

**T.C.**  
**GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**

**VERİ MADENCİLİĞİ TEKNİKLERİ İLE**  
**DÖVİZ KURU TAHMİNİ**

**Erol TUTAL**  
**YÜKSEK LİSANS TEZİ**  
**İŞLETME BÖLÜMÜ ANABİLİM DALI**

**GEBZE**

**2011**

**T.C.**  
**GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**

**VERİ MADENCİLİĞİ TEKNİKLERİ İLE**  
**DÖVİZ KURU TAHMİNİ**

**Erol TUTAL**  
**YÜKSEK LİSANS TEZİ**  
**İŞLETME BÖLÜMÜ ANABİLİM DALI**

**TEZ DANIŞMANI**  
**Doç. Dr. Hüseyin İNCE**

**GEBZE**  
**2011**



## YÜKSEK LİSANS TEZİ JÜRİ ONAY SAYFASI

G.Y.T.E. Sosyal Bilimler Enstitüsü Yönetim Kurulu'nun 20.06.2011 tarih ve 2011/27 sayılı kararıyla oluşturulan jüri tarafından 04.07.2011 tarihinde tez savunma sınavı yapılan **Erol TUTAL**'ın tez çalışması **İşletme** Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

### JÜRİ

ÜYE  
(TEZ DANIŞMANI) : Doç. Dr. Hüseyin İnce

ÜYE : Prof. Dr. Halit Keskin

ÜYE : Doç. Dr. Salih Zeki İmamoğlu

### ONAY

G.Y.T.E. Sosyal Bilimler Enstitüsü Yönetim Kurulu'nun ...../...../20... tarih ve ...../..... sayılı kararı.

İMZA/MÜHÜR

## ÖZET

### VERİ MADENCİLİĞİ TEKNİKLERİ İLE DÖVİZ KURU TAHMİNİ

**Erol TURAL**

Veri madenciliği genel olarak; büyük miktardaki veri içerisinde, gizli kalmış, değerli, kullanılabilir bilgilerin açığa çıkarılması olarak ifade edilebilir. Büyük boyutlardaki verilerden işimize yarayacak, geleceği daha iyi görebilmemizi ve tahminler yapabilmemizi sağlayacak bilgileri elde etmenin en kullanışlı yollarından biri de veri madenciliğidir. Veri Madenciliğinin yaygın olarak kullanıldığı alanlardan biri de döviz kuru tahminidir.

Çalışmada önce Veri Madenciliği süreci, işlevleri ve teknikleri tanıtılmıştır. Daha sonra döviz kuru tahmininde kullanılan Veri Madenciliği tekniklerinden Yapay Sinir Ağları ve Destek Vektör Makineleri kullanılarak örnek bir uygulama da yapılmıştır. Yapılan çalışmada, döviz kurlarının gelecekteki değerinin ve yönünün yüksek doğruluk oranı ile öngörülmesine çalışılmıştır.

Bu çalışmada kullanılan veriler, <http://www.oanda.com/currency/converter/> sitesinden elde edilmiştir. Çalışmada USD, EUR, GBP ve JPY döviz kurlarının 01.01.2001 ile 30.05.2011 tarihleri arasındaki değerleri kullanılmıştır.

**Anahtar kelimeler:** Veri Madenciliği, Döviz Kuru Tahmini, Yapay Sinir Ağları, Destek Vektör Makinesi, MAE, RMSE, NRMSE, MAPE

# **SUMMARY**

## **FORECASTING EXCHANGE RATES WITH DATA MINING TECHNIQUES**

**Erol TUTAL**

Data Mining, generally can be expressed as revealing hidden, valuable and useful knowledge from huge data stack. One of the usable ways of revealing knowledge for using, forecasting future and making predictions is Data Mining. One of the widely used field of Data Mining is Forecasting Exchange Rates.

In this thesis, firstly process, functions and techniques of data mining are described. Then, an application is done by using Artificial Neural Networks and Support Vector Machines which are the ways of data mining techniques for forecasting exchange rates. In application, forecasting of future values and trends of exchange rates in high accuracy rating is processed.

The datas which are used in this study are taken from the internet address of <http://www.oanda.com/currency/converter/>. In study the exchange rates of USD, EUR, GBP and JPY between the date of January 01, 2001 and May 30, 2011 are used.

**Key Words:** Data Mining, Forecasting Exchange Rates, Artificial Neural Networks, Support Vector Machines, MAE, RMSE, NRMSE, MAPE

## TEŐEKKÜR

Tezimin planlanmasında, araştırılmasında, yürütülmesinde ve oluşumunda ilgi ve yardımlarını esirgemeyen, engin bilgi ve tecrübelerinden yararlandığım değerli hocam ve danışmanım sayın Doç. Dr. Hüseyin İNCE' ye sonsuz teşekkürlerimi sunarım.

Tez çalışmam süresince yardımlarını esirgemeyen ve sonuçlandırılmasına büyük katkı ve imkânlar sağlayan değerli hocalarım Prof. Dr. Halit KESKİN ve Doç. Dr. Salih Zeki İMAMOĞLU ile tüm GYTE ailesine teşekkürlerimi sunarım.

En değerli hazinem aileme, eşim Sevim TUTAL ve emeği geçen tüm dostlarıma manevi hiçbir yardımı esirgemedен yanımda oldukları için tüm kalbimle teşekkür ederim.

Erol TUTAL

# İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
<b>ÖZET</b>	iv
<b>SUMMARY</b>	iv
<b>TEŞEKKÜR</b>	vi
<b>İÇİNDEKİLER DİZİNİ</b>	vii
<b>SİMGELER VE KISALTMALAR DİZİNİ</b>	viii
<b>ŞEKİLLER DİZİNİ</b>	ix
<b>TABLolar DİZİNİ</b>	xi
<b>1. GİRİŞ</b>	1
<b>2. VERİ MADENCİLİĞİ</b>	3
2.1. Veri Madenciliği Tanımı ve Araçları	3
2.2. Veri Madenciliği Süreci	6
2.3. Veri Madenciliği İşlevleri	8
2.4. Veri Madenciliği Teknikleri	10
2.4.1. Regresyon Modeli	10
2.4.2. Sınıflama Modeli	12
2.4.2.1. Karar Ağaçları	14
2.4.2.2. K- En Yakın Komşu Algoritması	14
2.4.2.3. Yapay Sinir Ağları	15
2.4.3. Kümeleme Modeli	16
2.4.4. Birliktelik Kuralı	17
<b>3. DÖVİZ KURU TAHMİNİNDE KULLANILAN TEKNİKLER</b>	19
3.1. Çalışmaya İlişkin Literatür Özeti	20
3.2. Yapay Sinir Ağları	24
3.3. Destek Vektör Regresyonu	32
<b>4. VERİ MADENCİLİĞİ UYGULAMASI</b>	39
4.1. Çalışmanın Tanıtılması	39
4.1.1. Çalışmanın Amacı ve Önemi	39
4.1.2. Çalışmanın Kapsamı	39
4.1.3. Çalışmada Kullanılan Değişkenler	39

4.2.	Döviz Kuru Tahmini ve Modellemesi	41
4.2.1.	Veri Seti	41
4.2.2.	Değişkenler ve Kriterler	41
4.2.3.	Modelleme ve Analiz Sonuçları	43
4.2.4.	YSA ve DVM ile Yapılan Tahminin Karşılaştırılması	49
<b>5.</b>	<b>SONUÇ</b>	<b>51</b>
	<b>KAYNAKLAR</b>	<b>53</b>
	<b>ÖZGEÇMİŞ</b>	<b>59</b>



## SİMGELER VE KISALTMALAR DİZİNİ

et al.	: ve diğerleri
vd.	: ve diğerleri
p	: sayfa
s	: sayfa
vb.	: ve benzeri
CRISP DM	: Cross-Industry Standard Process for Data Mining
GYA	: Geri Yayılım Algoritması
YSA	: Yapay Sinir Ağları
DVM	: Destek Vektör Makinesi
DVR	: Destek Vektör Regresyonu
ARIMA	: Auto Regressive Integrated Moving Average
MAE	: Mean Absolute Error - Ortalama Mutlak Hata
RMSE	: Root Mean Squared Error - Kök Hata Kareler Ortalaması
NRMSE	: Normalised Root Mean Squared Error - Düzeltilmiş Kök Hata Kareler Ortalaması
MAPE	: Mean Absolute Percentage Error - Ortalama Mutlak Yüzde Hata

## ŞEKİLLER DİZİNİ

<b><u>Şekil</u></b>	<b><u>Sayfa</u></b>
Şekil 1: CRISP DM Süreci	7
Şekil 2: Basit Bir Karar Ağacı	14
Şekil 3: Yapay Sinir Modeli	26
Şekil 4: Basit Bir Yapay Sinir Ağı	27
Şekil 5: Bir Geriye Yayılım Algoritması Ağı Modeli	31
Şekil 6: Destek Vektör Regresyon (DVM) 'un Grafik Gösterimi	33
Şekil 7: DVM Kullanarak USD Kuru Tahmini	44
Şekil 8: DVM Kullanarak EUR Kuru Tahmini	44
Şekil 9: DVM Kullanarak GBP Kuru Tahmini	45
Şekil 10: DVM Kullanarak JPY Kuru Tahmini	45
Şekil 11: YSA Kullanarak USD Kuru Tahmini	46
Şekil 12: YSA Kullanarak EUR Kuru Tahmini	47
Şekil 13: YSA Kullanarak GBP Kuru Tahmini	47
Şekil 14: YSA Kullanarak JPY Kuru Tahmini	48

## TABLÖLAR DİZİNİ

<b><u>Tablo</u></b>	<b><u>Sayfa</u></b>
Tablo 1: Veri Madenciliđi Araçları Kullanım Oranları	6
Tablo 2: Veri giriři ve Sinir ađının ađırlıkları için bařlangıç deđerleri	28
Tablo 3: DVM'ye Gre Hesaplanan Deđerler	46
Tablo 4: YSA'ya Gre Hesaplanan Deđerler	48
Tablo 5: DVM ve YSA'ya Gre Hesaplanan Deđerler	49

# 1. GİRİŞ

Veri madenciliği büyük veri yığını içerisinde, gizli kalmış, değerli, kullanılabilir bilgilerin açığa çıkarılması olarak ifade edilebilir. Günümüzde işletmelerin yoğun teknoloji ve bilgisayar kullanımlarının artmasıyla birlikte müşteri verileri elektronik ortamda tutulmaya başlanmış, büyük miktarlardaki veriyi işleme yeteneğine sahip teknolojilere gereksinim duyulmuştur. Teknolojinin gelişimiyle bilgisayar ortamında ve veritabanlarında tutulan veri miktarının artması Veri Madenciliğini zorunlu kılmıştır. Veri Madenciliği uygulamalarını gerçekleştirmek için çeşitli programlar geliştirilmiştir.

Veri Madenciliği tekniklerinin yaygın bir şekilde kullanıldığı alanlardan biri de döviz kuru tahminidir. Ekonomide işletmelerin planlama yapabilmeleri açısından geleceğin öngörülebilmesi çok önemli bir gereksinimdir. Özellikle son yıllarda dünyada ve ülkemizde yaşanan ekonomik krizlerden ve politik gelişmelerden dolayı meydana gelen piyasalardaki dalgalanmalar, yatırım araçlarının ve döviz kurlarının gelecekte alacağı değerini öngörülebilmesini önemli hale getirmiştir.

Uluslararası piyasalarda alışveriş ve bir yatırım aracı olarak kullanılan dövizin önemi, işlerini döviz üzerinden gerçekleştiren işletmeler için daha fazladır. Son yıllarda geleceği öngörmede; özellikle döviz kuru gibi doğrusal olmayan, birçok değişkene bağımlı zaman serilerinin tahmin edilmesinde veri madenciliği önemli bir araç olmuştur.

Yapay Sinir Ağları(YSA) ve Destek Vektör Regresyonu(DVR), çözümü güç ve karmaşık olan bir çok problemin çözümüne uygulanmış ve son derece başarılı sonuçlar elde edilerek, birçok alanda uygulanmaya başlanmıştır. Ekonomi alanında geleceği öngörmede; özellikle döviz kuru gibi lineer olmayan değişkenlerin, bağımlı zaman serilerinin tahmin edilmesinde önemli araçlar olmuşlardır.

Çalışmanın ilk bölümünde; veri madenciliği tanımı ve araçları genel olarak anlatılmış, veri madenciliği süreç, işlev ve teknikleri verilmiştir. İkinci bölümde; döviz ve döviz kuru hakkında genel bir bilgi verildikten sonra döviz kurunu tahmin

etmede kullanılan yöntemlerden Yapay Sinir Ağları ve Destek Vektör Regresyonu açıklanmıştır. Üçüncü bölümde; döviz kuru tahmini konusunda MATLAB kullanılarak gerçekleştirilen örnek bir uygulama sunulmuştur. Gerçekleştirilen uygulamada, USD, EUR, GBP ve JPY döviz kurlarının 10 yıllık verisi kullanılarak döviz kurlarının gelecekteki değerinin ve yönünün yüksek doğruluk oranı ile öngörülmesine çalışılmıştır. Dövizin gelecekteki değerinin ne olacağını, yükselip düşeceğini öngörülmesinde Yapay Sinir Ağları ve Destek Vektör Makinesi tekniklerinin etkinlikleri ve doğruluk oranları karşılaştırılarak incelenmiştir.

## 2. VERİ MADENCİLİĞİ

### 2.1. VERİ MADENCİLİĞİ TANIMI VE ARAÇLARI

Günümüzde sadece bilgiye ulaşmak değil, gerekli şartlar altında bilgi üretmek de önemli bir konu haline gelmiştir. Sürekli büyüyen sayısal veri ortamları arasından yararlı ve de gerekli olan bilgiye ulaşmayı sağlamak gerçek bir çaba haline gelmektedir. Veri madenciliği bu safhada öne çıkan ve önemi gitgide artan bir olgudur. (Öğüt,2005)

Bilgisayar ve iletişim teknolojilerindeki gelişmelere paralel olarak donanımın ucuzlaması verilerin uzun süre depolanabilmesini, dolayısıyla büyük kapasiteli veritabanların oluşmasını sağlamıştır. Bu nedenle büyük veritabanlarında istenilen anlamlı ve kullanılabilir bilgiye erişmek yeni bir disiplinin olan Veri Madenciliği'nin doğmasına sebep olmuştur. Verilerin çeşitli istatistiksel metotlarla analiz edilerek, kurumların karar verme sürecinin etkinliğini artırmasına ve yeni stratejiler geliştirmesine katkı sağlanmaktadır (Kaya ve Köymen,2008).

Veri madenciliği, genellikle büyük gözlemsel verilerin aralarında beklenmedik ilişkileri keşfederek ve bunları veri sahibi için hem anlaşılır hem de kullanılabilir bir biçimde yeni bir yöntemle özetleyerek analiz etme işlemidir. Veri madenciliği, makine öğrenimi, örüntü tanıma, istatistik, veritabanları ve büyük veritabanlarından bilginin çıkarılması konusunun ele alınması için görselleştirme tekniklerini bir araya getiren bilimler arası bir alandır. Veri madenciliği her geçen gün daha yaygın olarak kullanılmaktadır. Çünkü bu, şirketlere karlı modelleri ve veritabanlarındaki trendleri bulma gücünü vermektedir. Şirketler ve enstitüler, megabyte ve terabyte miktarındaki verileri toplamak için milyonlarca dolar harcamaktadırlar, fakat bu veri yığınları içinde derinlerde gizli kalmış değerli ve kullanılabilir bilginin avantajını elde edememektedirler. Yine de veri madenciliği uygulamaları daha yaygın oldukça, bu teknikleri uygulamayan şirketler geride kalma ve Pazar payını kaybetme tehlikesi

içindedirler. Çünkü rakipleri veri madenciliği kullanmaktadırlar ve rekabet gücünü elde etmektedirler (Larose,2005).

Teknolojinin gelişimiyle birlikte bilgi, internet ortamında bolca bulunmaktadır. Uygun ve faydalı bilginin elde edilmesi sürecinde kullanıcılar istenen bilgiyi ve kaynakları aramak, çıkarmak, süzmek, analiz etmek ve değerlendirmek için daha çok teknik ve otomatik yöntem kullanmak zorundadırlar. Veri yığınlarındaki devasa artış, veri analizi ve işlemesine benzeri görülmemiş bir ihtiyaç artışına sebep olmuştur. Veri madenciliği, otomatik ve ileri veri analizi için katalizör görevi üstlenmiştir. Veri madenciliği belirli , daha önce bilinmeyen ve faydalı olması beklenen bilginin geniş veri yığından elde edilmesi olarak tanımlanabilir. Veri madenciliği aşırı bilgi yüklemesini azaltır ve karar vermeyi güçlendirir. Bu sonuç, organizasyonlar tarafından toplanan birçok bilginin içindeki ilişki ve örüntüleri araştırma işlemi sonucunda faydalı bilginin süzülmesi ve çıkarılması ile elde edilir (Wang,2009).

ZDNET News Online Technology Magazine'e göre Veri Madenciliği gelecek on yılın en önemli devrimci gelişmelerinden biri olarak tahmin edilmiştir. Aslında MIT Technology Review, Veri Madenciliğini dünyayı değiştirecek on yeni teknolojiden biri olarak seçmiştir. Müşterilerden, ortaklardan ve tedarikçilerden sağlanan bilgilerin etkin bir şekilde analiz edilmesi birçok şirket için önem kazanmıştır. BridgeGate LLC'nin yönetici ortağı Duddley Brown'ın ifadesine göre birçok şirket veri yığını stratejisi oluşturdu ve şu anda bütün bu veriyle ne yapacaklarını incelemektedir. Gartner Group'a göre Veri Madenciliği, istatistik ve matematik teknikleri kadar model belirleme teknolojilerini kullanarak depolanmış büyük miktarda veriyi eleyerek anlamlı yeni bağıntıları, modelleri ve eğilimleri keşfetme işlemidir (Larose,2005).

Günümüzde veri madenciliği bir çok alanda kullanılmaktadır. Veri madenciliği operasyonel kararların alınmasında, stratejik ve politik kararların verilme süreçlerinde önemli bir yere sahiptir. Günümüzde özel sektör veya kamusal sektör, Müşteri İlişkileri Yönetimi (CRM) Kurumsal Kaynak Planlaması (ERP) gibi çeşitli uygulamalar ve teknikler vasıtasıyla veri madenciliği tekniğini kullanmaktadır.

