data analizidir.

Karar ağaçlarına örnek olarak aşağıdaki grafik verilebilir;

Bu karar ağacı yukarıdan aşağı doğru irdelendiğinde en üst katmandaki yaş, aile ile yaşamak ve profesyonel olma kutucukları temel nodlar olarak tanımlanmaktadır. Ara kutucuklar; yaş, çocuk, eve sahip olma ve kaldığı yer ara nodlar ve en alt kutucuklar bitiş nodlar olarak tanımlanmaktadır. Karar ağacı bittiği durumlarda bitiş nodundan elde edilen skorlar kullanılmaktadır. Örnek bir karar ağacı yapısı Şekil 4.1’de görülmektedir.



Şekil 4.1 : Örnek bir karar ağacı yapısının şematik gösterimi (Anderson, 2007).

Karar ağaçlarının diğer tekniklere göre çeşitli dezavantajları ve avantajları bulunmaktadır;

1. Belirli bir kural setinin oluşması durumunda, bu teknik hızlı ve kolay anlaşılabilir yüksek veya alçak risk kategorilerini tanımlayabilmektedir.
2. Bilgisayar programları ile kullanımı oldukça basittir. Değişkenlerin seçimi ve karar ağacı yapısının oluşturulmasıyla işlem sona ermektedir. Fakat çok fazla esnekliğe sahip bir modelleme süreci değildir. Değişkenler üzerinde çok fazla değişiklik yaparak sonuç üzerine nasıl etki edeceği hususunda yeterli bilgi edinilememektedir.
3. Küçük veri setlerinin kullanılması sonucunda sonuçların güvenilirliği konusunda tereddütler yaşanabilmektedir. Modelin güvenilirliği konusunda endişe edilmemesi için büyük veri setleriyle çalışmak gerekmektedir.
4. Basit karar ağaçlarında sonuçların irdelenmesi oldukça basittir. Fakat karar ağacı yapısı karmaşıklaştığı durumlarda sonuçların irdelenmesi oldukça zordur (Anderson, 2007).

4.3.5. K - En yakın komşu yaklaşımı

En yakın komşu tekniği, ilk defa Fix ve Hudges tarafından geliştirilen sınıflandırma probleminde standart, parametrik olmayan bir yaklaşımdır. Bu teknik K – en yakın

komşu yaklaşımı ilk defa Chatterjee ve Barcun, daha sonra Henley ve Hand tarafından uygulanmıştır. Bu tekniğin dayandığı mantık, herhangi iki başvurunun birbirinden ne kadar uzak olduğunu ölçmek için başvuru veri uzayında bir mesafeyi seçmeye dayanmaktadır. Geçmiş başvuruların örnekleme temsili standart alınır. Yeni bir başvuru temsili örneklemede k kadar yakın başvurular arasındaki (yeni başvurunun en yakın komsuları) iyi-kötü oranına dayanarak iyi veya kötü olarak sınıflanır (Thomas, 2002).

Bu yaklaşımı uygulamak için üç tane parametreye ihtiyaç vardır: mesafe, en yakın komsular setini oluşturan kaç tane başvuru sayısı olduğu (k), ve bir başvurunun iyi olarak sınıflanabilmesi için başvuru iyi oranının ne olması gerektiği. Normal olarak, eğer komsuların çoğunluğu iyi ise, başvuru iyi olarak sınıflandırılır. Aksi taktirde başvuru kötü olarak sınıflanır. Ortalama varolan maliyet M , ve iyiyi reddetmenin ortalama kayıp karı K olarak tanımlansın. Eğer en yakın komsuların en azından $M/M + K$ tanesi iyi ise, yeni bir başvuru iyi olarak sınıflandırılır. Eğer yeni bir başvurunun iyi olma olasılığı iyi olan komşuların oranı ise, bu kıstas, beklenen kaybı minimize edecektir.

Mesafenin seçimi oldukça önemlidir. Fukunaga ve Flick genel bir mesafe tanımı yapmıştır;

$$d(x_1, x_2) = (x_1 - x_2)^T A(x_a) (x_1 - x_2)^{1/2} \quad (4.6)$$

$A(x)$, $p \times p$ simetrik sonlu pozitif bir matristir. Eğer x 'e bağlı ise, $A(x)$ yerel mesafe olarak adlandırılır, eğer x 'ten bağımsız ise global mesafe olarak adlandırılır. Yerel mesafenin eksikliği, genelde uygun olmayan deneme setinin özelliklerini dikkate alır. Bu nedenle birçok araştırmacı global mesafeye odaklanır. KS'de en yakın komşu yaklaşımının en detaylı uygulaması Henley ve Hand tarafından yapılmıştır. Bu teknik ile öklit uzunluğu ve iyi ile kötüyü en iyi ayıran yönün uzunluğu karışımına odaklanılır. Eğer w ; p boyutlu yön vektörü ise, Henley ve Hand'in mesafe ifadesi şöyledir ;

$$d(x_1, x_2) = \{(x_1 - x_2)^T (I + D w^T) (x_1 - x_2)\}^{1/2} \quad (4.7)$$

KS'de lineer ve lojistik regresyon yaklaşımları kadar çok sık kullanılmamasına karşın, en yakın komşu tekniği gerçek uygulamalar için bazı önemli özelliklere sahiptir. Dinamik olarak yeni olaylar eklemeyerek deneme setini güncellemek çok kolaydır ve eklenenin iyi veya kötü olduğu bilindiğinde o olay örneklemeden kolayca

çıkarılabilir. İlk seferde iyi bir mesafe bulmak, skorkart oluşturmada regresyon tekniğiyle hemen hemen eşdeğer bir netice verir. Böylece çoğu uygulayıcı bu noktada ilerlemeyi durdurup geleneksel bir skorkart kullanmayı tercih ederler. Sınıflandırma ağacı yaklaşımı ile kıyasladığımızda, en yakın komsu yaklaşımı her bir başvuranın özelliği için bir skor üretmez. Uygulayıcılar için bir denge noktası belirler ve onların gerçekte sistemin ne yaptığını anlamalarını sağlar.

4.4. Kurulan Modellerin Kontrolü

Elde edilen skorkartların güvenilirliklerinin test edilmesi için çeşitli istatistiksel yöntem ve modellerden faydalanılmaktadır. Elde edilen sonuçlar doğrultusunda skorkartların uygulanabilirliği konusunda bir kanıya varılmaktadır.

4.4.1. Diverjans istatistiği

Diverjans istatistiği, parametrik bir istatistik olup her iki grubun normal dağıldığını varsayar. İki grup arasında ortalamaların farkının karesini ortalama varyanslarına böler. Kredi skorlama modellerinde iki grup (iyiler ve kötüler) olması doğrultusunda formülasyon şu şekilde oluşur.

$$D^2 = (\pi_G - \pi_B)^2 / ((\sigma_G^2 + \sigma_B^2)/2) \quad (4.8)$$

Diverjans istatistiği sürekli ölçeğe sahip olan sonuçlara uygulanabilir, fakat dağılımın şekli konusunda herhangi bir bilgi vermemektedir. Çalışmalarda bu istatistiğin kullanılması ile çok fazla karşılaşmamaktadır;

1. Sürekli karakteristiğe sahip çıktılar üzerinde kullanılmaktadır,
2. İki grubun da dağılımlarının normal dağılıma uyduğu varsayılmaktadır,

Bu istatistiğin tüm modeller için kullanılması uygun olmamasına karşın lojit ve probit modellerden elde edilen skor dağılımları için kullanılması uygundur (Anderson, 2007).

4.4.2. Yanlış sınıflandırma matrisi

Tahmin edici modelin ne kadar iyi çalıştığını öğrenmek için kullanılan bir matristir. Oldukça basit bir mantığa sahip olup doğru sınıflandırılmış örnekler yüzdesini göstermektedir. Matris kurulurken şu aşamalar izlenir;

- a. Kesim skorunun seçimi,

- b. Kesim skoru üzerinde kalanların iyi altında kalanların kötü olarak kabul edilmesi,
- c. Karşılıklı olarak beklenen iyi ve kötülerin gerçekteki iyi ve kötüler ile tabloya aktarılması,
- d. Her hücreye düşen örneklerin yüzdeleştirilmesi,
- e. Modelden çıkarılabilecek oranların elde edilmesi.

Kesim oranının seçilmesi, bu matriste önemli bir rol oynamaktadır. Burada kullanılan yöntemlerden bir tanesi kötü orana sahip örneklem kümesi sayısı kadar noktaya kesim noktası seçilmesidir.

Yanlış sınıflandırma maliyeti modelin uygunluğu hakkında bilgi verebilmesine karşın toplam yanlış sınıflandırma maliyet hesaplanmadığı sürece modelin uygunluğu konusunda yeterli derecede bilgi alınamamaktadır. Yanlış sınıflandırma matrisinin eski skorkard ile yerine geçecek skorkardın karşılaştırılması için uygun bir metot olduğu öne sürülmüştür.

4.4.3. Kullback diverjans ölçüsü

Kullback Diverjans Ölçüsü, iki dağılımın frekanslarının farkını ölçmek için kullanılmaktadır. Bu ölçüt kanıtların ağırlığına (Weight of Evidence) dayanmaktadır. Bu ağırlık algılanan risk ile elde edilen bilgiler arasında bağlantı kurmaya yarayan bir araçtır (Siddiqi, 2006).

$$W_i = \ln N_i P_i - \ln N_P \quad (4.9)$$

Yukarıdaki denklem ile hesaplanır. P = pozitif olup/kötü, N = P'in tersi yani oluşmama/iyi oranlarını göstermektedir. Formülün ilk kısmı, ilgili grup için odds oranını gösterir. İkinci kısım ise örneklem ile ana kütle arasındaki sabit oranı gösterir.

4.4.4. Kolmogrov-Smirnov istatistiği

Kredi skorum ve diğer disiplinlerde oldukça kullanılan bir istatistiktir. Kolmogrov-Smirnov(KS) istatistiği, parametrik olmayan bir ölçüt olup bir eğri ile hatayı tayin etmektedir. KS eğrisi görsel bir veri aracı olup skorkardın etkiliğini ölçmektedir.

KS istatistiği şu şekilde formülize edilebilir;

$$D_{KS} = \max \{ \text{abs}(c_p Y - c_p X) \} \quad (4.10)$$

KS istatistiği, en çok, modelin tahmin edici gücünü ölçmede ve iki dağılımın birbirinden farklılaşıp farklılaşmadığını tespit etmede kullanılmaktadır. Hipotez testleri de ayrıca kullanılabilir. DKS noktası $c/\sqrt{n}$ formülü ile bulunmaktadır.

C anlam derecesine göre değişmektedir, n ise örneklem sayısıdır. Çoğu durumda dağılımın normal olduğu kabul edilmekte ve c 0.005 anlamlılık derecesinde 1.36 olarak alınmaktadır. DKS kritik noktası n'nin yani örneklem sayısının değişimine oldukça hassastır. Eğer örneklem sayısı arttırılır ise örneklem hatası düşer ve test daha başarılı sonuç elde eder.

4.4.5. Gini katsayısı

Gini katsayısı gelir dağılımındaki eşitsizliği ölçmek amacıyla kullanılan bir ölçüm değeridir. İtalyan ekonomist Corrado Gini tarafından geliştirilen bu katsayı 0 ile 1 arasında çıkan bir değerdir. Mutlak eşitlik doğrusu ile Lorenz eğrisi altında kalan alanın, mutlak eşitlik doğrusu altında kalan üçgenin alanına oranlanması ile bulunur. Gini katsayısı skorkart modellerinde, modelin, iyi kötü müşterileri ayırt etme gücünü ölçmek için kullanılan oldukça yaygın ve güvenilir bir yöntemdir. Gini katsayısının formülü aşağıda denklemde verilmiştir (Thomas, 2002):

$$D = 1 - \sum_{i=1}^n ((c_p Y_i - c_p Y_{i-1}) (c_p X_i + c_p X_{i-1})) \quad (4.11)$$

Formülde yer alan $c_p N_i$ iyilerin kümülatif yüzdelerini, $c_p P_i$ de kötülerin kümülatif yüzdelerini göstermektedir.

Gini değerinin kaç olması kabul edilebilir bir değerdir? Bu sorunun cevabı, skor kart modelinin başvuru ya da davranış skorkart modeli olmasına göre farklılık gösterebilir. Başvuru skorlamasında %50 den büyük bir değer çıkması tatmin edici sonuçlar içerdiğini gösterir. Gini katsayısının %30 dan az olması ise sonuçların yeterince iyi olmadığını gösterir.

4.4.6. Chi-square (χ^2) testi

Chi-square testi, iki karakteristik arasında lineer ilişkiye bakarak bir p değeri hesaplar. Bu p değeri modelin güvenilirliğinin bir ölçütüdür. İki karakteristik

arasında benzerlikler veya farklılık bulunma olasılığının bir şans ile ortaya çıktığı görüşünün karşıtıdır. Gözlenen ve beklenen dağılımları karşılaştırarak teorilerin ve hipotezlerin değerlendirilmesinde kullanılır. Bu testin farklı çeşitleri mevcut olmasına karşın en çok kullanılan Pearson Chi-Square testidir.

$$\chi^2 = \sum_{i=1}^n ((O_i - E_i)^2 / E_i) \quad (4.12)$$

O= gözlenen frekans (sıklık)

E_i= i. her sınıf için beklenen frekans ifade etmektedir.

X²=0 ise mükemmel uyumu göstermekte iken, bu sayısının büyümesi dağılımların farklılaştığını göstermektedir. Bu rakamın çok yüksek olması dağılımlar arasında bir ilişki olmadığını gösterir.

Chi-square sayısı olasılığa dönüştürülüp p harfi olarak ele alınmaktadır. X² 0 iken p 100% X² yükseldiğinde p değeri de sıfıra yaklaşmaktadır (Anderson,2007).

4.4.7. Alıcı çalıştırma karakteristiği(ROC)

Alıcı çalıştırma karakteristiği (ROC) 1940’lı yıllarda radar operatörleri tarafından doğru ve yanlış sinyalleri fark edebilmek için kullanılmıştır. Daha sonraları psikoloji alanında kullanılmaya başlanmış ve günümüzde mühendislik, tıp ve kredi skorlama alanlarında kullanılmaktadır.

Alıcı çalıştırma karakteristiği, “sinyal bulma” teorisine dayanıp iki kavrama dayandırılmaktadır;

1. Duyarlılık: Doğruyu bulma yeteneğidir.
2. Özgüllük: Yanlış bulma yeteneği olarak adlandırılmaktadır (Anderson,2007).

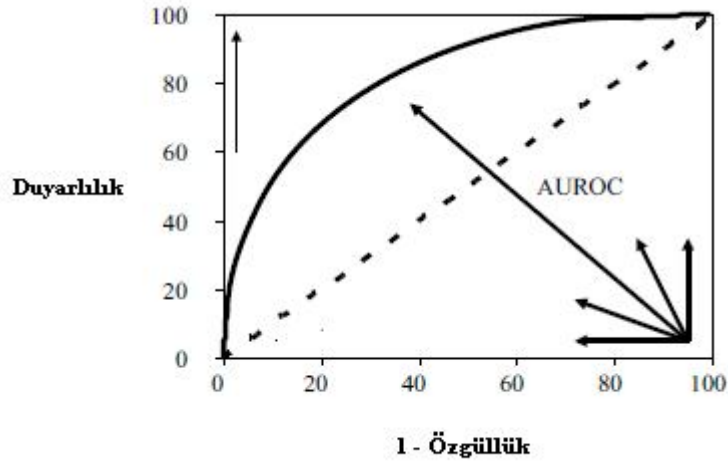
Alıcı çalıştırma karakteristiği eğrisi altında kalan alan 0 ile 1 arasında değer alır. Bu alan, modelin, hedef değişkende belirlenen grupları ne kadar iyi ayırabildiğinin bir göstergesidir.

Kredi skorlama çalışmalarına örnek bir sınıflama tablosu aşağıdaki çizelgede verilmiştir.

Çizelge 4.1: Sınıflandırma tablosu.

Tahmin Edilen	Gözlenen	
	y=0	y=1
y=0	a	B
y=1	c	D
Toplam	a+c	b+d

Tabloya göre duyarlılık $a/a+c$ özgüllük ise $d/(b+d)$ olarak hesaplanır. Belirlenen kesim noktasına göre bu değerler farklılık gösterebilir. Kesim noktası genellikle 0.50 olarak tercih edilmektedir. Alıcı çalıştırma karakteristiği eğrisi çizilirken duyarlılık y eksenine, $(1-\text{özgüllük})$ x ekseninde olacak şekilde çizilir. Şekil 4.2’de örnek bir ROC eğrisi gösterilmiştir.



Şekil 4.2: Alıcı çalıştırma karakteristiği(ROC)

Alıcı çalıştırma karakteristiği eğrisi, iyi ve kötü kredilerin birbirlerine karşı hareketlerine odaklanır. Örneğin müşterilerin iyi ya da kötü performanslı olma durumu, bağımlı değişken olarak kullanılan bir lojistik regresyon modelinde, müşterilerin gecikmeye düşme olasılıkları hesaplanmak istensin. Bu durumda, p , müşterinin kötü olma olasılığıdır. Duyarlılık ise lojistik regresyon sonucunda tahmin edilen grup üyelikleri içinde p ’nin doğru tahmin edilme oranıdır diyebiliriz. Özgüllük ise lojistik regresyon modeli sonucuna göre tahmin edilen ancak gerçekte gözlenmeyen durumların oranını ifade eder.

5. YAPAY SİNİR AĞLARI

5.1. Yapay Zeka

1950’li yılların başında ortaya çıkan “yapay zekâ” çalışmaları gelişen bilgisayarlar tarafından matematiksel formülize edilemeyen problemlerin çözümünde kullanılmaya başlanmıştır.

Yapay zekâ ile kurulan sistemlerin en temel özellikleri, olaylara ve problemlere çözüm üretirken veya çalışırken bilgiye dayalı karar verebilme özelliklerinin olması ve eldeki bilgiler ile olayları öğrenerek, sonraki olaylar hakkında kararlar verebilmeleridir.

Bilgisayar teknolojinin hızla gelişimi ile yapay zekâ sistemleri insan karar verme sürecine oldukça benzer hareket edebilmekte ve daha karmaşık problemlerde sonuçlar üretebilmektedir. Özellikle belirli bir algoritma ve formülasyon kullanılarak çözülemeyen problemlerin çözülmesi için yapay zeka sistemleri geliştirilmekte, problemin çözümü için geliştirilmiş bir algoritma mevcut ise geleneksel bilgisayar yöntemleri ile problem çözülebilmektedir.

Yapay zekâ sistemleri, literatürde zeki sistemler olarak adlandırılmaktadır. Burada en dikkat çekici nokta, bilgisayarların öğrenme yetileridir. Zeki sistemler için öğrenme “zaman içinde yeni bilgilerin keşfedilmesi yoluyla davranışların iyileştirme süreci” olarak tanımlanmaktadır. Zeki sistemlerde öğrenme şekilleri şu şekilde sıralandırılabilir (Öztemel, 2008);

- Alışkanlık yolu ile öğrenme,
- Görerek öğrenme,
- Talimatlardan öğrenme,
- Örneklerden öğrenme,
- Analoji yolu ile öğrenme,
- Açıklamalardan öğrenme,

- Deney yolu ile öğrenme,
- Keşfetmek yolu ile öğrenme.

