



**DALGACIK DÖNÜŞÜMÜ VE DESTEK  
VEKTÖR MAKİNELERİ İLE TAHMİN:  
BİST ÜZERİNE BİR UYGULAMA**

Muhammet Yasir ALTINTOP

Doktora Tezi

Danışman: Doç. Dr. Nezih TAYYAR

Uşak

Mart, 2016

DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE TAHMİN:

BİST ÜZERİNE BİR UYGULAMA

Muhammet Yasir ALTINTOP

DOKTORA TEZİ

İşletme Anabilim Dalı İşletme Bölümü

Danışman: Doç. Dr. Nezih TAYYAR

Uşak

Uşak Üniversitesi Sosyal Bilimler Enstitüsü

Mart, 2016

## DOKTORA TEZ ÖZETİ

### DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE TAHMİN: BIST ÜZERİNE BİR UYGULAMA

Muhammet Yasir ALTINTOP

İşletme Anabilim Dalı

Uşak Üniversitesi Sosyal Bilimler Enstitüsü, Mart 2016

Danışman: Doç. Dr. Nezh TAYYAR

Temelde sinyal analizinde kullanılan bir yöntem olan dalgacık dönüşümü ile veri madenciliği yöntemlerinden bir yapay öğrenme algoritması olan destek vektör makineleri yöntemlerinin birlikte finansal zaman serileri üzerinde kullanılmasının amaçlandığı bu çalışmada; teknik göstergeler, yurtdışı borsa endeksleri, makroekonomik göstergeler ve haftanın günleri gibi bağımsız değişkenler kullanılarak BIST 100 endeksinin yönü tahmin edilmektedir. Analizde, 15 Ekim 1997 tarihi ile 30 Aralık 2014 tarihleri arasındaki borsanın açık olduğu işgünlerini kapsayan 4296 günlük gözlem veri seti kullanılarak günlük, haftalık ve aylık veri setleri oluşturulmuştur.

Çalışmada günlük, haftalık ve aylık periyotlarda değişken seçme yöntemleri (temel bileşen analizi, k-en yakın komşuluk, korelasyon, lojistik regresyon, kısmi en küçük kareler, rastgele orman algoritması) yardımıyla çeşitli modeller oluşturulmuş, ayrıca tüm değişkenlerin yer aldığı modeller de oluşturulmuştur.

Çalışmanın amacı doğrultusunda; geliştirilen ham ve dalgacık modellerin borsa endeksi üzerindeki tahmin performansları ölçülerek dalgacık dönüşümünün destek vektör makinelerinin tahmin performansına etkisi incelenmiş ve dalgacık dönüşümünün destek vektör makinelerinin tahmin performansını belirgin şekilde artırdığı saptanmıştır.

Elde edilen analiz sonuçlarına göre; BIST 100 endeksinin yönünü en iyi tahmin eden modelin günlük veri setindeki tüm teknik gösterge, yurtdışı borsa endeksleri, makroekonomik göstergeler ve gün değişkenlerinden oluşan Dalgacık TUM55 modeli (%90,84) olduğu, bu modelin ham veri setindeki (%56,90) tahmin oranını yaklaşık %60 artırdığı tespit edilmiştir. Haftalık modellerde ise en yüksek tahmin performansı, lojistik regresyon ile seçilen H-LR<,20 ve H-LR<,05 dalgacık modelleriyle (%84,62) elde edilmiş; aylık modellerde en yüksek tahmin performansı ise k-en yakın komşuluk ile seçilen A-KNN20 ve lojistik regresyon ile seçilen A-LR<,10 dalgacık modellerinde (%74,36) gerçekleşmiştir. Ayrıca elde edilen modellerin tahmin performansları değerlendirildiğinde; günlük veri setinde değişken seçimi ile elde edilen değişkenler kullanılarak haftalık ve aylık modellerin oluşturulabileceği, haftalık ve aylık veri setlerinde ayrıca değişken seçimi yapılmasına gerek olmadığı sonucu elde edilmiştir. Diğer bir önemli sonuç ise, dalgacık dönüşümünün yön tahmininde ham veri setinde karşılaşılan azalış yönündeki sınıflandırma hatalarını azaltarak, artış ve azalışların birbirine yakın oranlarda sınıflandırılmasını sağlamasıdır.

**Anahtar kelimeler:** *Dalgacık Dönüşümü, Ayrık Dalgacık Dönüşümü, Destek Vektör Makineleri, Endeks Yön Tahmini.*



## **ABSTRACT**

### **FORECASTING WITH WAVELET TRANSFORM AND SUPPORT VECTOR MACHINES: AN APPLICATION ON BIST**

**Muhammet Yasir ALTINTOP**

**Department of Business Administration**

**Uşak University Social Sciences Institutes, March 2016**

**Advisor: Assoc. Prof. Dr. Nezir TAYYAR**

In this study where it is aimed to use wavelet transform, a method used essentially in signal analysis, and support vector machines method, a machine learning algorithm of data mining, together on financial time series. BIST 100 index direction is estimated by using independent variables such as technical indicators, overseas stock indices, macroeconomic indicators and days of week. In the analysis daily, weekly and monthly data sets have been established by using observation data set for 4296 days including the weekdays when the stock market is open between 15th October 1997 and 30th December 2014.

In the study, several models have been established via variable selection methods (principal component analysis, k nearest neighbors, correlation, logistic regression, partial least squares, random forest algorithm) in daily, weekly and monthly period and models including all variables have also been established.

In accordance with the aim of the study, the effect of wavelet transform on estimation performance of support vector machines has been analyzed by measuring estimation performance of the developed raw and wavelet models on stock market index and it has been identified that wavelet transform significantly increases the estimation performance of support vector machines.

According to the obtained analysis results, it has been identified that the best estimating model of BIST 100 index direction is Wavelet TUM55 model (90,84%) which is formed by all technical indicators, overseas stock indices, macroeconomic indicators and day variables in daily data set and this model increases the estimation rate in raw data set (56,90%) about 60 %. However, the highest estimation performance in weekly models has been provided by H-LR<, 20 and H-LR <, 05 wavelet models selected by logistic regression, on the other hand the highest estimation performance in monthly models has occurred in A-KNN20 model selected by k nearest neighbors and A-LR <,10 wavelet model (74,36 %) selected by logistic regression. In addition, when the estimation performance of the obtained models is considered, it has been found that weekly and monthly models can be established by using the variables obtained by the variable selection in daily data set and there is no need for extra variable selection in weekly and monthly data set. Another significant result is also the fact that by reducing the classification errors of decreases encountered in raw data set in direction estimation of wavelet transform, it provides the classification of increases and decreases in almost equal proportions.

**Keywords:** *Wavelet, Discrete Wavelet Transform, Support Vector Machines, Index Direction Forecast.*



UŞAK ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ

JÜRİ VE ENSTİTÜ ONAYI

İşletme Ana Bilim Dalı Doktora Programı 124005118 No'lu öğrencisi M.Yasir ALTINTOP'un "Dalgacık Dönüşümü Ve Destek Vektör Makineleri İle Tahmin: BIST Üzerine Bir Uygulama" adlı tezi 21 / 03 / 2016 tarihinde, aşağıdaki jüri tarafından Uşak Üniversitesi Sosyal Bilimler Enstitüsü Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca, Doktora tezi olarak değerlendirilerek kabul edilmiştir.