İstatistik ile olan bağlantısı, veri madenciliğini tıp ve ekonomi gibi bilim dalları için de vazgeçilmez hale getirmiştir (Öğüt,2005).

Veri Madenciliği tekniği pazarlama ve satış alanında hedef pazarların tespitinde, müşteri ilişkilerinin yönetiminde, veri kaynaklarını işlemek için müşteri kartı bilgilerinin kaydedilmesinde, müşteri şikayetlerinin incelenmesinde, e-ticarete, satış kampanyalarının verimlilik analizlerinin yapılmasında, reklamcılıkta, indirim kartları ve bonus tespitinde, karlılığın artırılması gibi birçok değişik alanda veri madenciliğinden yararlanılmaktadır. Ayrıca, astronomi, biyoloji, finans, sigorta gibi bir çok başka alanda da uygulanmaktadır. Son 20 yıldır Amerika Birleşik Devletleri'nde çeşitli veri madenciliği algoritmalarının gizli dinlemeden, vergi kaçakçılıklarının ortaya çıkarılmasına kadar çeşitli uygulamalarda da kullanıldığı görülmektedir (Altıntaş,2006).

Yüksek performanslı veya paralel işlemcili bilgisayar sistemleri kullanıldığında veri madenciliği araçları, “Hangi müşterilerin benim promosyon e-maillerime ilgisi var ve sebebi ne” gibi müşteri merkezli soruların cevaplandırılması amacıyla büyük miktardaki veritabanlarını analiz edebilir. (Wang,2009) Veri madenciliği algoritmalarının gelişimi ve uygulaması, güçlü yazılım araçlarının kullanımına ihtiyaç duymaktadır. Mevcut araçların sayısı geliştikçe en uygun aracın seçimi artarak zorlaşmaktadır. (Mikut and Reischl, 2011) Veri madenciliği araçları gelecekteki davranış ve trendleri tahmin ederek işletmelerin etkin kararlar almalarında büyük rol oynar. Aynı zamanda bu araçlar, çözümü çok zaman gerektiren işletme sorunlarına da yanıt bulabilmektedir (Thearling, <http://thearling.com/text/dmwhite/dmwhite.htm>).

Veri madenciliği araçları veritabanında depolanmış olan veriyi yararlı bilgiye dönüştürmek için kullanılır. Değişik modeller kullanan değişik yazılım programları değişik sonuçlar elde etmek için aynı veri üzerinde çalıştırılır. Bazen sonuç umulmadık şekilde olabilir, ama kuralların ve keşfedilmiş sebep ilişkilerinin çoğu trendleri karşılayabilmektedir (Wang,2009, p.258).



2010 Mayıs ayında 912 kişi ile yapılan bir ankete göre, son 12 ayda gerçek bir proje için veri madenciliği araçlarının kullanım oranları Tablo.1'de verilmiştir. Tablo.1'de de görüldüğü gibi en çok tercih edilen veri madenciliği araçları RapidMiner, R ,Excell ve Knime'dir.

Tablo.1. Veri Madenciliği Araçları Kullanım Oranları  
(<http://www.kdnuggets.com/polls/2010/data-mining-analytics-tools.html>,2010)

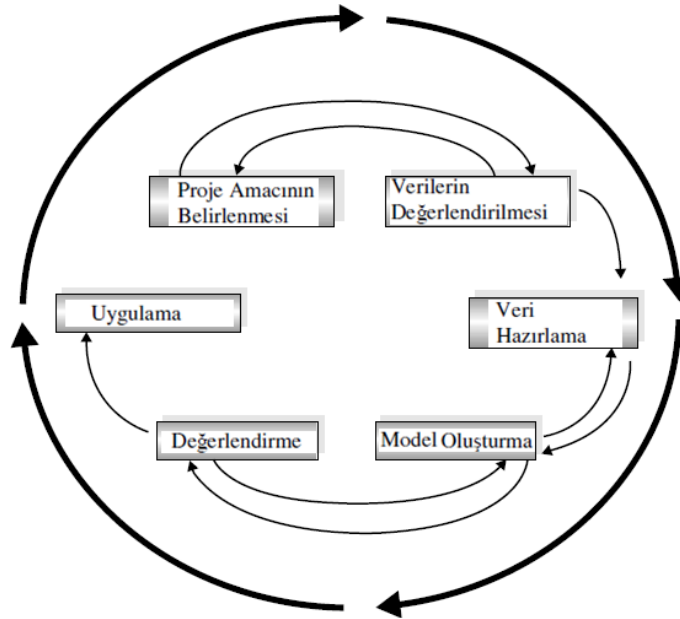
RapidMiner (345)	37.8%
R (272)	29.8%
Excel (222)	24.3%
KNIME (175)	19.2%
Your own code (168)	18.4%
Pentaho/Weka (131)	14.3%
SAS (110)	12.0%
MATLAB (84)	9.2%
IBM SPSS Statistics (72)	7.9%
Other free tools (67)	7.3%
IBM SPSS Modeler (former Clementine) (67)	7.3%
Microsoft SQL Server (63)	6.9%
Statsoft Statistica (57)	6.2%
Other commercial tools (56)	6.1%
SAS Enterprise Miner (50)	5.5%
Zementis (34)	3.7%
Orange (25)	2.7%
Oracle DM (19)	2.1%
KXEN (19)	2.1%
Diğer	8.1%

## 2.2. VERİ MADENCİLİĞİ SÜRECİ

Veri madenciliği süreçleri konusunda 1996 yılında DaimlerChrysler, SPSS, ve NCR'ı temsil eden bir analist grubu yoğun çalışmalar yaptılar ve sonuçta CRISP DM ismini verdikleri bir veri madenciliği standardı geliştirdiler.

## CRISP DM (Cross-Industry Standard Process for Data Mining)

CRISP DM, tescilli olmayan ve serbest kullanıma açık veri madenciliğini iş veya araştırma bölümünün genel problem çözme stratejisine uydurmak için bir işlem standardı sunmaktadır. Şekil.1'de verilen CRISP DM'e göre verilen bir veri madenciliği projesinin 6 basamaktan oluşan bir yaşam döngüsü vardır (Larose,2005):



Şekil.1. CRISP DM Süreci (Larose,2005)

**Proje amacının belirlenmesi :** Bu aşamada projenin hangi sektörde ne amaçla kullanılacağını, projenin sonunda neyin hedeflendiğinin, nelere ihtiyaç duyulduğunun, proje sonunda elde edilecek bilginin nasıl değerlendirileceğinin açıkça ortaya koyulması gerekir.

**Verilerin değerlendirilmesi :** Bu aşama verinin toplanmasıyla başlar. Veri analizcisi ve alan uzmanları açıklayıcı veri analizi gibi yöntemlerle veriyi tanımaya, kalitesi hakkında fikir sahibi olmaya çalışırlar.

**Verilerin hazırlanması :** Bu aşamada işlenmemiş verinin projede kullanılabilir duruma getirilmesi amaçlanır. Hatalı veya analizin yanlış

yönlenmesine neden olabilecek veriler temizlenir. Bu aşama en çok iş gücü gerektiren ve toplam süreç içinde en fazla zaman alan aşamadır.

**Modelleme** : Bir veri madenciliği problemi için birden fazla teknik kullanılabilir, problem için uygun olan teknik veya tekniklerin bulunabilmesi için birçok teknik oluşturulup bunların içinden en uygun olanlar seçilir.

**Değerlendirme** : Bu aşamada, daha önce oluşturulmuş olan model, uygulamaya koyulmadan önce son kez tüm yönleriyle değerlendirilir, kalitesi ve etkinliği ölçülür.

**Uygulama** : Kurulan ve geçerliliği kabul edilen model doğrudan bir uygulama olabileceği gibi başka bir uygulamanın alt parçası olarak da kullanılabilir. İşlenen veri kullanıcının anlayabileceği, karar alma sürecinde kullanılacak bir şekilde son kullanıcıya verilir.

## 2.3. VERİ MADENCİLİĞİ İŞLEVLERİ

Veri madenciliği, verilen veri setinden istenilen bilgileri elde etmek amacıyla çeşitli işlevleri görmektedirler. Halihazırdaki yaygın veri madenciliği işlevleri şunları içerir (Larose,2005; Agard and Kusiak,2004; Mitra et al,2002):

- 1. Tanımlama (Description)** : Bazen araştırmacı ve analizciler bir verideki modeli tanımlamanın yollarını kolayca bulmaya çalışırlar. Örneğin bir anketör işsiz kalan birinin mevcut başkanı yapılacak başkanlık seçiminde daha az destekleyeceğini kanıtlayabilir. Model ve trendlerin tanımlanması, bu model ve trendler için muhtemel açıklamaları önerir. Veri madenciliği modeli mümkün olduğunca açık olmalı. Yani veri madenciliği modelinin sonuçları açıklama ve tarif içeren açık bir model tanımlamalı.

- 2. Sınıflandırma (Classification) :** Bir verinin önceden tanımlanmış birçok sınıfa ayrılmasıdır. Sınıflandırma, daha önce görünmemiş veri kayıtlarına tarihi verilerden elde edilen bilgilere dayanarak etiket atanmasını içerir. Sınıflandırma işlevinin iş ve araştırma alanlarında kullanım örnekleri şunlardır: Bir kredi kartı işleminin sahte olup olmadığının belirlenmesi, mortgage işleminde kredi riskinin iyi veya kötü olduğunun değerlendirilmesi, belirli bir hastalığın olup olmadığının teşhis edilmesi, belirli bir mali ya da şahsi davranışın bir terörist eylem olup olmadığının belirlenmesi.
- 3. Kümeleme (clustering) :** Bir verinin benzerlik matrislerine ya da ihtimal yoğunluk modeline dayalı verilerin doğal gruplandırılması olan çeşitli kümelerden biriyle eşleştirilmesi ve benzer davranıştaki daha küçük alt yığınlara bölmektir. Kümeleme işlemi hedef bir değişkeni sınıflandırma ya da tahmin etmeye çalışmaz. Bunun yerine bir veri yığınının birbirine benzer alt gruplara ayrılması için çalışır. Kümeleme işlevinin iş ve araştırma alanlarında kullanım örnekleri şunlardır: Büyük pazarlama bütçeleri olmayan küçük sermayeli niş ürünlerinin hedef pazarlaması, muhasebe denetlemesi amacıyla, mali davranışları tehlikesiz ve şüpheli kategorilere ayırma, veri setinin yüzlerce niteliği olduğu zaman boyut indirgeme aracı olarak kullanılması.
- 4. Tahmin (Prediction):** Sınıflandırmaya benzer, farkı tahmindeki sonuçlar gelecekle ilgilidir. Tahmin işlevinin iş ve araştırma alanlarında kullanım örnekleri şunlardır: Gelecek üç ayın stoklama fiyatının tahmin edilmesi, hız limiti artırıldığında trafik ölüm oranlarının tahmin edilmesi, bir spor turnuvasındaki kazananın istatistiki karşılaştırmalara dayalı olarak tahmin edilmesi.
- 5. Birliktelik (Association) :** Hangi niteliğin birleşeceğinin bilinmesi işidir. İş dünyasında en yaygın olanı ilişki analizi ya da pazar sepet analizi olarak bilinen iki ya da daha fazla nitelik arasındaki ilişkinin belirlenmesi işlevidir. Birliktelik işlevinin iş ve araştırma alanlarında kullanım örnekleri şunlardır : bir şirketin yeni servisine abonelerinin olumlu cevap verenlerin oranının araştırılması, ebeveynleri iyi okuyucu olan çocukların kendilerine kitap okuma oranlarının belirlenmesi, haberleşme ağlarında bozulmanın tahmin edilmesi, bir

süpermarkette hangi ürünlerin bir arada satıldığıının ve hangi ürünlerinde hiçbir zaman bir arada satılmadığının bulunması, yeni bir ilaçta tehlikeli yan etkilerin ortaya çıkma oranlarının belirlenmesi.

Veri madenciliği işlevleri burada sayılanlarla sınırlı olmayıp yaygın olarak öne çıkanları tanımlanmışlardır.

## **2.4. VERİ MADENCİLİĞİ TEKNİKLERİ**

Bir veri madenciliği problemi için tek bir teknik kullanılabileceği gibi, problem için uygun olan teknik veya tekniklerin bulunabilmesi için birçok teknik oluşturulup bunların içinden en uygun olanlar seçilebilir. Veri madenciliğinde en çok kullanılan teknikler regresyon, sınıflandırma, kümeleme ve birliktelik kural modelidir.

### **2.4.1. Regresyon Modeli**

Regresyon, sürekli sayısal bir değişkenin, aralarında doğrusal ya da doğrusal olmayan bir ilişki bulunduğu varsayılan diğer değişkenler yardımıyla tahmin edilmesini içeren bir istatistik yöntemidir. Sınıflandırma ve regresyonun amacı tahmin için kullanılabilecek bir veri madenciliği modeli oluşturmaktır. İstatistik biliminin en önemli konularından birisini regresyon analizi oluşturmaktadır. Regresyon analizi bilimsel deneyler, matematik, finans, ekonomi, tıp gibi bilim alanlarında yoğun olarak kullanılmaktadır (Wang,2009, p.141).

Regresyon teknikleri veriyi bir fonksiyon olarak modellemeye çalışırlar. Bunu yaparken orijinal veri yerine bu fonksiyonu temsil eden katsayıların sadece küçük bir grubu depolanır. En basit regresyon tekniği veri dağıtımını bir doğrusal fonksiyon olarak modelleyen doğrusal regresyondur. Basitliğine ve veriler arasında karmaşık ilişkilerin yakalanmasına izin vermemesine rağmen bu teknik çoğunlukla kabul edilebilir sonuçlar üretmektedir. Ayrıca doğrusal olmayan regresyonlar da vardır. Bu

teknik doğrusal olana göre hesaplama yönüyle kayda değer şekilde daha karmaşıktır fakat daha geniş durumlara uygulanabilmektedir (Wang,2009, p.49).

Regresyon analizinin esası, gözlenen bir olayın değerlendirilirken, hangi olayların etkisi içinde olduğunun araştırılmasıdır. Regresyon analizi yapılırken, gözlenen değerler ve etkiye sahip olayların fonksiyon olarak ifadesi gerekmektedir. Bu amaçla kurulan modele regresyon modeli denmektedir. Bağımsız değişken olarak bir değişken kullanılırsa basit regresyon, birden fazla değişken kullanılırsa, bu regresyon modellerine çoklu regresyon modeli denmektedir. Regresyon analizi, bir bağımlı değişken ile bir veya birden fazla bağımsız değişken arasındaki ilişkiyi sayısal hale dönüştürmek için kullanılan istatistiksel analiz yöntemidir. Regresyon analizinin kullanım amacı değişkenler arasındaki ilişkinin niteliğini saptamaktır. Regresyon analizinde amaç her bağımsız değişkenin bağımlı değişkendeki değişmeye katkısının hesaplanması, dolayısıyla tahmin değişkenlerinin değerinden hareketle bağımlı değişkenin değerinin tahmin edilmesidir (Özçınar,2006).

## Regresyon Yöntemleri

Regresyon yöntemleri, bağımsız değişken sayısına bağlı olarak basit doğrusal regresyon ve çoklu regresyon olarak ikiye ayrılırlar.

**1. Basit Doğrusal Regresyon :** Aralarında doğrusal ilişki bulunan bir bağımlı ve bir bağımsız değişken arasındaki ilişkiyi çözümleyen regresyon analizi türüdür. Basit doğrusal regresyon modeli şu şekilde denklem olarak verilmiştir. Denklemde  $Y_i$  bağımlı değişken,  $X_i$ , bağımsız değişkendir (Sykes,1992).

$$Y_i = b_0 + b_1 X_i + e_i \quad (1.1)$$

$X_i$  : Bağımsız değişken

$b_0$  : Bağımlı Değişken, doğrunun y eksenini kestiği nokta

$b_1$  : Regresyon katsayısı, bağımsız değişkeni pozitif yönde etkileyen değer

$e_i$  : Gözlenemeyen hata değeri

Regresyon analizinin amacı veriden elde edilen bilgiye dayanarak gözlenemeyen  $b_0$  ve  $b_1$  değerleri için bir tahmin üretmektir. Burada  $e_i$  hata değerini ihmal edecek olursak bir doğru denklemi meydana gelir ve denklem  $Y_i = b_0 + b_1 X_i$  haline gelir. Bu doğruda  $b_0$  doğrunun dik eksenini kestiği nokta  $b_1$  ise eğimi ifade eder. Sonuç olarak regresyon analizi ile doğrusal doğrunun nasıl oluştuğu tahmin edilmeye çalışılır. Regresyon analizi  $e_i$  hata değerinin sistematik bir özelliği olmadığını ve ortalama 0 (sıfır)'a eşit olduğunu kabul eder.

**2. Çoklu Regresyon :** Çoklu regresyon analizi birden fazla bağımsız değişkenlerden yola çıkarak bağımlı değişkenin tahmin edilmesi ve hangi bağımsız değişkenlerin bağımlı değişken üzerinde anlamlı bir etkiye sahip olduğunun bulunması için kullanılır. Regresyon analizinde örnek veri noktalarını en iyi temsil eden çizgi ya da düzlem bulunmaya çalışılır. Çoklu doğrusal regresyon modeli şu şekilde denklem olarak verilmiştir :

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_p X_{pi} + e_i \quad (1.2)$$

$b_0$  : Bağımlı değişken

$X_i$  : Bağımsız değişken

$b_1, b_2, \dots, b_p$  : Regresyon katsayıları, bağımsız değişkeni etkileyen değerler

$e_i$  : Hata değeri

Burada artık bir doğrudan söz edemeyiz. (p) sayısı kadar boyut vardır ve bu çok boyutlu sistem dikkate alınır. Çoklu regresyon modeli  $X_{1i}, X_{2i}, \dots, X_{pi}$  girdi değerlerine bağlı olarak  $Y_i$  değerinin tahmin edilmesini sağlar. Modelde yer alan regresyon katsayıları en küçük kareler yöntemi kullanılarak hesaplanır (Sykes, 1992).