Bilgisayara bilgilerin öğretilmesi için belirli paradigmlar ve yaklaşımlar geliştirilmiştir. Bunlardan önce çıkanlar şu şekildedir;

Sembol işleme Yöntemi: Bu sistemler, geleneksel yapay zeka teknolojisine dayanan öğrenme sistemleridir. Belirli formatlarda bilgisayar sunulan bilgilere dayanarak muhakeme yolu ile öğrenme gerçekleştirilir

Yapay Sinir Ağları: Bu sistemler örneklerden genellemeler yaparak öğrenirler.

İstatistiksel Örüntü Tanıma: Bu tür öğrenmede, bir verisetinin istatistiksel özellikleri ve dağılımı incelenerek veri hakkına genellemeler yapmak söz konusudur.

Genetik Algoritma ve Evrimsel Programlama: Bu paradigmda, bir problemin çözümü için, başlangıçta çözümleri rastgele atamak ve bu çözümlerden yeni çözümler üretmek ve daha sonra yeni çözümlerden daha iyi ve de yeni çözümler üretmek amaçlanmaktadır.

Vaka tabanlı Öğrenme: Bu teknoloji, örneklerden de öte, vakalara bakarak öğrenme esasına dayanmaktadır (Öztemel, 2008).

5.2. Yapay Sinir Ağları Yönteminde Temel Tanımlar

5.1.1. Yapay Sinir Ağları Yönteminin Tanımı

Beynin en küçük öğeler, sinir hücreleridir. Sinir hücreleri, bilgi işleme faaliyetini tek başına değil, bir grup oluşturarak sağlamaktadırlar. Bu oluşturulan gruba sinir ağı adı verilmektedir. Yapay sinir ağı, sinir hücresinin ve sinir ağının bilgisayar ile taklit edilmesi ile ortaya çıkan bir bilgi işleme sistemidir. Yapay sinir ağları birbiri ile bağlantılı olan, paralel bağlanan, her biri kendi bilgi işleme yeteneği ve belleği bulunan dağıtılmış sistemlerdir. Yapay sinir ağları, insan beynindeki yapay nöronlardan elde edilen benzer yapıdaki sistemlerdir. Yapay sinir ağları, biyolojik nöronlara dayanılarak geliştirilmiş bir matematiksel model ya da ölçümleme modelidir.

Yapay sinir ağları, çok değişkenli ve değişkenler arasında karmaşık, karşılıklı etkileşimin bulunduğu veya tek bir çözüm kümesinin bulunmadığı durumlarda başarılı sonuçlar üreten bir yapay zekâ teknolojisidir. Yapay sinir ağları, insan

beyninin çalışma ve düşünebilme özelliklerinden yola çıkılarak geliştirilmiş olan bir bilgi işlem teknolojisidir. Yapay sinir ağları, çok sayıda girdinin doğrusal olmayan sistemlerde başarılı bir şekilde kullanılmasını sağlayan, girdi-çıkı arasındaki ilişki örneklerinin modellenmesine dayanan bir yöntemdir. Yapay sinir ağlarının her biri, sürecin bir parçası olan nöronlardan oluşan, biyolojik sinir ağlarının fonksiyonları modellenerek geliştirilmiş sistemlerdir. Yapay sinir ağları yöntemi, bir sorunu çözmek için gerekli sayıdaki örnekle geçmişteki verilere dayanarak, modeller geliştiren istatistiksel yöntemdir (Bigus,1996).

Yapay sinir ağları bilgi işleme sistemi olup biyolojik sinir ağları gibi çalışmaktadır. Yapay sinir ağları insan sinir sisteminin matematiksel olarak modellenmesi vasıtasıyla geliştirilmiştir ve belirli varsayımları mevcuttur (Fausett,2004);

- Bilgi işlem, nöron adı verilen elemanlar vasıtasıyla olmaktadır.
- Sinyaller, nöronların aralarında bulunan bağlantılar vasıtasıyla ilerlemektedir.
- Her bağlantı, belirli bir ağırlığa sahip olup bu ağırlıklar sinyaller ile çarpılmaktadır.
- Her nöron bir aktivasyon fonksiyonuna sahip olup bu aktivasyon formülü vasıtasıyla çıktı sinyalini oluşturmaktadır.

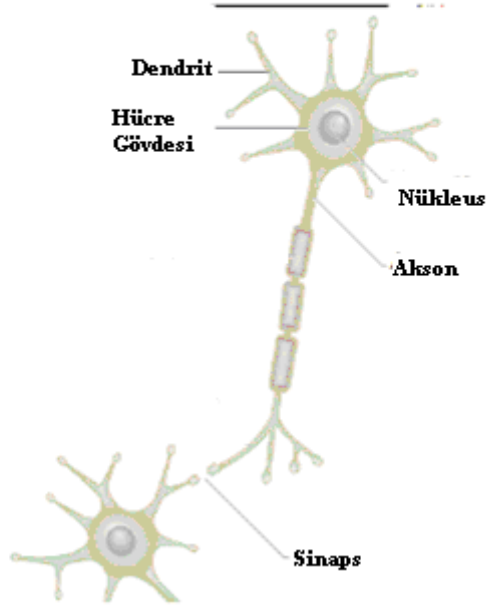
Yapay sinir ağları şu şekillerde karakterize olmaktadır;

- nöronlar arasındaki bağlantılar,
- bağlantılar arasındaki bağlantıların belirlenmesinde (eğitme, öğretme, algoritma),
- aktivasyon formülüdür (Fausett,2004).

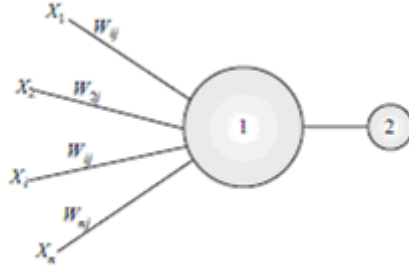
5.2.1. Biyolojik sinir ağları

İnsan beyni yaklaşık olarak farklı 1011 çeşit nöron(sinir hücresi) sahiptir. Bir nöronda, nükleus adı verilen yapı dendritler vasıtasıyla diğer nöronlar ile bağlantıyı sağlamaktadır. Nükleus aksonlar vasıtasıyla karmaşık kimyasal reaksiyonları elektriksel olarak nörona iletmektedir. Nöron tarafından gönderilen sinyal, eşik değere eşit veya eşik değerinin üstünde olursa nöronun sabit derece ve sürede elektrik sinyali göndermesini tetiklemektedir. Bu yolla nöronlar arasında mesajlar

ilerlemektedir (Dalkılıç, 2010). Şekil 5.1’de insan sinir hücresi yapısı, şekil 5.2’de yapay sinir ağları hücre yapısı gösterilmiştir.



Şekil 5.1 : Biyolojik sinir ağının gösterimi (Fausett,2004).



Şekil 5.2 : Yapay sinir ağının sistematik gösterimi (Fausett,2004).

5.3. Yapay Sinir Ağlarındaki Süreçler

5.3.1. Girdiler

Girdiler, bir yapay sinir hücresine dış dünyadan gelen bilgilerdir. Yapay sinir ağlarında girdiler işlenmemiş verilerdir. Yapay sinir ağları, girdi verilerine göre çıktı oluşturmaktadır. Girdilerin ağırlıklarına göre öğrenme işlevi yerine getirilerek, eğitim aşaması gerçekleştirilir (Öztemel, 2008).

5.3.2. Ağırlıklar

Ağırlıklar, gelen bilgilerin hücre üzerindeki etkisini belirleyen değerlerdir. Bilgiler, bağlantılar üzerindeki ağırlıklar üzerinden hücreye girmekte ve ağırlıklar yapay sinirde girdi olarak kullanılacak değerlerin göreceli kuvvetini (matematiksel katsayısını) göstermektedirler. Yapay sinir ağı içinde girdilerin hücreler arasında iletimini sağlayan tüm bağlantıların farklı ağırlık değerleri bulunur. Böylelikle ağırlıklar her bir işlemci elemanın her girdisi üzerinde etki yapmış olur. Ağırlıklar değişken veya sabit değerler olabilirler.

5.3.3. Aktivasyon fonksiyonu

Ağırlıklandırılmış girdilerin toplamı bir aktivasyon formülü kullanılarak çıktı değeri elde edilmektedir. Çoğu durumda aktivasyon işleminden geçen ağırlıklandırılmış girdi değerlerinin çıktısı -1 ile 1 arasında veya 0 ile 1 arasında değerler almaktadır. Çoğu çalışmada kullanılan aktivasyon formülleri aşağıdaki çizelgedeki gibidir (Bigus,1996). Örnek aktivasyon fonksiyonlarına aşağıdaki çizelgede yer verilmiştir.

Çizelge 5.1: Aktivasyon fonksiyonları

Aktivasyon Fonksiyonu	Matematiksel Tanımı
Lineer	$f(x) = x$
Lojistik Sigmoid	$f(x) = 1 / 1 + \exp(-x)$
Hiperolik Tanjant	$f(x) = \tanh(x)$
Gaussian	$f(x) = \exp(-x^2/2\sigma^2)$

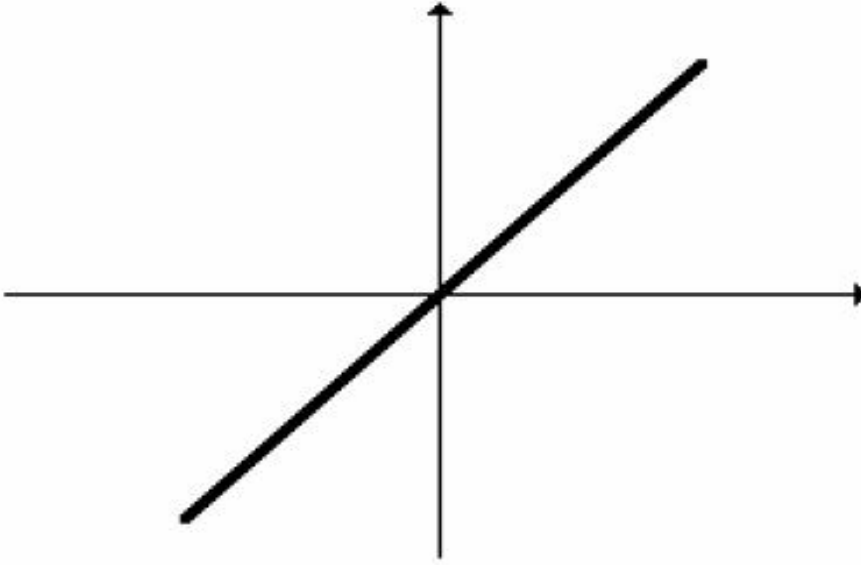
Yapay sinir ağından gerçek değerler olarak çıktı elde edilmesi isteniyorsa lineer aktivasyon formülü kullanılması uygun görülmektedir. Çok ara katmanlı yapay sinir ağlarında her ara katmanda lineer aktivasyon formülünün kullanılması lineer olmayan problemlerin hesaplanmasında sorun yarattığı için diğer aktivasyon formüllerinin kullanılması önerilmektedir.

5.3.3.1 Doğrusal aktivasyon fonksiyonu

Doğrusal problemlerin çözümünde kullanılan bu fonksiyon, gelen net girdileri doğrudan hücre çıkışı olarak vermektedir. Matematiksel olarak

$$F(s) = S$$

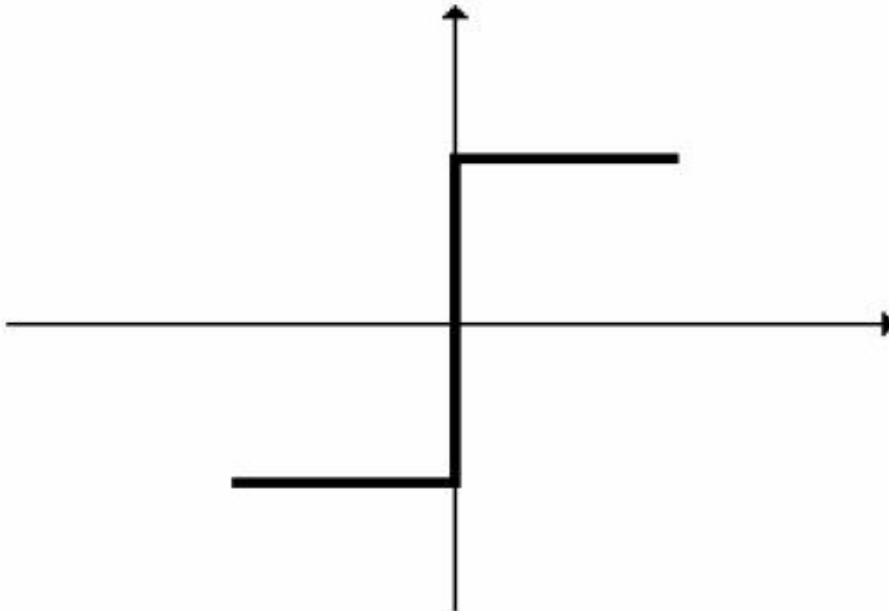
şeklinde tanımlanmaktadır.



Şekil 5.3 : Doğrusal aktivasyon fonksiyonunun şekilsel gösterimi.

5.3.3.2. Adım fonksiyonu

Gelen net girdi değerinin belirlenen bir eşik değerinin altında ya da üstünde olmasına göre hücrenin çıktısı 1 veya 0 değerlerini almaktadır. Adım fonksiyonunun şekilsel gösterimi aşağıdaki gibidir:



Şekil 5.4 : Adım fonksiyonunun şekilsel gösterimi.

5.3.3.2. Sigmoid aktivasyon fonksiyonu

Sigmoid aktivasyon fonksiyonu, türevi alınabilir, sürekli ve doğrusal olmayan bir fonksiyon olması nedeniyle uygulamada en çok kullanılan aktivasyon fonksiyonudur. Bu fonksiyon, net girdinin her değeri için 0 ile 1 arasında bir değer üretmektedir ve formülü şu şekildedir.

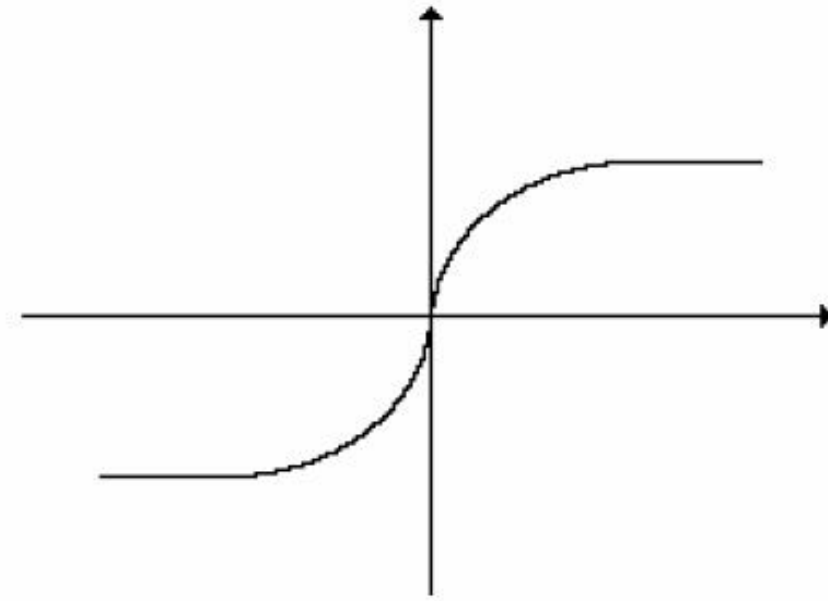
$$Y = F(s) = 1 / 1 + e^{-s} \quad (5.3)$$

5.3.3.4. Hiperbolik tanjant fonksiyonu

Hiperbolik tanjant fonksiyonu, gelen net girdinin tanjant fonksiyonundan geçirilmesi ile hesaplanmaktadır ve sigmoid aktivasyon fonksiyonunun farklı bir çeşididir. Sigmoid aktivasyon fonksiyonunda çıktı 0 ile 1 arasında bir değer alırken, hiperbolik tanjant fonksiyonunda çıktı -1 ile 1 arasındadır ve şu şekilde hesaplanır.

$$y = F(s) = e^s + e^{-s} / e^s - e^{-s} \quad (5.4)$$

Bu fonksiyonun şekilsel gösterimi aşağıdaki gibidir;



Şekil 5.5 : Hiperbolik tanjant fonksiyonunun şekilsel gösterimi.

5.3.4. Çıktı

Çıktılar, son süreç elemanlarıdır. Çıktı, aktivasyon fonksiyonu tarafından belirlenmektedir. Her nöron kendi çıktısını diğer bir nörona girdi olarak göndermektedir. Bir nörondan sadece bir çıktı değeri oluşmaktadır.

Yapay sinir ağında çıktı, problemin çözümüdür. Örneğin kredi başvurusunun değerlendirilmesinde çıktı olarak kredi değerlemesi; olumlu ya da olumsuz olarak sonuçlanacaktır. Yapay sinir ağlarında girdi gibi çıktılar da sayısal değerlerden oluşmaktadır. +1 olumlu, 0 olumsuz değerlendirilmektedir. Yapay sinir ağlarının amacı çıktı değerinin hesaplanmasıdır.

5.4. Yapay Sinir Ağlarının Sınıflandırılması

Yapay sinir ağlarının sınıflandırılması çeşitli şekillerde yapılabilmesine karşın genel olarak karakteristikleri, öğrenme paradigmalarına yani ağların eğitilme aşamasında kullanılan yöntem ve ağdaki bilgi akışının şekline göre oluşmaktadır.

5.4.1. Öğrenme paradigmalarına göre sınıflandırılması

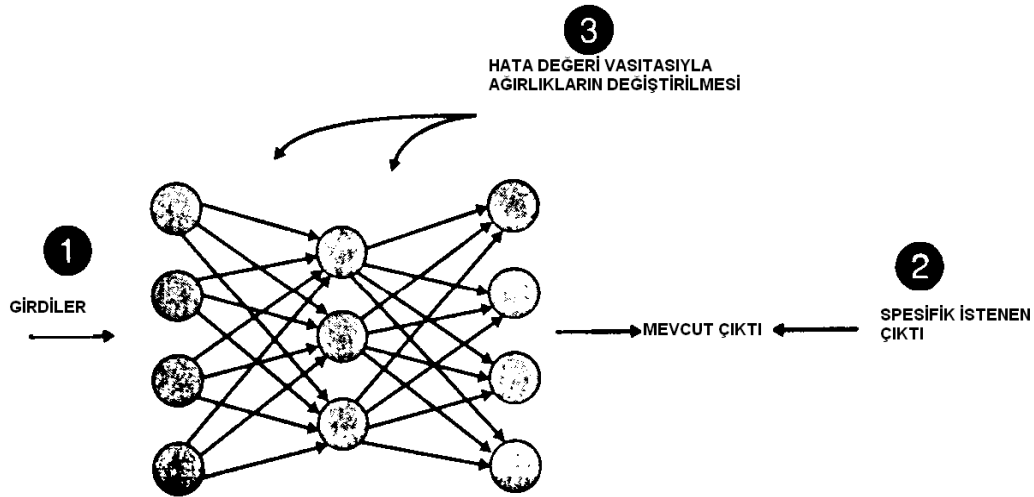
Yapay sinir ağlarını sınıflandırılmada en çok kullanılan yöntemlerden bir tanesi öğrenme paradigmalarına göre sınıflandırmaktır. En temel olarak 3 çeşit öğrenme paradigması tanımlanılabilir;

- Danışmanlı öğrenme,
- danışmansız öğrenme,
- takviyeli öğrenmedir.