Jüri	Adı Soyadı	İmza
Danışman	: Doç. Dr. Nezh TAYYAR	
Üye	: Prof. Dr. Suat ŞAHİNLER	
Üye	: Doç. Dr. A. Şahin GÖRMÜŞ	
Üye	: Prof. Dr. Ali ELEREN	
Üye	: Yrd.Doç.Dr.Murat GÜNDÜZ	

Doç. Dr. Asım ARI  
Enstitü Müdürü

## ÖNSÖZ

Bu çalışmanın ortaya çıkması sürecinde bilgi, tecrübe ve bana destekleriyle yanımda olan ve tezimde emeği geçen danışmanım Doç. Dr. Nezih TAYYAR hocama; araştırmamın her safhasında fikirleri ile çalışmalarına yön veren ve yardımlarını esirgemeyen saygıdeğer Prof. Dr. Suat ŞAHİNLER ve Doç. Dr. A. Şahin GÖRMÜŞ hocalarıma; manevi desteklerinden dolayı aileme ve desteklerini esirgemeyen değerli dostlarıma teşekkürü bir borç bilirim.

Ayrıca; daima her konuda yanımda olup, sabırla destekleyen ve bana katlanan sevgili eşim ve sevgili oğluma sonsuz teşekkür ederim.

Muhammet Yasir Altıntop

## ÖZGEÇMİŞ

### Kişisel Bilgiler

Adı Soyadı : Muhammet Yasir ALTINTOP

Doğum Yeri ve Tarihi : Uşak - 1979

Lisans : Uludağ Üniversitesi Fen Edebiyat Fakültesi

Yüksek Lisans : Uşak Üniversitesi Sosyal Bilimler Enstitüsü

Bildiği Yabancı Diller : İngilizce

### İş Deneyimi

Çalıştığı Kurumlar : Celal Bayar Üniversitesi (2012 - ... )

Milli Eğitim Bakanlığı (2004 - 2012)

### İletişim

E-posta Adresi : yasir.altintop@cbu.edu.tr

## İÇİNDEKİLER

DOKTORA TEZ ÖZETİ .....	iii
ABSTRACT .....	v
JÜRİ VE ENSTİTÜ ONAYI .....	vii
ÖNSÖZ .....	viii
ÖZGEÇMİŞ .....	ix
İÇİNDEKİLER .....	x
TABLolar LİSTESİ .....	xiii
ŞEKİLLER LİSTESİ .....	xix
KISALTMALAR .....	xxii
GİRİŞ .....	1
1. BÖLÜM: DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE HİSSE SENEDİ PİYASALARINDA TAHMİN ÇALIŞMALARI .....	4
1.1. DALGACIK DÖNÜŞÜMÜ KULLANILARAK YAPILAN ÇALIŞMALAR .....	4
1.2. DESTEK VEKTÖR MAKİNELERİ KULLANILARAK YAPILAN YÖN TAHMİNİ ÇALIŞMALARI .....	9
1.3. DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE TAHMİN ÇALIŞMALARI .....	12
1.4. DİĞER ENDEKS TAHMİN ÇALIŞMALARI (XU100) .....	14
2. BÖLÜM: DALGACIK DÖNÜŞÜMÜ, DESTEK VEKTÖR MAKİNELERİ VE ÖZNİTELİK SEÇİMİNDE KULLANILAN YÖNTEMLER .....	19
2.1. DALGACIK DÖNÜŞÜMÜ YÖNTEMİ .....	19
2.1.1. Dalgacık Analizi .....	19
2.1.2. Fourier Dönüşümü (FD) .....	22
2.1.3. Kısa Zaman Fourier Dönüşümü (KZFD) .....	30
2.1.4. Dalgacık Dönüşümü (DD) .....	33
2.1.4.1. Sürekli Dalgacık Dönüşümü (SDD) .....	36
2.1.4.2. Dalgacıkların Ölçeklenmesi ve Kaydırılması .....	38
2.1.4.3. Ayrık Dalgacık Dönüşümü (ADD) .....	39
2.1.4.4. Dalgacık Serileri .....	47
2.1.4.5. Dalgacık Çeşitleri .....	48
2.1.4.5.1. Haar Dalgacığı .....	49
2.1.4.5.2. Daubechies Dalgacıkları .....	50
2.2. DESTEK VEKTÖR MAKİNELERİ .....	52
2.2.1. İstatistiksel Öğrenme Teorisi .....	54
2.2.1.1. VC-Boyutu .....	55
2.2.1.2. Ampirik Risk Minimizasyonu .....	56
2.2.1.3. Yapısal Risk Minimizasyonu .....	57
2.2.2. Optimal Ayırıcı Hiperdüzlem .....	59
2.2.2.1. Doğrusal Olarak Ayrılabilir Örüntüler için Optimal Hiperdüzlem .....	59
2.2.2.1.1. Kanonik Hiperdüzlemlerin Yapısı .....	61
2.2.2.1.2. Optimal Hiperdüzlem Oluşturma .....	61
2.2.2.2. Doğrusal Olarak Ayrılamayan Örüntüler için Optimal Hiperdüzlem .....	66
2.2.3. Doğrusal Olmayan Destek Vektör Makineleri .....	70

2.2.3.1.	Yüksek Boyutlu Bir Uzaya Eşleme (Çekirdek Hilesi) .....	70
2.2.3.2.	Mercer Teoremi .....	73
2.2.3.3.	Çekirdek Fonksiyonları .....	73
2.2.3.3.1.	Doğrusal Çekirdek Fonksiyonu .....	73
2.2.3.3.2.	Polinom Çekirdek Fonksiyonu .....	74
2.2.3.3.3.	Radyal Tabanlı Çekirdek Fonksiyonu .....	74
2.2.3.3.4.	Sigmoid Çekirdek Fonksiyonu .....	74
2.3.	DEĞİŞKEN (ÖZNİTELİK) SEÇİMİNDE KULLANILAN YÖNTEMLER .....	74
2.3.1.	Temel Bileşen Analizi (PCA) .....	75
2.3.2.	K-En Yakın Komşuluk (KNN) .....	76
2.3.3.	Korelasyon .....	77
2.3.4.	Lojistik Regresyon (LR) .....	78
2.3.5.	Kısmi En Küçük Kareler (PLS) .....	79
2.3.6.	Rastgele Orman Algoritması (Randomforest) .....	79
3.	BÖLÜM: BORSA ENDEKSLERİ VE TEKNİK GÖSTERGELER .....	82
3.1.	BORSA İSTANBUL (BIST) VE ENDEKSLER .....	82
3.1.1.	Borsa İstanbul (BIST) .....	82
3.1.2.	BIST Endeksleri .....	82
3.2.	TEKNİK GÖSTERGELER .....	84
3.2.1.	Open-High-Low-Close-Volume .....	84
3.2.2.	Aroon Osilatörü (Aroon Oscillator - AO) .....	84
3.2.3.	Bollinger Bantları (Bollinger Bands - BB) .....	86
3.2.4.	Chaikin Osilatörü (Chaikin Oscillator - CO) .....	87
3.2.5.	Chaikin'in Para Akım Göstergesi (Chaikin's Money Flow - CMF) .....	87
3.2.6.	Chaikin'in Dalgalanma Göstergesi (Chaikin's Volatility - CV) .....	88
3.2.7.	Chande Momentum Osilatörü (Chande Momentum Oscillator - CMO) .....	88
3.2.8.	Denge İşlem Hacmi Göstergesi (On Balance Volume - OBV) .....	88
3.2.9.	Göreceli Güç Endeksi (Relative Strength Index - RSI) .....	89
3.2.10.	Hareketli Ortalamalar (Moving Averages - MA) .....	89
3.2.10.1.	Basit Hareketli Ortalama (Simple Moving Average - SMA) .....	90
3.2.10.2.	Üssel Hareketli Ortalama (Exponential Moving Average - EMA) .....	90
3.2.10.3.	Ağırlıklı Hareketli Ortalama (Weighted Moving Average - WMA) .....	91
3.2.11.	Hareketli Ortalamalar Yakınlaşma - Uzaklaşma Göstergesi (Moving Average Convergence Divergence - MACD) .....	91
3.2.12.	Mal Kanal Endeksi (Commodity Channel Index - CCI) .....	92
3.2.13.	Momentum Göstergesi (Momentum Indicator - MI) .....	92
3.2.14.	Ortalama Doğruluk Aralığı (Average True Range - ATR) .....	93
3.2.15.	Para Akışı Endeksi (Money Flow Index - MFI) .....	94
3.2.16.	Parabolik SAR (Parabolic SAR) .....	94
3.2.17.	Stokastik Osilatörü (Stochastic Oscillator - SO) .....	95
3.2.18.	Stokastik Momentum Endeksi (Stochastic Momentum Index - SMI) .....	96
3.2.19.	Williams'in %R Göstergesi (Williams' %R) .....	96
3.2.20.	Yönel Hareket Göstergesi (Directional Movement Indicator - DMI) .....	97