#### 2.4.2. Sınıflandırma(Classification) Modeli

Sınıflandırma, bir veri setini birbirini dışlayan gruplara ayırma işlemidir. Bu işlemde bir grubun üyeleri mümkün olduğunca birbirine yakındır ve gruplar ise birbirlerinden mümkün olduğunca uzaktır. (Thearling,

<http://theartling.com/text/dmwhite/dmwhite.htm>) Genel olarak söylemek gerekirse, sınıflandırma bir objeyi özelliklerine göre bir kategoriye atama işlemidir. Veri madenciliğinde sınıflandırma, görünmeyen veri objelerinin önceden tanımlanmış sınıflara ayrılmasında kullanılan bir model (ya da işlevi) öğrenmek için önceden sınıflandırılmış veri objelerinin bir dizi analizini ifade eder, yani veri nesnesini daha önceden belirlenen sınıflardan biriyle eşleştirme sürecidir. Sınıflandırma gelecekteki veriyi sınıflandırmak için veriden bir model çıkaran bir veri analizi yöntemidir. Sınıflandırma ve regresyonun ikisi de tahmin için kullanılmasına rağmen, aralarındaki fark, sınıflandırmada bağımlı değişken kategoriktir, regresyonda ise sayısaldır. Yani sınıflandırma kategorik değerleri tahmin ederken, regresyon süreklilik gösteren değerlerin tahmin edilmesinde kullanılır (Wang,2009, p.141-147).

Sınıflandırma işlemi verileri özelliklerine göre sınıflara ayıran bir tekniktir. Bu teknikte asıl önemli olan veritabanına yeni bir veri eklendiğinde bu verinin hangi sınıfa dahil edileceğine karar vermektir. Sınıf etiketi olmayan veri, sınıf etiketi bilinen bir veri dizisi ile analizden geçirilerek kontrol edilir ve sınıflandırılır. Elde edilen modeller, sınıflandırma ağaçları (IF-THEN), karar ağaçları, matematiksel formüller ve yapay sinir ağları gibi çeşitli yapılarla gösterilmektedirler (Demiralay,2005).

Sınıflandırmada, bir kategorik hedef değişken vardır. Örneğin gelir grubu 3 sınıfa ya da kategoriye ayrılabilir: Yüksek gelir, orta gelir ve düşük gelir. Farz edelim ki araştırmacı veritabanında olmayan kişilerin gelir gruplarını bu kişinin diğer özelliklerine (yaş, cinsiyet, iş) dayanarak sınıflandırmak istiyor. İşte bu işlev sınıflandırma işlevidir ve veri madenciliği metotlarına ve tekniklerine tam olarak uygundur. (Larose,2005, p.95)

Sınıflandırmanın birçok uygulaması vardır: Hasta veritabanından hastalık belirtilerine dayalı olarak hastalığın tespit edilmesi, terörist eylemleri tespit etmek amacıyla kredi kartı işlemlerini analiz etmek, el yazısı örneklerinden harf ve rakamları tanımak, ilaç keşfetmede bileşik yapısına dayalı olarak aktif bileşikleri

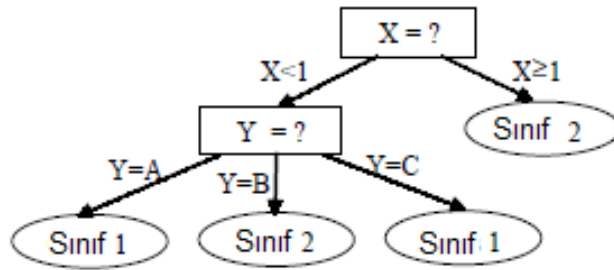


aktif olmayanlardan ayırt etmek. Sınıflandırma modelini öğrenmek için kullanılan örnekler grubuna alıştırma veri seti (training data set) denir (Larose,2005, p:14).

Sınıflandırmada kullanılan başlıca teknikler şunlardır (Wang,2009, p.144);

#### 2.4.2.1. Karar Ağaçları

En popüler sınıflandırma metotlarından biridir, veriden bir karar ağacı oluşturur. Bir karar ağacı, her düğümün bir nitelik üzerindeki testi temsil ettiği bir ağaç yapısından oluşan tahmin metodudur. Burada her ayrılan dal bir testin sonucunu temsil eder ve her yaprak düğüm bir sınıf ya da sınıf bölümü ile etiketlenir.



Şekil.2. Basit Bir Karar Ağacı (Wang,2009, p.144)

Şekil.2'de yapısı gösterilen basit bir karar ağacında her nesne alabileceği değerlere göre kökten yapraklara doğru bir yön izler ve bu şekilde sınıflandırılır. Tipik bir karar ağacı öğrenme algoritması yukarıdan aşağı geri dönüşümlü böl ve parçala stratejisini bir karar ağacı oluşturmak için adapte eder. Eğitim verisinin tamamı için kök düğümünden başlamak üzere seçilen kriterlere göre aldığı değerleri baz alarak veri iki ya da daha fazla alt kümelere ayrılır. Her bir alt küme için yavru düğüm oluşturulur ve işlem her yavru düğümdeki veriler için tekrar edilir. Son verme kriterine ulaşıncaya kadar işleme devam edilir. Birçok karar ağacı öğrenme algoritması vardır. Bunlar bilgi kazanma, kazanma oranı vs. gibi nitelik seçme kriterlerine göre ayrılırlar (Wang,2009, p.144).

#### 2.4.2.2. K-En Yakın Komsu (k-nearest neighbour) Algoritması:

K En Yakın Komşu Sınıflandırma Algoritması bilinmeyen örnekleri alıştırma verisindeki k En Yakın Komşulardan en yaygın sınıfa göre sınıflandırır. Bu algoritma bütün örneklerin n-boyutlu uzaydaki noktalara karşılık geldiğini kabul eder. N-boyutlu gelecek uzayında Euclid denklemine göre bir komşunun en kısa mesafesi varsa en yakın olduğu varsayılır. K=1 olduğu zaman bilinmeyen örnek alıştırma setindeki en yakın komşusuna göre sınıflandırılır. K-en yakın komşu metodu bütün alıştırma örneklerini depolar ve öğrenmeyi yeni bir örneğin sınıflandırılmasına ihtiyaç duyulana kadar erteler. Bu tip öğrenme örnek tabanlı ya da tembel öğrenme olarak adlandırılır (Wang,2009, p.144).

K En Yakın Komşu Sınıflandırma sezgisel, uygulaması kolay ve pratik olarak verimlidir. Bu sınıflandırma tekniği sınıflandırılacak her yeni örnek için yeni hedef fonksiyona uygun farklı tahmin üretebilir. Bu, hedef fonksiyonun karmaşık olduğu durumlara avantaj sağlar, fakat daha az karmaşık lokal tahminlerin birleşimi olarak tanımlanabilir. Yine de yeni örneklerin sınıflandırma maliyeti bütün hesaplamaların sınıflandırma aşamasında yapılıyor olmasından dolayı yüksek olabilir (Wang,2009, p.144).

#### **2.4.2.3. Yapay Sinir Ağları:**

Yapay Sinir Ağları biyolojik sinir sisteminin taklit edilmesi, matematiksel olarak modellenmesi çabalarının bir sonucu olarak ortaya çıkmıştır. Biyolojik sinir sistemindeki bilinen yapılar ve işlevleri yapay sinir ağlarında matematiksel modellerle, biyolojik sinir sistemindeki eşleniklerinin görevlerini yerine getirecek şekilde modellenmiştir. Yapay Sinir Ağlarında öğrenmeyi sağlamak için ağırlık fonksiyonları kullanılmaktadır, insanın deneme yanılma yoluyla öğrenmesi yapay sinir ağlarında yinelemeli eğitim sayesinde gerçekleştirilmektedir (Larose,2005,p.128).

Yapay Sinir Ağı, birbirine içten bağlı birkaç katman hesap biriminden (sinir ağı ya da düğüm) oluşur. Her birim basit bir fonksiyonu hesaplar. Her katmanın girdisi

bir önceki katmanın çıktısıdır. Birimler arası her bir bağlantıya bir ağırlık atanır. Her bir katmanda birimler arasında paralel hesaplama yapılabilir. İlk katmandaki girdiyi alan birimler giriş birimleri, son katmandaki ağırlık çıktısını üreten birimler de çıkış birimleri olarak adlandırılır (Wang,2009, p.144).

### **2.4.3. Kümeleme(Clustering) Modeli**

Kümeleme yöntemi, denetimsiz sınıflama modeli olarak da bilinmektedir. Kümeleme heterojen veri kümelerini veri karakteristikleri bakımından homojen gruplara bölme, bir başka deyişle diğer gruplardan çok farklı fakat üyeleri benzeyen grupları bulma işidir (Özçınar,2006). Kümeleme analizi, nesnelerin alt gruplara toplanmasını sağlayan bir işlemdir. Böylece nesnelere, örneklenen kitle özelliklerini iyi yansıtan temsil gücüne sahip olurlar. Sınıflamanın aksine, yeniden tanımlanmış sınıflara dayalı değildir. Kümeleme, bir denetimsiz öğrenme (unsupervised learning) yöntemidir (Altıntaş,2006).

Kümeleme kayıtların, gözlemlerinin, benzer nesnelerin sınıflarının gruplandırılması ile ilgilidir. Bir küme birbirine benzer kayıtlar topluluğudur ve bu kümedekiler diğer kümedekilere benzemezler. Kümelemeyi sınıflandırmadan ayıran nokta hedef değişkenin olmayışıdır. Kümeleme işlemi sınıflandırmaya ya da hedef değişken tahmin etmeye çalışmaz. Bunun yerine kümeleme algoritması bütün veri grubunu görece homojen alt gruplara ve kümelere ayırmaya çalışır. Bu durumda küme içindeki benzerlik maksimum, küme dışındaki kayıtlarla benzerlik ise minimum durumuna getirilir. Kümeleme yöntemi veri madenciliği sürecinde başlangıç adımı olarak uygulanır. İş alanında kümeleme yöntemi hedef pazarlama, muhasebe denetimi gibi alanlarda kullanılmaktadır (Larose,2005).

Tipik bir kümeleme modeli şu aşamalardan oluşur: Model kurgulama (ilerideki parçalama ve seçimi içerecek şekilde), veri alanına uygun model yakınlık ölçüsünün tanımlanması, kümeleme, veri ayırımı ve çıktının doğrulanması (Wang,2009, p.159).

Temel kümeleme algoritmaları şu şekilde sayılabilir (Wang,2009, p.159);

**1. Hiyerarşik Kümeleme :** Küme hiyerarşisi ya da dendrogram olarak bilinen kümeler ağacı inşa eder. Her bir küme düğümü yavru kümeler içerir.

**2. Ayırma (Partitioning) Kümelemesi:** Bir veri yığınının bir kümeyi temsil eden ayırımlara ayrıldığı algoritmadır. Bu kümeleme algoritmasında aynı kümedeki nesnelere birbirine benzemesi ve farklı kümelerdeki nesnelere ise birbirinden ayrı olması sağlanır.

**a) K-Means Algoritması:** Bilimsel ve endüstriyel alanda çok kullanılan bir kümeleme algoritmasıdır. Bu algoritmada önce belli sayıda nesne seçilir. Başlangıçta bu nesnelere her birinin ayrı bir kümeyi temsil ettiği farz edilir. Diğer objeler bu nesnelere ortalamalarına olan benzerliklerine göre kümelendirilir.

**b) K-Medoids Algoritması :** Bu algoritmada küme, noktalarından herhangi biriyle temsil edilir. Küme içindeki ortalama değeri almak yerine en merkeze yerleştirilmiş bir değer alınır. Temsili değerler (medoid) rastgele seçilir ve diğer nesnelere bu seçilen nesnelere benzerliğine göre kümelendirilir.

#### **2.4.4. Birliktelik Kural Modeli (Association Rule Model)**

Birliktelik kural modeli bir işlevsel veritabanındaki nesnelere nasıl bir arada gruplandırılacağı ile ilgilidir. Çoğunlukla market sepet analizi olarak bilinir çünkü marketlerde tezgahlar tarafından sıklıkla bir araya konan maddelerin analizine benzetilebilir. İstatistik açısından bir değişken kümesi içinde ilişkinin keşfi için yarı otomatik bir tekniktir. Birliktelik kural modeli, çok büyük sayı uygulamalarında, referans sistemlerinde, müşteri ilişkileri yönetiminde (CRM), promosyon uygulamalarında ve çapraz satış yöntemlerinde yaygın olarak kullanılır (Wang,2009, p.59).

Birliktelik kuralları, büyük veri kümeleri arasında benzerlik ve birliktelik ilişkileri bulurlar. Toplanan ve depolanan verinin sürekli büyümesi yüzünden,

şirketler veritabanlarındaki birliktelik kurallarını ortaya çıkarmak istemektedirler. Birliktelik kurallarının kullanıldığı en tipik örnek market sepeti uygulamasıdır. Bu işlem, müşterilerin yaptıkları alışverişlerdeki ürünler arasındaki birliktelikleri bularak müşterilerin satın alma alışkanlıklarını analiz etmektedir. Bu tip birlikteliklerin keşfedilmesi, müşterilerin hangi ürünleri bir arada aldıkları bilgisini ortaya çıkarır ve market yöneticileri de bu bilgi ışığında daha etkili satış stratejileri geliştirebilirler. Örneğin bir müşteri süt satın alıyorsa, aynı alışverişte sütün yanında ekmek alma olasılığı nedir? Bu tip bir bilgi ışığında rafları düzenleyen market yöneticileri ürünlerindeki satış oranını arttırabilmektedirler (Altıntaş,2006, s.36).

Örneğin bir A ürününü satın alan müşteriler aynı zamanda B ürününü de satın alıyorsa, bu durum Birliktelik Kuralı ile şu şekilde gösterilir:

$$A \Rightarrow B \text{ [destek} = \%2, \text{ güven} = \%60]$$

Buradaki %2 destek değeri, analiz edilen tüm alışverişlerin %2'sinde A ile B ürünlerinin birlikte satıldığını göstermektedir. %60 oranındaki güven değeri ise A ürününü satın alan müşterilerin %60'ının aynı alışverişte B ürününü de satın aldığını göstermektedir (Altıntaş,2006,s.37).

Büyük veri tabanlarında birliktelik kuralları bulunurken, şu iki işlem basamağı takip edilir; 1- Sık tekrarlanan öğeler bulunur: Bu öğelerin her biri en az, önceden belirlenen minimum destek sayısı kadar sık tekrarlanırlar. 2- Sık tekrarlanan öğelerden güçlü birliktelik kuralları oluşturulur: Bu kurallar minimum destek ve minimum güven değerlerini karşılamalıdır (Larose,2005,p.184).

### 3. DÖVİZ KURU TAHMİNİNDE KULLANILAN TEKNİKLER

Genel olarak yabancı ülkelerin para birimine döviz, iki ülkenin para birimleri arasındaki orana ise döviz kuru denilmektedir. Diğer bir deyişle döviz kuru, bir yabancı paranın milli para cinsinden fiyatıdır. En başlangıçta forex (foreign exchange rate) yani yabancı döviz kuru ödemeler dengesi ile belirlenirdi. Ödemeler dengesi bir ülke için sadece makbuzlarının ve uluslararası alışverişlerde ödemelerin listelenmesi yöntemi idi. Denge temel olarak malların ithal ve ihraç edilmesiyle belirlenirdi. Bu yüzden o zamanda Forexi tahmin etmek çok zor değildi. Daha sonra her para birimi için faiz oranları ve diğer arz-talep faktörleri daha etkili olmaya başladı. Artan forex ticareti ve nakit ile bonolardaki spekülasyonlar forex tahminindeki zorluğa katkıda bulundular (Yao et al, 1996).

Nominal döviz kuru, iki para biriminin birbirine göre nispi fiyatı anlamına gelir. Reel döviz kuru, yabancı ülkelerde üretilen malların yurtiçinde üretilen mallar cinsinden görece fiyatını yansıtan ve uluslararası rekabeti ölçmek için yaygın olarak kullanılan bir ekonomik göstergedir. Reel döviz kuru, ekonomik birimlerin üretim ve tüketim kararlarının yurtiçinde ve yurtdışında üretilen mallar arasında dağılımını etkileyerek, cari işlemler dengesi üzerinde belirleyici rol oynamaktadır. Reel efektif döviz kuru, nominal efektif döviz kurunun fiyat endeksleriyle düşürülmüş halidir. Uluslararası Para Fonu (IMF), tarafından da kullanılan bu tanıma göre reel efektif döviz kuru, ilgili ülkenin fiyat düzeyinin dış ticaret yaptığı ülkelerin fiyat düzeylerine oranının ağırlık geometrik ortalaması olarak hesaplanmaktadır (Yıldız, 2006).

1973'te sanayileşmiş ülkelerin sabit yabancı döviz kuru sistemini bırakıp değişken döviz kuru sistemini uygulamaya başlamalarından bu yana araştırmacılar döviz kuru hareketlerini açıklamaya uğraşmaktadırlar. Bunun sonucu olarak binlerce araştırmacı ve uzman tarafından birçok tahmin metodu geliştirilmiştir. Teknik ve temel analizler finans alanında çokça bilinen öne çıkan tahmin metodları arasındadır. Yabancı döviz kurları çok ilişkili faktörler tarafından etkilenir. Bu faktörler, ekonomik, politik ve hatta psikolojik olabilir. Bu faktörlerin etkileşimi çok karmaşık

bir tarzda olabilir. Bu yüzden, yabancı döviz kurundaki değişimi tahmin etmek genellikle çok zordur (Yao et al, 1996).

Döviz kuru tahmini, modern zaman serisi tahmininin dikkat çekici uygulamalarından biridir ve birçok iş ve finans enstitülerinin başarısı için çok önemlidir. Kurlar doğal olarak düzensiz, değişken ve belirleyici bir şekilde kaotiktir. Bu durumlarda yapılan bir genel kabule göre geçmiş veri bu davranışların hepsini içerir. Sonuç olarak geçmiş veri tahmin işleminin temel girdisidir. Birçok öngörü teknikleri bulunmaktadır. Bunlar temel olarak çeşitli kabullere, matematiksel fonksiyonlara ve özel model parametrelerine dayanarak geliştirilmişlerdir. Fakat daha iyi bir sonuç için verilen bir öngörü görevinde en uygun tekniği bulmak önemlidir (Kamruzzaman et al,2003).