Bu üç öğrenme şeklini detaylı açıklamadan önce bir yorumları yapılabilir; Danışmanlı öğrenme, anneden yeni bir iş öğrenmeye benzetilebilir. Her denemenin ardından gözlemci tarafından geri bildirimler yapılarak yapılan işin daha iyi olması için çabalanır. Danışmansız öğrenmece bir yığın dokümanın verilmesi ve bu yığınlar arasından bir şema çıkarılması isteğine dayandırılabilir. Takviyeli öğrenme, gerçek hayattaki bir işe sahip olma gibidir. Birçok iş için tercih yapılması gerekir ve tercihler doğru yapıldıkça sonunda ödül vardır. Her şey sana bağlıdır, başarılı ve başarısız tercihlerde geri bildirimler yapılır fakat hangi seçimlerin doğru veya yanlış olduğu senin tarafından belirlenir (Kamruzzaman, Begg, Sarker, 2006).

5.4.1.1. Danışmanlı öğrenme

Danışmanlı öğrenme örnekler ile programlamaya benzetilmektedir. Bu yaklaşımda yapay sinir ağına bir örnek veya problem verilir ve yapay sinir ağından bir sonuç çıkarması veya sınıflandırma yapması beklenir. Bu aşamada, danışman tarafından geri bildirimler ile çıkan çıktılar değerlendirilir. Öğrenme algoritması çıkan çıktılar ile istenen çıktılar arasındaki farkı değerlendirir. Bu aşamadan sonra ağ kendi içinde yapay sinir ağlarının ağırlıklarını tekrar istenen çıktı değeri doğrultusunda tekrar yapılandırır. Bu sayede istenen çıktı değerine olabildiğince yaklaşılmaya çalışılır. Fakat bu süreç uzun sürebilmektedir. Bazı durumlarda yapay sinir ağının istenen çıktı değerine ulaşabilmesi veya yaklaşabilmesi için 20 belki 100 defa örnek gösterilmesi gerekebilmektedir (Kamruzzaman, Begg, Sarker, 2006). Danışmanlı öğrenme modelinin şematik gösterimine şekil 5.6'deyar verilmiştir.



Şekil 5.6 : Danışmanlı öğrenme modeli.

Eğer gerçek sonuç ve örnekler biliniyorsa, danışmanlı öğrenmenin ne gibi faydaları vardır, sorusu sorulabilir. Eğer birçok örnek mevcut ve veri tabanı çok büyük ise yapay sinir ağı danışmanlı öğrenme uygulanarak eğitilir, ekstra bir programlamaya gereksinim olmadan aynı sonuçları kendisi bulabilmektedir.

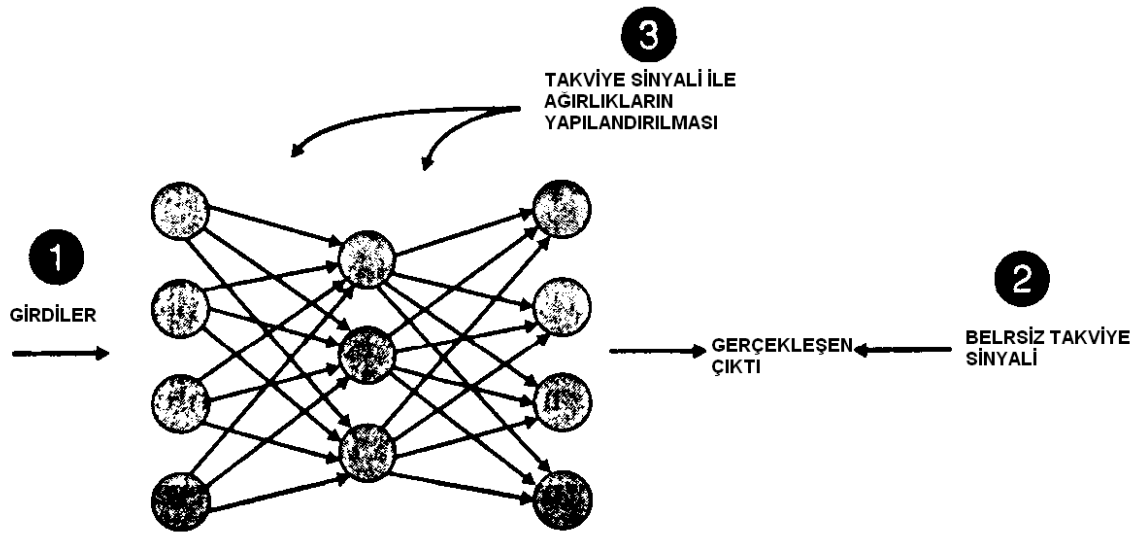
5.4.1.2. Danışmansız öğrenme

Danışmansız öğrenme, birçok verinin olduğu ve cevapların doğru bilinmediği durumlarda tercih edilmektedir. Burada sorunun cevabı bilinmezken, soru bilinmektedir. “Veriler arasında bir bağlantı mevcut mudur?” gibi sorulara cevap

aranmaktadır. Burada istenen, yapay sinir ağından istenen verinin şablonuna bakması ve benzer şablonlarda olanları aynı grup içinde kümelemesidir. Danışmansız öğrenme ile yapay sinir ağları bu görevi yapabilmektedirler. Burada verinin yapay sinir ağına doğru bir şekilde beslenmesi ve yapay sinir ağının verileri önem derecelerine göre ayırt etmesi gerekmektedir. Ayırma işlemi yapıldıktan sonra istenen analizler yapay sinir ağları vasıtasıyla yapılabilmektedir (Kamruzzaman, Begg, Sarker, 2006).

5.4.1.3. Takviyeli öğrenme

Takviyeli öğrenmede, problemler veya örnekler bilinirken doğru cevap bilinmemektedir. Bir oyunda bir hamle yapıldığı ve karşı tarafın da karşılık olarak bir hamle yaptığı düşünüldüğünde 10 veya 20 hamle sonunda, iki taraftan biri kazanıp diğeri kaybettiği düşünüldüğünde, burada elde edilen bilgi yapay sinir ağlarından takviye sinyali olarak adlandırılmaktadır. Belirli kararlar sonucunda başarılı olup olmadığımız ortaya çıkmaktadır. Takviyeli öğrenme ile kurulan yapay sinir ağlarının zamana bağlı değişen problemlerin çözümünde etkili olduğu görülmektedir. Takviyeli öğrenme modelinin şematik gösterimine şekil 5.6 yer verilmiştir.



Şekil 5.7 : Takviyeli öğrenme modeli.

Eğer geri bildirim bilgisi kesin olarak biliniyorsa yapay sinir ağının danışmanlı öğrenme ile eğitilmesi takviyeli öğrenmeye göre daha kısa ve daha az masraflı olacaktır.

5.4.1.2. Yapay sinir ağlarında öğrenme kuralları

Öğrenme kuralı; girdi değerlerine dayanarak ağırlıkları değiştiren bir matematik denklemdir. Öğrenme kuralı, ağıla ilave veri girdikçe sinir ağlarının daha güvenilir çıktılar üretmesine imkan verir. Öğrenme kurallarından bazıları şunlardır:

5.4.1.2.1. Hebbian öğrenme kuralı

Bu kurala göre, eğer bir nöronun girdi ve çıktısı aktif diğer bir ifade ile aynı işaretli ise, nöronlar arası ağırlık artar, dolayısıyla nöronlar arasındaki ilişki güçlenir, aksi takdirde çıktı aktif değil, ancak girdi aktif ise ağırlık azalır. Hebbian kuralı, diğer tüm öğrenme algoritmalarının temeli sayılır. Günümüzde en çok bilinen yapay sinir ağı öğrenme algoritması olan “(hata) geri yayılım algoritması”nın temeli de bu kurala dayanmaktadır.

5.4.1.2.2. Delta kuralı

Delta Kuralı, gerçek çıktı ile istenen çıktı arasındaki farkı ifade eden ağıl hatasını minimize etmek için bağlantı ağırlıklarının sürekli değiştirilmesi ilkesine dayanır. Ayrıca bu kural, Windrow-Hoff öğrenme veya en küçük ortalamalar karesi olarak da adlandırılır.

5.4.1.2.3. Gradyan iniş kuralı

Bu kural, Delta kuralına benzer, hatta aktivasyon fonksiyonunun türevi bağlantı ağırlıklarına uygulanmadan önce, Delta hatasını düzeltmek için kullanılır. Giriş verilerinin güçlü bir modelden çıkarılmadığı uygulamalarda, bu kural özellikle önemlidir. Bu kurallara dayanarak geliştirilen çeşitli optimizasyon (öğrenme) algoritmaları ve tekniklerine ilişkin detaylı bilgi ilerdeki bölümlerde verilmektedir.

5.4.2. Yapay sinir ağlarının yapılarına göre sınıflandırılması

5.4.2.1. İleri beslemeli ağlar

İleri beslemeli bir ağıda işlemci elemanlar (İE), genellikle katmanlara ayrılmışlardır. Her bir katmandaki hücreler sadece bir önceki katmanın hücrelerince beslenir. İleri

beslemeli YSA'da, hücreler katmanlar şeklinde düzenlenir ve bir katmandaki hücrelerin çıkışları, bir sonraki katmana ağırlıklar üzerinden giriş olarak verilir. Giriş katmanı dış ortamdan aldığı bilgileri hiçbir değişikliğe uğratmadan orta (gizli) katmandaki hücrelere iletir. Bilgi, orta ve çıkış katmanında işlenerek ağ çıkışı belirlenir. İşaretler, giriş katmanından çıkış katmanına doğru tek yönlü bağlantılarla iletilir. İE'ler bir katmandan diğer bir katmana bağlantı kurarlarken, aynı katman içerisinde bağlantıları bulunmaz. İleri beslemeli ağlara örnek olarak, Çok Katmanlı Algılayıcı (Multi Layer Perceptron –MLP) ve Doğrusal Vektör Parçalama (Linear Vector Quantization – LVQ) ağları verilebilir (Alavala,2002).

5.4.2.2. Geri beslemeli ağlar

Bu çeşit sinir ağlarının dinamik hafızaları vardır ve bir andaki çıkış, hem o andaki hem de önceki girişleri yansıtır. Bundan dolayı, özellikle tahmin uygulamalarında kullanılırlar. Bu ağlar çeşitli tipteki problemlerin tahmininde oldukça başarı sağlamışlardır. Bu ağlara örnek olarak; Hopfield, Düzenleyici Harita (Self Organizing Map – SOM), Elman ve Jordan ağları verilebilir (Alavala,2002).

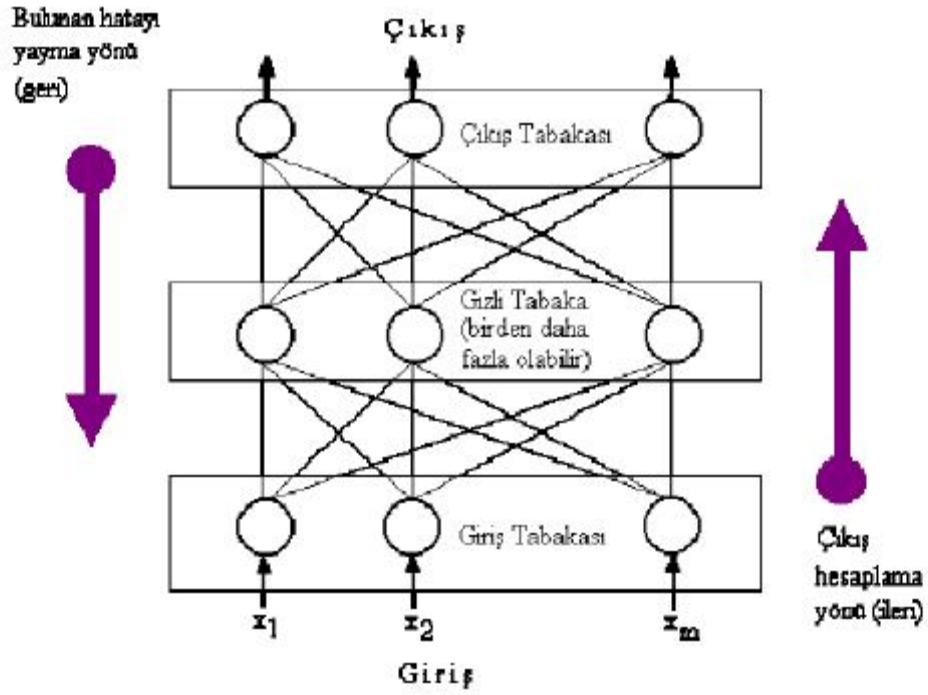
5.5. Çok Katmanlı Algılayıcılar

Çok katmanlı algılayıcı yapay sinir ağları, mühendislik uygulamalarında en çok kullanılan sinir ağı modellerinden biridir. Birçok katmanlı algılayıcı (perceptron modeli; bir giriş, bir veya daha fazla ara ve bir çıkış katmanından oluşur. Bir katmandaki bütün işlem elemanları bir üst katmandaki bütün işlem elemanlarına bağlıdır. Bilgi akışı ileri doğru olup geri besleme yoktur. Bu nedenle ileri beslemeli sinir ağı modeli olarak adlandırılmaktadır.

Ağ yapısında, giriş katmanında herhangi bir bilgi işleme yapılmaz. Buradaki işlem elemanı sayısı tamamen uygulanan problemin giriş nöron sayısına bağlıdır. Ara katman sayısı ve ara katmandaki işlem elemanı sayısı ise deneme-yanılma yolu ile bulunur. Çıkış katmanındaki eleman sayısı ise yine uygulanan probleme dayanılarak belirlenir.

Bu karar ağacı yukarıdan aşağı doğru irdelendiğinde en üst katmandaki yaş, aile ile yaşamak ve profesyonel olma kutucukları temel nodlar olarak tanımlanmaktadır. Ara kutucuklar; yaş, çocuk, eve sahip olma ve kaldığı yer ara nodlar ve en alt kutucuklar bitiş nodlar olarak tanımlanmaktadır. Karar ağacı bittiği durumlarda bitiş nodundan

elde edilen skorlar kullanılmaktadır. Çok katmanlı yapay sinir ağlarının çalışma prensibine şekil 5.7’de yer verilmiştir.



Şekil 5.7 : Çok katmanlı algılayıcı.

Çok katmanlı algılayıcı yapay sinir ağlarında ağa bir örnek gösterilir ve örnekler sonucunda beklenen sonuçlar ağa verilir(danışmanlı öğrenme). Örnekler giriş katmanına uygulanır, ara katmanda işlenir ve çıkış katmanından da çıkışlar elde edilir. Kullanılan eğitim algoritmasına göre, ağın çıkışı ile arzu edilen çıkış arasındaki hata geriye doğru yayılarak, hata minimuma düşünceye kadar ağın ağırlıkları değiştirilir.

5.5. Yapay Sinir Ağlarında Geliştirici Unsurlar

5.5.1. Öğrenme hızı

Öğrenme katsayısı, yani öğrenme hızı faktörü, eğitim örüntüsü ile ağın çıktı örüntüsünü birbirine yaklaştırmak için ağırlıkların ayarlanmasında kullanılmaktadır. Kullanıcı tarafından seçilen öğrenme hızı çok düşük olduğu durumlarda durumunda ağırlıklarda değişimin çok yavaş olacak ve dolayısıyla ağ çok yavaş öğrenecektir. Hızın büyük olması durumunda ağırlıklarda büyük miktarda değişimlere neden

olabilir, bu durumda ağın performansını azaltıcı etkide bulunduğu gözlenmektedir (Dalkılıç, 2010).

5.5.2. Momentum katsayısı

Momentum katsayısı ağı öğrenmesi esnasında, yerel bir optimum noktaya takılıp kalmaması için ağırlık değişim değerinin belirli bir oranda bir sonraki değişime eklenmesini sağlar. Momentum katsayısının kullanılması, ani sıçramaları ortadan kaldırma eğilimi gösterecektir, ancak her zaman işe yaramayabilir ve hatta yakınsamaya zarar verebilir. Önceden de ifade edildiği gibi, momentum katsayısı bir önceki parametre değişiminin belli bir oranının her iterasyonda bir sonraki parametre değişimine eklenmesi ile gerçekleşir. Çoğunlukla deneme-yanılma ile belirlense de uygulamada 0.9 civarındaki bir momentum katsayısının iyi sonuçlar verdiği görülmüştür (Dalkılıç, 2010).

5.5.3. Gizli katman sayısı

Yapay sinir ağlarında en önemli noktalardan bir tanesi gizli katmanda kaç nöron olacağına karar vermektir. Gizli katman nöron sayısı, öğrenme sırasında bellekte fazla bilgi barındırmak haricinde öğrenme işleminin daha iyi yapılmasını da sağlamaktadır.

Gizli katman nöron sayısı bulunması için çeşitli yöntemler öne sürülmüştür. Bunlar $2n+1$, $2n$, n , $n/2$ gibi girdi katmanı nöron sayısına oranlı gizli katman sayılarıdır.

Çelebi ve Bayraktar, yaptıkları çalışmalarda gizli katmandaki nöron sayısının belirlenmesinde genel bir kuralın bulunmadığına dikkat çekmişlerdir.

Gizli katman, girdi katmanının aldığı ağırlandırılmış veriyi probleme uygun bir fonksiyonla işleyerek bir sonraki katmana iletir. Bu katmanda gereğinden az nöron kullanılması girdi verilerine göre daha az hassas çıktı elde edilmesine sebep olur. Gereğinden çok sayıda nöron kullanılması durumunda aynı ağda yeni tip veri gruplarının işlenmesinde zorluklar ortaya çıkar ve uyum aşımı olur.

Gizli katman sayısının aşağıdaki formül ile bulunduğu çalışmalar da mevcuttur (Bayru,2007).

$$N_s = (\frac{1}{2}(N_g + N_c) + \sqrt{N_d}) / N_b \quad (5.5)$$

Burada

N_s : Gizli katmandaki nöron sayısını,
 N_g : Girdi katmanındaki nöron sayısını,
 $N_ç$: Çıktı katmanındaki nöron sayısını
 N_d : Gözlem sayısını
 N_b : Katman sayısını göstermektedir.

Çalışmaların çoğunda gizli katman sayısının deneme yoluyla bulunduğu görülmektedir. Çalışmada hem deneme yoluyla bulunan gizli katman nöron sayısı hem de formül ile elde edilen gizli tabaka nöron sayılı yapay sinir ağları birbiri ile karşılaştırılacaktır.

5.5.4. Hata toleransı

Yapay sinir ağlarının en iyi sonucu yakaladıklarına ilişkin kesin bir bilgi yoktur. Bu nedenle yapay sinir ağı kullanan kişiler ağı performansını ölçerken “e” kadar bir hatayı kabul etmektedirler. Bu hataya “Hata toleransı ” adı verilmektedir.