4. BÖLÜM: DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE TAHMİN: BIST ÜZERİNE BİR UYGULAMA.....	99
4.1. ÇALIŞMANIN AMACI VE KAPSAMI.....	99
4.2. UYGULAMANIN MODELİ.....	99
4.3. VERİLERİN TOPLANMASI VE HAZIRLANMASI.....	101
4.4. UYGULAMA AŞAMALARININ BELİRLENMESİ.....	105
4.4.1. Kullanılan Yazılım ve Paketler.....	106
4.4.2. Model Parametrelerinin Belirlenmesi ve Çapraz Geçerleme.....	107
4.4.3. Sınıflandırıcıların Performans Değerlendirmesi (Sınıflandırma Matrisi).....	109
4.4.4. Uygulama Aşamalarının Şema İle Gösterimi.....	110
4.5. ANALİZDE KULLANILAN DEĞİŞKENLER.....	112
4.5.1. Makroekonomik Göstergeler.....	112
4.5.2. Teknik Gösterge Değişkenleri.....	112
4.5.3. Yurtdışı Borsa Endeks Verileri.....	114
4.5.4. Gün Değişkenleri.....	114
4.5.5. Bağımlı Değişken ve Hesaplanması.....	115
4.6. GÜNLÜK VERİLERDE DEĞİŞKEN SEÇİMİ.....	115
4.6.1. Günlük Veriler ile Endeks Yönü Tahmini.....	119
4.6.2. Haftalık Veriler ile Endeks Yönü Tahmini.....	138
4.6.3. Aylık Veriler ile Endeks Yönü Tahmini.....	154
4.7. HAFTALIK VERİLERDE DEĞİŞKEN SEÇİMİ VE ENDEKS YÖNÜ TAHMİNİ.....	170
4.8. AYLIK VERİLERDE DEĞİŞKEN SEÇİMİ VE ENDEKS YÖNÜ TAHMİNİ.....	183
5. BÖLÜM: TARTIŞMALAR VE YORUMLAR.....	200
6. BÖLÜM: SONUÇ VE ÖNERİLER.....	213
EKLER.....	218
KAYNAKÇA.....	221



## TABLolar LİSTESİ

Tablo 1 : Borsa İstanbul Hisse Senedi Endeksleri .....	83
Tablo 2 : BIST 100 Endeksi Artış ve Azalış Gözlem Sayıları.....	103
Tablo 3 : Eğitim ve Test Verileri Artış ve Azalış Gözlem Sayıları .....	104
Tablo 4 : Sınıflandırma Matrisi.....	109
Tablo 5 : Makroekonomik Değişkenler ile İlgili Temel İstatistikler .....	112
Tablo 6 : Teknik Gösterge Değişkenleri ile İlgili Temel İstatistikler .....	113
Tablo 7 : Endeks Değişkenleri ve İlgili Temel İstatistikler .....	114
Tablo 8 : Gün Değişkenleri ile İlgili Temel İstatistikler .....	114
Tablo 9 : Günlük Veri Setinde Temel Bileşen Analizi .....	116
Tablo 10 : LR ile Anlamlılık Düzeyleri (Günlük) .....	118
Tablo 11 : Sınıflandırma Matrisi (Günlük TBA6 Modeli) .....	120
Tablo 12 : Sınıflandırma Matrisi (Günlük Dalgacık TBA6 Modeli).....	121
Tablo 13 : Sınıflandırma Matrisi (Günlük KOR<,50 Modeli).....	121
Tablo 14 : Sınıflandırma Matrisi (Günlük Dalgacık KOR<,50 Modeli) .....	122
Tablo 15 : Sınıflandırma Matrisi (Günlük KOR<,80 Modeli).....	123
Tablo 16 : Sınıflandırma Matrisi (Günlük Dalgacık KOR<,80 Modeli) .....	123
Tablo 17 : Sınıflandırma Matrisi (Günlük KOR>,80 Modeli).....	124
Tablo 18 : Sınıflandırma Matrisi (Günlük Dalgacık KOR>,80 Modeli) .....	125
Tablo 19 : Sınıflandırma Matrisi (Günlük KNN7 Modeli).....	125
Tablo 20 : Sınıflandırma Matrisi (Günlük Dalgacık KNN7 Modeli) .....	126
Tablo 21 : Sınıflandırma Matrisi (Günlük KNN20 Modeli).....	127
Tablo 22 : Sınıflandırma Matrisi (Günlük Dalgacık KNN20 Modeli) .....	127
Tablo 23 : Sınıflandırma Matrisi (Günlük PLS17 Modeli).....	128

Tablo 24 : Sınıflandırma Matrisi (Günlük Dalgacık PLS17 Modeli) .....	129
Tablo 25 : Sınıflandırma Matrisi (Günlük PLS20 Modeli).....	129
Tablo 26 : Sınıflandırma Matrisi (Günlük Dalgacık PLS20 Modeli) .....	130
Tablo 27 : Sınıflandırma Matrisi (Günlük RF6 Modeli).....	131
Tablo 28 : Sınıflandırma Matrisi (Günlük Dalgacık RF6 Modeli) .....	131
Tablo 29 : Sınıflandırma Matrisi (Günlük LR<,20 Modeli) .....	132
Tablo 30 : Sınıflandırma Matrisi (Günlük Dalgacık LR<,20 Modeli).....	133
Tablo 31 : Sınıflandırma Matrisi (Günlük LR<,10 Modeli) .....	133
Tablo 32 : Sınıflandırma Matrisi (Günlük Dalgacık LR<,10 Modeli).....	134
Tablo 33 : Sınıflandırma Matrisi (Günlük LR<,05 Modeli) .....	135
Tablo 34 : Sınıflandırma Matrisi (Günlük Dalgacık LR<,05 Modeli).....	135
Tablo 35 : Sınıflandırma Matrisi (Günlük LR<,01 Modeli) .....	136
Tablo 36 : Sınıflandırma Matrisi (Günlük Dalgacık LR<,01 Modeli).....	137
Tablo 37 : Sınıflandırma Matrisi (Günlük TÛM55 Modeli).....	137
Tablo 38 : Sınıflandırma Matrisi (Günlük Dalgacık TÛM55 Modeli) .....	138
Tablo 39 : Sınıflandırma Matrisi (H-TBA6 Modeli) .....	139
Tablo 40 : Sınıflandırma Matrisi (Dalgacık H-TBA6 Modeli).....	139
Tablo 41 : Sınıflandırma Matrisi (H-KOR>,80 Modeli).....	140
Tablo 42 : Sınıflandırma Matrisi (Dalgacık H-KOR>,80 Modeli) .....	141
Tablo 43 : Sınıflandırma Matrisi (H-KNN7 Modeli).....	141
Tablo 44 : Sınıflandırma Matrisi (Dalgacık H-KNN7 Modeli) .....	142
Tablo 45 : Sınıflandırma Matrisi (H-KNN20 Modeli).....	143
Tablo 46 : Sınıflandırma Matrisi (Dalgacık H-KNN20 Modeli) .....	143
Tablo 47 : Sınıflandırma Matrisi (H-PLS17 Modeli) .....	144