Döviz kurunun belirlenmesi için temelde 3 değişik kur politikası vardır: **1. Serbest Kur Sistemi:** Ülke para biriminin diğer tüm para birimlerine karşı kontrolsüz bırakılması, **2. Sabit Kur Sistemi;** Para birimi değerinin başka bir ülkenin para birimine ya da bir araca sabitlenmesi, **3. Karma Sistem:** Para birimi değerinin serbest dalgalanmasına izin verilirken aşırı dalgalanmalarda müdahale edilmesi (Yıldız ,2006).

Forex tahmin metodu uygulamasının iki temel adımı vardır : Veri serisinin analiz edilmesi ve veri serisine en uygun tahmin metodunun seçilmesi. Nakit piyasasındaki karı arttırmak için tüccarlar tarafından daha da iyi tahmin metotları kullanılmaktadır. Günümüzde tüccarlar piyasanın geleceği hakkında bilgi sahibi olabilmek için tek bir tekniğe bağlı kalmayıp, değişik işaretleri yakalayabilmek için çok değişik teknikler kullanılmaktadır (Yao et al, 1996).

### 3.1. ÇALIŞMAYA İLİŞKİN LİTERATÜR ÖZETİ

Jon Faust ve arkadaşları, standart döviz tahmin modellerinin gerçek zamanlı tahmin performansını değişik veri örnekleri kullanarak test etmişlerdir. Bu çalışmalarında birçok veri örneği test edilerek veri revizyonlarının döviz kuru

tahminleri üzerinde yaptığı etkiler incelenmiş ve tamamen revize edilmiş veri yerine orijinal veri kullanarak yapılan tahminin daha iyi sonuç verdiği bulunmuştur (Faust et al., 2003).

Nathan Lael Joseph, özellikle alternatif tahmin ve tanımlama metotları uygulandığında değişik veri aralıkları için döviz kuru doğruluğunun tahmini konusunda ilgilenmiştir. Çalışmada yedi döviz kurundaki değişimler incelenerek eşbütünleşmenin tespit edilme olasılığı artırılmıştır. Çalışma sonucunda elde edilen en önemli sonuç iyi bir model tanımlamanın çok iyi bir tahmin performansı manasına gelmediğidir (Joseph, 2001).

Adrian Trapletti ve arkadaşları, gün içi verilerini kullanarak USD–DEM, USD–JPY ve DEM–JPY döviz kuru üçgeni için eşbütünleşme analizi sunmuşlardır. Çalışmalarında örsen dışı tahmin doğrulama ölçülerine bağlı olarak bir geri yayılım modeli tahmin edildi ve hesaplandı. Çalışma sonunda döviz kurları ve zaman aralığı arasındaki eşbütünleşme ilişkisinin kullanılmasının tahmin sonuçlarını güçlendirdiği tespit edilmiştir. Çalışmada ayrıca yüksek sıklıktaki veriler için tipik dönemsel oynaklığa sebep olan zaman aralığının değiştirilmesinin ticaret stratejisinin sonuçlarını güçlendirdiği tespit edilmiştir (Trapletti et al., 2002).

Cheol S. Euna ve Sanjiv Sabherwal, önde gelen ticari bankaların döviz kuru tahminindeki performansını tesadüf hareket modeli kullanarak değerlendirmişlerdir. Bu araştırmalarında Amerikan Dolarının(USD) İngiliz Sterlini(GBP), Alman Markı(DEM), İsviçre Frankı ve Japon Yenine göre kurlarının bir senelik tahmini incelenmiştir. Çalışmalarında elde ettikleri buldular : 1. Çoğu banka Japon Yeni hariç diğer üç döviz kuru için Tesadüf Hareket Modeli'nin daha iyi performans gösterdiğine dair bazı kanıtlar göstermişlerdir, 2. Tahmin edilenle gerçek döviz kuru değişimi arasındaki kusurlu ilişki bankaların tahmindeki hatalarının kaynak sebebidir, 3. Yerli bankalar ülkenin döviz kurlarını bilgi asimetrisi önererek diğer bankalardan daha doğru tahmin etmekte, 4. Birçok bankanın tahminleri bandwagon tipi etki göstermektedir, yani çoğu banka hız tahmincisidir, halihazırdaki kur değişimlerini tahmin etmeye meyillidirler. Çalışmada şu ilginç sonuç da ortaya çıkmıştır : genel temayüle aykırı olan muhalif bankalar, diğer bankalardan daha iyi



performans göstermektedirler (Euna and Sabherwal,2002).

Fang-Mei Tseng ve arkadaşları, Fuzzy ARIMA ya da FARIMA ismini verdikleri bir metot geliştirmişler ve bunu NTD (Yeni Tayvan Doları) ve USD (Amerikan Doları) arasındaki döviz kurunu tahmin etmek için kullanmışlardır. Önerdikleri metodun iyi tahmin yapmakla kalmayıp aynı zamanda karar vericilere en iyi ve en kötü durumu da sunduğunu örneklerle açıklamışlardır. Önerdikleri Fuzzy ARIMA'nın ARIMA'dan daha iyi performans gösterdiğini, daha tatmin edici olduğunu, daha az gözlem gerektirdiğini ve daha uygun bir tahmin aracı olduğunu açıklamışlardır (Tseng et al.,2001).

Martin Martens, çalışmasında gün içi değerlerin günlük oynaklığı tahmin etmek için önemli bilgiler içerip içermediğini araştırmaktadır. Günlük oynaklık tahminleri gün içi aralıklarına ait çeşitli oynaklık tahminlerinden oluşturulur. Bu durum günlük oynaklıkta çok iyi sonuç veren DEM/USD ve YEN/USD döviz kurlarında görülmektedir. Çalışmada yapılan simülasyon deneyinde gün içi dönemsel oynaklık örnekleri açıkça modellenirse ortalamada sadece günlük oynaklık tahminlerindeki marjinal gelişmelere öncülük edeceği bulunmuştur (Martens,2001).

T. Davis Jefferson ve arkadaşları, Kanada- Amerikan Dolar kuru verilerine uygulanan çeşitli sinir ağı modelleri sunmuşlardır. Çalışmada geri yayılım, modüler, radyal temelli fonksiyonlar, Doğrusal Vektör Nicelemesi, Fuzzy ARTMAP ve Genetik Takviye Öğrenimi Ağları incelenmiştir. Amaç günlük değerlerde değişim yönlerinin tahmin edilmesinde bu ağların performansının karşılaştırılmasıdır. Yapılan testler sonucunda Sinir ağlarının diğer ağlara üstünlükleri tespit edilmiştir (Jefferson et al.,2001).

Jingtao Yao ve Chew Lim, Tan Sinir Ağı Modelinin yabancı döviz kuru tahmini için uygun olduğuna dair deneysel kanıt sunmuşlardır. Bu çalışmalarında çok fazla piyasa verileri ya da bilgi kullanmadan faydalı tahmin yapılabileceğini ve basit teknik göstergelerle örnek veri dışında iyi kağıt kar elde edilebileceğini sonuçlarla göstermişlerdir. Yaptıkları gözlemler Sinir Ağları'nın her ne kadar bazı çözümleri

gereken problemleri olsa bile finansal tahmin için uygun olduğu sonucunu vermiştir (Yao and Tan,2000).

Mark T. Leung ve arkadaşları, Genel Gerileme Sinir Ağı(GRNN) isimli özel sinir ağı yapısının tahmin edilebilirliğini incelemişler ve performansını Çok Katmanlı Geri beslemeli Ağ (MLFN), çok değişkenli transfer fonksiyonu ve tesadüfü hareket modelleri gibi diğer birçok tahmin teknikleri ile karşılaştırmışlardır. Çalışmalarında Genel Gerileme Sinir Ağı (GRNN)'nin sadece daha yüksek bir tahmin doğruluğu olması değil aynı zamanda değişik döviz kurları için istatistiksel olarak diğer modellerden daha iyi olduğu sonucunu elde etmişlerdir (Leung et al.,2000).

Zoran Vojinovic ve arkadaşları, göreceli yeni olan ve yabancı döviz zaman serisi işleminin tahmin edilmesinde kullanılabilecek veri madenciliği tekniklerinden birini açıklamışlardır. Araştırma bu tip teknikleri zor olan gerçek dünya veri setlerine teşhir ederek geliştirilme ve uygulamasına katkıda bulunmayı amaçlamaktadır. Sonuçlar, günlük (Amerikan Doları) \$US/ (Yeni Zelanda Doları) \$NZ kapanış döviz kurlarının tahmini için Radyal Temel Fonksiyon Sinir Ağı modelinin kestirilmesi, açıkça geleneksel doğrusal geri beslemeli modelin tahmininden hem yönsel değişim hem de döviz kuru tahmininin kendisi yönüyle daha iyi olduğunu göstermiştir. Çalışma ayrıca model girdi sayısı, gizli katman sinir sayısı ve eğitim veri setinin tahmin doğruluğu üzerindeki etkisini incelemiştir (Vojinovic et al.,2001).

Francesco Lisi ve Rosa A. Schiavo, ortak veri setleri ve değişkenleri için Sinir Ağları ve Karmaşık Modellerinin karşılaştırmalı değerlendirmesini sunmuşlar ve aynı deneysel şartlarda şanstan daha iyi tahmin üretilip üretemeyeceği doğrulamışlardır. Çalışmada tahmin performansı, elde edilen tahminin istatistiksel önemi bazında olduğu kadar iyi bilinen normalleştirilmiş Ortalama Kare Hatası (NMSE) bazında da ölçülmüştür. Elde edilen deneysel sonuçlar göstermiştir ki, Sinir Ağları Karmaşık Modellerle NMSE'ye göre karşılaştırıldığında, esas olarak iki model de Tesadüfi Hareket Hipotezine dayanan modellere göre daha iyi işlemektedir. Çalışmada bulunan bir diğer sonuç da, istatistiki açıdan Sinir ağları ile yapılan tahminler karmaşık modellere göre yapılanlara eşit sonuçlar vermiştir. Birçok durumda istatistikli açıdan ikisi de Tesadüfi Hareket Modelinden daha iyi sonuç

vermiştir (Lisi and Schiavo,1999).

Mona ve Hassan E. El Shazly dört döviz kuru (Sterlin, Mark, Yen ve İsveç Frankı) için üç aylık spot kurları tahmin etmek için sinir ağlarını ve genetik öğrenmeyi birleştiren bir hibrit sistem geliştirmişlerdir. Ağların tahminleri ileri dönük kurlarla karşılaştırılmış ve doğruluk derecelerine ve kur hareketlerinin yön değişimlerini tahmin etme yeteneklerine göre değerlendirilmiştir. Sonuçta Sinir Ağları ile döviz kurlarındaki değişimleri tahmin etmenin daha kolay olduğu bilgisine ulaşılmıştır (El Shazly and El Shazly,1999).

### 3.2. YAPAY SİNİR AĞLARI

Yapay Sinir Ağları biyolojik sinir sisteminin taklit edilmesi, matematiksel olarak modellenmesi çabalarının bir sonucu olarak ortaya çıkmıştır. Biyolojik sinir sistemindeki bilinen yapılar ve işlevleri yapay sinir ağlarında matematiksel modellerle, biyolojik sinir sistemindeki eşleniklerinin görevlerini yerine getirecek şekilde modellenmiştir. Yapay Sinir Ağlarında öğrenmeyi sağlamak için ağırlık fonksiyonları kullanılmaktadır, insanın deneme yanılma yoluyla öğrenmesi yapay sinir ağlarında yinelemeli eğitim sayesinde gerçekleştirilmektedir. YSA'lar aralıksız çıktı ürettiğinden dolayı gayet doğal olarak öngörü ve tahmin için kullanılabilirler. (Larose,2005,p.128) YSA'lar yeni ve iddialı bir hesaplama teknolojisidir ve birçok finansal uygulamanın dinamiklerini keşfetmek için yeni bir yol önermektedir (Yao et al, 1996).

Yapay Sinir Ağları sahip olduğu hesaplama gücü ile farklılık göstermektedir. Öne çıkan özellikleri paralel dağıtık yapısı ve öğrenme ve bu sayede genelleştirme yeteneğidir. Bu özellikleri Yapay sinir Ağları'nı karmaşık problemleri çözmede başarılı kılmaktadır. Yapay Sinir Ağları'nın kullanımı şu özellik ve imkanları sunmaktadır (Haykin,1999):

**1. Doğrusal Olmama :** Doğrusal Olmama özellikle fiziki mekanizma altyapısı doğrusal değilse çok önemli bir özelliktir.

**2. Giriş – Çıkış Eşleştirme :** Ağ eldeki problem için giriş – çıkış eşleme oluşturarak örneklerden öğrenir.

**3. Uyumluluk :** Sinir ağlarının ağırlıklarını çevresel koşullardaki değişimlere uyumlulaştırma yeteneği vardır. Özel bir çevre için çalışmaya öğretilmiş bir sinir ağı çevresel koşullarda küçük değişiklikler olması durumunda kolaylıkla yeniden öğretilir.

**4. Açık Cevap :** Model sınıflandırma durumunda, Yapay Sinir Ağları sadece belli bir modelin seçimi hakkında bilgi sağlamak için değil aynı zamanda karar vermede güven hakkında bilgi sağlamak için tasarlanırlar.

**5. İçerik Bilgisi :** Bilgi sinir ağının tam yapı ve aktivasyon durumu tarafından sunulur.

**6. Hata Toleransı :** Bir sinir ağı potansiyel olarak hata toleransını kapsar, ya da sağlam bir hesaplama yeteneğine sahiptir.

**7. VLSI(Very-Large-Scale-Integrated) Uygulanabilirliği:** Sinir ağlarının paralel yapısı onu belli görevlerin hesabında oldukça hızlı yapmaktadır. Bu özellik onu VLSI(Çok büyük ölçekli entegre) teknolojilere uygulanabilir yapmaktadır.

**8. Analiz ve Dizayn Bütünlüğü :** Sinir ağlarının bütün uygulamalarında bütünlük ve tek düzenlik vardır.

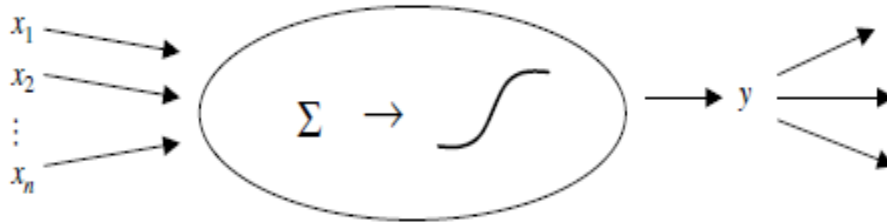
**9. Biyolojik Sinirlere Benzerlik :** Sinir ağları beyindeki sinirlerden esinlenerek icat edilmişlerdir (Haykin,1999).

Yapay Sinir Ağları, sahip olduğu paralellik, hata toleransı, öğrenilebilirlik ve gerçekleştirme kolaylığı gibi özelliklerinden dolayı alışlagelmiş bilgi işleme yöntemlerinden farklılık göstermektedir. Bu özellikleri itibari ile de diğer

alışlagelmiş hesaplama yöntemlerine göre daha başarılı sonuçlar üretebilmektedirler (Yıldız ,2006).

Yapay Sinir Ağları, Matematik, İstatistik, Fizik, Bilgisayar Bilimleri ve mühendislik gibi bilim alanları ile modelleme, işaret işleme, karakter tanıma, zaman seri analizi, robotik sistemler, finans alanında vb. alanlarda kullanılmaktadır. Yapay Sinir Ağları zaman serilerinin gelecekteki değerinin tahmin edilmesinde gelişmekte olan ve ulaştığı başarılı sonuçlardan dolayı umut vadeden bir yöntemdir (Haykin,1999; Kovalerchuk and Vityaev, 2005; Bishop,2006; Velagic et al, 2008).

Şekil 3’de temel bir yapay sinir ağı nöronu görülmektedir. Burada ( $X_i$ ) girdileri, bu girdilerin tabi tutuldukları işlevler ve ( $Y$ ) çıktısı görülmektedir. Girdiler( $X_i$ ) yukarı yönlü sinirlerinden (ya da veri setlerinden) toplanır ve toplama( $\Sigma$ ) gibi daha sonra başka sinirlere aşağı yönlü kanallara ayrılan  $y$  çıktısını üretmek için kendileri de (genellikle doğrusal olmayan) aktivasyon fonksiyonlarına girdi olan fonksiyonların birleşimi aracılığıyla bir araya getirilir (Larose,2005,p.128).

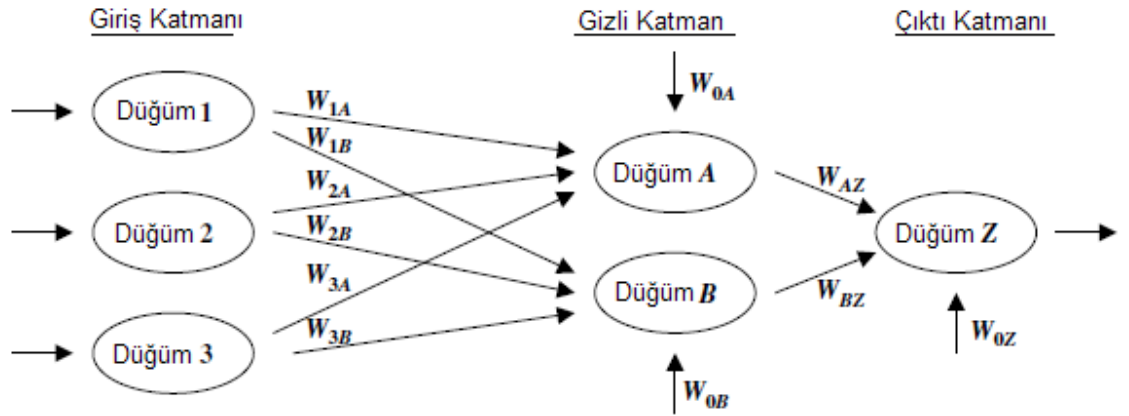


Şekil 3. Yapay Sinir Modeli (Larose,2005,p.128)

YSA kullanmanın avantajlarından biri onların gürültülü(karışık) verilerde sağlam oluşlarıdır. Çünkü ağ her bağlantıya atanmış ağırlıkları olan birçok düğüm(yapay sinir) içerir ve veri setindeki bilgilendirici olmayan (bazen de hatalı) örneklerle çalışmayı öğrenebilir. Ne yazık ki şunu görmeliyiz ki, uzman olmayanlar için anlaşılır olan sezgisel kurallar üreten karar ağaçlarının aksine, yapay sinir ağları insan algılayışına nispeten donuktur. Ayrıca, YSA’lar genellikle karar ağaçlarından birçok saati aşan daha uzun süre eğitime ihtiyaç duyarlar (Larose,2005,p.128).