Herhangi bir hata toleransının altındaki noktada ağı öğrenmiş olduğu kabul edilir. Bu noktalara lokal çözümler denilmektedir. En iyi çözüm olmamalarına rağmen kabul edilebilir hata düzeyinin altında bir hataya sahip olduğundan kabul edilebilir çözümler olarak ele alınabilirler.

6. KREDİ SKORLAMA LİTERATÜR ÇALIŞMALARI

Literatürde kredi skorlama çalışmaları, konusu ve konumu itibarıyla çeşitli çalışmalarda yer almıştır. Finansal kuruluşlarda öne çıkan 2 çeşit skorlama sorunu mevcuttur. Birincisi, yeni kredi başvurusunda bulunan müşterilere kredi verilip verilmeyeceği konusunda cevap vermek için geliştirilen kredi skorlama metotları, ikincisi de mevcut müşterilerin durumlarını irdelemek amacıyla kullanılan davranış skoru metotlarıdır.

1980'ler itibarıyla Amerika'da kredi kartı kullanımının ve kredi kartı taleplerinin artması doğrultusunda her bir başvuruyu irdelemek ile sorumlu danışman veya görevlilerin bu iş yükünün altından kalkmakta zorlanmaları ile bankalar kredi skorlama metotlarını kullanmaya başlamışlardır.

2000'li yılların başından itibaren kredi skorlama metotları olarak en çok kullanılan lineer diskriminant analizi ve lojistik regresyon yanında çeşitli diğer teknikler araştırmaların konusu olmuş ve bu istatistiksel yöntemler ile karşılaştırılmaya başlanmıştır. İstatistiksel metotların çeşitli nedenlerle eleştiri konusu olması literatür çalışmalarında diğer yöntemlerin kullanılması konusunda destekleyici olmaktadır. Lineer diskriminant analizinde, iyi ve kötü kredi gruplarının kovaryans matrislerinin eşit olmamasının modelin performansını düşürücü etkiye sahip olduğu öne sürülmüştür. Lojistik regresyonun kredi skorlama metodu olarak kullanılmasında değişkenlerin herhangi bir dağılıma uymaları gibi bir varsayım bulunmaz, fakat literatürde, lojistik regresyon yönteminde kullanılan değişkenler arasında lineer bağlantı olacak şekilde dizayn edilmesi, kredi skorlama problemleri için eleştiri kaynağı oluşturmıştır.

Literatür çalışmaları incelendiğinde, 2000'li yılları ile birlikte uzman sistemler aracılığıyla kredi skorlama çalışmalarının yoğunluk kazandığı görülmektedir. Yapay sinir ağları ile kredi skorlama modelleri üretme konusu, 2000'li yılların başında oldukça yoğun bir şekilde irdelenmiştir. Daha sonraları, karar ağaçları, çok değişkenli uyarlanabilir regresyon uzanımları, çoklu vektör makineleri ve kaba setler gibi yöntemler kredi skorlama yöntemleri olarak kullanılmıştır.

Yapay sinir ağıları ile kredi skorlama modelleri, üzerinde oldukça çalışılmış bir konu olarak gözükmektedir. Kredi skorlama modeli olarak yapay sinir ağlarının en büyük özellikleri; değişkenler arasında herhangi bir varsayıma gereksinim olmaması, ikili (binary) sonuçlar için uygun bir model olmasıdır. Bu yöntemin en büyük eleştirisi aldığı konular; en iyi yapay sinir ağının elde edilmesinin uğraştırıcı olması ve bir kredi talebinin ret veya kabul edilmesinde yapının herhangi bir bilgi vermemesidir. Bu nedenle yapay sinir ağları siyah kutu (“Black Box”) olarak adlandırılmaktadır.

Yapılan literatür çalışmalarında istatistiksel ve istatistiksel olmayan yöntemlerin karşılaştırılması için 1. tip hata, 2. tip hata ve doğruluk oranı diye adlandırılan 3 parametre kullanıldığı görülmektedir.

1.tip hata = Temerrüde düşmemiş kredilerin model sonucunda ret edilme olasılığıdır.

2.tip hata = Temerrüde düşmüş kredilerin model sonucunda kabul edilme olasılığıdır.

Doğruluk oranı= Modelden doğru tahmin edilen kredilerin tüm kredilere oranıdır.

6.1. Yapay Sinir Ağları Üzerine Yapılan Çalışmalar

West tarafından yapılan çalışmada, çeşitli yapay sinir ağları kullanılarak elde edilen kredi skorlama modelleri ile 3 ayrı verisetinde modeller oluşturulmuştur. Kullanılan yapay sinir ağları; ileri beslemeli yapay sinir ağı, çok katmanlı algılayıcı, uzman ağların karışımı, yarıçapsal tabanlı ağlar, ileri beslemeli yapay sinir ağları olup bu modellerden elde edilen sonuçlar lineer diskriminant analizi, lojistik regresyon ve karar ağaçları ile karşılaştırılmıştır. Çalışmada iki ayrı veriseti kullanılmıştır. 1000 örneklem sayısı (700 iyi kredi-300 kötü kredi) ve 24 değişken sayısı ile Alman veriseti ve 690 örneklem sayısı (307 iyi kredi-383 kötü kredi) ve 14 değişken ile Avusturya kredileri verisetidir. Yapay sinir ağlarının yapısı, gizli nöron sayısı, momentum katsayısı ve öğrenme katsayısı gibi bilgilere çalışmada yer verilmemiştir. Elde edilen sonuçlarda, genel hata fonksiyonunda yapay sinir ağları ile elde edilen sonuçların iyi olduğu gözlenmiş ve en düşük hata fonksiyonuna uzman ağların karışımı yapay sinir ağı ile elde edilmiştir. Lojistik regresyon ile elde edilen model hata fonksiyonu sıralamasında ikinci sırada yer alıp çok katmanlı algılayıcıya oranla daha başarılı bir performans sergilemiştir. Çalışmada, yapay sinir ağlarından yarıçapsal tabanlı ağı, lineer diskriminant analizi ve karar ağacı (CART) ile elde edilen sonuçların bu veri setlerinde kötü olduğu vurgulanmıştır (West, D. (2000)).

Chen ve Huang tarafından yapılan çalışmada, yapay sinir ağı, lineer diskriminant analizi ve karar ağacı (CART) kredi skorlama metotları karşılaştırılmıştır. Çalışmada geri beslemeli yapay sinir ağı kullanılıp en iyi ağ yapısı elde edilmeye ve yapay sinir ağının kötü kredi olarak tanımladığı kredileri neden bu şekilde tanımladığı anlaşılmaya çalışılmıştır. Çalışmada kullanılan veriseti “UCI Repository of Machine Learning Databases”’den alınmış olup 307 iyi kredi ve 383 kötü krediden oluşan bir verisetidir. Değişken sayısı 15 olup çeşitli gizli katman sayıları, momentum katsayıları ve öğrenme katsayıları denenerek en iyi ağ yapısı bulunmaya çalışılmıştır. En iyi ağ yapısına 15 girdi katman, 28 gizli katmanlı ve 1 çıktı katmanlı ağ yapısı ile ulaşılmıştır. Elde edilen yapay sinir ağı ile lineer diskriminant analizi ve karar ağacı sonuçları karşılaştırıldığında yapay sinir ağının iyi kredileri ayırt etmede, bu iki yöntemle göre daha kötü bir performans gösterdiği fakat kötü kredileri ayırt etmede daha başarılı olduğu görülmüştür. Yapay sinir ağından elde edilen sonuçlarda kabul edilmeyen krediler genetik algoritma yardımıyla analiz edilmiş ve hangi değişkenlerin kredi başvurusunun kabul edilmemesinde daha önemli olduğu anlaşılmaya çalışılmıştır. Müşteri tercihi ve algılanan maliyet arasında bir ilişki kurulmaya çalışılmıştır (Chen, M. C. ve Huang S. H. (2003)).

Malhotra tarafından yapılan çalışmada iki teknik birleştirilerek kredi skorlama yöntemi geliştirilmeye çalışılmıştır. Bulanık yapay sinir ağı ANFİS ile diskriminant analizi yoluyla elde edilen kredi skorlama modelleri karşılaştırılmıştır. Veriseti 790 adet (500 iyi kredi- 290 kötü kredi) gözlem içerip yapay sinir ağının eğitimi için 500 kredi sınama için 290 kredi kullanılmıştır. Elde edilen sonuçlarda, iyi kredilerin belirlenmesinde diskriminant analizinin, ANFİS yapay sinir ağına göre daha başarılı olduğu fakat genel ve kötü kredi belirlemede yapay sinir ağının daha iyi bir performans sergilediği görülmüştür (Malhotra, R. ve Malhotra, D.K. (2002)).

Yapay sinir ağı ile yapılan başka bir çalışmada, Kim ve Sohn, yapay sinir ağı ile kredi skorlama modeli kurmuş ve elde edilen sonuçlar ile 4 alt müşteri grubu tanımlamışlardır. Bu gruplardan yola çıkılarak çeşitli yönetsel yorumlarda bulunulmuştur. Yapay sinir ağının türü ve ağ yapısı hakkında bilgi verilmemiştir. Veriseti “UCI Repository of Machine Learning Databases”’den alınmış olup 700 adet iyi, 300 adet kötü krediyi içerdiği bilinmektedir. Elde edilen gruplar; iyi kredi olup iyi kredi seçilenler, kötü kredi olup kötü kredi seçilenler, iyi kredi olup kötü

kredi seçilenler, kötü kredi olup iyi kredi seçilenlerdir. Grupların özelliklerine göre yönetimsel değerlendirmeler yapılmıştır (Kim, Y. S. ve Sohn, S.Y. (2004)).

Malhotra tarafından yapılan diğer bir çalışmada, geri beslemeli yağay sinir ağı ile çoklu diskriminant analizi, kredi skorlama modelleri olarak karşılaştırılmıştır. Çalışmada veriseti, 12 farklı kredi kuruluşundan alınan 1078 adet krediden oluşmaktadır. Bunlardan 700 adedi, yapay sinir ağı için eğitim setinde kullanılmış, 378 adeti sınaama aşamasında kullanılmıştır. Çalışmada ağların öğrenme katsayısı 0.02 ve momentum katsayısı 0.01 olarak alınmıştır. Çalışmada yapay sinir ağının öğrenme aşamasında 7 alt örnek grubu farklı döngü ve farklı gizli katman sayıları ile eğitilmiş ve bu 7 gruptan elde edilen sonuçların ortalaması alınmıştır. Elde edilen sonuçlarda birinci tip hata, ikinci tip hata ve genel hata düzeyinde geri beslemeli yapay sinir ağının, çoklu diskriminant analizine göre daha iyi bir performans sergilediği görülmüştür (Molhotra, R. ve Malhotra, D.K. (2003)).

Abdou, Pointon ve El-Masry tarafından yapılan çalışmada, gerçek veri setleri kullanılarak diskriminant analizi, lojistik regresyon, probalistik yapay sinir ağı ve çok katmanlı algılayıcı gibi tekniklerle kredi skorlama modelleri oluşturulmuş ve performansları karşılaştırılmıştır. Çalışmada dikkat edilen unsurlar; literatürde probalistik yapay sinir ağları ile çok fazla kredi skorlama modelinin oluşturulmaması ve gerçek verisetleriyle çalışılmamasıdır. Mevcut veriler Mısırdaki yerel bir bankadan alınmış ve 581 adettir. 433 adet iyi kredi ve 148 adet kötü kredi mevcuttur, 20 adet değişken ile modeller kurulmuştur. Çalışmada verisetinin %80'i eğitim seti olarak kullanılmış, %20'si sınaama veriseti olarak kullanılmıştır. Çalışmada kullanılan 20 değişken 12 değişkene düşürülmüştür. Bunun nedeni bazı değişkenlerin özdeş değerler içermesidir. Önce, her bir yapay sinir ağı metodu 20 kere çalıştırılmış ve en iyi 5 adet sinir ağı seçilmiştir. Daha sonra, bu yapay sinir ağları farklı adet nöron sayıları denenerek modeller oluşturulmuştur. Elde edilen değerlerde, klasik istatistiksel metotlar arasında en iyi performansı lojistik regresyon, en iyi yapay sinir ağlarında çok katmanlı algılayıcı elde etmiştir. Farklı nöronlu çalışmalar lojistik regresyon ile elde edilen değerlerden daha yüksek olmasına karşın çok katmanlı algılayıcıdan daha düşüktür (Abdou, H., Pointon, J. ve El-Masry, A. (2008)).

6.2. İleri Düzey İstatistiksel ve Uzman Sistemler Üzerine Yapılan Çalışmalar

Ong, Huang, Tzeng tarafından yapılan çalışmada kredi skorlama yöntemi olarak kullanılan yapay sinir ağlarının değişkenler arasında lineerlik, herhangi bir lineerlik olma veya olmama gibi bir durumdan etkilenmemesine rağmen küçük veri setlerinde etkili bir şekilde yapıyı iyi modelleyemediği öne sürülmüştür. Bu açıdan bu çalışmada genetik algoritmalar vasıtasıyla kredi skorlama yöntemi geliştirilmiştir. Bu model karar ağaçları, lojistik regresyon ve yapay sinir ağlarından çok katmanlı algılayıcı ile karşılaştırılmıştır. Veriseti “UCI Repository of Machine Learning Databases” alınmış olup, iki ayrı verisetinde yapılan çalışmalardan birinci veriseti Avurturya kredi bilgileri içermekte olup 383 örnekten 307 tanesi iyi kredi 76 tanesi kötü kredidir. Bağımsız değişken sayısı 14’tür. Alman veriseti 700 adet kredi kaydından oluşmakta, 300’ü kötü kredidir. 20 adet değişken mevcuttur. Çalışmada hata oranlarının hesaplanması için 5 ayrı alt küme test edilmiş ve ortalama hata oranı o modelin hata oranını yansıtmıştır. Yapılan çalışmada en üstün performansı gösteren model, genetik algoritma ile elde edilen kredi skorlama modelidir. Yapay sinir ağları, lojistik regresyona göre daha iyi performans göstermesine rağmen aralarındaki fark oldukça azdır (Ong, C.S., Huang, J.J. ve Tzeng, G.H. (2005)).

Tsa, , Wu tarafından yapılan çalışmada iki konu üzerinde durulmuştur. Birincisi tekli sınıflandırmaya göre çoğul sınıflandırma metoduyla eğitilen çok katmanlı algılayıcı yapay sinir ağının, kredi skorlama modeli olarak daha iyi bir performans gösterebileceğidir. İkincisi ise çoğu çalışmada, birinci tip hata ve ikinci tip hata olarak adlandırılan kredi skorlama modelini değerlendirmesinin de kullanılan hata fonksiyonlarının kullanılmaması olup, bu çalışmada kullanılmasıdır. Çok katmanlı yapay sinir ağları üzerinde yapılan çalışmada yapay sinir ağları birden fazla ağın gruptandırılarak eğitilmesi ile elde edilen sonuçlar ile normal olarak eğitilen çok katmanlı yapay sinir ağları karşılaştırılmak istenmiştir. Bu çalışma da üç farklı veriseti üzerinden yapay sinir ağları oluşturulmuştur. Bunlar Avustralya 690 kredi (307 iyi, 383 kötü kredi) 14 değişken sayısı, Almanya 1000 kredi (700 iyi, 300 kötü kredi) ve 20 değişken sayısı, Japonya 690 kredi (307 iyi, 383 kötü kredi) ve 15 değişken sayısı veri setleridir. Verisetlerinin %70’i eğitim aşaması için, %30’u sınav aşaması için kullanılmıştır. Her veriseti için çok katmanlı algılayıcı ağının eğitiminde farklı epochlar (50, 100, 200, 300), farklı gizli katman sayıları (8,16,24,32) kullanılmıştır. Elde edilen sonuçlarda, çoklu yapay sinir ağların normal

eđitilen yapay sinir ađlarına oranla daha iyi bir performans verdiđi gözlemlenmiřtir. Bunun nedeni olarak, eđitim verisetinin küçük parçalara bölünmesinin gruplandırılarak eđitim yapılan ađ yapısını kötü oranda etkilediđi gözlemlenmiř ve ikili (binary) sınıflandırma problemlerinde tekli sınıflamanın daha etkili olduđu söylenmiřtir (Tsai, C. F. ve Wu, J.W. (2008)).

Bellotti ve Crook tarafından yapılan alıřmada, oklu vektör makineleri kredi skortlama metodu olarak kullanılmıřtır. 2000’li yıllar itibarıyla ortaya ıkmıř olan bir algoritma olup sınıflama problemlerinde kullanılmaya bařlanmıřtır. Kredi skortlama tekniđi olarak kullanılmaya bařlanması 2000’li yılların bařı ile olmuřtur. Bu alıřmada, farklı oklu vektör makineleri algoritmaları, klasik k-en yakın komřu, lojistik regresyon ve lineer diskriminant analizi ile karřılařtırılmıřtır. alıřmada, 2004 yılında kullandırılan 2500 adet veriseti ile alıřılmıř ve kredi kullandırım tarihinden itibaren 3 ile 12 ay ierisinde batan krediler kötü krediler olarak ele alınmıřtır. Elde edilen sonularda ok büyük farklılıklar gözetilmeksizin bazı oklu vektör makineleri algoritmalarının iyi performans verdiđi gözlemlense bile lojistik regresyon ile elde edilen deđerler ile farklılařma ok yüksek olmamıřtır (Bellotti, T. ve Crook J. (2009)).

Lee, Chiu, Chou, Lu tarafından yapılan alıřmada amalanan regresyon karar ađaları (CART) ve ok deđerkenli uyarlanabilir regresyon uzanımları(MARS) ile kredi skortlama yöntemi geliřtirilmesi ve elde edilen sonuların eřitli teknikler ile karřılařtırılmasıdır. Kullanılacak olan CART ve MARS yöntemlerinin lineer diskriminant analizi, lojistik regresyon ve yapay sinir ađlarına göre avantajları olarak 3 nokta belirtilmiřtir.

- 1) Deđerkenler arasında bir iliřki bulunması gibi varsayımlarda bulunmamaktadır.
- 2) Modellerin kurulmasının ardından mevcut deđerkenler arasından önemli olanlar ayırt edilebilmekte ve eđitim süreci iin ok uzun süreye gereksinim bulunmamaktadır.
- 3) Sonulanan sınıflandırma modelleri kolay bir řekilde yorumlanabilir.

Yapılan alıřmada, Tayvan’dan elde edilen 8000 âdet kredi bilgisi kullanılmıřtır. 4000 âdeti modellerin oluřumunda, 2000 âdeti skortlamanın belirlenmesinde ve 2000 âdeti sınama ařamasında kullanılmıřtır. Yapay sinir ađları modeli iin, geri dönüřüm algoritmalı yapay sinir ađı kullanılmıř, ara katmandaki nöron sayıları ve öđrenme

oranı çeşitli denemeler yapılarak en iyi yapay sinir ağı modeli elde edilmeye çalışılmıştır. Yapay sinir ağı yapısı için çeşitli gizli katman sayısı ve öğrenme oranı denenmiştir. En iyi sonuç 20 ara katmanlı ve öğrenme oranı 0.005 olarak elde edilmiştir. Yöntemlerin karşılaştırılmasında sınıflama ve regresyon karar ağaçları ve çok değişkenli uyarlanabilir. Regresyon uzanımlarının en iyi sonuç verdiği görülmüştür. Yapay sinir ağının hata 1 terimi için en iyi sonucu verdiği görülürken hata 2 terimi için en iyi sonucu çok değişkenli uyarlanabilir regresyon uzanımları vermiştir. Doğruluk oranında en yüksek skoru sınıflama ve regresyon karar ağaçları elde etmiştir (Lee, T. S., Chiu, C. C., Chou, Y. C. ve Lu C. J. (2006)).