Tablo 48 : Sınıflandırma Matrisi (Dalgacık H-PLS17 Modeli).....	145
Tablo 49 : Sınıflandırma Matrisi (H-PLS20 Modeli) .....	145
Tablo 50 : Sınıflandırma Matrisi (Dalgacık H-PLS20 Modeli).....	146
Tablo 51 : Sınıflandırma Matrisi (H-RF6 Modeli) .....	147
Tablo 52 : Sınıflandırma Matrisi (Dalgacık H-RF6 Modeli).....	147
Tablo 53 : Sınıflandırma Matrisi (H-LR<,20 Modeli).....	148
Tablo 54 : Sınıflandırma Matrisi (Dalgacık H-LR<,20 Modeli) .....	149
Tablo 55 : Sınıflandırma Matrisi (H-LR<,10 Modeli).....	149
Tablo 56 : Sınıflandırma Matrisi (Dalgacık H-LR<,10 Modeli) .....	150
Tablo 57 : Sınıflandırma Matrisi (H-LR<,05 Modeli).....	151
Tablo 58 : Sınıflandırma Matrisi (Dalgacık H-LR<,05 Modeli) .....	151
Tablo 59 : Sınıflandırma Matrisi (H-LR<,01 Modeli).....	152
Tablo 60 : Sınıflandırma Matrisi (Dalgacık H-LR<,01 Modeli) .....	153
Tablo 61 : Sınıflandırma Matrisi (H-TÜM55 Modeli) .....	153
Tablo 62 : Sınıflandırma Matrisi (Dalgacık H-TÜM55 Modeli).....	154
Tablo 63 : Sınıflandırma Matrisi (A-TBA6 Modeli) .....	155
Tablo 64 : Sınıflandırma Matrisi (Dalgacık A-TBA6 Modeli).....	155
Tablo 65 : Sınıflandırma Matrisi (A-KOR>,80 Modeli).....	156
Tablo 66 : Sınıflandırma Matrisi (Dalgacık A-KOR>,80 Modeli) .....	157
Tablo 67 : Sınıflandırma Matrisi (A-KNN7 Modeli).....	157
Tablo 68 : Sınıflandırma Matrisi (Dalgacık A-KNN7 Modeli) .....	158
Tablo 69 : Sınıflandırma Matrisi (A-KNN20 Modeli).....	159
Tablo 70 : Sınıflandırma Matrisi (Dalgacık A-KNN20 Modeli) .....	159
Tablo 71 : Sınıflandırma Matrisi (A-PLS17 Modeli) .....	160

Tablo 72 : Sınıflandırma Matrisi (Dalgacık A-PLS17 Modeli).....	161
Tablo 73 : Sınıflandırma Matrisi (A-PLS20 Modeli) .....	161
Tablo 74 : Sınıflandırma Matrisi (Dalgacık A-PLS20 Modeli).....	162
Tablo 75 : Sınıflandırma Matrisi (A-RF6 Modeli) .....	163
Tablo 76 : Sınıflandırma Matrisi (Dalgacık A-RF6 Modeli).....	163
Tablo 77 : Sınıflandırma Matrisi (A-LR<,20 Modeli).....	164
Tablo 78 : Sınıflandırma Matrisi (Dalgacık A-LR<,20 Modeli) .....	165
Tablo 79 : Sınıflandırma Matrisi (A-LR<,10 Modeli).....	165
Tablo 80 : Sınıflandırma Matrisi (Dalgacık A-LR<,10 Modeli) .....	166
Tablo 81 : Sınıflandırma Matrisi (A-LR<,05 Modeli).....	167
Tablo 82 : Sınıflandırma Matrisi (Dalgacık A-LR<,05 Modeli) .....	167
Tablo 83 : Sınıflandırma Matrisi (A-LR<,01 Modeli).....	168
Tablo 84 : Sınıflandırma Matrisi (Dalgacık A-LR<,01 Modeli) .....	169
Tablo 85 : Sınıflandırma Matrisi (A-TÜM55 Modeli) .....	169
Tablo 86 : Sınıflandırma Matrisi (Dalgacık A-TÜM55 Modeli).....	170
Tablo 87 : Haftalık Veri Setinde Temel Bileşen Analizi .....	171
Tablo 88 : LR ile Anlamlılık Düzeyleri (Haftalık) .....	173
Tablo 89 : Sınıflandırma Matrisi (Haftalık TBA6 Modeli) .....	174
Tablo 90 : Sınıflandırma Matrisi (Haftalık Dalgacık TBA6 Modeli).....	175
Tablo 91 : Sınıflandırma Matrisi (Haftalık KNN20 Modeli).....	176
Tablo 92 : Sınıflandırma Matrisi (Haftalık Dalgacık KNN20 Modeli) .....	176
Tablo 93 : Sınıflandırma Matrisi (Haftalık PLS20 Modeli).....	177
Tablo 94 : Sınıflandırma Matrisi (Haftalık Dalgacık PLS20 Modeli).....	178
Tablo 95 : Sınıflandırma Matrisi (Haftalık RF6 Modeli) .....	178

Tablo 96 : Sınıflandırma Matrisi (Haftalık Dalgacık RF6 Modeli).....	179
Tablo 97 : Sınıflandırma Matrisi (Haftalık LR<,20 Modeli).....	180
Tablo 98 : Sınıflandırma Matrisi (Haftalık Dalgacık LR<,20 Modeli).....	180
Tablo 99 : Sınıflandırma Matrisi (Haftalık LR<,10 Modeli).....	181
Tablo 100 : Sınıflandırma Matrisi (Haftalık Dalgacık LR<,10 Modeli).....	182
Tablo 101 : Sınıflandırma Matrisi (Haftalık LR<,05 Modeli).....	182
Tablo 102 : Sınıflandırma Matrisi (Haftalık Dalgacık LR<,05 Modeli).....	183
Tablo 103 : Aylık Veri Setinde Temel Bileşen Analizi.....	184
Tablo 104 : LR ile Anlamlılık Düzeyleri (Aylık).....	186
Tablo 105 : Sınıflandırma Matrisi (Aylık TBA5 Modeli).....	187
Tablo 106 : Sınıflandırma Matrisi (Aylık Dalgacık TBA5 Modeli).....	188
Tablo 107 : Sınıflandırma Matrisi (Aylık KOR>,80 Modeli).....	189
Tablo 108 : Sınıflandırma Matrisi (Aylık Dalgacık KOR>,80 Modeli).....	189
Tablo 109 : Sınıflandırma Matrisi (Aylık KNN20 Modeli).....	190
Tablo 110 : Sınıflandırma Matrisi (Aylık Dalgacık KNN20 Modeli).....	191
Tablo 111 : Sınıflandırma Matrisi (Aylık PLS20 Modeli).....	191
Tablo 112 : Sınıflandırma Matrisi (Aylık Dalgacık PLS20 Modeli).....	192
Tablo 113 : Sınıflandırma Matrisi (Aylık RF7 Modeli).....	193
Tablo 114 : Sınıflandırma Matrisi (Aylık Dalgacık RF7 Modeli).....	193
Tablo 115 : Sınıflandırma Matrisi (Aylık LR<,20 Modeli).....	194
Tablo 116 : Sınıflandırma Matrisi (Aylık Dalgacık LR<,20 Modeli).....	195
Tablo 117 : Sınıflandırma Matrisi (Aylık LR<,10 Modeli).....	195
Tablo 118 : Sınıflandırma Matrisi (Aylık Dalgacık LR<,10 Modeli).....	196
Tablo 119 : Sınıflandırma Matrisi (Aylık LR<,05 Modeli).....	197