YSA’ların bir sakıncası bütün nitelik değerlerinin hatta kategorik değişkenler

de dahil standart yapıda kodlanması yani değerlerin 0 ile 1 arasında olması gereğidir. Çok fazla içerikten oluşan girdiler için bunu 0 ve 1 arasında değerlendirmek zorluğa sebep olmaktadır. Örneğin bir veri medeni hali diye bir nitelik içerdiğini varsayalım. Bu niteliğin içeriği boşanmış, evli, ayrılmış, bekar, dul ve bilinmeyen olabilir}. Bu durumda bu değerleri YSA'ya uydurabilmek için 0 ve 1 arası değerlerle ifade etmek gerekir (0.0, 0.2, 0.4, 0.6, 0.8 ve 1). YSA'ların sadece sayısal değerleri kabul ediyor oluşu, içeriğin gerçek anlamın dışında kodlanmasına sebep olabilir. Çıktı dikkate alındığında ise görülecektir ki YSA çıktı düğümleri her zaman 0 ve 1 arasında sürekli olan değerleri çıktı olarak verir (Larose,2005,p.128).



Şekil.4. Basit Bir Yapay Sinir Ağı (Larose,2005,p.128)

Şekil.4'teki gibi bir Yapay Sinir Ağı katmanlı, ileri beslemeli, tamamen birbirine bağlı yapay sinir ağlarından ya da düğümlerden oluşur. Ağın ileri besleme özelliği ağı tek yönlü akışla sınırlar ve döngüye ya da çevrime izin vermez. YSA iki ya da daha fazla katmandan oluşur. Fakat çoğu 3 katmandan oluşur : Giriş katmanı, gizli katman ve çıkış katmanı. Birden fazla gizli katman olabilir. Fakat birçok ağ çoğu amaç için yeterli olan sadece bir gizli katman içerir. YSA'ların tam olarak bağlı olması demek, bir katmandaki her düğüm, kendi katmanındaki düğümlere değil bir sonraki katmandaki her düğüme bağlıdır demektir. Her düğüm arasındaki bağlantının kendisine atanan bir ağırlığı(W) vardır. Başlangıçta bu ağırlıklara 0 ve 1 arasında değişen rastgele değerler atanır (Larose,2005,p.128).

$x_0 = 1.0$	$W_{0A} = 0.5$	$W_{0B} = 0.7$	$W_{0Z} = 0.5$
$x_1 = 0.4$	$W_{1A} = 0.6$	$W_{1B} = 0.9$	$W_{AZ} = 0.9$
$x_2 = 0.2$	$W_{2A} = 0.8$	$W_{2B} = 0.8$	$W_{BZ} = 0.9$
$x_3 = 0.7$	$W_{3A} = 0.6$	$W_{3B} = 0.4$	

Tablo.2. Veri girişi ve Sınır ağınlının ağırlıkları için başlangıç değerleri  
(Larose,2005,p.128)

Örnek olarak Tablo2.'deki verileri kullanarak gizli katman düğümlerinin ve çıkış katmanı düğümlerinin yapısını oluşturalım. İlk önce bir kombinasyon fonksiyonu (genellikle Toplam  $\Sigma$ ) düğüm girişlerinin bir doğrusal kombinasyonunu ve bağlantı ağırlıkları için “net” olarak isimlendireceğimiz tek bir sayısal değer üretir. Yani verilen bir  $j$  düğümü için

$$net\ j = \sum_i W_{ij} x_{ij} = W_{0j} x_{0j} + W_{1j} x_{1j} + \dots + W_{Ij} x_{Ij} \quad (2.1)$$

Burada  $x_{ij}$   $j$  düğümüne  $i$ 'nci giriş değerini sembolize eder.  $W_{ij}$ ,  $j$  düğümüne  $i$ 'nci giriş değerine atanan ağırlığı sembolize eder ve  $j$  düğümüne  $I + 1$  giriş vardır. Şunu da eklemek gerekir ki  $x_1, x_2, \dots, x_I$  yukarı doğru giden düğümlerden girişleri sembolize eder.  $x_0$  regresyon modelindeki sabit faktöre benzeyen sabit bir giriş sembolize eder ve bu kural olarak sürekli  $x_0 = 1$  değerini alır.

A düğümü için,

$$\begin{aligned} net_A &= \sum_i W_{iA} X_{iA} = W_{0A}(1) + W_{1A} X_{1A} + W_{2A} X_{2A} + W_{3A} X_{3A} \\ &= 0.5 + 0.6(0.4) + 0.8(0.2) + 0.6(0.7) = 1.32 \end{aligned}$$

Daha sonra bir aktivasyon fonksiyonu hesaplanır. En çok kullanılan aktivasyon fonksiyonu sigmoid fonksiyonudur.

$$y = \frac{1}{1+e^{-x}} \quad (2.2)$$

Burada  $e$  doğal logaritmanın temelidir ve yaklaşık değeri 2.718281828'dir.

Yani A düğümünde aktivasyon giriş değeri olarak  $net_A = 1.32$  alınacaktır ve çıkış değeri olarak  $y = 1/(1 + e^{-1.32}) = 0.7892$  üretilecektir.

Aynı şekilde B düğümü için de hesap yapılır:

$$\begin{aligned} net_B &= \sum_i W_{iB} X_{iB} = W_{0B}(1) + W_{1B} X_{1B} + W_{2B} X_{2B} + W_{3B} X_{3B} \\ &= 0.7 + 0.9(0.4) + 0.8(0.2) + 0.4(0.7) = 1.5 \end{aligned}$$

$$\text{Sonra; } f(net_B) = \frac{1}{1+e^{-1.5}} = 0.8176$$

Z düğümü A ve B düğümlerinden gelen çıktıları birleştirir :

$$\begin{aligned} net_Z &= \sum_i W_{iZ} X_{iZ} = W_{0Z}(1) + W_{AZ} X_{AZ} + W_{BZ} X_{BZ} \\ &= 0.5 + 0.9(0.7892) + 0.9(0.8176) = 1.9461 \end{aligned}$$

$$f(net_Z) = \frac{1}{1+e^{-1.9461}} = 0.8750$$

Bu 0.8750 değeri Yapay Sinir Ağının çıkış değeridir ve ilk gözlemdeki hedef değişkenin tahmini değerini temsil eder.

Doğrusal olmayan model çeşidi olan Yapay Sinir Ağları, döviz kuru tahmininde güçlü bir alternatif olduğu ispatlanmıştır. (Jasic and Wood, 2003), (Tenti, 1996) ve (Qi and Zhang, 2001) YSA'ların döviz kuru tahmininde başarılı bir şekilde kullanılabileceğini yaptıkları çalışmalarla göstermişlerdir. YSA, özellikle karmaşık, düzensiz, bağıntısız ve parçalı bilgiler için doğru çözümler bulmaya uygun bir metottur. YSA'nın döviz kuru tahmininde bir araç olarak seçilmesinin nedeni, bazı önemli ve çekici özelliklerinin olmasıdır: 1) Bunlardan ilki, klasik modellere dayalı olan tahmin metotlarının zıddına, YSA çok az kısıtlayıcı varsayıma sahiptir ve veri tipine kolayca adapte edilebilir. YSA'nın bu özelliği bazı finansal tahminlerin yürütülmesinde temel verinin elde edilemediği durumlarda çok istenen bir durumdur. 2) YSA geliştirilebilir, 3) YSA genel bir fonksiyonel yapısı vardır (Kadılar vd.,

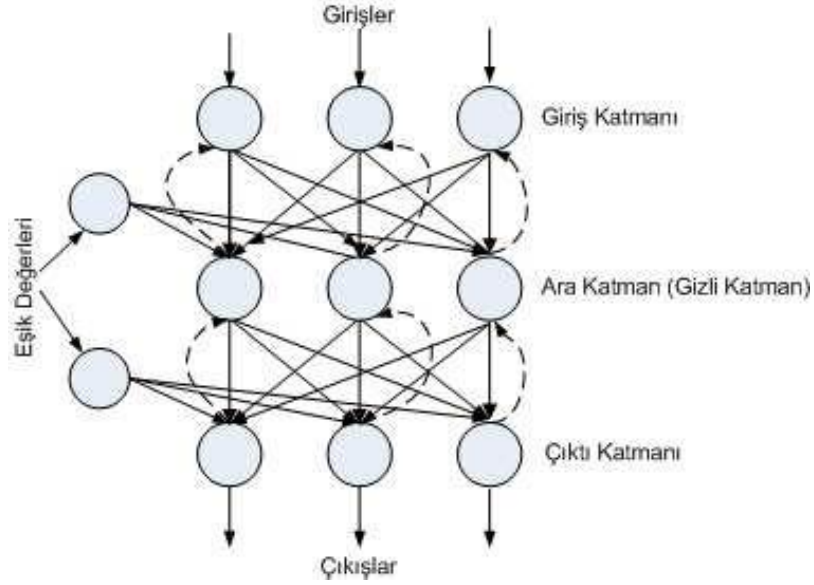


2009).

YSA'ların döviz kuru tahmininde kullanılması ile ilgili Jefferson T. Davis, Athanasios Episcopos ve Sannaka Wettimuny yaptıkları çalışmada Kanada-Amerikan Dolar kuru verilerine uygulanan çeşitli sinir ağı modelleri sunmaktadırlar. Yaptıkları testler sonucunda Sinir Ağlarının diğer ağlara üstünlükleri tespit edilmiştir.( Jefferson et al.,2001) Jingtao Yao ve Chew Lim Tan, Sinir Ağı Modelinin yabancı döviz kuru tahmini için uygun olduğuna dair deneysel kanıt sunmaktadırlar. (Yao and Tan,2000) Francesco Lisi ve Rosa A. Schiavo, Sinir Ağları ve Karmaşık Modeller ile NMSE'ye göre karşılaştırarak iki modelin de Tesadüfi Hareket Hipotezine dayanan modellere göre daha iyi işlediğini tespit etmişlerdir. (Lisi and Schiavo, 1999) Prof Dr. Mona R. El Shazly ve Prof Dr. Hassan E. El Shazly makalelerinde dört döviz kuru için üç aylık spot kurları tahmin etmek için sinir ağlarını ve genetik öğrenmeyi birleştiren bir hibrit sistem geliştirmişler ve Sinir Ağları ile döviz kurlarındaki değişimleri tahmin etmenin daha kolay olduğu bilgisine ulaşmışlardır (El Shazly and El Shazly, 1999).

## **Geri Yayılım (Back Propagation) Algoritması**

Geri yayılım algoritması adını hatayı yansıtarak yayma biçiminden alır. Bu algorithmada ulaşılan sonuç ve hedeflenen sonuç arasındaki fark yani hata, tüm ağırlıklara yansıtılır. Temel amaç hedef çıktı ile ulaşılan çıktı arasındaki hatayı en aza indirmektir. Geriye Yayılım Algoritması özellikle finansal zaman serilerinin gelecekteki davranışlarını tahmin etmede tercih edilmektedir. Geriye Yayılım Algoritmasının dezavantajları yavaş olması ve basit bir problemin çözümünde bile yapay sinir ağının eğitiminin binlerce yineleme gerektirmesidir. Bu algoritmanın başarısı ağırlık katsayılarının ilk değerlerine, momentum ve öğrenme katsayısı değerlerinin seçimine bağlıdır (Özçınar,2006).



Şekil 5. Bir Geriye Yayılım Algoritması Ağı Modeli (Yıldız, 2006, s.89)

Şekil.5'te verilen bir geri yayımlı ağı modeli, giriş, gizli ve çıkış olmak üzere 3 katmandan oluşmakla birlikte, problemin ve veri setinin özelliklerine göre gizli katman sayısı ve yapay sinir hücresi sayısı fazla olabilir. GYA'da gizli katmanda ve çıkış katmanında bir de eşik hücresi (bias neuron) bulunur (Yıldız, 2006, s.89).

$$w_{pq,k}(N + 1) = w_{pq,k}(N) - \eta_{p,q} \delta_{pq,k} \Phi_{p,j}(I) + \mu \Delta w_{pq,k}(N) \quad (2.3)$$

Bu denklemde;

w : ağırlık değeri,

p, q : hücre numarası,

k : katman sayısı,

N : devir sayacı,

$\delta$  : geriye yayılacak hata terimi,

$\mu$  : momentum,

$\Phi(I)$  : ağı bulduğu sonuçtur.

GYA ağlarında bir katmandan diğer bir katmana aradaki katmanları atlamadan sırayla geçilebilir. Girdi olarak kullanılan veriler her bir katmandaki sinir hücrelerinde işlenerek katmandan katmana aktarılarak çıktı katmanına kadar iletilir. Sonuçta elde edilen çıktılar hedef çıktılar ile karşılaştırılarak aradaki fark yani "hata"

hesaplanır. Hesaplanan hata değerleri her çıktı düğümüne karşılık gelen katmandaki düğümlere geri gönderilir. Böylece ara katmanlardaki düğümler toplam hatanın sadece kendine ait olarak hesaplanan kısmını içerir. Bu şekilde hataya bağlı olarak her katmandaki ağırlıklar yeniden güncellenir. Bu işleme hatanın önceden belirlenmiş olan kabul edilebilir sınırlar içerisinde kalmasına kadar devam edilir ve böylece ağırlık eğitimi tamamlanmış olur (Yıldız, 2006, s.90).

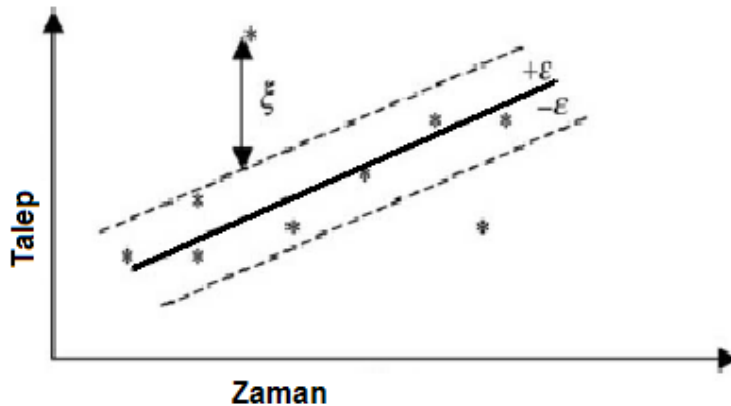
### **3.3. DESTEK VEKTÖR REGRESYONU-DVR (SUPPORT VECTOR REGRESSION)**

Destek Vektör algoritması Rusya'da Viladimir Vapnik tarafından 1960'lı yıllarda geliştirilen, Genelleştirilmiş Dikey Algoritmasının bir doğrusal olmayan algoritmasıdır. Destek Vektör Makineleri (Support Vector Machines) (DVM) ileri yönde beslemeli, öğrenme teorisine dayanan bir sinir ağı algoritmasıdır. Destek Vektör Makinesi(DVM) çok kolay formülasyondan oluşan ve daha esnek bir öğrenim modeli sunmaktadır. DVM'lerin artan popülerliği sağlam teorik yapısı ve bilimsel olarak geniş kullanım alanına sahip olmasına bağlıdır (Smola and Schölkopf, 2003; Levis and Papageorgiou ,2005; Chuang and Lee, 2009).

DVM'ye üç özelliği ile ilgi çekmektedir: 1) geleneksel sınıflandırıcılara kıyasla, yüksek sınıflandırma doğruluğu ve çok iyi genelleme yetkinliği, 2) çok az kontrol parametresi içermesi nedeniyle mimari tasarım için daha az çalışmanın yeterli olması 3) öğrenme problemini doğrusal olarak kısıtlanmış kuadratik programlama yöntemleriyle çözme olasılığıdır. DVM temel avantajları şunlardır: 1. DVM'ler yapısal risk minimizasyonunun bir formunu uygulamaktadırlar 2. Konveks bir kuadratik programlama problemidir 3. Elde edilen sınıflandırıcı tamamen kendi destek vektörleri cinsinden ve kernel fonksiyonu tipinden belirlenebilir 4. Öğrenme tekniği; eğitim kümesindeki nispeten çok az veri noktasıyla bile iyi genelleme yapması. DVM'nin en önemli özellikleri uygulama kolaylığı ve güçlü olmalarıdır (Tolun, 2005).

DVM'lerin döviz kuru tahmini konusunda çeşitli çalışmalar yapılmıştır. (Cao et al., 2005) yaptıkları çalışmada DVM'lerin tahmin gücünün olduğunu ispat etmişler, DVM'lerin finans zaman serisi tahmininde kullanılabileceğini göstermişlerdir. (Pai et al.,2006) DVM'leri doğrusal olmayan regresyon ve zaman serisi problemlerinin çözümünde başarıyla uygulamışlardır. (Cao and Tay,2001) yaptıkları deneyler sonucunda, DVM'lerin döviz kuru tahmininde kullanımının avantajlı olduğu sonucuna ulaşmışlardır.

Yeni araştırmalar, veriden öğrenme için yeni ve güçlü bir teknik olarak ortaya çıkan ve özellikle daha iyi bir performans için sınıflandırma ve regresyon problemlerini çözen Destek Vektör Makinesine yönlendirilmiştir. Benzersiz sonuca ulaşmadaki genelleştirme yeteneği avantajlarından dolayı DVM'ler araştırmacıların dikkatini çekmiştir ve text sınıflandırma, görüntü tanıma veri madenciliği ve bioinformatik gibi birçok uygulamada kullanılmıştır. Vapnik'in duyarsızlık kayıp fonksiyonunu duyurmasıyla DVM'nin regresyon modeli olan Destek Vektör Regresyonu - DVR (Support Vector Regression) doğrusal olmayan tahmin problemlerinin çözümü için artan bir ilgi çekmektedir. DVR makine endüstrisinin ürün değer tahmini, motor dayanım tahmini, rüzgar hız tahmini ve finansal zaman serisi tahmini gibi birçok zaman serisi tahmin probleminde başarılı bir şekilde uygulanmıştır (Lu et al.,2009). Şekil.6'da Destek Vektör Regresyonunun grafik gösterimi verilmiştir.