Lee, Chiu, Lu tarafından yapılan çalışmada, hibrid bir yapı olarak lineer diskriminant analizi ile kredi skorlama modelinin ilk aşamada kurulması ve bu modelden elde edilen anlamlı değişkenler geriye yayılım algoritmali yapay sinir ağı kurulmuştur. Çalışmada, Tayvan'daki özel bir bankadan elde edilen kredi bilgileri kullanılmıştır. 6000 adet kredi bilgisi içinden 4000 (%66) tanesi eğitim aşamasında, 2000 (%33) tanesi onaylama aşamasında kullanılmıştır. Bağımsız değişken sayısı 9 adettir. Çalışmada çeşitli ağ yapıları, ara katmanda gizli nöron sayısı (17-18-19-20-21) ve öğrenme oranı (0.02-0.04-0.06) değiştirilerek en iyi ağ yapısı bulunmaya çalışılmıştır. Geri yayımlı yapay sinir ağında en iyi yapı 9 girdi katmanı nöron sayısı, 19 ara gizli katman nöron sayısı ve 1 çıktı katmanı olarak 0.04 öğrenme oranıyla elde edilmiştir. Değişkenler olarak müşteri yaşı, adı, cinsiyeti, medeni durumu, eğitim statüsü, ikamet durumu, işi, yıllık geliri, kredi limiti, ikamet durumları alınmış hibrit modelde 7 adet değişken kullanılmıştır. Bunlar; cinsiyet, oturduğu yer, kredi limiti, yıllık geliri, ikamet durumu, yaş ve eğitim statüsüdür. En iyi hibrid ağ yapısı 7 girdi katmanı nöron sayısı, 16 ara gizli katman nöron sayısı ve 1 çıktı katmanı olarak 0.06 öğrenme oranıyla elde edilmiştir. Kullanılan geri beslemeli yapay sinir ağının iyi kredilerin seçiminde lojistik regresyon ve diskriminant analizine oranla daha başarılı olduğu fakat kötü kredilerin ayırt edilmesinde daha başarısız olduğu bilgisi elde edilmiştir. Hibrid yapı ise lineer diskriminant analizi ve lojistik regresyona göre daha iyi 1. ve 2. tip hata oranına sahip olup yalnızca 1.tip hata oranında geri yayımlı yapay sinir ağından daha kötü bir performans göstermiştir (Lee T.S., Chiu, C.C. ve Lu, C.J. (2002)).

Sustersic , Mramor , Zupan tarafından yapılan çalışma da amaçlanan, çok büyük ve çok değişkenli verisetlerinde, uygun değişkenlerin seçimi ve yapay sinir ağı ile kredi

skorlama modeli oluşturmaktır. Çalışmada, değişkenlerin seçiminde genetik algoritma yöntemi kullanılmış ve kredilerin iyi ve kötü krediler olarak ayırt edilmesi Kohonen ağı yöntemi tarafından yapılmıştır. Daha sonra elde edilen veriseti geri yayımlı yapay sinir ağı ve lojistik regresyon yöntemleri ile ayrı olarak kredi skorlama modelleri oluşturulmuştur. Verisetinde 84 adet değişken olduğu bilinmekte fakat verisetinin büyüklüğü bilinmemektedir. Genetik algoritma ve kohonen ağı ile yapılan çalışmalar ile değişken sayısı 21'e düşürülmüş, 383 adet eğitim veriseti ve 183 adet sınav veriseti çalışmada kullanılmıştır. Geri yayımlı yapay sinir ağından ve lojistik regresyon modellerinden elde edilen sonuçlarda 1. tip hata için lojistik regresyon ile elde edilen sonuçların daha iyi bir performans gösterdiği fakat 2. tip hata ve genel doğruluk derecesi açısından geri yayımlı algoritmalı yapay sinir ağından daha iyi bir performans gösterdiği görülmüştür (Sustersic, M., Mramor, D. ve Zupan J. (2009)).

Wang , Ma , Huang , Xu tarafından yapılan çalışmada karar ağaçları çeşitli gruplama teknikleri ile kredi skorlama modeli oluşturulup lineer regresyon, lineer diskriminant analizi, çok katmanlı algılayıcı ile oluşturulan kredi skorlama modelleri ile karşılaştırılmıştır. Gruplama tekniklerinin kullanılmasındaki neden, gürültü (noise) verilerin etkisini azaltıp daha iyi bir tahmin gücü olan model üretmektir. Gruplama (ensemble) tekniklerinin ardında yatan fikir, her bir bileşen modelinin özelliklerini, veri kümesinde bulunan farklı örnekleri yakalamak için kullanmaktır. Çalışmada 2 ayrı veriseti kullanılmıştır. Avusturya verisetinde 14 değişken ve 690 kredi mevcut olup, bunların 307'si iyi 382 kötü kredidir. Alman verisetinde 20 adet değişken ve 1000 adet örnek mevcut olup 700 adet iyi kredi ve 300 adet kötü kredidir. Yapay sinir ağı yapısı hakkında bilgi verilmemiştir. Her iki veriseti ile elde edilen sonuçlarda, gruplama metodu ile elde edilen karar ağaçlarının normal karar ağacına göre hem 1.tip hatada hem de 2.tip hatada daha başarılı olduğu görülmüştür. Gruplama metodu ile elde edilen karar ağaçlarının toplam doğruluk oranında diğer metotlar ile karşılaştırıldığında daha iyi bir performans gösterdiği görülmüştür (Wang, G., Ma, J., Huang L. ve Xu K. (2011)).

6.3. Türkiye’de Kredi Skorlama Üzerine Yapılan Çalışmalar

Kredi skorlama modellerinin geliştirilmesinde farklı parametrik ve parametrik olmayan metotlar kullanılmaktadır. Bu konu üzerinde ülkemizde de çeşitli akademik

alışmalar yapılmaktadır. En yaygın olarak kullanılan kredi skortama metodu olarak lojistik regresyon metodu ne ıkmaktadır.

Yapılan yksek lisans tezlerinde elde edilen veri setleri, lojistik regresyon ile modellenmiř ve elde edilen modellerin doęrulukları sınanmıřtır. Eliuz tarafından yapılan alışmada ticari bir bankadan elde edilen veriseti, 10 baęımsız deęiřken iermektedir. Elde edilen sonular bakımından alıcı alıřtırma karakteristięi (ROC) eęrisi kullanılmıř ve elde edilen sonular doęrultusunda modelin ayırt etme gcnn kabul edilebilir olduęu grlmřtr (Eliuz, E. (2009)).

Kasapoęlu tarafından yapılan alışmada, veriseti, ticari bir bankadan elde edilmiř olup 11 adet baęımsız deęiřken ile lojistik regresyon modeli oluřturulmuřtur. İyi kredilerde iyi kredi tahmin etme oranı, iyi kredilerde kt kredi tahmin oranı, kt kredilerde kt kredi olma oranı ve kt kredilerde iyi kredi olma oranı bu alışmada verilmiřtir. Ayrıca alıcı alıřtırma karakteristięi (ROC) eęrisi modelin anlam gcnn aıklayabilmek iin kullanılmıřtır. Elde edilen sonularda, Gini katsayısı 0,95 bulunarak modelin ayırt etme gcnn ok iyi olduęu saptanmıřtır (Kasapoęlu, B. (2009)).

zdemir tarafından yapılan farklı bir skorkart alıřmasında, veriseti, ticari bir bankadan alınmıř olup 15 adet baęımsız deęiřken iermektedir. Elde edilen sonularda, modelin Gini katsayısı %62 bulunarak modelin ayırt etme gcnn iyi olduęu sylenmiř ve alıcı alıřtırma karakteristięi eęrisi ile sonular irdelenmeye alıřılmıřtır (zdemir, H. (2010)).

Yksek lisans tezlerinde, genel olarak retilen kredi skortama modellerinde lojistik regresyon metodunun kullanıldıęı gzlemlenmektedir. Farklı metotlar kullanılarak lojistik regresyonun performansının karřılařtırılması makaleler vasıtasıyla literatr alıřmalarında yapılmıřtır. inko tarafından yapılan alışmada, bir finansal kuruluř tarafından elde edilen verisetleri ile lojistik regresyon, diskriminant analizi, karar aęacı ve yapay sinir aęları metotları ile modelleme alıřmaları yapılmıřtır. alışmada kullanılan yapay sinir aęı tr ve yapısı hakkında bilgi verilmemiřtir. Elde edilen sonularda karar aęacı, 1.tip hata ve doęru sınıflama sayısı bakımından en iyi performans gsteren metot olarak grlmřtr. 2.tip hata oranı bakımından en iyi performans gsteren metot yapay sinir aęı olmuřtur (inko, M. (2006)).

İnce ve Aktan tarafından yapılan çalışmada, regresyon karar ağacı ve çok katmanlı algılayıcı yapay sinir ağları ile kredi skorlama modelleri oluşturulmuştur. Daha sonra çeşitli gruplama tekniklerinden yararlanılarak farklı modeller kurulup sonuçlar karşılaştırılmıştır. Veriseti ticari bir finans kuruluşu tarafından elde edilmiş olup çok katmanlı algılayıcı yapay sinir ağı regresyon karar ağacına göre daha üstün bir sınıflandırma yapmıştır. Çeşitli gruplama metotları kullanılarak elde edilen sonuçlarda, eğitim seti için karar ağacı ve yapay sinir ağlarına göre doğru sınıflandırma oldukça başarılı yapılmasına rağmen, sınama seti için yapılan doğru sınıflandırma oranında çok iyi bir performans gösterememiştir (İnce, H. ve Aktan, B. (2010)).

7. LOJİSTİK REGRESYON VE YAPAY SİNİR AĞLARI İLE KREDİ SKORLAMA MODELLERİ

Çalışmada, 2008 yılında ticari bir bankaya ait tüketici kredisi için başvuran müşterilere ait bilgiler kullanılarak lojistik regresyon ve yapay sinir ağları ile skorkartlar oluşturulmuştur. Modellerde bağımlı değişken kredi verilip verilmemesi olup 10 adet bağımsız değişken modellere dâhil edilmiştir. Kredi kartı skorklama ve tüketici kredi skorklama modellerinin finansal kuruluşlar tarafından oldukça yaygın olarak kullanılmasının nedeni, bankaların bu ürünlerini kullanan müşteri sayılarının milyonları geçen sayıya ulaşmasıdır. Bu doğrultuda, bu çalışmada tüketici kredi bilgileri doğrultusunda skorkartlar oluşturulmuştur.

Literatür çalışmalarında kullanılan verisetleri genel olarak ortak ulaşılabilir, açık veritabanları üzerinden elde edilmiştir. İrdelenen çalışmaların çok azında finans kuruluşlarından elde edilen veriler ile skorkartlar oluşturulmuştur. Bunlar Abdou, H., Pointon, J. ve El-Masry, A. (2008), Bellotti, T. ve Crook J. (2009), Lee, T. S., Chiu, C. C., Chou, Y. C. ve Lu C. J. (2006), Lee T.S., Chiu, C.C. ve Lu, C.J. (2002) tarafından yapılan çalışmalardır. Çalışmada ticari bir bankaya ait veri seti kullanılmıştır.

Kredi skorklama çalışmaları üzerine yapılan çalışmalarda çok farklı yöntemlerin kullanılmasının en büyük nedenleri, problemin karışıklığından ve bu karışık problemin kredi verici kuruluşlar tarafından hayati derecede önemli olmasından kaynaklanmaktadır. Kredi verici kuruluşların en büyükleri olan bankalar tarafından kredilerin iyi veya kötü verilme süreci, kredi skorlamadan elde edilen sonuçlara bakılarak değerlendirilmektedir. Yapılan literatür çalışmalarında, finansal kuruluşlar bakımından skorklama yöntemlerinden elde edilen sonuçlarda en büyük riskin, skorklama sürecinde kötü kredilerin iyi kredi olarak tahmin edilmesi olduğu görülmüştür. Literatürdeki çalışmalarda yapay sinir ağları ile elde edilen modellerde, 2.tip hata olarak adlandırılan kötü kredileri iyi kredi olarak sınıflandırılması hata oranının düştüğü görülmüştür. Çalışmada kullanılan lojistik regresyon ve yapay sinir ağları yöntemleri ile bu yüzdeler karşılaştırılmıştır.

Literatür çalışmalarında, lojistik regresyon yönteminden elde edilen ile yapay sinir ağlarından elde edilen sonuçların karşılaştırılmasında farklı yorumlar yapılmış hangi yöntemin daha üstün olduğu sorusu tam olarak cevaplandırılmamıştır. Bunun nedeni, yapılan çalışmalarda kullanılan farklı veri setlerinin varlığı ve yapay sinir ağları için farklı ağ yapılarının seçilmiş olmasıdır.

Yapay sinir ağları ile kredi skorlama modeli oluşturulan çalışmalarda, genel olarak, çok katmanlı algılayıcı ve geri beslemeli yapay sinir ağlarının kullanıldığı görülmüştür. Kullanılan eğitim seti ve sınama setleri oranlarının genel olarak %70 eğitim seti için kullanıldığı, %30 sınama aşaması için kullanıldığı kabul görmüştür. Khasman tarafından yapılan çalışmada, en iyi ağ yapılarına %50 eğitim, %50 sınama verilerinden ulaşıldığı öne sürülmüştür (Khasman, A. (2010)). Elde edilen veri setlerinin göreceli olarak az olması nedeniyle, %70'lik eğitim setlerinin modellerde kullanıldığını düşünülmektedir. Çalışmada öğrenme-sınama verisetleri oranları değiştirilerek yapay sinir ağı performansının artırılması amaçlanmıştır.

Yapay sinir ağı yapısında genel olarak incelenen konular; gizli katman nöron sayısı, momentum katsayısı, öğrenme katsayısıdır. Genel olarak kabul görmüş bir ağ yapısı yani gizli katman sayısı, öğrenme katsayısı ve momentum katsayısı literatürde bulunmamaktadır. Bu oranlar deneme yanılma yolu ile değiştirilerek en iyi ağ yapısına ulaşılmaya çalışılmaktadır. Çalışmada, öncelikle fõrmülleştirme ile gizli katman nöron sayısı bulunmuş ve daha sonra sabit momentum ve öğrenme katsayısında deneme yanılma yolu ile farklı bir gizli katman nöron sayısı elde edilmiştir. Elde edilen bu gizli katman nöron sayıları için farklı momentum katsayısı ve öğrenme katsayılarında sinir ağları test edilerek en iyi ağ yapısı elde edilmeye çalışılmıştır.

Çalışmada, ticari bir bankadan 2008 yılı itibarıyla tüketici kredileri bilgileri alınacak olup lojistik regresyon ve çok katmanlı yapay sinir ağları metotları karşılaştırılacaktır. Yapay sinir ağlarının küçük verisetlerini modellemede başarısızlığı öne sürülmektedir. Buna karşın daha düşük öğrenme-sınama seviyelerinde (%50) kredi skorlama modellerinde daha başarılı olduğu görülmüştür. Çalışmada öncelikle %70 öğrenme, %30 sınama aşamaları ile iki yöntem kullanılarak model oluşturulacak ve daha sonra bu verisetlerinde öğrenme ve sınama oranları farklılaştırılarak en iyi öğrenme-sınama yapısı bulunmaya çalışılacaktır.

7.1. Modelde Kullanılan Değişkenler

Lojistik regresyon modelinde kullanılan değişkenler, 10 adet bağımsız değişken ve 1 adet bağımlı değişkenden oluşmaktadır. Bağımsız değişkenler; çocuk sayısı, meslek, müşteri ilişki süresi, cinsiyet, eğitim durumu, gelir-kredi oranı, istenilen taksit sayısı, otomatik ödeme talimatlı geçen gün sayısı ve müşteri yaşıdır. Bağımlı değişken, 12 ay içinde kredinin temerrüde düşmesidir. Çizelge 7.1’de bağımsız değişkenler ve kodlamaları gösterilmiştir.

Çizelge 7.1 : Bağımlı ve bağımsız değişkenler ve kodları.

KATEGORİK DEĞİŞKENLER		KODLAMA
ÇOCUK SAYISI		
	0	1
	1	2
	2	3
	3	4
	4+	5
MESLEK		
	1	
	2	
	3	
	4	
MUSTERİ İLİSKİ SÜRESİ (AY)		
	0	1
	1-9	2
	10-24	3
	25+	4
İKAMET DURUMU		
	Kendi Evi	1
	Kiracı veya Ailesi ile Birlikte	2
	Diğer	3
CİNSİYET		
	Bekar Bayan	1
	Evli Erkek - Evli Bayan	2
	Bekar Erkek	3
EĞİTİM DURUMU		
	Üniversite Mezunu	1
	Ortaokul mezunu	2
	Eğitimsiz - İlkokul Mezunu - Lise Mezunu - Master/Doktora Mezunu	3
SÜREKLİ DEĞİŞKENLER		
GELİR KREDİ TUYARI		
İSTENİLEN TAKSİT SAYISI		
OTOMATİK ÖDEME TALİMATLI GEÇEN GÜN SAYISI		
YAS		

Kredi müşterilerinin çocuk sayılarına bakıldığında, çocuk sayısı arttıkça kredilerin temerrüt oranlarının da arttığı görülebilmektedir. Bu açıdan, müşteri çocuk sayısı modele değişken olarak dâhil edilmiştir. Çizelge 7.2’de çocuk sayısı frekans sayılarına yer verilmiştir.

Çizelge 7.2 : Çocuk sayısı frekans tablosu.

ÇOCUK SAYISI	FREKANS					
	İYİ	İyiler İçerisinde %	%	KÖTÜ	Kötüler İçerisinde %	%
0	747	67,85	69,62	326	43,82	30,38
1	167	15,17	52,35	152	20,43	47,65
2	82	7,45	46,86	93	12,50	53,14
3	39	3,54	41,94	54	7,26	58,06
4+	66	5,99	35,68	119	15,99	64,32
Toplam	1101	100,00	59,67	744	100,00	40,33

Bağımsız değişkenler içerisine, meslek, kategorik bir değişken olarak dâhil edilmiştir. Birbiriyle aynı karakteristiğe sahip olduğu düşünülen meslek grupları aşağıdaki çizelgede gruplandırılmıştır.

Çizelge 7.3 : Meslek grupları ve kodlanmaları.