Tablo 120 : Sınıflandırma Matrisi (Aylık Dalgacık LR<,05 Modeli).....	197
Tablo 121 : Sınıflandırma Matrisi (Aylık LR<,01 Modeli).....	198
Tablo 122 : Sınıflandırma Matrisi (Aylık Dalgacık LR<,01 Modeli).....	199
Tablo 123 : Günlük Modeller Artış ve Azalış Sınıflandırma Oranları .....	200
Tablo 124 : Haftalık Modeller Artış ve Azalış Sınıflandırma Oranları .....	201
Tablo 125 : Aylık Modeller Artış ve Azalış Sınıflandırma Oranları .....	202
Tablo 126 : Değişkenlerin Modellerde Yer Alma Sıklıkları.....	208
Tablo 127 : Modellerdeki Değişken Sayıları .....	210
Tablo 128 : XU100 Endeksi Yön Tahmini Çalışmaları.....	211

## ŞEKİLLER LİSTESİ

Şekil 1 : Zaman, Frekans ve Dalgacık Uzayı Gösterimi.....	20
Şekil 2 : Durağan ve Durağan Olmayan Fonksiyonların Zaman ve Frekans Uzayı Gösterimi.....	23
Şekil 3 : Örnek Fonksiyon Grafiği.....	25
Şekil 4 : Fourier Serisi Terimlerinin Toplam Grafiği.....	27
Şekil 5 : Fourier Katsayıları Grafiği.....	29
Şekil 6 : $ft = cost + \cos 10t + \cos (100t)$ Fonksiyonunun Fourier Dönüşümü. ...	29
Şekil 7 : Birbirinden Farklı İki Fonksiyonun Fourier Dönüşümü.....	30
Şekil 8 : Kısa Zaman Fourier Dönüşümü.....	31
Şekil 9 : Kısa Zaman Fourier Dönüşümünün Sinyali Taraması .....	33
Şekil 10 : Sonsuz Uzunluğa Sahip Kosinüs Dalgasının Bir Parçası ile Sonlu Uzunluğa Sahip Dalgacık.....	34
Şekil 11 : Gerilmiş (Ölçeklendirilmiş) Sinüsoid ve Dalgacık.....	34
Şekil 12 : Dalgacık Dönüşümü Yardımıyla Sinyalin Ayırıştırılması.....	35
Şekil 13 : Orjinal Sinyal ile Dalgacığın SDD Tipi Karşılaştırılması .....	36
Şekil 14 : SDD'de Sinyal ve Süreksizliğin Dalgacık Gösterimi .....	38
Şekil 15 : Zaman-Frekans Düzleminin Kritik Örnekleme .....	40
Şekil 16 : Daubechies 12 Dalgacık ve Ölçekleme Fonksiyonları ile Alçak Geçiren ve Yüksek Geçiren Filtreleri.....	44
Şekil 17 : Ayrık Dalgacık Dönüşümünün Dört Seviyeli Ayırıştırılması.....	45
Şekil 18 : 2014 Yılı BIST 100 Endeksi Günlük Kapanış Değerlerinin Haar ve Daubechies 4 Dalgacıkları Yardımıyla Bileşenlerine Ayırıştırılması.....	46
Şekil 19 : Yaygın Kullanılan Bazı Dalgacık Çeşitleri.....	48
Şekil 20 : Haar Ortalama Dönüşümünün Grafik Gösterimi.....	50
Şekil 21 : Haar Ortalama Dönüşümünün Yaklaşım ve Detay Katsayıları.....	50

Şekil 22 : Daubechies 4 Dalgacıđı için Zaman Uzayında Filtre Katsayılarıyla Gösterimi.....	51
Şekil 23 : Daubechies 4 Dalgacıđı ve Ölçekleme Fonksiyonu.....	51
Şekil 24 : Daubechies Ana Dalgacıđının Fraktal Öz Benzerliđi.....	52
Şekil 25 : Model Seçimi.....	55
Şekil 26 : $R^2$ 'de Üç Noktanın Doğru ile Ayrılması Durumları .....	58
Şekil 27 : Doğrusal Ayrılabilir Örüntüler için Optimal Hiperdüzlem ve Destek Vektörleri .....	61
Şekil 28 : İki Boyutlu Uzayda Doğrusal Olarak Ayrılamayan Durum .....	67
Şekil 29 : Doğrusal Olmayan Eşleme .....	71
Şekil 30 : Çekirdek Hilesi.....	72
Şekil 31 : 2 Boyutta Temel Bileşen Analizi.....	76
Şekil 32 : Rastgele Orman Genel Yapısı .....	80
Şekil 33 : BIST 100 Endeksi Günlük Kapanış Grafiđi .....	102
Şekil 34 : BIST 100 Endeksi Haftalık ve Aylık Kapanış Grafikleri.....	103
Şekil 35 : db4 Dalgacıđının Bileşenleri .....	105
Şekil 36 : 5-Kat Çapraz Geçerleme.....	108
Şekil 37 : Uygulama Aşamaları .....	111
Şekil 38 : Günlük Veri setinde KNN .....	117
Şekil 39 : Günlük Veri setinde PLS .....	119
Şekil 40 : Haftalık Veri setinde KNN .....	172
Şekil 41 : Haftalık Veri setinde PLS.....	173
Şekil 42 : Aylık Veri setinde KNN .....	185
Şekil 43 : Aylık Veri setinde PLS.....	186
Şekil 44 : Günlük Modeller Doğru Sınıflandırma Oranları .....	204



Şekil 45 : G-Haftalık Modeller Doğru Sınıflandırma Oranları.....	204
Şekil 46 : G-Aylık Modeller Doğru Sınıflandırma Oranları.....	205
Şekil 47 : Haftalık Modeller Doğru Sınıflandırma Oranları .....	206
Şekil 48 : Aylık Modeller Doğru Sınıflandırma Oranları .....	207



## KISALTMALAR

ADD	: Ayrik Dalgacık Dönüşümü (Discrete Wavelet Transform)
AO	: Aroon Oscillator (Aroon Osilatörü)
ARIMA	: Auto-Regressive Integrated Moving Average (Oto regresif Hareketli Ortalama)
ARM	: Ampirik Risk Minimizasyonu (Empirical Risk Minimization)
ATR	: Average True Range (Ortalama Doğruluk Aralığı)
BB	: Bollinger Bantları (Bollinger Bands)
BIST	: Borsa İstanbul
BOVESPA	: São Paulo Stock Exchange - Brezilya Endeksi
CAC	: Fransa Borsası Endeksi
CCI	: Commodity Channel Index (Mal Kanal Endeksi)
CMF	: Chaikin's Money Flow (Chaikin'in Para Akım Göstergesi)
CMO	: Chande Momentum Oscillator (Chande Momentum Osilatörü)
CO	: Chaikin Oscillator (Chaikin Osilatörü)
CV	: Chaikin's Volatility (Chaikin'in Dalgalanma Göstergesi)
DAX	: Deutscher Aktien Index - Almanya Endeksi
DD	: Dalgacık Dönüşümü (Wavelet Transform)
DJIA	: Dow Jones Industrial Average - ABD Endeksi
DMI	: Directional Movement Indicator (Yönel Hareket Göstergesi)
DVM	: Destek Vektör Makineleri (Support Vector Machines)
EGARCH	: Exponential Generalized AutoRegressive Conditional Heteroskedasticity (Üssel Genelleştirilmiş Oto regresif Koşullu Değişen Varyans)
EMA	: Exponential Moving Average (Üssel Hareketli Ortalama)
FD	: Fourier Dönüşümü (Fourier Transform)
FTSE	: Financial Times Stock Exchange Index - İngiltere Endeksi
GA	: Genetik Algoritma (Genetic Algorithm)
GARCH	: Generalized AutoRegressive Conditional Heteroskedasticity (Oto regresif Koşullu Değişen Varyans)