Şekil.6. Destek Vektör Regresyon (DVM) 'un Grafik Gösterimi. (\* 'lar öğrenim noktalarını sembolize etmektedir.) (Lu et al.,2009)

Öğrenim verisi şu şekilde olmak üzere (Levis and Papageorgiou,2005),

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \in R^Z \times R$$

$x_t$ , t zaman periyodundaki giriş vektörüdür,

$y_t$ , her  $x_t$  için atanmış müşteri talebidir.

Burada amaç  $F(x_t)$  regresyon fonksiyonunu bulmaktır.

$$F(x_t) = w^T x_t + \beta \quad (2.4)$$

$$w, x_t \in R^Z, \beta \in R$$

İstatistiki öğrenme teorisi, genelleştirme yapısında regresyon fonksiyonu elde etmek için model karmaşıklığı ve öğrenme hata toleransının kontrol edilmesini içermektedir. Öğrenim hata toleransını kontrol etmek için  $\varepsilon$ -duyarsız kayıp fonksiyonu  $\xi$  şu şekilde oluşturulur (Levis and Papageorgiou,2005):

$$|\xi|_\varepsilon = \max(0, |F(x_t) - y_t| - \varepsilon) \quad (2.5)$$

$\varepsilon$ -duyarsız kayıp fonksiyonu  $\varepsilon$ 'den küçük hataların dikkate alınmadığını göstermektedir. Fakat grafikte  $\varepsilon$ ' den büyük sapmalar da  $\varepsilon$ -duyarsız tüpünün dışında kalan öğrenim noktaları olarak gösterilmiştir. Sonuçta DVR fonksiyonu Problem  $\varepsilon - SVR$  olarak şu şekilde ifade edilir :

$$\begin{aligned} \min_{w, \beta, \xi_t, \xi_t^*} & \frac{1}{2} \|w\|^2 + c \cdot \sum_{t=1}^N (\xi_t + \xi_t^*) \\ y_t - w^T x_t - \beta & \leq \varepsilon + \xi_t \quad \forall t = 1, \dots, N \\ w^T x_t + \beta - y_t & \leq \varepsilon + \xi_t^* \quad \forall t = 1, \dots, N \\ \xi_t & \geq 0 \quad \forall t = 1, \dots, N \\ \xi_t^* & \geq 0 \quad \forall t = 1, \dots, N \end{aligned} \quad (2.6)$$

Fonksiyondaki ilk terim model karmaşıklığını (düzlüğü), ikinci terim hata toleransını( ya da model doğruluğunu) ifade eder. C parametresi düzlük ve hata

toleransı arasındaki denge noktasını belirleyen pozitif bir değerdir.  $\xi_t$  ve  $\xi_t^*$ ,  $\varepsilon$ -duyarsız tütünün aşağısı ve yukarısındaki mutlak sapmaları sembolize eder. Lagrangean fonksiyonu şu şekilde oluşturulur (Levis and Papageorgiou,2005):

$$\begin{aligned}
 L = & \frac{1}{2} \|w\|^2 + C \cdot \sum_{t=1}^N (\xi_t + \xi_t^*) & (2.6) \\
 & - \sum_{t=1}^N \lambda_t \cdot (\varepsilon + \xi_t - y_t + (w^T x_t + \beta)) \\
 & - \sum_{t=1}^N \lambda_t^* \cdot (\varepsilon + \xi_t^* + y_t - (w^T x_t + \beta)) \\
 & - \sum_{t=1}^N \mu_t \cdot \xi_t - \sum_{t=1}^N \mu_t^* \cdot \xi_t^*
 \end{aligned}$$

Optimum Karush–Kuhn–Tucker (KKT) noktasındaki şartlar L'nin kısmi türevlerini 0 olmaya zorlamaktadır.

$$\begin{aligned}
 \frac{\partial L}{\partial w} = 0 & \Rightarrow w - \sum_{t=1}^N (\lambda_t - \lambda_t^*) \cdot x_t = 0 & (2.7) \\
 \frac{\partial L}{\partial \beta} = 0 & \Rightarrow \sum_{t=1}^N (\lambda_t - \lambda_t^*) = 0 \\
 \frac{\partial L}{\partial \xi_t} = 0 & \Rightarrow C - \lambda_t - \mu_t = 0 \quad \forall t = 1, \dots, N \\
 \frac{\partial L}{\partial \xi_t^*} = 0 & \Rightarrow C - \lambda_t^* - \mu_t^* = 0 \quad \forall t = 1, \dots, N
 \end{aligned}$$

Son elde ettiğimiz eşitliği önceki eşitlikte yerine koyduğumuzda çift optimizasyon problemini(Problem D) elde ederiz:

$$\begin{aligned}
 \max_{\lambda_t, \lambda_t^*} & \frac{1}{2} \sum_{r=1}^N \sum_{t=1}^N (\lambda_r - \lambda_r^*) \cdot (\lambda_t - \lambda_t^*) \cdot x_r^T x_t \\
 & - \varepsilon \sum_{t=1}^N (\lambda_t + \lambda_t^*) + \sum_{t=1}^N y_t \cdot (\lambda_t - \lambda_t^*)
 \end{aligned}$$

$$\begin{aligned} \sum_{t=1}^N (\lambda_t - \lambda_t^*) &= 0 \\ 0 \leq \lambda_t &\leq C \quad \forall t = 1, \dots, N \\ 0 \leq \lambda_t^* &\leq C \quad \forall t = 1, \dots, N \end{aligned}$$

Çift problemin çözümü sonucunda  $w$  optimum vektör ve  $F(x_t)$  regresyon fonksiyonunu elde ederiz:

$$\begin{aligned} w &= \sum_{t=1}^N (\lambda_t - \lambda_t^*) \cdot x_t \\ F(x_t) &= w^T x_t + \beta \end{aligned} \tag{2.8}$$

Sonuçta  $F(x_t)$  regresyon fonksiyonu şu hale gelir:

$$F(x_t) = \sum_{t=1}^N (\lambda_t - \lambda_t^*) \cdot x_t^T x_t + \beta \tag{2.9}$$

$$\lambda_t \cdot (\varepsilon + \xi_t - y_t + (w^T x_t + \beta)) = 0 \quad \forall t = 1, \dots, N$$

$$\lambda_t^* \cdot (\varepsilon + \xi_t^* + y_t - (w^T x_t + \beta)) = 0 \quad \forall t = 1, \dots, N$$

$$(C - \lambda_t) \cdot \xi_t = 0 \quad \forall t = 1, \dots, N$$

$$(C - \lambda_t^*) \cdot \xi_t^* = 0 \quad \forall t = 1, \dots, N$$

Problem P şu şekilde tanımlanır:

$$\min_{\beta, \xi, \xi^*} \sum_{t=1}^N (\xi_t + \xi_t^*) \tag{2.10}$$

$$\begin{aligned}
y_t - \sum_{r=1}^N (\lambda_r - \lambda_r^*) \cdot x_r^T \cdot x_t - \beta &\leq \varepsilon + \xi_t \quad \forall t = 1, \dots, N \\
\sum_{r=1}^N (\lambda_r - \lambda_r^*) \cdot x_r^T \cdot x_t + \beta - y_t &\leq \varepsilon + \xi_t^* \quad \forall t = 1, \dots, N \\
\xi_t &\geq 0 \quad \forall t = 1, \dots, N \\
\xi_t^* &\geq 0 \quad \forall t = 1, \dots, N
\end{aligned}$$

Regresyon fonksiyonu şu hale getirilir:

$$F(x_t) = \sum_{r=1}^N (\lambda_r - \lambda_r^*) \cdot \Phi(x_r^T) \cdot \Phi(x_t) + \beta \quad (2.11)$$

Gelecek uzayında K Kernel Fonksiyonu şu şekilde sembolize edilir:

$$K(x_r, x_t) = \Phi(x_r^T) \cdot \Phi(x_t) \quad (2.12)$$

K Kernel Fonksiyonu denklemden yerine konulduğunda:

$$F(x_t) = \sum_{r=1}^N (\lambda_r - \lambda_r^*) \cdot K(x_r, x_t) + \beta \quad (2.13)$$

Çift ve temel model fonksiyonu doğrusal olmayan destek vektör regresyonu şu şekilde oluşturulur :

$$\begin{aligned}
\max_{\lambda_t, \lambda_t^*} & - \frac{1}{2} \sum_{t=1}^N \sum_{r=1}^N (\lambda_t - \lambda_t^*) \cdot (\lambda_r - \lambda_r^*) \cdot K(x_t, x_r) \\
& - \varepsilon \sum_{t=1}^N (\lambda_t + \lambda_t^*) + \sum_{t=1}^N y_t \cdot (\lambda_t - \lambda_t^*)
\end{aligned}$$



$$\sum_{t=1}^N (\lambda_t - \lambda_t^*) = 0$$

$$0 \leq \lambda_t \leq C \quad \forall t = 1, \dots, N$$

$$0 \leq \lambda_t^* \leq C \quad \forall t = 1, \dots, N$$

$$\min_{\beta, \xi, \xi^*} \sum_{t=1}^N (\xi_t + \xi_t^*)$$

$$y_t - \sum_{r=1}^N (\lambda_r - \lambda_r^*) \cdot K(x_t, x_r) - \beta \leq \varepsilon + \xi_t \quad \forall t = 1, \dots, N$$

$$\sum_{r=1}^N (\lambda_r - \lambda_r^*) \cdot K(x_t, x_r) + \beta - y_t \leq \varepsilon + \xi_t^* \quad \forall t = 1, \dots, N$$

$$\xi_t \geq 0 \quad \forall t = 1, \dots, N$$

$$\xi_t^* \geq 0 \quad \forall t = 1, \dots, N$$

## 4. VERİ MADENCİLİĞİ UYGULAMASI

### 4.1. ÇALIŞMANIN TANITILMASI

#### 4.1.1. Çalışmanın Amacı ve Önemi

Yapılan çalışmada, USD, EUR, GBP ve JPY döviz kurlarının gelecekteki değerinin ve yönünün, Yapay Sinir Ağları ve Destek Vektör Regresyonunun kullanılması ile yüksek doğruluk oranıyla öngörülmesine çalışılmıştır.

Çalışma sayesinde kullanılan tekniklerin döviz kuru tahmininde doğruluk oranları tespit edilecek ve bu alanda yapılmış uygulamalara katkı sağlanmış olacaktır.

#### 4.1.2. Çalışmanın Kapsamı

Çalışmada USD, EUR, GBP ve JPY döviz kurlarının 10 yıllık günlük döviz kuru değeri baz alınmış ve YSA ve DVM modellemesi kullanılarak tahmin geliştirilmiştir. Çalışmada MATLAB programı kullanılarak modelleme oluşturulmuştur. DVM ve YSA ile bulunan döviz kuru tahminleri MAE, RMSE, NRMSE ve MAPE metrikleri ile doğruluk değerleri açısından karşılaştırılmıştır.

Çalışmada döviz kuru tahmininde Veri Madenciliği tekniklerinden Yapay Sinir Ağları ve Destek Vektör Regresyonu kullanılması konusu incelenmiştir. Yapılan çalışmada, döviz kurlarının gelecekteki değerinin ve yönünün yüksek doğruluk oranı ile öngörülmesine çalışılmıştır.

#### 4.1.3. Çalışmada Kullanılan Değişkenler

**Basit Hareketli Ortalama (Simple Moving Average) :**

Döviz kuru tahmininde kullanılan en yaygın ortalama değerdir. Belirli periyottaki kurların basit ortalamasıdır. n adet kurun basit hareketli ortalamasını almak için şu formül kullanılır:

$$SMA = \frac{\sum_{i=1}^n close_i}{n} \quad (2.15)$$

En çok kullanılan zaman aralıkları 5, 10 ve 50 günlük aralıklardır. Biz de uygulamada 5,10 ve 50 günlük SMA'ları kullandık.

### **Bant Geniřlięi (Band Width) :**

Band width, döviz kurlarının salınım yaptığı üst zirve ve en alt düzey arasındaki bölgedir. Band width, döviz kurlarındaki gürültü diye tabir edilen aşırı oynaklıkları belirli bir aralıkta tutup grafiksel deęişimin daha kolay yorumlanmasına imkan tanır.

$$\text{Upper Band} = SMA + 2x \text{SD(Standart Deviation)}$$

$$\text{Lower Band} = SMA - 2x \text{SD(Standart Deviation)}$$

$$\text{Band Width} = \text{Upper Band} - \text{Lower Band}$$

Yatırımcılar alım satım yapacakları zamanı band width'e göre belirlerler. Band Width'in en altında yani Lower Band'da alım yaparlar, Band Width'in en üstünde yani Upper band'de satış yaparlar. (Bollinger,2010)

### **Deęişim Oranı (Rate of Change) :**

Deęişim Oranı, iki tarih arasında kurların deęişim yüzdesini verir. Eđer deęişim oranı (-) çıkarsa kurlar düşüyor demektir. Deęişim oranı (+) çıkarsa kurlar artıyor demektir. Deęişim oranı 0 çıkarsa kurlar sabit kalmış demektir.

$$\text{Değişim Oranı: } \frac{\text{Son Kur} - \text{İlk Kur}}{\text{İlk Kur}} \times 100 \quad (2.16)$$

### **KAMA (Kaufman Adaptive Moving Average) :**

Perry Kaufman 1995 yılında Adaptive moving average konusunda çalışmış ve kendi geliştirdiği belirli gürültü göstergesi olan mükemmellik oranı ER (Efficiency Ratio)' ye endekslemiş olan KAMA metodunu geliştirmiştir. (Kaufman, 1995) Kurlardaki oynaklık fazla olduğu zaman ER, 0'a yakın olacaktır. Kurlardaki oynaklık az olduğu zaman ER, 1'e yakın olacaktır. (Angeli, 2006)

$$\text{smoothing constant (c)} = [\text{ER} * (\text{fastest} - \text{slowest}) + \text{slowest}]^2 \quad (2.17)$$

$$\text{KAMA}_0 = \text{KAMA}_{-1} + c * (\text{C}_0 - \text{KAMA}_{-1}) \quad (2.18)$$

$$\text{fastest} = 0.667$$

$$\text{slowest} = 0.0645$$

## **4.2. DÖVİZ KURU TAHMİNİ VE MODELLEMESİ**

### **4.2.1. Veri Seti**

Bu çalışmada kullanılan veriler, <http://www.oanda.com/currency/converter/> internet sitesinden elde edilmiştir. Çalışmada USD, EUR, GBP ve JPY döviz kurlarının 01.01.2001 ile 30.05.2011 tarihleri arasındaki 3.802 adet değeri vcf formatında alınmıştır. Daha sonra vcf dosyaları excell formatına dönüştürülerek kullanıma hazırlanmıştır.

### **4.2.2. Değişkenler ve Kriterler**

Excell formatında hazırlanan verilerden 5 günlük, 10 günlük ve 50 günlük SMA(Simple Moving Avarage)'lar hesaplanmıştır. Band Width, Rate of Change ve KAMA değerleri hesaplanmıştır. Bunlar daha sonra MATLAB programında girdi olarak kullanılmıştır.

MATLAB'da DVM ve YSA algoritmaları kullanılarak tahmini döviz kurları bulunmuş ve bunlar için grafikler çizilerek gerçek döviz kuru ile tahmin edilen döviz kurlarının görsel olarak karşılaştırılması sağlanmıştır. Ayrıca MAE, RMSE, NRMSE ve MAPE metrikleri bulunmuştur. Bulunan metrikler yapılan tahminlerin doğruluk değerini göstermiştir.