MESLEKLER	
1	AKADEMİSYEN, ASKER, BANKACI, DANIŞMAN, DİPLOMAT (KONSOLOS), DİŞHEKİMİ, DOKTOR, ECZACI, KAPTAN, KİMYAGER/LABORANT, MÜFETTİŞ, NOTER, ÖĞRETMEN/ĞİT MEN, PİLOT, VETERİNER, İŞÇİ (KAMU), SM/YMM, DENETÇİ, BİYOLOG, DİNADAMI, HEMŞİRE, TİBBİ MÜESSİL, EMEKLİ
2	AVUKAT, GÜVENLİK MEMURU, HAKİM/SAVCI, HUKUK MÜŞAVİRİ, KUYUMCU, MEMUR, MİMAR/MÜHENDİS, SEKRETER, SPORCU, SİGORTACI, TAHSİLDAR, TERZİ, ŞİRKET ORTAĞI, EVHANIMI, BELEDİYE ÇALIŞANI
3	AŞÇI, AYAKKABICI, BİLGİSAYARCI (PROGRAMCI), GALERİ SAHİBİ, KUAFÖR/GÜZELLİK UZMANI, MUHASEBECİ, MÜTEAHHİT, OTO KİRALAMACI, RANTİYE, SANATÇI (MÜZİSYEN), SANATÇI (RESSAM), TAŞIYICI, TEKNİSYEN, TESİSATÇI, TURİZMCİ, YÖNETİCİ ASİSTANI, ZABIT AMEMURU, ŞOFÖR, İŞÇİ (ÖZEL), SANAYİCİ, DESİNATÖR, SİSTEM ANALİSTİ, HİZMETLİ, SPOKER, SUNUCU
4	BİLGİSAYARCI (DONANIMCI), ÇİFTÇİ, EMLAKÇI, ESNAF, BELİRLİ FİRMAYÖNETİCİSİ, GAZETECİ, REKLAM VE HALKLA İLİŞKİLER, HOSTES/HOST, İŞLETMECİ, KONF EKSYONCU, MİT MENSUBU, PAZARLAMACI, POLİS, SATIŞ ELEMANI, SERBEST MESLEK, SERBEST MUHASEBECİ, DİĞER FİRMAYÖNETİCİSİ, TERCÜMAN/REHBER, SAĞLIK MEMURU, GARSON/SERVİS ELEMANI, TEMİZLİK GÖREVLİSİ, ÇAYCI

Modellerin geliştirilmesinde kullanılacak verilerde iyi-kötü müşteriler için meslek dağılımı aşağıdaki şekilde oluşmuştur.

Çizelge 7.4 : Meslek frekans tablosu.

MESLEK	FREKANS					
	İYİ	İyiler İçerisinde %	%	KÖTÜ	Kötüler İçerisinde %	%
1	275	24,98	48,08	297	39,92	51,92
2	528	47,96	59,26	363	48,79	40,74
3	129	11,72	71,27	52	6,99	28,73
4	169	15,35	84,08	32	4,30	15,92
Toplam	1101	100,00	59,67	744	100,00	40,33

Burada da görüldüğü üzere farklı meslek gruplarında temerrüt oranları farklılık göstermektedir.

Müşteri ilişki süresi ile ilişki süresi arttıkça temerrüt oranlarının azaldığı çizelge 7.5’de görülmektedir.

Çizelge 7.5 : Müşteri ilişki süresi frekans tablosu.

MÜŞTERİ İLİŞKİ SÜRESİ	FREKANS					
	İYİ	İyiler İçerisinde %	%	KÖTÜ	Kötüler İçerisinde %	%
1	578	52,50	49,96	579	77,82	50,04
2	111	10,08	62,01	68	9,14	37,99
3	131	11,90	79,39	34	4,57	20,61
4	281	25,52	81,69	63	8,47	18,31
Toplam	1101	100,00	59,67	744	100,00	40,33

Cinsiyet bakımından örneklem grubu incelendiğinde farklı cinsiyet ve medeni durumlarda farklı temerrüt oranlarının ortaya çıktığı görülmektedir. Bekâr erkek kredi müşterilerinin temerrüt etme oranı en yüksek olurken, bekâr bayanların en düşük olduğu çizelge 7.6’da görülmektedir.

Çizelge 7.6 : Cinsiyet frekans tablosu.

CİNSİYET	FREKANS					
	İYİ	İyiler İçerisinde %	%	KÖTÜ	Kötüler İçerisinde %	%
Bekar Bayan	163	14,80	74,09	57	7,66	25,91
Evli Erkek - Evli Bayan	763	69,30	60,13	506	68,01	39,87
Bekar Erkek	175	15,89	49,16	181	24,33	50,84
Toplam	1101	100,00	59,67	744	100,00	40,33

Eğitim durumu bakımından bakıldığında iyi-kötü müşteri ayrımı yapılmasında eğitim durumu arttıkça kredilerin temerrüt etme olasılıklarının çizelge 7.7’de azaldığı söylenebilir.

İkamet durumu açısından örneklem grubuna bakıldığında, örneklem grubunun yaklaşık %75’lik kısmının kiracı veya ailesi ile birlikte oturduğu görülmektedir.

Temerrüt etme olasılığı bakımından ev sahibi olan gruptaki kişilerin kredilerinin en düşük temerrüt etme olasılığına sahip oldukları çizelge 7.8’de görülmektedir.

Çizelge 7.7 : Eğitim durumu frekans tablosu.

EĞİTİM DURUMU	FREKANS					
	İYİ	İyiler İçerisinde %	%	KÖTÜ	Kötüler İçerisinde %	%
Üniversite Mezunu	248	22,52	66,85	123	16,53	33,15
Ortaokul mezunu	66	5,99	60,00	44	5,91	40,00
Eğitimsiz - İlkokul Mezunu - Lise Mezunu - Master/Doktora Mezunu	787	71,48	57,70	577	77,55	42,30
Toplam	1101	100,00	59,67	744	100,00	40,33

Çizelge 7.8 : İkamet durumu frekans tablosu.

İKAMET DURUMU	FREKANS					
	İYİ	İyiler İçerisinde %	%	KÖTÜ	Kötüler İçerisinde %	%
Kendi Evi	162	14,71	64,80	88	11,83	35,20
Kiracı veya Ailesi ile Birlikte	873	79,29	60,92	560	75,27	39,08
Diğer	66	5,99	40,74	96	12,90	59,26
Toplam	1101	100,00	59,67	744	100,00	40,33

7.2. Lojistik Regresyon Modeli

Lojistik regresyon modeli kurulması için toplamda 1845 adet müşteri başvuru bilgileri kurulmuştur. Modelde kullanılan bağımlı değişken, 2008 yılı sonu itibarıyla temerrüde düşmedir. Elde edilen verisetinin %70’i ile lojistik regresyon modeli elde edilmiştir. %’30’u modelin doğrulanması aşamasında kullanılacak veri setine ilişkin sayısal veriler çizelge 7.9’da gösterilmektedir.

Çizelge 7.9 : Veri seti test ve öğrenme adet ve yüzdeleri.

Ağırlandırıklandırılmamış Durum	Adet	Yüzde
Seçilen Örnekler	1311	71,1
Analize Dahil Edilenler	0	,0
Eksik Toplam	1311	71,1
Seçilmeyen Örnekler	534	28,9
Toplam	1845	100,0

Mevcut verisetinde kategorik deęişkenler; çocuk sayısı, meslek, müşteri ilişki süresi, eğitim durumu, medeni durumu ve ikamet durumudur. Çizelgede kategorik deęişkenlere ilişkin kodlamalara yer verilmiştir.

Çizelge 7.10 : Kategorik deęişkenlerin kodlanması.

		Frekans	Parametrelerin Kodlanması			
			(1)	(2)	(3)	(4)
COCUK_SAYISI	1	771	1,000	,000	,000	,000
	2	208	,000	1,000	,000	,000
	3	126	,000	,000	1,000	,000
	4	68	,000	,000	,000	1,000
	5	138	,000	,000	,000	,000
MESLEK	1	398	1,000	,000	,000	
	2	639	,000	1,000	,000	
	3	134	,000	,000	1,000	
	4	140	,000	,000	,000	
MUSTERI_ILISKI_SURESI	1	822	1,000	,000	,000	
	2	125	,000	1,000	,000	
	3	115	,000	,000	1,000	
	4	249	,000	,000	,000	
IKAMET_DURUMU	1	179	1,000	,000		
	2	1009	,000	1,000		
	3	123	,000	,000		
CINSIYET	1	157	1,000	,000		
	2	901	,000	1,000		
	3	253	,000	,000		
EGITIM_DURUMU	1	250	1,000	,000		
	2	79	,000	1,000		
	3	982	,000	,000		

Mevcut tabloda genel istatistik skoru, analiz dışında kalan deęişkenlerin chi-square deęerini vermektedir. %5 anlamlılık düzeyine göre bu deęer anlamlıdır. Buradan elde edilen sonuçlarda, modele girmeyen deęişkenlerden bir veya birden fazlasının dâhil edilmesinin, modelin tahmin edici gücünü arttırıcı yönde etki edeceğini göstermektedir.

Modele giren deęişkenlerin sayısal deęişkenlerine bakıldığında en önemli deęişkenin gelir/kredi oranı olduğu görülmektedir. Exp(B) deęerine bakıldığında, gelir kredi oranındaki 1’lik bir artışın iyi kredi olma şansını %48,5 arttırdığı görülmektedir. Çizelge 7.11’de kategorik deęişkenler arasında müşteri ilişki süresinin iyi kredi olma olasılığını en fazla etkileyen kategorik deęişken olduğu görülmektedir.

Çizelge 7.11 : Modele dahil edilmeyen değişkenler.

		Score	Serbestlik Derecesi	Anlamlılık Düzeyi
Step 0	Değişkenler	COCUK_SAYISI	83,297	4 ,000
		COCUK_SAYISI(1)	74,654	1 ,000
		COCUK_SAYISI(2)	7,052	1 ,008
		COCUK_SAYISI(3)	7,416	1 ,006
		COCUK_SAYISI(4)	8,696	1 ,003
		MESLEK	58,999	3 ,000
		MESLEK(1)	31,335	1 ,000
		MESLEK(2)	,034	1 ,853
		MESLEK(3)	8,811	1 ,003
		MUSTERI_ILISKI_SURESI	110,632	3 ,000
		MUSTERI_ILISKI_SURESI(1)	97,833	1 ,000
		MUSTERI_ILISKI_SURESI(2)	1,479	1 ,224
		MUSTERI_ILISKI_SURESI(3)	13,293	1 ,000
		IKAMET_DURUMU	15,948	2 ,000
		IKAMET_DURUMU(1)	3,309	1 ,069
		IKAMET_DURUMU(2)	1,253	1 ,263
		CINSIYET	29,587	2 ,000
		CINSIYET(1)	8,932	1 ,003
		CINSIYET(2)	4,714	1 ,030
		EGITIM_DURUMU	8,717	2 ,013
		EGITIM_DURUMU(1)	7,175	1 ,007
		EGITIM_DURUMU(2)	,816	1 ,366
		GELIR_KREDI_TUTARI	,020	1 ,888
		ISTENILEN_TAKSIT_SAYISI	2,660	1 ,103
		OTOMATIK_ODEMEN_GUN_SAYISI	51,976	1 ,000
		YAS	,026	1 ,872
	Genel İstatistik	204,551	20	,000

Aşağıdaki çizelgede önemli istatistiklerden biri Wald istatistiğidir. Bu istatistik katsayıların sıfıra eşit olup olmadığını test etmektedir. Buradaki temel hipotez H_0 : katsayının sıfıra eşit olması ve alternatif hipotez H_a : katsayının sıfırdan farklı olması şeklinde kurulur. Significant değeri <0.05 olması durumunda H_0 reddedilir. Bir bağımsız değişkenin katsayısının sıfırdan farklı olması halinde, bu bağımsız değişkenin bağımlı değişkenin tahmininde katkısı olduğu anlamına gelmektedir.

Çizelge 7.12 : Modele dahil edilen değişkenler.

Modele Giren Değişkenler							
		B	Standart	Wald İstatistiği	Serbestlik Derecesi	Anlamlılık Düzeyi	Exp(B)
Step 1a	COCUK_SAYISI			10,780	4	,029	
	COCUK_SAYISI(1)	-,694	,232	8,943	1	,003	,500
	COCUK_SAYISI(2)	-,559	,237	5,571	1	,018	,572
	COCUK_SAYISI(3)	-,494	,263	3,521	1	,061	,610
	COCUK_SAYISI(4)	-,171	,322	,282	1	,595	,843
	MESLEK			31,774	3	,000	
	MESLEK(1)	1,256	,263	22,813	1	,000	3,512
	MESLEK(2)	,763	,257	8,808	1	,003	2,145
	MESLEK(3)	,389	,311	1,563	1	,211	1,476
	MUSTERI_ILISKI_SURESI			52,615	3	,000	
	MUSTERI_ILISKI_SURESI(1)	1,388	,204	46,151	1	,000	4,007
	MUSTERI_ILISKI_SURESI(2)	,810	,269	9,090	1	,003	2,248
	MUSTERI_ILISKI_SURESI(3)	,475	,290	2,678	1	,102	1,608
	IKAMET_DURUMU			5,545	2	,063	
	IKAMET_DURUMU(1)	-,629	,270	5,430	1	,020	,533
	IKAMET_DURUMU(2)	-,424	,220	3,704	1	,054	,655
	CINSIYET			14,199	2	,001	
	CINSIYET(1)	-,550	,241	5,204	1	,023	,577
	CINSIYET(2)	-,601	,161	13,891	1	,000	,548
	EGITIM_DURUMU			1,251	2	,535	
	EGITIM_DURUMU(1)	-,025	,174	,022	1	,883	,975
	EGITIM_DURUMU(2)	-,298	,267	1,250	1	,264	,742
	GELIR_KREDI_TUTARI	,396	,199	3,932	1	,047	1,485
	ISTENILEN_TAKSIT_SAYISI	,002	,006	,128	1	,721	1,002
	OTOMATIK_ODEMEN_GUN_SAYISI	-,002	,001	4,525	1	,033	,998
	YAS	,015	,007	4,496	1	,034	1,015
	Sabit	-1,197	,520	5,293	1	,021	,302

Lojistik regresyon ile elde edilen model doğrultusunda lojistik regresyon fonksiyonu şu şekilde elde edilmiştir:

$$Y = -1,197 + \text{COCUK_SAYISI}(1) * -0,694 + \text{COCUK_SAYISI}(2) * -0,559 + \text{COCUK_SAYISI}(3) * -0,494 + \text{COCUK_SAYISI}(4) * -0,171 + \text{MESLEK}(1) * 1,256 + \text{MESLEK}(2) * 0,763 + \text{MESLEK}(3) * 0,389 + \text{MUSTERI_ILISKI_SURESI}(1) * 1,388 + \text{MUSTERI_ILISKI_SURESI}(2) * 0,810 + \text{MUSTERI_ILISKI_SURESI}(3) * 0,475 + \text{IKAMET_DURUMU}(1) * -0,629 + \text{IKAMET_DURUMU}(2) * -0,51 + \text{CINSIYET}(1) * -0,550 + \text{CINSIYET}(2) * -0,601 + \text{EGITIM_DURUMU}(1) * -0,550 + \text{EGITIM_DURUMU}(2) * -0,601 + \text{GELIR_KREDI_TUTARI} * 0,369 + \text{ISTENILEN_TAKSIT_SAYISI} * 0,002 + \text{OTOMATIK_ODEMEN_GUN_SAYISI} * -0,002$$

Elde edilen regresyon formülünden müşteri skoru ve ardından aşağıdaki formül vasıtasıyla kredinin temerrüt etme olasılığı hesaplanacaktır. $P(Y)$ değerinin 0,5’den büyük olması durumunda model kredinin temerrüt ettiği sonucuna ulaşacak, $P(Y)$ değerinin 0,5’den küçük olması durumunda kredinin temerrüt etmediği sonucuna ulaşılacaktır.

$$P(Y) = 1/1 + e^{-(a+bX_1+cX_2+....+nX_N+u)} \quad (7.1)$$

7.2.1. Lojistik regresyon modelinin anlamlılığının test edilmesi

Hosmer & Lemeshow istatistiği, modelin verilere uyum sağlayıp sağlamadığı konusunda karar vermede kullanılır. Bu istatistik, anlamlılık seviyesi 0,05 altında ise zayıf bir uyumu göstermektedir. Çizelge 7.13’de görüldüğü gibi modelimizde anlamlılık seviyesi 0,136 çıkmasından dolayı modelin tatmin edici şekilde verilere uyum gösterdiği söylenebilir.

Çizelge 7.13 : Hosmer ve Lemeshow Testi.

Hosmer and Lemeshow Test			
Step	Chi-square	Serbestlik Derecesi	Anlamlılık Düzeyi
1	12,355	8	0,136

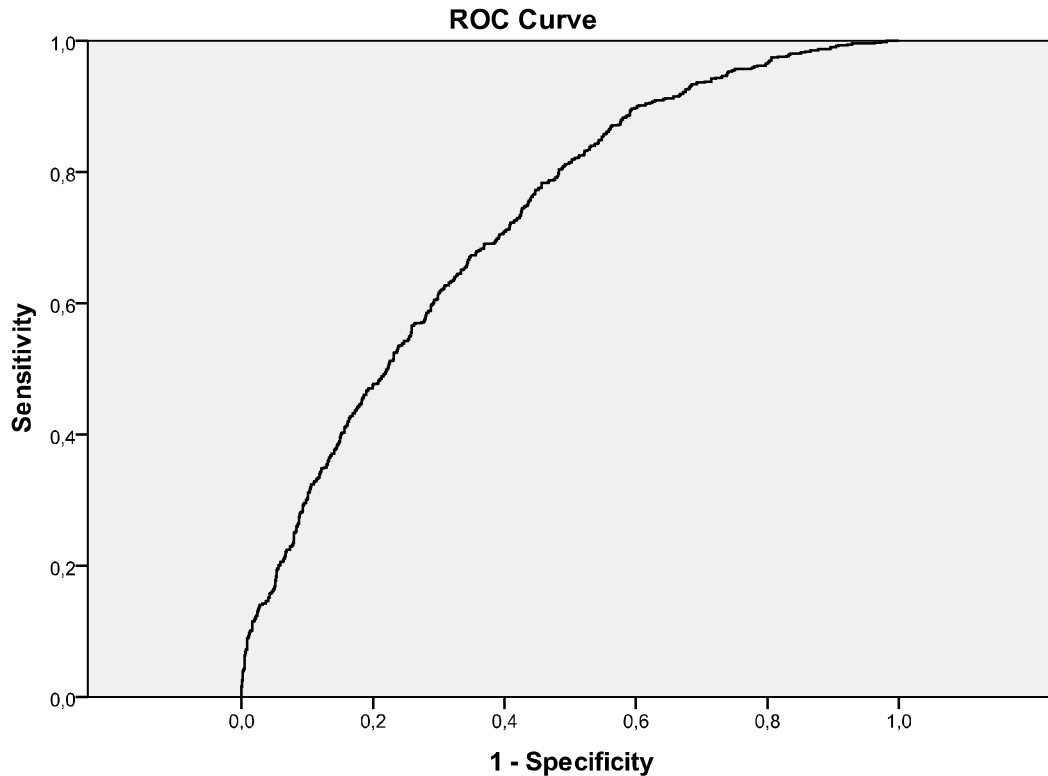
Tablodan elde edilen sonuçlarda eğitim aşamasında toplam doğruluk yüzdesi 67.1 olmuştur. Toplam 883 adet iyi krediden 616’sı doğru olarak bilinmiş, 167 adedi kötü kredi olarak bulunmuştur. Toplam 528 adet kötü krediden 264 âdeti kötü kredi olarak 264 adeti iyi kredi olarak bulunmuş ve eğitim setinde kötü kredileri bilme yüzdesi %50 olduğu çizelge 7.14’de görülmektedir.