GS	: Grid Search (Izgara Modeli)
GSMH	: Gayri Safi Milli Hasıla
HSI	: Hang Seng Index - Hong Kong Endeksi
İMKB	: İstanbul Menkul Kıymetler Borsası
KKT	: Karush-Kuhn-Tucker
KNN	: K-Nearest Neighbour (K-En Yakın Komşuluk)
KOSPI	: Korean Stock Exchanges - Güney Kore Endeksi
KZFD	: Kısa Zaman Fourier Dönüşümü (Short Time Fourier Transform)
LR	: Logistic Regression (Lojistik Regresyon)
MA	: Moving Averages (Hareketli Ortalamalar)
MACD	: Moving Average Convergence Divergence (Hareketli Ortalamalar Yakınlaşma-Uzaklaşma Göstergesi)
MFI	: Money Flow Index (Para Akışı Endeksi)
MI	: Momentum Indicator (Momentum Göstergesi)
NASDAQ	: National Association of Securities Dealers Automated Quotations - ABD Endeksi
NIKKEI	: Nihon Keizai Shimbun Index - Japonya Endeksi
NYSE	: New York Stock Exchange - ABD Endeksi
OBV	: On Balance Volume (Denge İşlem Hacmi Göstergesi)
ODKA	: Orta Doğu Kuzey Afrika (Middle East North Africa)
PCA	: Principal Component Analysis (Temel Bileşen Analizi)
PLS	: Partial Least Squares (Kısmi En Küçük Kareler)
PSO	: Particle Swarm Optimization (Parçacık Sürü Optimizasyonu)
RSI	: Relative Strength Index (Göreceli Güç Endeksi)
SDD	: Sürekli Dalgacık Dönüşümü (Continuous Wavelet Transform)
SMA	: Simple Moving Average (Basit Hareketli Ortalama)
SMI	: Stochastic Momentum Index (Stokastik Momentum Endeksi)
SO	: Stochastic Oscillator (Stokastik Osilatörü)
SP500	: Standart & Poors 500 Index - ABD Endeksi
SSE	: Shanghai Stock Exchange - Çin
VC	: Vapnik-Chervonenkis
WMA	: Weighted Moving Average (Ağırlıklı Hareketli Ortalama)
YRM	: Yapısal Risk Minimizasyonu (Structural Risk Minimization)
YSA	: Yapay Sinir Ağları (Artificial Neural Networks)

## GİRİŞ

İnsanođlu için geleceđi öngörebilmek ve ona göre pozisyon alıp kendini geleceđin bilinmezliđine karşı hazır ve güvende hissetme arzusu vazgeçilmezdir. Geleceđi öngörebilmenin temel veri kaynađı ise geçmiş tecrübelerden ortaya çıkan bilgidir. Girdileri ve çıktıları bilinen halihazırdaki veri ile gelecekteki durumları tahmin etme, dođa bilimlerinin yanında mühendislik ve sosyal bilimler için de ilgi çeken ve önem arz eden konulardandır. Finans alanında da finansal varlıkların gelecekteki deđerlerinde nasıl bir deđişimin olabileceđini tahmin etmek ve bu öngörülere göre pozisyon almak, öngörülerin dođruluđu oranında pozitif getiri ortaya koyacaktır. Bu tür tahminler için bir insanın zihninde birçok ve derinlikli veriyi aynı anda harmanlayıp bu veri setinden sađlıklı bir öngörü ortaya koyabilmesi çok olanaklı görülmemektedir.

Ancak, teknolojinin ve dolayısıyla bilgisayar destekli sistemlerin gelişmesiyle son yirmi yılda gözle görülür bir ilerleme kat edilen çeşitli veri madenciliđi metotları ile çok büyük veri setleri kullanılarak geçmiş veriler üzerindeki örüntü ve ilişkiler tespit edilebilmektedir. Makine öğrenmesi teknikleriyle sisteme öğretilen bu örüntü ve ilişkiler kullanılarak gelecekte meydana gelebilecek benzer veya benzer olmayan durumlarda ne gibi sonuçlar ortaya çıkabileceđi belirli bir dođruluk oranıyla tahmin edilebilmektedir. Son yıllarda yapılan çalışmalarda bu tahminlerdeki dođruluk oranlarının nasıl artırılabilceđi ve bunun için hangi yöntemlerin nasıl kullanılacağı tartışılmaktadır. Girdi olarak kullanılan verinin niteliđinin yanında, birden fazla yöntemin birlikte kullanıldıđı hibrit sistemlerin tahmin performansı üzerinde olumlu etkisinin olduđu yönünde araştırmacıların fikir birliđi içinde olduđu görülmektedir.

Matematiksel temelli bir dönüşüm olarak dalgacık dönüşümü, sinyal analizinde bir sinyaldeki gürültüyü ve temel yapıyı birbirinden ayırmada kullanılırken, son zamanlarda zaman serilerinde de kullanımının arttıđı

görülmektedir. Finansal zaman serileri de birer sinyal olarak ele alındığında, dalgacık dönüşümü ile serinin kaba yapısı ve serideki detaylar ayrı ayrı analiz edilebilecektir.

Buradan hareketle bu çalışmanın amacı; dalgacık dönüşümü ve makine öğrenmesi yöntemlerinden destek vektör makineleri ile finansal zaman serisi olarak ele alınan borsa endeksinin yönünü tahmin edebilecek bir model geliştirmek, geliştirilen modelin borsa endeksi üzerindeki tahmin performansını ölçmek ve dalgacık dönüşümünün destek vektör makinelerinin tahmin performansına etkisini incelemektir.

Altı bölümden oluşan çalışmanın ilk bölümünde literatürde finansal zaman serilerinde dalgacık dönüşümü kullanılan çalışmalar ve destek vektör makineleri yöntemi kullanılan çalışmalar verilmekte, ardından dalgacık dönüşümü ve destek vektör makineleri yönteminin birlikte kullanıldığı sınırlı sayıdaki yabancı literatürdeki çalışmaya yer verilmektedir. Ayrıca çeşitli yöntemlerle XU100 borsa endeksi yön tahmini yapılan çalışmalar bu bölümde özetlenerek literatürdeki eksiklikler saptanmaktadır.

İkinci bölümde ise dalgacık dönüşümü, destek vektör makineleri ve öznitelik seçiminde kullanılan yöntemler hakkında bilgi verilmektedir. Öncelikle dalgacık dönüşümü yöntemi başlığı altında dalgacık dönüşümünün tarihi gelişimi ve dalgacık dönüşümüne temel oluşturan Fourier dönüşümü üzerinde durulmakta ve kısa zaman Fourier dönüşümüne değinilmektedir. Ardından, dalgacık dönüşümünün teorik temelleri ile birlikte dalgacıkların ölçeklenmesi ve kaydırılması, sürekli dalgacık dönüşümü ve ayrık dalgacık dönüşümü verilmektedir. Ayrıca, zaman serileri analizlerinde en çok kullanılan dalgacık çeşitleri detaylı olarak incelenmektedir.

Destek vektör makineleri başlığı altında istatistiksel öğrenme teorisinden bahsedilerek destek vektör makineleri yönteminin temel çalışma prensibi olan yapısal risk minimizasyonu incelenmektedir. Daha sonra optimal ayırıcı hiperdüzlemler verilerek doğrusal olarak ayıramayan örüntüler için optimal hiperdüzlemlere değinilmekte, buradan doğrusal olmayan destek vektör makineleri yöntemine geçiş yapılarak çekirdek hilesinden bahsedilmektedir. Destek vektör

makineleri yönteminin çekirdek fonksiyonları detaylı olarak incelenerek doğrusal, polinom, radyal tabanlı ve sigmoid çekirdek fonksiyonları ele alınmaktadır.