**MAE(Mean Absolute Error) - Ortalama Mutlak Hata:** Tahmin edilen ve gerçek değer arasındaki farkın ortalamasıdır, yani ortalama tahmin hatasıdır. Formülü şu şekildedir: (Willmott et al.,2005)

$$MAE = \frac{\sum_{i=1}^n |x_i - y_i|}{n} \quad (3.1)$$

$x_i$ : Gerçek Değer,  $y_i$ : Tahmini Değer

**RMSE(Root Mean Squared Error) - Kök Hata Kareler Ortalaması:** Kök Hata Kareleri Ortalaması sayısal tahminler için en çok kullanılan doğrulama ölçülerinden biridir. Bu değer, tahmini değer ve onun karşılık gelen gerçek değer arasındaki farkın karelerinin ortalaması alınarak hesaplanır. Formülü şu şekildedir: (Willmott et al.,2005)

$$RMSE = \frac{\sqrt{\sum_{i=1}^n (x_i - y_i)^2}}{n} \quad (3.2)$$

$x_i$ : Gerçek Değer,  $y_i$ : Tahmini Değer

**NRMSE(Normalised Root Mean Squared Error) - Düzeltilmiş Kök Hata Kareler Ortalaması:** Düzeltilmiş Kök Hata Kareleri Ortalaması, tahmini değer ve onun karşılık gelen gerçek değer arasındaki farkın karelerinin toplamının gerçek değerlerin karelerinin toplamının karekökünün 100 ile çarpımıdır. Bulunan değer % değeridir.. Formülü şu şekildedir: (Fryza et al.,2003)

$$\text{NRMSE \%} = 100 \times \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i)^2}} \quad (3.3)$$

$x_i$ : Gerçek Değer,  $y_i$ : Tahmini Değer

**MAPE(Mean Absolute Percentage Error) - Ortalama Mutlak Yüzde Hata:** Ortalama Mutlak Yüzde Hata, tahmini değer ve onun karşılık gelen gerçek değer arasındaki farkın gerçekte değere bölümünün ortalamasının 100 ile çarpımıdır. Bulunan değer % değeridir.. Formülü şu şekildedir: (Li et al.,2007)

$$\text{MAPE} = 100 \times \frac{\sum_{i=1}^n \frac{|x_i - y_i|}{y_i}}{n} \quad (3.4)$$

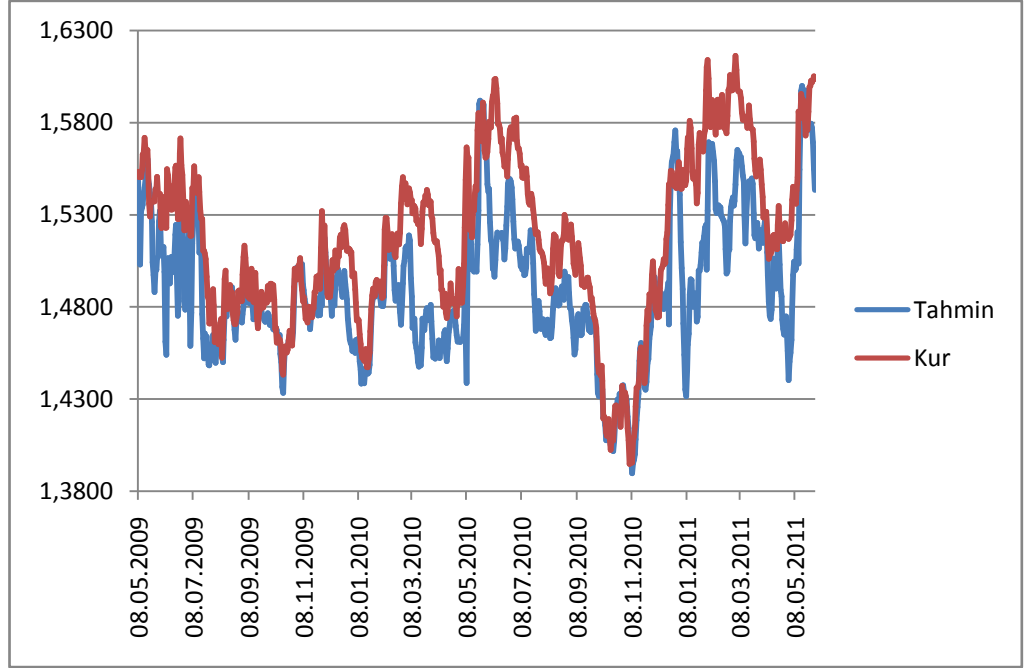
$x_i$ : Gerçek Değer,  $y_i$ : Tahmini Değer

#### 4.2.3. Modelleme ve Analiz Sonuçları

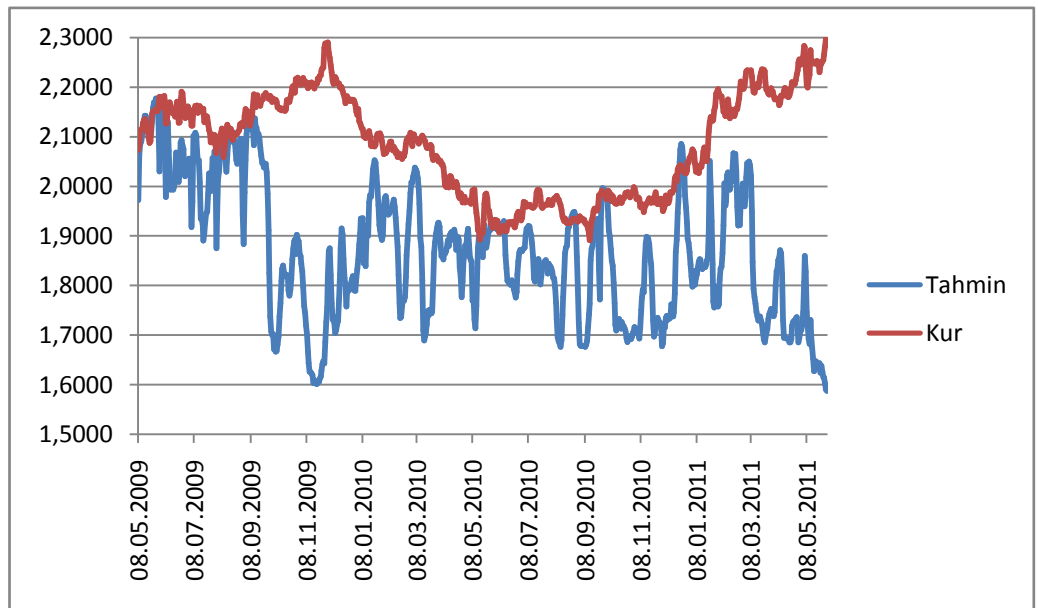
Önce USD, EUR, GBP ve JPY döviz kurlarının 01.01.2001 ile 30.05.2011 tarihleri arasındaki 3.802 adet değeri vcf formatında alınmıştır ve bu dosyalar excell formatına dönüştürülmüştür. Excell formatında hazırlanan verilerden 5 günlük, 10 günlük ve 50 günlük SMA(Simple Moving Avarage)'lar hesaplanmıştır. Band Width, Rate of Change ve KAMA değerleri hesaplanmıştır.

**Destek Vektör Makinesi (DVM) ile Döviz Kuru Tahmini:**

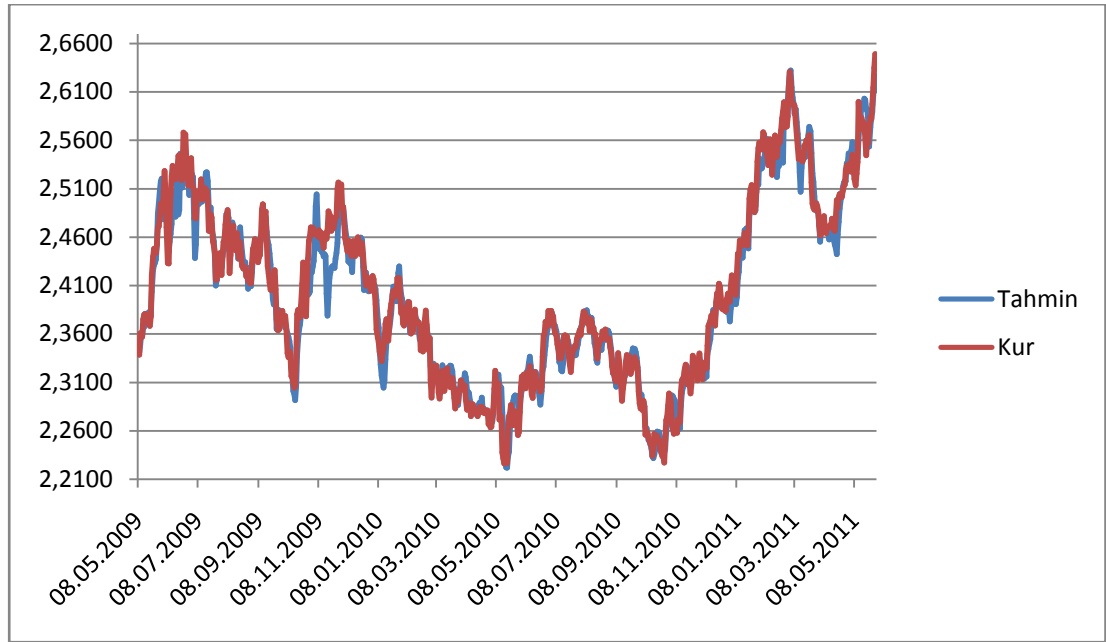
MATLAB'da önce DVM algoritması kullanılarak döviz kurları tahmini yapılmıştır ve bunlar için gerçek döviz kurları ile tahmini döviz kurları karşılaştırılması grafikler çizilerek yapılmıştır. Şekil.7,8,9 ve 10'da DVM kullanılarak elde edilen USD, EUR, GBP ve JPY döviz kuru tahminleri gerçek verilerle birlikte grafik olarak verilmiştir.



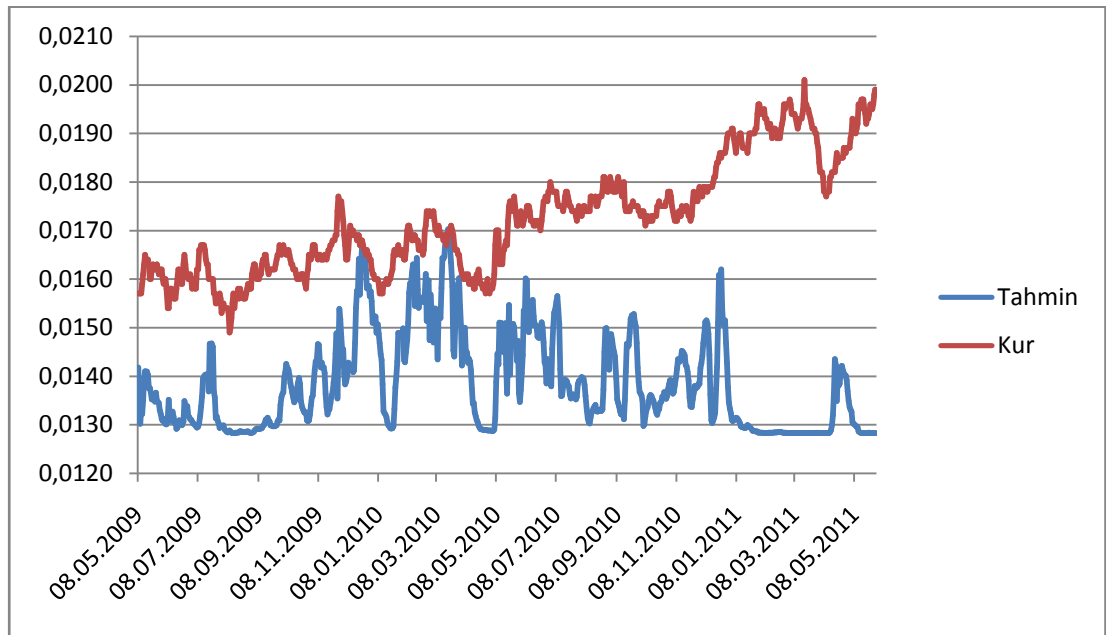
Şekil.7. DVM Kullanarak USD Kuru Tahmini



Şekil.8. DVM Kullanarak EUR Kuru Tahmini



Şekil.9. DVM Kullanarak GBP Kuru Tahmini



Şekil.10. DVM Kullanarak JPY Kuru Tahmini

Tablo3.'de DVM kullanılarak elde edilen döviz kuru tahminleri ve gerçek veriler kullanılarak bulunan MAE,RMSE,NRMSE ve MAPE değerleri verilmiştir.

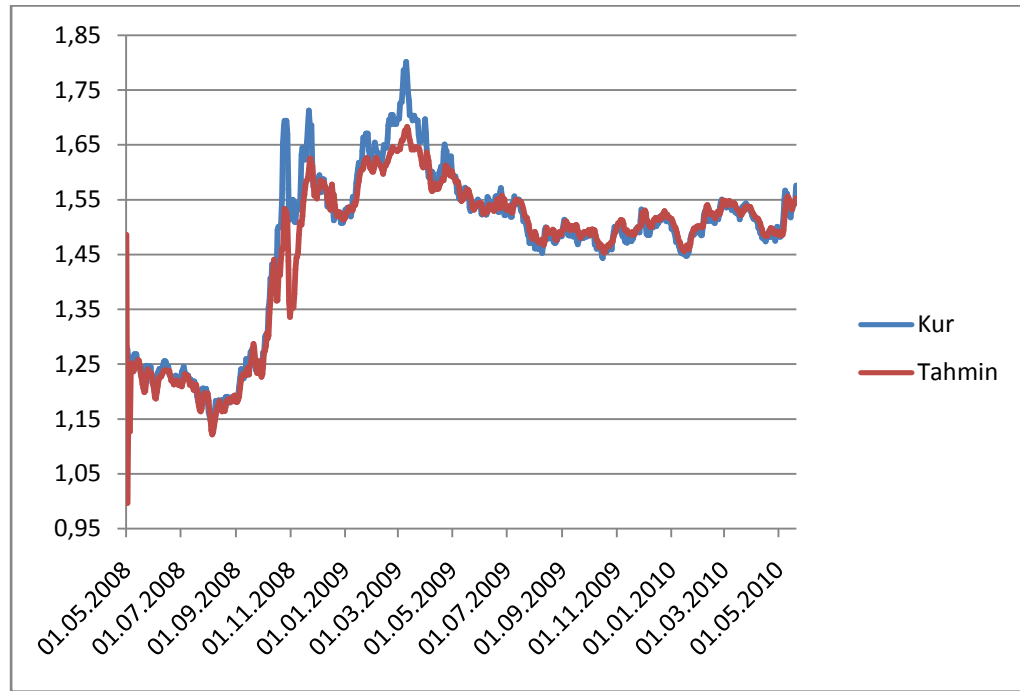


Tablo.3. DVM'ye Göre Hesaplanan Değerler

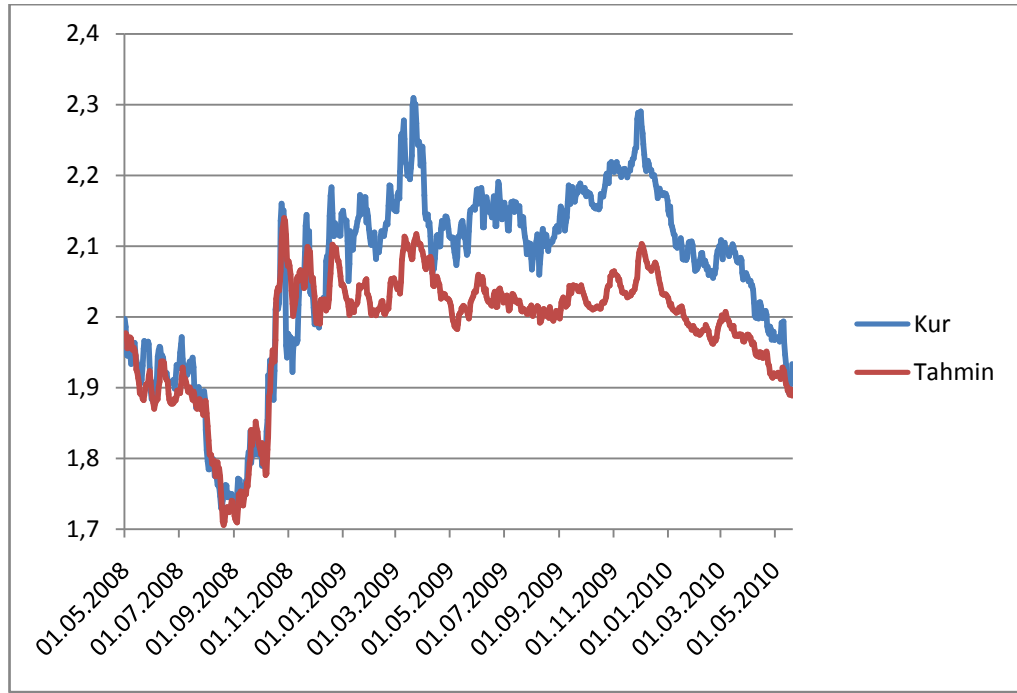
Değişkenler Kurlar	MAE	RMSE	NRMSE %	MAPE %
<b>USD</b>	0.02806	0.03653	2.40761	1.82365
<b>EUR</b>	0.21813	0.27629	13.23055	10.24596
<b>GBP</b>	0.01260	0.01725	0.71709	0.52056
<b>JPY</b>	0.00340	0.00376	21.78205	19.31547

### Yapay Sinir Ağları (YSA) ile Döviz Kuru Tahmini:

MATLAB'da önce YSA algoritması kullanılarak döviz kurları tahmini yapılmıştır ve bunlar için gerçek döviz kurları ile tahmini döviz kurları karşılaştırılması grafikler çizilerek yapılmıştır. Şekil.11,12,13 ve 14'de YSA kullanılarak elde edilen USD, EUR, GBP ve JPY döviz kuru tahminleri gerçek verilerle birlikte grafik olarak verilmiştir.



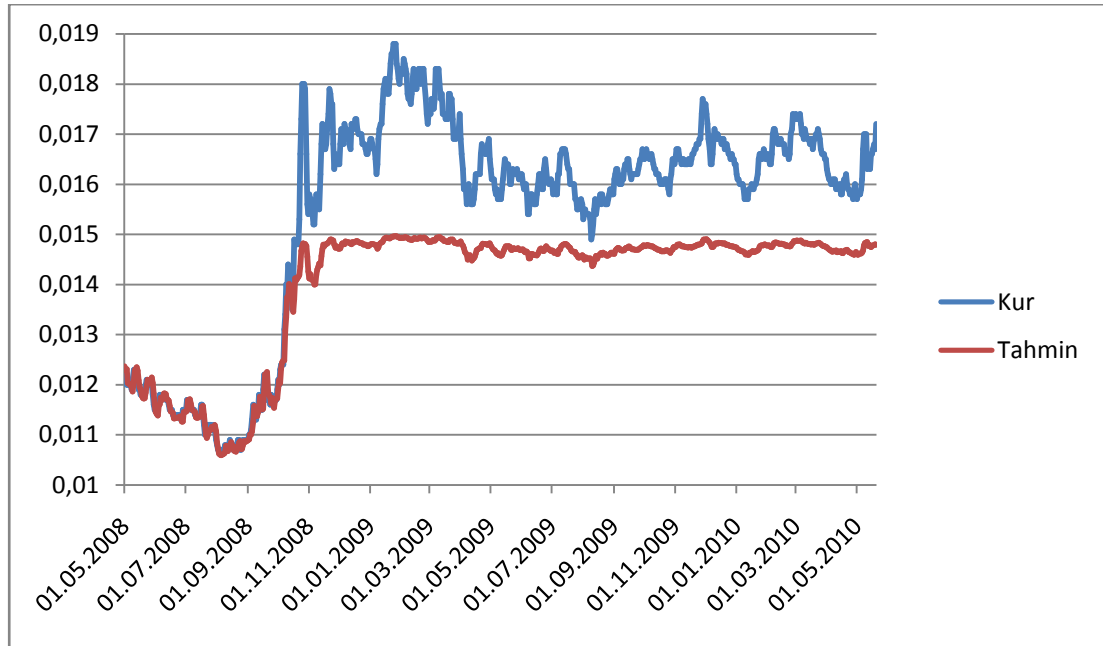
Şekil.11. YSA Kullanarak USD Kuru Tahmini



Şekil.12. YSA Kullanarak EUR Kuru Tahmini



Şekil.13. YSA Kullanarak GBP Kuru Tahmini



Şekil.14. YSA Kullanarak JPY Kuru Tahmini

Tablo.4.'de YSA kullanılarak elde edilen döviz kuru tahminleri ve gerçek veriler kullanılarak bulunan MAE, RMSE, NRMSE ve MAPE değerleri verilmiştir.