Çizelge 7.14 : Sınıflandırma Tablosu.

Sınıflandırma Tablosu								
	Gözlenen		Tahmin edilen					
			Eğitim			Test		
			TEMERRUT_DURU MU		Doğru Yüzde si	TEMERRUT_DURU MU		Doğru Yüzde si
			0	1		0	1	
Step 1	TEMERRUT_DURU MU	0	616	167	78,7	240	78	75,5
		1	264	264	50,0	111	105	48,6
	Genel Yüzde				67,1			64,6

Sınama verisetinde toplam doğru yüzdesi %64,6’dır. 318 iyi krediden 240 iyi kredi olarak 78’i kötü kredi olarak sınıflandırılmıştır. 216 kötü krediden 105 tanesi kötü kredi 111 tanesi iyi kredi olarak sınıflandırılmıştır.

Modelin değerlendirilmesinde alıcı çalıştırma karakteristiği eğrisi kullanılabilir. “Alıcı çalıştırma karakteristiği(ROC) eğrisi” modelin ayırt etme gücünü gösteren bir istatistiktir. Regresyon sonrasında her bir gözlem için elde edilen temerrüt olasılığı değerinin gerçekleşen temerrüt durumunu ne ölçüde açıklayabildiğini gösterir. Modelin ROC eğrisi şekilde görülmektedir:



Diagonal segments are produced by ties.

Şekil 7.1 : Lojistik regresyon modeli ile elde edilen alıcı çalıştırma karakteristiği(ROC) eğrisi.

Alıcı çalıştırma karakteristiği eğrisinin altında kalan alan % 72.5 olarak bulunmuştur. Gini katsayısı literatürde modelin ayırt etme gücünü göstermek için kullanılmaktadır. Aşağıdaki şekilde formülize edilmektedir.

$$|2 * ROC| - 1$$

Gini=0 ise modelin ayırtıcılığı yoktur. Gini 0.4 ile 0.6 arasında ise modelin kabul edilebilir bir ayırtırma gücü, 0.6 ile 0.8 arasında ise iyi bir ayırtırma gücü ve 0.8'den büyük ise çok iyi bir ayırtırma gücü var demektir.

Elde ettiğimiz modelde alıcı çalıştırma karakteristiği(ROC) değeri %72.5 olmasından dolayı Gini katsayısı şu şekilde oluşur:

$$|2 * 0,725| - 1 = 0.45$$

Elde edilen değer 0.4 ile 0.6 arasında olmasından dolayı modelin kabul edilebilir bir ayırtırma gücüne sahip olduğu söylenebilir.

7.3. Yapay Sinir Ağları ile Kredi Skorlama Modelinin Kurulması

Çalışmada çok katmanlı algılayıcı yapay sinir ağı kullanılmıştır. Literatür çalışmaları incelendiğinde kredi skorlama modelleri olarak en çok kullanılan yapay sinir ağı yapısı olmasından dolayı çalışmada kullanılmıştır. Kredi skorlama modeli olarak bu ağların kullanılmasındaki nedenler, öngörü ve sınıflandırma işlemleri için oldukça uygun olması ve doğrusal olmayan yapılar için oldukça kullanışlı olmalarıdır. Yapay sinir ağının mimarisi sürecinde, 10 adet girdi katmanı ve 1 adet çıktı katmanından oluşan bir yapı oluşturulacaktır. Çalışmada yapay sinir ağları ile model kurmak için SPSS 18.00 programı kullanılmıştır.

7.3.1. Yapay sinir ağlarının yapısı

Çok katmanlı yapay sinir ağlarının geliştirilmesi için 2 yöntem kullanılmıştır. Aktivasyon fonksiyonu sigmoid seçilmiş ve online eğitim tipi seçilmiştir. Bu sürecin ardından önce %70 (1311 adet kredi) öğrenme, %30 (534 adet kredi) deneme verisetleri ayrılmış ve deneme yoluyla en iyi sonucu üretecek olan gizli katman nöron sayısı bulunmaya çalışılmıştır. En iyi ağ yapısı 10 girdi katmanı nöron sayısı, 21 gizli katman nöron sayısı ve 1 çıktı katmanı nöron sayısı ile elde edilmiştir. 2. aşamada formülden elde edilen gizli katman nöron sayısı bulunmuştur.

Gizli katman nöron sayısını belirlemek için aşağıdaki formül kullanılmıştır;

$$N_s = (\frac{1}{2} (N_g + N_n) + \sqrt{N_d}) / N_b \quad 7.1$$

Burada

N_s : Gizli katmandaki nöron sayısını,

N_g : Girdi katmanındaki nöron sayısını,

N_n : Çıktı katmanındaki nöron sayısını

N_d : Gözlem sayısını

N_b : Katman sayısını göstermektedir.

Çalışmada uygulanan veriler ile bu sayı şu şekilde hesaplanmaktadır.

$$(0.5 * (11+1) + \sqrt{1844}) / 3 = 16.3$$

Çalışmada 21 gizli katmanlı ve 16 gizli katmanlı çok katmanlı algılayıcı yapay sinir ağları için en iyi performans veren ağ yapısına ulaşmak amacıyla çeşitli momentum ve öğrenme katsayılarında denemeler yapılarak optimum ağ yapıları elde edilmeye çalışılmıştır. Yapılan çalışmalarda, öğrenme katsayısı 0.012'nin üstünde olduğu durumlarda, yapay sinir ağının yakınsayamamasından dolayı tahmin etme performansının düştüğü gözlemlenmiştir. Momentum katsayısının 0.8 ile 0.99 arasında kullanılması tavsiye edilmiştir. Çalışmada öğrenme katsayıları olarak 0.001, 0.005, 0.008 ve 0.012 ile 0.8 ve 0.9 momentum katsayıları denenecek ve en iyi yapay sinir ağı yapıları elde edilmeye çalışılacaktır. Tüm ağlarda kullanılan aktivasyon fonksiyonu sigmoid aktivasyon fonksiyonu olup online eğitim tipi seçilmiştir. Her ağ yapısı için geri yayımlı yapay sinir ağı 5 kere çalıştırılıp ortalama doğruluk derecelerine bakılacaktır. Bu aşamada sinir ağlarında öğrenme ve sınav aşamaları 71:29 olarak düzenlenmiştir. Sonuçlara çizelge 7.15'de yer verilmiştir:

Çizelge 7.15 : 16 gizli katmanlı algılayıcın yapay sinir ağı momentum katsayıları ve öğrenme katsayıları karşılaştırılması.

Öğrenme Katsayısı		Momentum Katsatasyısı	
		0.8	0.9
	0.001	65,39%	66,12%
	0.005	66,29%	66,17%
	0.008	66,28%	66,21%
	0.012	65,46%	65,75%

16 gizli katmanlı çok katmanlı algılayıcı için yapılan sınamalarda, en yüksek doğruluk derecesine sahip yapı olarak öğrenme katsayısı 0.005 ve momentum katsayısı 0.8 olan yapı bulunmuştur. Sonuçlara çizelge 7.16de yer verilmiştir:

Çizelge 7.16 : 21 gizli katmanlı algılayıcın yapay sinir ağı momentum katsayıları ve öğrenme katsayıları karşılaştırılması.

Öğrenme Katsayısı		Momentum Katsatasyısı	
		0.8	0.9
	0.001	66,20%	66,46%
	0.005	66,63%	65,81%
	0.008	65,79%	65,67%
	0.012	65,91%	66,54%

21 gizli katmanlı çok katmanlı algılayıcı için yapılan sınamalarda en yüksek doğruluk derecesine sahip yapı olarak öğrenme katsayısı 0.005 ve momentum katsayısı 0.8 olan yapı bulunmuştur.

Farklı öğrenme ve sinama oranları olarak kullanılacak çok katmanlı algılayıcı yapay sinir ağlarının yapıları şu şekilde oluşmuştur;

- 1- 10 girdi, 16 gizli, 1 çıktı katmanından oluşan, öğrenme katsayısı 0.005, momentum katsayısı 0.8 olan çok katmanlı algılayıcı yapay sinir ağı,
- 2- 10 girdi, 21 gizli, 1 çıktı katmanından oluşan, öğrenme katsayısı 0.005, momentum katsayısı 0.8 olan çok katmanlı algılayıcı yapay sinir ağlarıdır.

7.3.2. Yapay sinir ağlarından elde edilen sonuçlar

Tabloda lojistik regresyon ile elde edilen sonuçlara ve en iyi ağ yapılarına sahip çok katmanlı algılayıcıların beşer deneme ardından ortalama tahmin değerlerine yer verilmiştir;

Çizelge 7.17 : Yapay sinir ağları sınıflandırma tabloları.

Sınıflandırma Tablosu									
	Gözlenen		Tahmin edilen						Toplam Doğru Yüzdesi
			Eğitim			Test			
			TEMERRUT DURUMU		Doğru Yüzdesi	TEMERRUT DURUMU		Doğru Yüzdesi	
			0	1		0	1		
Lojistik Regresyon	TEMERRUT_DURUMU	0	616	167	78,7	240	78	75,5	
		1	264	264	50	111	105	48,6	
	Genel Yüzde				67,1			64,6	66,40
10-16-1 Çok Katman Algılayıcı Yapay Sinir Ağı	TEMERRUT_DURUMU	0	610	173	77,9	233,8	84,2	73,5	
		1	255,2	272,8	51,7	109,6	106,4	49,3	
	Genel Yüzde				67,3			63,7	66,29
10-21-1 Çok Katman Algılayıcı Yapay Sinir Ağı	TEMERRUT_DURUMU	0	616,4	166,6	78,7	236,2	81,8	74,3	
		1	257,8	270,2	51,2	109,4	106,6	49,4	
	Genel Yüzde				67,6			64,2	66,63

Yukarıdaki tabloda lojistik regresyon ve yapay sinir ağlarından elde edilen sonuçlar karşılaştırılmıştır. Toplam doğru yüzdesi olarak 10-21-1 ağ yapılı 0.05 öğrenme katsayılı ve momentum katsayısı 0.8 olan çok katmanlı yapay sinir ağının en iyi performansı gösterdiği görülmektedir. Lojistik regresyon, toplam doğru yüzdesinde ikinci sırada yer alırken, ara katman sayısının formülüzasyon yolu ile bulunması sonucunda elde edilen 10-16-1 ağ yapılı çok katmanlı algılayıcı yapay sinir ağının en düşük toplam doğru yüzdesine sahip olduğu çizelge 7.18’de görülmektedir.

Çizelge 7.18 : Kredi skorlama modellerinin karşılaştırılması.

	1.Tip Hata(%)	2.Tip Hata(%)
Lojistik Regresyon	22,3	50,4
10-16-1 Çok Katman Algılayıcı Yapay Sinir Ağı	23,4	49,0
10-21-1 Çok Katman Algılayıcı Yapay Sinir Ağı	22,6	49,4

1. tip hata ve 2. tip hata oranlarına bakıldığında doğru kredilerin belirlenmesinde en yüksek doğruluk yüzdesine sahip modelinin lojistik regresyon yöntemi ile elde edildiği görülmektedir. Finansal kuruluşlar açısından daha fazla zarara neden olan 2. tip hata açısından modeller karşılaştırıldığında yapay sinir ağları ile elde edilen modellerin daha düşük yüzde ile bu hata tipinde lojistik regresyona üstünlük sağladıkları görülmüştür. En düşük yüzdede elde edilen ağ yapısı 10-16-1 ağ yapısıdır. Kötü kredilerin tespit edilmesinde yapay sinir ağlarının daha başarılı olduğu görülmüştür.

7.3.3. Yapay sinir ağı öğrenme-test oranları karşılaştırılması

Bu aşamada elde edilen en iyi ağ yapılarının, farklı sına-öğrenme oranlarında veriseti bölünerek yapay sinir ağlarının performansına etkisi incelenmeye çalışılacaktır. Gizli katman nöron sayısı formülleştirme yolu ile bulunan ve daha sonra momentum ve öğrenme katsayıları deneme yanılma yöntemiyle bulunan 10-16-1 çok katmanlı algılayıcı yapay sinir ağı ilk olarak irdelenmiştir. Elde edilen sonuçlara çizelge 7.19’da yer verilmiştir.

Çizelge 7.19: Yapay sinir ağı öğrenme- test oranları karşılaştırılması.

Öğrenme-Test%	Öğrenme Seti			Test Aşaması			Toplam Doğru Payı	ROC Değeri
	1.tip Hata	2.tip Hata	Doğru Sınıflama	1.tip Hata	2.tip Hata	Doğru Sınıflama		
90-10	30,10%	41,90%	65,20%	26,90%	42,20%	66,30%	65,31%	69,60%
80-20	24,90%	47,20%	65,90%	27,90%	51,50%	63,40%	65,42%	72,20%
70-30	11,40%	61,70%	68,30%	16,00%	63,40%	64,80%	67,12%	74,70%
60-40	20,90%	50,30%	67,40%	24,70%	54,80%	62,90%	65,64%	74,10%
50-50	18,00%	56,60%	66,70%	18,50%	59,20%	64,80%	65,75%	72,30%
40-60	26,30%	36,60%	69,30%	30,90%	40,90%	65,20%	66,83%	74,80%
30-70	13,30%	70,40%	63,60%	12,50%	68,70%	64,90%	64,50%	68,80%

Elde edilen bulgulardan görüldüğü üzere, en yüksek doğru kredi bulma yüzdesine sahip öğrenme-sınama aşamasının verisetinin, %70’nin öğrenme aşamasından geçirilmesi, %30’nun sınama aşamasından geçirilmesi sonucu olduğu gözlemlenmektedir. %70 öğrenme % 30 sınama aşamasında iyi kredilerin oldukça iyi ayırt edildiği fakat kötü kredilerin ayırt edilmesinde başarılı bir performans gösterilmediği görülmektedir. %40 öğrenme %60 test aşamasından geçirilen çok katmanlı yapay sinir ağının kötü kredileri ayırt etmede oldukça başarılı olduğu görülmüştür. Alıcı çalıştırma karakteristiği değerlerinin karşılaştırılmasında en yüksek alıcı çalıştırma karakteristiği değerine sahip ağ yapısının %40 öğrenme ve %60 sınama aşamasından geçirilen ağ yapısı olduğu görülmektedir. %70 öğrenme ve %30 sınama aşamasında geçirilen ağ yapısı alıcı çalıştırma karakteristiği değerinde 2. sıradadır.

Toplam doğruluk payında en yüksek değeri, %70 öğrenme %30 sınama aşamasından geçirilen yapay sinir ağının elde ettiği görülmüştür. Buna karşın kötü kredilerin tespit edilmesinde daha başarılı sonuç veren ve toplam doğru tahmin etme yüzdesinde 2. sırada yer alan, %40 öğrenme %60 sınama aşamasından geçirilen ağın performansının iyi olduğu gözlemlenmiştir.

Farklı öğrenme ve sinama aşamaları, 10 girdi katman nöron sayılı, 21 gizli katman nöron sayılı ve 1 çıktı katmanlı nöronlardan oluşan çok katmanlı algılayıcı yapay sinir ağı için denemiştir. Daha önceki deneyler veriseti için elde edilen optimum değerler olarak momentum katsayısı 0.8 ve öğrenme katsayısı 0.05 olarak farklı öğrenme-sinama veri setleri modellenmiştir. Elde edilen bulgular aşağıda gösterilmiştir.

Çizelge 7.20 : Yapay sinir ağları öğrenme- test oranları karşılaştırılması.

Öğrenme-Test%	Öğrenme Seti			Test Aşaması			Toplam Doğru Payı	ROC Değeri
	1.tip Hata	2.tip Hata	Doğru Sınıflama	1.tip Hata	2.tip Hata	Doğru Sınıflama		
90-10	17,50%	57,50%	66,60%	19,20%	55,40%	64,70%	66,40%	72,10%
80-20	20,80%	48,50%	67,20%	24,80%	51,50%	65,40%	67,10%	72,90%
70-30	22,10%	50,40%	66,50%	25,50%	53,20%	63,30%	65,58%	71,90%
60-40	20,70%	57,30%	64,70%	17,90%	53,80%	67,30%	65,75%	70,50%
50-50	21,00%	51,50%	66,90%	21,90%	56,30%	64,00%	65,47%	73,20%
40-60	24,10%	39,90%	69,20%	27,40%	46,30%	65,20%	66,78%	75,30%
30-70	20,80%	52,20%	66,50%	21,20%	56,90%	64,40%	65,04%	71,70%

Elde edilen bulgulardan en yüksek doğru sınıflandırmanın % 80 öğrenme ve % 20 sinama aşamasından geçirilmesi sonucunda elde edildiği görülmüştür. Öğrenme setinde en düşük 1.tip hataya sahip yapı, % 90 öğrenme % 10 eğitim seti ile oluşturulan yapıdır. Öğrenme setinde en düşük 2. tip hataya sahip ağ yapısı, %40 öğrenme % 60 sinama aşamasından geçirilmesi ile yapay sinir ağının elde ettiği görülmüştür. Öğrenme verisetinin azaltılması ve sinama verisetinin artırılması sonucunda % 50 öğrenme-% 50 sinama ve % 40 öğrenme-% 60 sinama aşamalarından geçirilen yapay sinir ağının alıcı çalıştırma karakteristiği değerleri de yükselmiştir.

7.4. Elde Edilen Sonuçların Yorumlanması

Ticari bir banka tarafından elde edilen veriseti için lojistik regresyon ve yapay sinir ağları metotları ile skorkatlar oluşturulmuştur. İki yöntemin karşılaştırılmasında toplam doğruyu bilme başarı yüzdesinde, 10 giriş katmanlı 21 ara katmanlı 1 çıktı katmanlı yapay sinir ağının lojistik regresyon yöntemine göre çok az farkla daha başarılı olduğu görülmüştür. Elde edilen sonuçlarda 1. tip hata ve 2. tip hata oranlarının karşılaştırılmasında yapay sinir ağlarının, 2.tip hata tespitinde yani temerrüte düşmüş kredilerin tespitinde daha başarılı oldukları görülmüştür. Finans kuruluşları açısından 1. tip hata yani iyi kredilerin ret edilmesi, 2.tip hataya oranla çok daha az maliyetlidir. Skorkartlardan elde edilecek sonuçlarda iyi kredilerin kabul edilmemesinden ziyade kötü kredilerin tespit edilemeyip kabul edilmeleri sonucunda

temerrüde düşmeleri finans kuruluşlarının daha fazla emek harcanmasına ve daha fazla zarar etmesine yol açmaktadır.