Öznitelik seçiminde kullanılan yöntemler başlığı altında ise modelleri oluşturan değişkenleri belirlemede kullanılan temel bileşen analizi, k-en yakın komşuluk, korelasyon, lojistik regresyon, kısmi en küçük kareler ve rastgele orman algoritması değişken seçimi yöntemleri detaylıca irdelenmektedir.

Üçüncü bölümde ise çalışmada yön tahmini yapılan Borsa İstanbul ve barındırdığı endeksler hakkında bilgi verilmektedir. Ayrıca teknik göstergeler başlığı altında literatürde sıklıkla kullanılan teknik göstergelerin hesaplanmaları ve yorumlanmaları ile ilgili detaylı bilgi verilmiştir.

Dördüncü bölümde, önce çalışmanın amacı ve kapsamı verilmekte ardından uygulamanın modeli tanıtılarak, verilerin toplanması, hazırlanması ve verilerle ilgili bilgi verilmektedir. Uygulama aşamalarının belirlenmesi, kullanılan yazılım ve paketler, model parametrelerin belirlenmesinde kullanılan metot olan çapraz geçirme ve sınıflandırma matrisleri hakkında verilen bilgilerin akabinde uygulama aşamaları şema ile de gösterilmektedir.

Ardından analizde kullanılan bağımsız değişkenler, yurtdışı borsa endeksleri, teknik göstergeler, makroekonomik göstergeler ve gün değişkenleri olarak gruplandırılarak bu değişkenler hakkında istatistiki bilgiler verilmektedir. BIST100 endeksinin yönünü gösteren bağımlı değişkenin hesaplanmasından da kısaca bahsedilerek, günlük, haftalık ve aylık verilerde değişken seçimi ile oluşturulan ham ve dalgacık modellerin ayrı ayrı destek vektör makineleri sınıflandırıcıları ile elde edilen tahmin performansları ele alınmaktadır.

Beşinci bölümde ise, dördüncü bölümde elde edilen çalışmanın bulguları toplu olarak değerlendirilmekte ve sonuçları yorumlanmaktadır. Ayrıca elde edilen bulgular ile diğer çalışmaların sonuçları arasındaki benzerlikler ve farklılıklar tartışılmaktadır.

Son bölümde ise çalışmada elde edilen sonuçlar ayrıntılı şekilde açıklanarak çeşitli öneriler ortaya konulmaktadır.

# 1. BÖLÜM: DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE HİSSE SENEDİ PİYASALARINDA TAHMİN ÇALIŞMALARI

## 1.1. DALGACIK DÖNÜŞÜMÜ KULLANILARAK YAPILAN ÇALIŞMALAR

Dalgacık dönüşümünün zaman serileri üzerine uygulanmasını konu alan ilk çalışmalardan Morettin (1996), zaman serilerinin analizinde Fourier analizinden dalgacık analizine geçişi incelemekte ve dalgacık dönüşümünün kullanılabilirliğini ortaya koymaktadır. Yine benzer şekilde Priestley (1996) dalgacıklar ve zaman bağımlı spektral analiz adlı çalışmasıyla dalgacık dönüşümünün zaman serilerine uygulanabilirliği ile ilgili temelleri ortaya koymuştur.

Ramsey ve Zhang (1997)'in Ekim 1992 ile Eylül 1993 tarihleri arasındaki Japon yeni ve Alman markının Amerikan dolarına göre kurlarının kullanıldığı çalışmada, dalgacık analizi uygulanarak döviz kurunda sadece alçak frekanslarda belirli bir yapı tespit edilmiştir.

Ramsey ve Lampart (1998)'in çalışmalarında, bazı ekonomik değişkenlerin ortogonal ayrışımının üretilmesi için dalgacık dönüşümü kullanılarak, altı farklı zaman ölçeği üzerinde gelir ile para arzı arasındaki ilişki Granger nedensellik analizi ile incelenmiş; alçak frekanslı ölçeklerde, gelir para arzının Granger nedeni iken yüksek frekanslı ölçeklerde, para arzı gelirin Granger nedeni olarak saptanmıştır.

Ramsey (1999)'in dalgacıkların ekonomik ve finansal verilerin analizine katkısını incelediği literatür çalışması ve dalgacıkların zaman serisi tahmininde kullanımına yönelik çeşitli yollar ortaya koyan Schlüter ve Deuschle (2010) gibi

çalışmalar, dalgacık dönüşümünün zaman serilerinde tahmin performanslarına pozitif katkı yapacağını ileri sürmektedir.

Nason ve Von Sachs (1999), istatistiksel zaman serileri analizinde dalgacık dönüşümünün rolünü inceledikleri çalışmalarında, zaman serisi olarak Elektrokardiyografi (EKG) sinyallerinden faydalanmış ve dalgacık dönüşümünün faydalarını tartışmışlardır.

Abramovich, Bailey ve Sapatinas (2000)'in çalışmalarında, dalgacık dönüşümü temelleriyle birlikte ele alınmış ve istatistiksel uygulamaları tartışılarak regresyon tahmininde olduğu gibi zaman serisi analizinde de kullanılabileceği ileri sürülmüştür.

Norsworthy, Li ve Gorener (2000)'in dalgacıkların kullanımını mühendislik alanından finansal uygulamalara ve zaman serilerine aktarmak amacıyla yaptıkları çalışmada, SP500 endeksini temsilen seçilen 99 hisse senedinin 1998 yılı günlük getirileri kullanılarak, 2,4,8,16 ve 32 günlük olmak üzere beş farklı frekans için beta katsayıları tahmin edilmiştir. Çalışmada dalgacık analizinin yüksek frekanslı hareketlerde hisse getirilerindeki piyasa etkisini kanıtladığını, hisse getirilerindeki piyasa etkisinin arındırılmasıyla beklenmedik olayların etkisinin tanımlanmasında verimli olacağı belirtilmiştir.

Gençay, Selçuk ve Whitcher (2001), dalgacık dönüşümü yöntemini döviz kuru volatilitelerinin hesaplanmasında kullandıkları çalışmalarında, Aralık 1986 ile Aralık 1996 tarihleri arasındaki Japon yeni ve Alman markının Amerikan dolarına göre kur verisi kullanılmış, döviz kuru volatiliteleri arasındaki ilişkinin gün içi ölçeklerde en düşük olduğu, ancak günlük ölçeğe doğru kademeli artış gösterdiği sonucu elde edilmiştir.

Lee (2002)'nin gelişmiş piyasalar (ABD, Almanya ve Japonya) ile Orta Doğu Kuzey Afrika (MENA) bölgesindeki gelişmekte olan piyasalar (Türkiye ve Mısır) arasındaki uluslararası yayılma etkilerinin incelendiği çalışmasında, önceki çalışmalarda kullanılan GARCH tipi modeller yerine ayrık dalgacık dönüşümü kullanılmıştır. Ampirik bulgulara göre, volatiliteler gibi fiyatın da yayılma etkisinin



gelişmiş piyasalardan MENA bölgesindeki piyasalara doğru olduğu ancak tersinin doğru olmadığı tespit edilmiştir. Ayrıca Türkiye piyasalarının gelişmiş piyasalara entegre olduğu saptanmıştır.

Kim ve In (2003)'ün Granger nedenselliği, spektral analiz ve dalgacık analizinin kullanıldığı çalışmada, en kısa zaman ölçeğinde çoğu finansal değişkenin Amerikan sanayi üretimi ile ilişki gösterdiği, orta zaman ölçeğinde karışık sonuçlar verdiği, uzun zaman ölçeğinde ise çoğu değişkende anlamlı ilişkiler olduğu saptanmış ve sanayi üretiminin finansal değişkenlerin Granger nedeni olduğu tespit edilmiştir.