Tablo.4. YSA'ya Göre Hesaplanan Değerler

Değişkenler Kurlar	MAE	RMSE	NRMSE %	MAPE %
USD	0.02141	0.03860	2.62675	1.47194
EUR	0.08525	0.09936	5.00600	4.23372
GBP	0.01777	0.02549	1.06481	0.74063
JPY	0.00144	0.00171	12.11029	9.81208

#### 4.2.4. YSA ve DVM ile Yapılan Tahminin Karşılaştırılması

USD, EUR, GBP ve JPY döviz kurlarının 10 yıllık verilerine göre YSA ve DVM algoritmalarıyla bulunan döviz kuru tahminlerine dayanılarak hesaplanan MAE, RMSE, NRMSE ve MAPE ölçütleri Tablo.5'de verilmiştir.

Tablo.5. DVM ve YSA'ya Göre Hesaplanan Değerler

	DVM				YSA			
	MAE	RMSE	NRMSE %	MAPE %	MAE	RMSE	NRMSE %	MAPE %
USD	0.02806	0.03653	2.40761	1.82365	0.02141	0.03860	2.62675	1.47194
EUR	0.21813	0.27629	13.23055	10.24596	0.08525	0.09936	5.00600	4.23372
GBP	0.01260	0.01725	0.71709	0.52056	0.01777	0.02549	1.06481	0.74063
JPY	0.00340	0.00376	21.78205	19.31547	0.00144	0.00171	12.11029	9.81208

MAE ve RMSE gerçek döviz kuru ile tahmin edilen döviz kuru arasındaki hata oranını, NRMSE ve MAPE ise hata yüzdelerini vermektedir. MAE ve RMSE değerleri hata miktarının büyüklüğünü ifade ettiği için MAE ve RMSE değerlerinin az olması o algoritmanın başarı şansını artırmaktadır. Aynı şekilde NRMSE ve MAPE değerleri hata yüzdelerini verdiği için bu değerlerin de düşük olması algoritmaların başarısını artırmaktadır. Tablo.5 incelendiğinde:

**USD döviz kuru için,** YSA ve DVM ile elde edilen MAE ve RMSE değerleri birbirine yakın bulunmuştur. Aynı şekilde NRMSE ve MAPE yüzde değerleri de birbirine yakın çıkmıştır.

**EUR döviz kuru için,** YSA ve DVM ile elde edilen MAE ve RMSE değerleri karşılaştırıldığında YSA'nın daha başarılı olduğu gözlenmiştir. Aynı şekilde NRMSE ve MAPE yüzde değerleri de karşılaştırıldığında YSA'nın daha düşük yüzde hatalar verdiği dolayısıyla daha başarılı tahmin ürettiği tespit edilmiştir.

**GBP döviz kuru için,** YSA ve DVM ile elde edilen MAE ve RMSE değerleri karşılaştırıldığında DVM'nin daha başarılı olduğu gözlenmiştir. Aynı şekilde NRMSE ve MAPE yüzde değerleri de karşılaştırıldığında DVM'nin daha düşük yüzde hatalar verdiği dolayısıyla daha başarılı tahmin ürettiği tespit edilmiştir.

**JPY döviz kuru için,** YSA ve DVM ile elde edilen MAE ve RMSE değerleri karşılaştırıldığında YSA'nın daha başarılı olduğu gözlenmiştir. Fakat NRMSE ve MAPE yüzde değerleri de karşılaştırıldığında YSA'nın daha düşük yüzde hatalar verdiği dolayısıyla daha başarılı tahmin ürettiği tespit edilmiştir.

Görüldüğü üzere YSA ve DVM'lerin başarısı döviz kurlarına göre değişiklik göstermektedir. Fakat genel olarak söylemek gerekirse iki algoritma da başarılı ve birbirine yakın değerler vermektedir. Dolayısıyla döviz kuru tahmininde YSA ve DVM'ler rahatlıkla tercih edilebilir özelliktedir.

## 5. SONUÇ

Teknolojinin geliřimiyle bilgisayar ortamında ve veritabanlarında tutulan veri miktarının artması Veri Madenciliğini zorunlu kılmıřtır. Veri Madencilięi uygulamalarını gerekleřtirmek iin eřitli bilgisayar programları geliřtirilmiřtir. Yapılan arařtırmalara gre en ok tercih edilen veri madencilięi araları RapidMiner, R ,Excell ve Knime'dir.

Bir lkenin para biriminin bařka bir lkenin para birimine gre deęeri olan dviz kurunun, etki eden faktrlerin ok fazla olması nedeniyle zellikle Trkiye gibi ekonomisi deęiřken olan lkelerde gelecekteki deęerlerini tahmin edebilmek ok da kolay deęildir. Oysa ekonomide iřletmelerin planlama yapabilmeleri aısından geleceęin ngrlebilmesi ok nemli bir gereksinimdir. zellikle son yıllarda dnyada ve lkemizde yařanan ekonomik krizlerden ve politik geliřmelerden dolayı meydana gelen piyasalardaki dalgalanmalar, yatırım aralarının ve dviz kurlarının gelecekteki alacaęı deęerin ngrlebilmesini nemli hale getirmiřtir. İřte bu noktada dviz kuru tahmini alanında Veri Madencilięi tekniklerinin kullanılması byk nem arz etmektedir.

Yapay Sinir Aęları(YSA) ve Destek Vektr Regresyonu(DVR), zm g ve karmařık olan bir ok problemin zmne uygulanmıř ve son derece bařarılı sonular elde edilerek, birok alanda uygulanmaya bařlanmıřtır. YSA ve DVR'ler ekonomi alanında geleceęi ngrmede, zellikle dviz kuru gibi lineer olmayan deęiřkenlerin, baęımlı zaman serilerinin tahmin edilmesinde nemli aralar olmuřlardır.

Bu alıřmada temel olarak, Yapay Sinir Aęları ve Destek Vektr Makineleri kurularak zaman serisi seklinde olan dviz kurlarının gelecekteki deęerleri yksek doęruluk oranıyla tahmin edilmeye alıřılmıřtır. Yapılan tez alıřmasında USD, EUR, GBP ve JPY dviz kurlarının 10 yıllık verisi kullanılmıř ve Simple Moving Avarage(SMA)'lar hesaplanmıřtır. Band Width(BW), Rate of Change(ROC) ve Kaufman Adaptive Moving Average(KAMA) deęerleri de hesaplanarak MATLAB

programında girdi olarak kullanılmıştır. Daha sonra YSA ve DVM algoritmalarıyla bulunan döviz kuru tahminlerine dayanılarak MAE, RMSE, NRMSE ve MAPE ölçütleri hesaplanmıştır.

MAE ve RMSE gerçek döviz kuru ile tahmin edilen döviz kuru arasındaki hata oranını, NRMSE ve MAPE ise hata yüzdeleri vermektedir. MAE ve RMSE değerleri hata miktarının büyüklüğünü ifade ettiği için MAE ve RMSE değerlerinin az olması o algoritmanın başarı şansını artırmaktadır. Aynı şekilde NRMSE ve MAPE değerleri hata yüzdeleri verdiği için bu değerlerin de düşük olması algoritmaların başarısını artırmaktadır.

Sonuçta USD, EUR, GBP ve JPY döviz kurlarının YSA ve DVM algoritmalarına göre bulunan tahmin değerleri ve bu değerlere dayanarak bulunan MAE, RMSE, NRMSE ve MAPE ölçütleri göz önüne alındığında YSA ve DVM'lerin başarısının döviz kurlarına göre değişiklik gösterdiği tespit edilmiştir. Fakat genel olarak söylemek gerekirse uygulanan iki algoritma da kendi içinde tutarlı, başarılı ve birbirine yakın değerler vermiştir. Bu nedenle YSA ve DVM ile yapılacak analizlerin, döviz kuru tahmininde karar vermeyi destekleyici bir yöntem olarak rahatlıkla kullanılabileceği sonucuna ulaşılmıştır..

## KAYNAKLAR

AGARD, B. & KUSIAK, A. (2004) , Data mining based methodology for the design of product families, Taylor and Francis Group

ALTINTAŞ, Tamer (2006), Veri Madenciliği Metotlarından Olan Kümeleme Algoritmalarının Uygulamalı Etkinlik Analizi, Sakarya Ün. Fen Bil. Ens.

ANGELI, Alessandro (2006), Smartview, Market Technician: The Journal of the STA

BISHOP, C. M (2006), Pattern Recognition and Machine Learning, Springer

BOLLINGER, John (2010), Using Bollinger Bands, Stocks & Commodities V. 10:2 (47-51)

BROCKWELL, P.J. & DAVIS, R.A. (2003). Introduction to Time networks Series and Forecasting. Springer: New York

CAO, Ding-Zhao & PANG Su-Lin & BAI, Yuan-Huai (2005), Forecasting Exchange Eates Using Support Vector Machines, Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou

CAO, Lijuan & TAY, Francis E.H. (2001), Financial Forecasting Using Support Vector Machines Neural Computing & Applications, Springer

CHUANG, Chen-Chia & LEE, Zne-Jung (2009), Hybrid robust support vector machines for regression with outliers, Elsevier



- DEMİRALAY, Meral (2005), Hiyerarşik kümeleme metotları ile veri madenciliği uygulamaları,
- EL SHAZLY, Mona R. & EL SHAZLY, Hassan E. (1999), Forecasting Currency Prices Using a Genetically Evolved Neural Network Architecture, International Review of Financial Analysis,
- EUNA, Cheol S. & SABHERWAL, Sanjiv (2002), Forecasting exchange rates: Do banks know better?, Global Finance Journal
- FAUST, Jon & ROGERS, John H. and WRIGHT, Jonathan H. (2003), Exchange rate forecasting: the errors we've really made, Journal of International Economics
- FRYZA, Tomas & HANUS, Stanislav (2003), Video Signals Transparency in Consequence of 3-D DCT Transform, Institute of Radio Electronics, Brno University of Technology
- HAMILTON, J.D. (1994). Time Series Analysis. Princeton University Press: Princeton, NJ
- HAYKIN, Simon (1999), Neural Networks: A Comprehensive Foundation, Pearson Prentice Hall
- JASIC, Teo & WOOD, Douglas (2003), Neural Network Protocols and Model Performance, Neurocomputing 55
- JEFFERSON, T. Davis & EPISCOPOS, Athanasios (2001), Predicting Direction Shifts on Canadian-US Exchange Rates with Artificial Neural Networks, Int. J. Intell. Sys. Acc. Fin.Mgmt.,

JOSEPH, Nathan Lael (2001), Model Specification and Forecasting Foreign Exchange Rates with Vector Autoregressions, Journal of Forecasting

KADILAR, Cem & ŞİMŞEK, Muammer & ALADAĞ, Çağdaş Hakan (2009), Forecasting The Exchange Rate Series With ANN: The Case of Turkey, Ekonometri ve İstatistik Sayı:9

KAMURAZZAMAN, Joarder & SARKER, Ruhul A. & AHMAD, Iftexhar (2003), SVM Based Models for Predicting Foreign Currency Exchange Rates

KAUFMAN, Perry J. (1995), Smarter Trading: Improving Performance in Changing Markets, McGraw-Hill, Inc.

KAYA, Halil & KÖYMEN, Kemal (2008), Veri Madenciliği Kavramı ve Uygulama Alanları, Doğu Anadolu Bölgesi Araştırmaları

KDNUGGETS (May 2010),Data Mining / Analytic Tools Used Poll  
<http://www.kdnuggets.com/polls/2010/data-mining-analytics-tools.html>

KOVALERCHUK, Boris & VITYAEV, Evgenii (2005), Data Mining for Financial applications, Data Mining and Knowledge Discovery Handbook

LAROSE, Daniel T. (2005), Discovering Knowledge in Data : An Introduction to Data Mining

LEUNG, Mark T. & CHEN An-Sing & DAOUK, Hazem (2000), Forecasting exchange rates using general regression neural Networks, Computers & Operations Research 27

LEVIS, A. A. & PAPAGEORGIU, L. G. (2005), Customer Demand Forecasting via Support Vector Regression Analysis, Institution of Chemical Engineers

- LI, Hongxing & FANG, Zhaoben & ZHAO, Dongming (2007), GBP/USD Currency Exchange Rate Time Series Forecasting Using Regularized Least-Squares Regression Method, Proceedings of the World Congress on Engineering 2007 Vol II.
- LISI, Francesco & SCHIAVO, Rosa A. (1999), A comparison between neural networks and chaotic models for exchange rate prediction, Computational Statistics & Data Analysis,
- LU, Chi-Jie & LEE, Tian-Shyug & CHIU, Chih-Chou (2009), Financial time series forecasting using independent component analysis and support vector regression, Elsevier
- MARTENS, Martin (2001), Forecasting daily exchange rate volatility using intraday returns, Journal of International Money and Finance,
- MIKUT, Ralf & REISCHL, Markus (2011), Data Mining Tools, Advanced Review
- MITRA, Sushmita & PAL, Sankar K. & MITRA, Pabitra (2002), Data Mining in Soft Computing Framework: A Survey, IEEE Transactions on Neural Networks,
- ÖĞÜT, Sertaç (2005), Veri Madenciliği Kavramı ve Gelişim Süreci, Yeditepe Üniversitesi İletişim Fakültesi Görsel İletişim Tasarımı Bölümü
- ÖZÇINAR, Hüseyin (2006), KPSS Sonuçlarının Veri Madenciliği Yöntemleriyle Tahmin Edilmesi, Pamukkale Ün. Fen Bil. Ens. Bilg. Müh.,
- PAI, Ping-Feng & LIN, Chih-Shen & HONG, Wei-Chiang & CHEN, Chen-Tung (2006), A Hybrid Support Vector Machine Regression for Exchange Rate Prediction, Information and Management Sciences Volume 17

- QI, M. & ZHANG, G.P. (2001). An investigation of model selection criteria for neural network time series forecasting, *European Journal of Operational Research*, 132
- SMOLA, Alex J. & SCHÖLKOPF, Bernhard (2003), A Tutorial on Support Vector Regression, Neurocolt Technical Report
- SYKES, Alan O. (1992), An Introduction to Regression Analysis, The Inaugural Coase Lecture
- TENTI, Paolo (1996). Forecasting Foreign Exchange Rates Using Recurrent Neural Networks, Taylor & Francis
- THEARLING, Kurt An Introduction to Data Mining, <http://thearling.com/text/dmwhite/dmwhite.htm>
- TOLUN, Seda (2008), Destek Vektör Makineleri: Banka Başarısızlığın Tahmini Üzerine Bir Uygulama, İstanbul Ün. Sos. Bil. Ens.
- TRAPLETTI, Adrian & GEYER, Alois (2002), Forecasting Exchange Rates using Cointegration Models and Intra-day Data, *Journal of Forecasting*
- TSAY, R.S. (2005). *Analysis of Financial Time Series*. Wiley: New York, NY
- TSENG, Fang-Mei & TZENG, Gwo-Hsiung & YU, Hsiao-Cheng & YUAN, Benjamin J.C. (2001), Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems*,
- VELAGIC, Jasmin & OSMIC, Nedim & LACEVIC, Bakir (2008), Neural Network Controller for Mobile Robot Motion Control, *World Academy of Science, Engineering and Technology*

- VOJINOVIC, Zoran & KECMAN, Vojislav & SEIDEL, Rainer (2001), A Data Mining Approach to Financial Time Series Modelling and Forecasting, Int. J. Intell. Sys. Acc. Fin.Mgmt
- WANG, John (2009), Encyclopedia of Data Warehousing and Mining, Montclair State University, USA
- WEI, W.W.S. (2005). Time Series Analysis : Univariate and Multivariate Methods, 2nd Edition. Addison Wesley: New York, NY
- WILLMOTT, Cort J. & MATSUURA, Kenji (2005), Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate Research Vol. 30: 79-82
- YAMAN, Kemal & SARUCAN, Ahmet & ATAK, Mehmet & AKTÜRK, Nizami (2001), Dinamik Çizelgeleme İçin Görüntü İşleme ve ARIMA Modelleri Yardımıyla Veri Hazırlama, Gazi Üniv. Müh. Mim. Fak. Der. Cilt 16, No 1
- YAO, Jingtao & POH, Hean- Lea & JASIC, Teo (1996), Foreign Exchange Rates Forecasting With Neural Networks, Department of Information Systems and Computer Science, National University of Singapore
- YAO, Jingtao & TAN, Chew Lim (2000), A case study on using neural networks to perform technical forecasting of forex, Neurocomputing 34
- YILDIZ, Ömür (2006), Döviz Kuru Tahmininde Yapay Sinir Ağlarının Kullanımı, Eskişehir Osmangazi Ün. Sos. Bil. Ens.

## ÖZGEÇMİŞ

1968 Bursa doğumlu. İlk ve orta öğreniminden sonra lise öğrenimini Bursa Tophane Teknik Lisesi Makine Bölümünde tamamladı. Daha sonra Yıldız Teknik Üniversitesi Bilgisayar Bilimleri Mühendisliği Bölümünü okudu. 1993 yılında Üniversite'den mezun olduktan sonra bilgisayar yazılımı alanında çeşitli firmalarda analist programcı ve proje lideri olarak çalıştı. 1998 yılında bir ortağı ile birlikte bilgisayar firması kurdu ve firmada genel müdürlük görevini yürüttü. 2004 yılında firmadaki ortaklık görevini bıraktıktan sonra sırasıyla Kocaeli doğalgaz firması İZGAZ'da Bilgi İşlem Müdürlüğü, Botaş International Ltd. firmasında Genel Müdüre bağlı olarak Finans Direktörlüğü ve Üsküdar Belediyesi'nde Bilgi İşlem Müdürlüğü görevlerini yürüttü. Evli ve 1 erkek 2 kız çocuğu vardır.