Yapay sinir ağları kullanılmasında en önemli zorluklardan bir tanesi, verisetine uygun olarak ağ yapısının seçilmesidir. Bu çalışmada literatürde kredi skorlama modellerinde kullanılan çok katmanlı algılayıcı yapay sinir ağı ile çalışılmıştır. Ağ yapısı yani giriş, ara ve çıktı katmanları sayıları, momentum ve öğrenme katsayısı gibi parametreler deneme yanılma yolu ile elde edilmiştir. Buradaki zorluk hangi yapay sinir ağının seçileceğinin tespit edilmesinin ardından spesifik bir ağ yapısının mevcut olmamasıdır. Ara katman sayısı bulmak için 2 ayrı yöntem izlenmiştir. Birincisi girdi, çıktı ve gözlem sayısına bağlı olarak formüleştirmeye dayalı bir ara katman sayısı ve ikincisi yöntem deneme yanılma yolu ile elde edilen ara katman sayısıdır. Girdi ara ve çıktı katmanının belirlenmesinin ardından çeşitli öğrenme ve momentum hızlarında ağlar eğitilmiştir. Deneme yolu ile elde edilen çok katmanlı algılayıcı 10-21-1 yapısında yapay sinir ağının formüleştirmeye yolu ile elde 10-16-1 çok katmanlı algılayıcı yapay sinir ağına oranla daha başarılı kredi sınıflandırması yaptığı görülmüştür.

Literatür çalışmalarında, yapay sinir ağları üzerinde kredi skorlama yaklaşımlarında öğrenme ve sına veriseti, genel olarak %70 öğrenme-%30 sına aşamalarından oluşarak ağ yapısı oluşturulmaktadır. Çalışmada öğrenme ve sına veriseti yüzdeleri değiştirilerek ağların performansları karşılaştırılmıştır. Öncelikle formüleştirmeye-deneme yöntemi ile 10 giriş katmanlı, 16 ara katmanlı ve 1 çıktı katmanlı yapay sinir ağı denenmiş, elde edilen sonuçlarda %70 öğrenme %30 sına aşamasından geçen ağın en yüksek toplam doğru kredileri tespit ettiği görülmüştür. Aynı durum sadece deneme yoluyla elde edilen 10 giriş katmanlı 21 ara katman ve 1 çıktı katmanlı yapay sinir ağı için de söz konusudur. Fakat finansal riskin en yüksek olduğu alıcı çalıştırma karakteristiklerinde, her iki yapay sinir ağında en yüksek performansı gösteren öğrenme-sına aşamalarını %40 öğrenme %60 sına aşaması oluşturmuştur.

Kredi skorlama teknikleri olarak literatürde bir çok metodoloji kullanılmaktadır. Bunlardan en yaygınları ikili sonuçları ayırt etmede başarılı performans gösteren lojistik regresyon modeli ve yapay sinir ağlarıdır. Son yıllarda birçok yeni metod kullanılmasına karşın özellikle hibrid yapıların başarılı oldukları görülmektedir. 2008 yılına ait tüketici kredisi bilgilerinden yola çıkılarak, lojistik regresyon ve çok

katmanlı algılayıcı sinir ağlarından elde edilen sonuçlar doğrultusunda literatürle paralel olarak toplam doğru kredi bulma performansları çok farklılık göstermemiştir. Kredi veren kuruluşlar tarafından 1.tip hata oranına oranla daha fazla zarara maruz kaldıkları 2. tip hata oranında, yapay sinir ağlarının daha düşük düzeyde bir hata oranına sahip oldukları görülmüştür.

Yapay sinir ağları, performansları açısından ağ yapıları çok önemli bir yer tutmaktadır. Bazı çalışmalar, belirli bir formülleştirme kullanarak ağ yapısındaki ara katman sayısını tespit edip öğrenme ve momentum hızlarını deneme yanılma yoluyla bulmaktadırlar. Diğer çalışmalarda, ara katman sayısı da deneme yanılma yoluyla bulunmaktadır. Mevcut veriseti, her iki yöntem için denemiş ve sadece deneme yanılma yolu ile elde edilen yapay sinir ağının, formülleştirme-deneme yolu ile elde edilen ağ yapısına göre daha iyi performans gösterdiği görülmüştür.

Modelde kullanılan veriseti yapay sinir ağı ağ yapıları sabit tutularak, farklı öğrenme sına aşamalarından geçirilerek performansları incelenmiştir. Literatürde kabul gören %70 öğrenme %30 sına verisine oranla, %40 öğrenme %60 sına aşamasından geçen ağların, 2. tip hata tespitlerinde daha başarılı oldukları görülmüştür.

Kredi skorlama modeli olarak yapay sinir ağlarının kullanılabileceği ve çeşitli şekillerde performanslarının artırılabilceği görülmüştür. Yapay sinir ağları, sonuçları konusunda bilgi vermemesi finansal kuruluşlar açısından sorun durdurulabilmektedir. Model sonucunda, kredi talebi kabul edilmeyen kişi hakkında yapay sinir ağları bilgi vermemektedirler. Ayrıca çeşitli kümeleme ve hibrid yöntemlerle, yapay sinir ağlarının kredi skorlama modeli olarak performanslarının iyileştirilebileceği konusu irdelenebilir.

8. SONUÇLAR

Amerika’da ortaya çıkan ve bütün Dünya’ya yayılan 2008 yılı krizi sonrasında bankaların yeniden yapılandırılması söz konusu olmuş devletler bazı bankalara maddi veya bankaların belirli bir kısım hisselerini satın alarak destek olmuşlardır. Krizin Amerika’da ortaya çıkmasına karşın etkisi neredeyse tüm Dünya’da ve ülkemizde hissedilmiştir. Bu süreçte 2000’li yıllarda Türkiye’de yapılan çeşitli bankacılık uygulamaları sayesinde krizin etkisinin çok büyük hissedilmediği görülmüştür.

Krizin ardından ekonomik olarak bir toparlanma olduğu gözlenmiş olsa da farklı görüşler ortaya atılmıştır. Bu krizin 1. dalga olduğu ve daha farklı etkiler ile daha derin krizlerin ortaya çıkabileceği savunulmuştur. Günümüzde Avrupa’da yaşanan borç krizinin çözilememesi durumunda, 2008 krizine oranla daha ciddi bir global kriz yaşanması durumunun söz konusu olacağı dile getirilmektedir.

Finans dünyasının da globalleşmeye maruz kalması ile Türkiye’de finansal kuruluşların en önemlileri olan bankalar üzerindeki denetim mekanizmaları arttırılmış, diğer yandan cari açık sorununu dile getirilerek kredi hacmindeki artış ve verilen kredi çeşitleri üzerine devlet kurumları dikkat çekmektedirler. Kredi konusunda tutunulan bu hassas tavır neticesinde banka karlarında düşüşler gözlenmiş ve bankalar alternatif yollardan karlarını arttırmaya çalışmışlardır. Bankaların sundukları hizmet karşılığında müşterilerinden aldıkları masraflarda bu dönemde bir artış meydana gelmiştir. Ayrıca bankalar, bu dönemde, müşteri seçiminde yüksek kar elde edebilecekleri ürünlere yoğunlaşmışlardır.

Bankalar kullandıkları krediler doğrultusunda belirli bir risk ile karşı karşıya kalmaktadırlar. Kredinin temerrüt durumu olarak adlandırılan bu durumda, temerrüt eden krediler ile ilgilenmek amacıyla, bankaların içinde farklı departmanlar bulunmaktadır. Temerrüt eden kredilerin çokluğu bir banka açısından hem finansal anlamda zarara neden olur hem de itibarı üzerinde negatif etkiye sebep olmaktadır..

Kredi ve risk kavramlarının ortaya çıkmasından itibaren, bu riski çeşitli şekillerde bertaraf etmeye çalışılmaktadırlar. Kredi skorlama yöntemleri, bu düşünceden yola

çıkılarak kullanılmaya başlanılmıştır. Yaklaşık 50 yıl önce ortaya çıkan bu yaklaşımlarda, iyi kredi ve kötü kredi çeşitli tekniklerle tespit edilmeye çalışılmıştır. Literatürde, istatistiksel ve istatistiksel olmayan yöntemlerin kullanılması sonucunda bu ayırım yapılmaya çalışılmaktadır. Kredi skorlama teknikleri ile insan gücüne duyulan ihtiyaç kısmi olarak azaltılmış ve müşteri kredi taleplerine daha hızlı cevap verilmeye başlanmıştır.

Kredi skorlama teknikleri, matematiksel model geliştirme teknikleriyle paralel olarak gelişmiştir. Problemin yapısı itibarıyla ilk kullanılmaya başlandıklarında, lineer regresyon ve diskriminant analizi gibi metotlar kredi skorlama yöntemi olarak kullanılmış, zaman ilerledikçe ileri düzey algoritmaların ve bilgisayarların gelişmesi sonucunda çok karmaşık modeller kredi skorlama çalışmaları için kullanılmıştır.

Çalışma kapsamında ticari bir bankadan elde edilen verisetleri doğrultusunda farklı iki yöntemle kredi skorlama modelleri oluşturulmaya çalışılmıştır. Bu yöntemlerden ilki, lojistik regresyon ile kredi skorlama modelidir. 1980'lerden itibaren lojistik regresyon, kredi skorlama modeli olarak kullanılmış ve kabul görmüştür. İkili sonuçlara uygun yapıda çıktı vermesi, diğer istatistiksel yöntemlerin aksine girdi veri setleri için herhangi bir varsayıma sahip olmaması gibi özellikleri bu yöntemin kredi skorlama çalışmalarında kullanılmasını kolaylaştırmıştır. İkinci yöntem, uzman sistemler olarak adlandırılan, yani bilgisayar programları aracılığıyla kullanılan yöntemlerden biri olan yapay sinir ağları yöntemleridir. Yapay sinir ağı lineer olmayan problemlerin çözümünde gösterdikleri başarı doğrultusunda kredi skorlama metodu olarak kullanılmış ve başarılı sonuçlar elde edildiği gözlenmiştir. Yapay sinir ağlarında karşılaşılan en büyük problem, optimum ağ mimarisinin bulunması aşamadır. Bu aşama momentum, öğrenme katsayıları ve gizli katman nöron sayısının bulunmasını içermektedir. Belirli bazı bulgular doğrultusunda bu değerlere ulaşılmaya çalışılmıştır.

Çalışmanın uygulama kısmının birinci aşamasında, ticari bir banka tarafından elde edilen tüketici kredi veriseti için lojistik regresyon ile kredi skorlama modeli kurulmuş ve sonuçlar irdelenmiştir. İkinci aşamada, yapay sinir ağlarının, ağ mimarisi deneme yanılma yolu ile bulunmuş, ardından gizli nöron sayısını veren bir formülasyon kullanılarak iki çeşit çok katmanlı yapay sinir ağı elde edilmiştir. Momentum ve öğrenme katsayıları deneme yanılma yoluyla elde edilmeye çalışılıp, en iyi ağ yapısına ulaşılmaya çalışılmıştır. Birinci tip hata, ikinci tip hata ve toplam

hata yüzdesi bakımından her iki yöntem ile elde edilen değerler karşılaştırılmış ve genel doğruluk yüzdesi bakımından yapay sinir ağlarının lojistik regresyona göre daha iyi performans verdiği görülmüştür. Finansal kuruluşlar açısından daha fazla maliyete sonuç olan ikinci tip hata oranında, literatür çalışmalarına paralel olarak lojistik regresyon yöntemine oranla yapay sinir ağlarında elde edilen ikinci tip hata oranının daha düşük olduğu görülmüştür.

Uygulamada en son aşamada elde edilen 2 yapay sinir ağı ağ yapısı, farklı öğrenme ve sınaama verisetlerinde denenmiştir. Burada test edilmek istenen, literatür çalışmalarında yoğun olarak kullanılan yüksek öğrenme ve düşük sınaama aşamalarından ziyade düşük öğrenme yüksek sınaama aşamalarında yapay sinir ağlarının performanslarıdır. Elde edilen sonuçlarda, düşük öğrenme ve yüksek sınaama oranlarında, yapay sinir ağlarının başarılı sonuçlar verebileceği görülmüştür.

KAYNAKLAR

- Abdou, H., Pointon, J. ve El-Masry, A.** (2008). Neural nets versus conventional techniques in credit scoring in egyptian banking. *Expert Systems with Applications*, **35**, 1275-1292
- Alavala, C.R.** (2003). *Fuzzy logic and neural Networks: Basic concept & appliations*, New Age International Publisher.
- Aloğlu, Z.A.** (2005). Bankacılık sektörünün karşılaştığı riskler ve bankacılık krizler üzerindeki etkileri. *Uzman Yeterlilik Tezi*, Türkiye Cumhuriyet Merkez Bankası Bankacılık ve Finansal Kuruluşlar Genel Müdürlüğü, Ankara.
- Anderson, R.** (2007). *The Credit Scoring Toolkit*, Oxford University Press
- Ataçoğlu, H.** (2006). Kredi riski takibi, “Sorunlu krediler ve erken uyarı sistemleri”. *Doktora Tezi*, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü İktisat Ana Bilim Dalı, İstanbul.
- Bayru, P.** (2007). Elektronik basında tüketici tercihleri analizi: Yapay sinir ağları ile lojit modelin performans değerlendirilmesi, *Doktora Tezi*, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Ana Bilim Dalı, İstanbul.
- Bankalar Kanunu:** Türkiye Bankalar Birliği, İstanbul, 1999.
- Bankacılık Düzenleme ve Denetleme Kurumu,** Finansal Piyasalar Raporu, Sayı:22, Haziran 2011.
- Bellotti, T. ve Crook J.** (2009). Support vector machines for credit scoring and discovery of significant features. *Expert Systems with Applications*, **36**, 3302-3308.
- Bigus, J. P.** (1996). *Data Mining with Neural, Networks*McGraw-Hill.
- Chen, M. C., Huang S. H.** (2003). Credit scoring and reject instances reassigning through evolutionary computation techniques, *Expert Systems with Applications*, **24**, 433-441.
- Çinko, M.** (2006). Kredi kartı değerlendirme tekniklerinin karşılaştırılması, İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi, 143-153.
- Dalkılıç, N.** (2010). Muhasebe mesleği sorumluluk sigortalarında yapay sinir ağları yöntemi ile risk değerlemesi. *Doktora Tezi*, Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü İşletme Bilim Dalı, Kütahya.
- Eliuz, E.** (2009). Tüketici kredilerinde risk yönetimi ve bir skorkart modeli önerisi, *Yüksek Lisans Tezi*, Marmara Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Anabilim Dalı, İstanbul.
- Fausett, L.**, “Fundamentals of Neural Networks : Architectures, Algorithms and applications”,

- Finanbank Management Trainee Eğitim Notları ,Temel Krediler, İstanbul , 2011.
- Finansbank, Kredi Türleri ve Talimatları, İstanbul ,2011.
- Haykin, S.** (1996). *Neural networks: A comprehensive foundation*, Prentice-Hall Second Edition.
- İnce, H. ve Aktan, B.** (2010). Kredi kartı taleplerin değerlendirmesinde grup ve bireysel kredi puanlama modellerinin karşılaştırılmalı bir analizi, BDDK Dergisi Cilt:4 Sayı:1, Ankara.
- Khasman, A.** (2010). Neural network for credit risk evaluation : Investigation of different neural models and learning schemes. *Expert Systems with Applications*, **37**, 6233-6239.
- Kim, Y. S. ve Sohn, S.Y.** (2004). Managing loan customers using misclassification patterns of credit scoring model. *Expert Systems with Applications*, **26**, 567-573.
- Lee T.S., Chiu, C.C. ve Lu, C.J.** (2002). Credit scoing using the hybrid neural discriminant technique . *Expert Systems with Applications*, **23**, 245-254.
- Lee, T. S., Chiu, C. C., Chou, Y. C. ve Lu C. J.** (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis*, **50**, 1113-1130.
- Kasapoğlu, B.** (2009). Kredi riskinin hesaplanmasında skortlama yaklaşımı, *Yüksek Lisans Tezi*, İstanbul Teknik Üniversitesi Sosyal Bilimler Enstitüsü İktisat Anabilim Dalı, İstanbul.
- Kamruzzaman, J., Begg, R. ve Sarker, R.** (2006). *Artificial neural networks in finance and manufacturing*, Idea Group Publishing.
- Molhotra, R. ve Malhotra, D.K.** (2002). Differentiating between good credits and bad credits using neuro-fuzzy systems. *European Journal of Operational Research* **136**, 190-221.
- Molhotra, R. ve Malhotra, D.K.** (2003). Evaluating consumer loans using neural network. *The International Journal of Management Science Omega*, **31**, p.83-96.
- Ong, C.S., Huang, J.J. ve Tzeng, G.H.** (2005). Building credit scoring models using genetic programming. *Expert Systems with Applications*, **29**, 41-47.
- Öker, A.** (2007). Ticari Bankalarda Kredi ve Kredi Riski Yönetimi – Bir Uygulama. *Doktora Tezi*, Marmara Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı, 2007, İstanbul.
- Öztemel, E.** (2008). *Yapay Sinir Ağları*, Papatya Yayıncılık, İstanbul.
- Özdemir, H.** (2010). Lojistik regresyon ile kredi skortlama ve bir uygulama, *Yüksek Lisans Tezi*, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, İstanbul.
- Siddiqi, N.** (2006). *Credit Risk Scorecards*, John Wiley & Sons, Inc., Hoboken, New Jersey.

- Sustersic, M., Mramor, D. ve Zupan J.** (2009). Consumer credit scoring models with limited data. *Experts Sysems with Application*, **36**, 4736-4744.
- Tekindağ, F.C.** (2010). Kredi risk yönetimi aracı olarak kobi kredileri için alternatif kredi skortlama model önerisi ve uygulama. *Doktora Tezi*, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir.
- Thomas, L.C., Edelman, D.B. ve Crook, J.N.** (2002). *Credit Scoring and Its Applications*, Society For Industrial and Applied Mathematics, Philedelphia.
- Tsai, C. F. ve Wu, J.W.** (2008). Using neural network ensembles for bankruptcy prediction and credit scoring. *Expert Systems with Applications*, **34**, 2639-2649.
- Ünsal, A. ve Dumans S.**” Türkiye’deki Bankaların Performanslarının Temel Bileşenler Yaklaşımı ile Karşılaştırmalı Analizi”.
- Wang, G., Ma, J., Huang L. ve Xu K.** (2011). Two Credit Scoring Models Based on Dual Stratgy Ensemble Trees, *Knowledge Based Systems* **26**, 61-68.
- West, D.** (2000). Neural network credit scoring models. *Computer Operations Research* **27**, 1131-1152.
- Yıldırım, O.** (2007).Türk bankacılık sektörünün temel sorunları ve sektörde yaşanan mali risker, Anadolu Üniversitesi, Alındığı tarih: 08.08.2011,.
- URL-1:**<http://muhasebeturk.org/ecopedia/394-k/4310-kaydi-para-nedir-ne-demek-anlami.html> , alındığı tarih : 10.10.2011.
- URL-2:**www.econturk.org/Turkiyeekonomisi/oguzbanka.doc

ÖZGEÇMİŞ



Ad Soyad: Burak Donel

Doğum Yeri ve Tarihi: 15/10/1984

Adres: Papatyah sokak Baykal Apartmanı Çiftelhavuzlar
İstanbul

E-Posta: budonel@hotmail.com

Lisans: Yıldız Teknik Üniversitesi Metalurji ve Malzeme
Mühendisliği

Yüksek Lisans :

Mesleki Deneyim ve Ödüller: Finansbank Krediler Analitik Yetkili Yardımcısı