Gençay, Selçuk ve Whitcher (2005) yaptıkları bir diğer çalışmada, sistematik risk hesabı için önerdikleri yöntemlerinde, farklı zaman ölçekleri dalgacıklar analizi ile incelenmiş, farklı zaman ölçekleri ve farklı frekanslar için farklı sistematik riskler tespit edilmiştir. Sonuç olarak da, portföy getirisi ve sistematik risk arasındaki ilişkinin dalgacık ölçeği arttıkça daha güçlü hale geldiği ifade edilmektedir.

Yamada (2005)'nin Japonya borsası sanayi endekslerinin 1983 ile 2003 yılları arasındaki verilerini kullanarak sistematik risk tahmini yaptığı çalışmada dalgacık dönüşümü kullanılmış, elde ettikleri bulgulara göre çoğu endüstride kısa ve uzun periyotların sistematik riski benzerken bazılarında ise tamamen farklı olduğu tespit edilmiş ve dalgacık bazlı sistematik risk tahmininin tümünde olmasa da bazı endüstrilerde yararı olduğu ortaya konulmuştur.

Gallegati (2005)'nin çalışmada, dalgacık analizi ile gelişmekte olan piyasalardan beş büyük Orta Doğu Kuzey Afrika hisse senedi piyasalarının (Mısır, İsrail, Ürdün, Fas ve Türkiye) Haziran 1997 ile Mart 2005 tarihleri arasındaki verileri ayrık dalgacık dönüşümü ile farklı zaman ölçeklerinde incelenmiş, ayrıca SP ve Eurostoxx gibi gelişmiş piyasalar ile zaman-ölçek ilişkileri incelenerek MENA hisse senedi piyasalarının hem bölgesel hem de uluslararası piyasalarla entegre olmadığı tespit edilmiştir.

Li, Shi ve Li (2006)'nin dalgacık bazlı göstergeler ve genetik algoritma ile finansal zaman serisi tahmini yaptıkları çalışmada, DJIA endeksinin 07 Nisan 1969

ile 09 Nisan 1981 tarihleri arasındaki 3035 işlem günü verisinden 1900 günü eğitim, 1135 günü ise test veri seti olarak kullanılmıştır. Dalgacık tabanlı göstergelerin tahmin performansının (max %85,39) teknik göstergelerin tahmin performansı ile (max %71,40) karşılaştırıldığı analizlerde dalgacık dönüşümünün tahmin performansını %20'ye kadar artırdığı saptanmıştır.

Yaptıkları bir diğer çalışmada yine dalgacık analizini kullanan Kim ve In (2007), tahvil gelirleri ile hisse senedi fiyatları arasındaki ilişkiyi G7 ülkeleri üzerinden inceleyerek farklı zaman ölçekleri ve farklı frekanslar için ülkelere göre değişen farklı şekilde ilişkiler saptamışlardır. Sonuç olarak, dalgacıklar analizi yardımıyla tahvil gelirleri ve hisse senedi fiyatlarındaki değişimin, Japonya hariç diğer ülkelerde birlikte hareket etmediği ortaya konulmuştur.

Buraya kadar verilen yabancı literatürde yer alan çalışmalara ek olarak, yine yabancı literatürde yer bulan ancak ülkemiz ekonomik ve finansal zaman serilerine dalgacık dönüşümü uygulanarak yapılan ilk çalışmada Gençay vd. (2002) tarafından enflasyon ve parasal genişleme verilerine dalgacık dönüşümü uygulanarak Granger nedensellik testi yapılmıştır. Sonuç olarak; Arjantin, İsrail ve Türkiye'de parasal genişlemenin enflasyonun Granger nedeni olduğu belirtilmektedir.

Ozun ve Cifter (2007)'in gelişmekte olan ülkelere Türkiye örneği olarak yaptıkları çalışmada, dalgacık analizi yardımıyla gayri safi milli hasıla (GSMH) ile enerji tüketimi arasındaki nedensellik incelenmiş, GSMH ile enerji tüketimi arasındaki dalgacık korelasyonunun uzun vadede 5-8 yıllık zaman aralığında maksimum olduğu görülmüş, böylece GSMH'nın enerji tüketimini maksimum 5-8 yıl sonra etkilediği sonucu elde edilmiştir.

DiSario, Saraoglu, McCarthy ve Li (2008)'in çalışmalarında, İMKB 100 endeksi verileri kullanılarak dalgacık dönüşümüyle volatilitedeki uzun dönem bağımlılık yapısı incelenmiş ve serilerdeki birbirinden uzak gözlemlerin birbiriyle ilişkisi dalgacıklar analizinden yararlanılarak elde edilmiştir. Yine dalgacık dönüşümünden yararlanan bir diğer çalışmada Aktan, Mabrouk, Ozturk ve Rhaiem (2009) İMKB verilerini kullanarak sistematik risk tahmini yaptıkları görülmektedir.

Dalgacık dönüşümünün finansal zaman serilerine uygulanması ile ilgili az sayıdaki Türkçe çalışmadan ilkinde Özün ve Çifter (2006), dalgacıklar analizi yöntemiyle "gösterge niteliğindeki yıllık bileşik faiz oranlarındaki değişimlerin bankaların hisse senedi getirileri üzerindeki etkisini" incelemektedir. Referans hazine bonosunun bir yıllık bileşik faiz oranı İMKB Bankacılık Endeksine ait 2 Ocak 2002 ile 18 Ağustos 2006 tarihleri arasındaki günlük verilerden elde edilen zaman serileri kullanılarak yapılan analizde elde edilen ampirik bulgular, faiz riskinin bankaların hisse senedi getirilerinde önemli bir etken olduğu ve bu etkinin ölçüğe göre değiştiğini göstermektedir. Çalışmada, endeks üzerinde faiz oranının kısa dönemde 16 güne kadar etkin olduğu, uzun dönemde ise 16 günden 63 güne kadar etkisinin tespit edilmediği ancak 64. günden itibaren tekrar başlayan etkinliğinin tespit edildiği belirtilmiştir.

Gürsakar (2009)'ın çalışmasında, dalgacık analizinin finansal zaman serileri analizi için de uygun bir yöntem olduğunu ortaya koyma amacı doğrultusunda Jensen tarafından geliştirilen ayrık dalgacık dönüşümü kullanılarak İMKB'deki uzun dönem bağımlılık yapısı tahmin edilmektedir. Dalgacık analizinin uzun dönem bağımlılık yapısı tahminlerinde daha etkin ve tutarlı sonuçlar verdiği görülmüş, dolayısıyla finansal analizlerde de dalgacık analizinin kullanılabileceği ifade edilmiştir.

Çifter (2010)'in çalışmasında, riske maruz değer öngörüsünde dalgacık bazlı uç değer teorisi geliştirilmiş, İMKB 100 endeksinin 15 Ocak 2001 ile 20 Mart 2009 tarihleri arasındaki günlük periyotta 2048 gözlem kullanılmış ve önerilen dalgacık bazlı uç değer teorisi modelleri diğer modellerle karşılaştırılmıştır. Çalışmada, dalgacık bazlı uç değer teorisi modellerinin daha iyi öngörü performansı sağladığı tespit edilmiştir.

Erdemir (2012)'in çalışmasında, SP500 Endeksi üzerindeki Avrupa tipi SPX al-sat opsiyonları içsel oynaklığının, gelecek piyasa volatilitelerini tahmin performansı incelenmiş, dalgacık hesaplamalarının uzun vadede al-sat opsiyonları oynaklıklarının piyasanın realize edilen oynaklığı için karşılaştırılabilir bilgi içerdiğini göstermekte olduğu ve kısa vadede sat opsiyon oynaklığının tahmin becerisinin daha iyi olduğu ifade edilmiştir.

Hem yabancı hem de yerli literatürden elde edilen yukarıdaki çalışmalar incelendiğinde dalgacık dönüşümü yönteminin yerel ekonomik ve finansal zaman serilerine uygulanması konusunda yeterli çalışma olmadığı görülmektedir. Ayrıca dalgacık dönüşümünün sistematik risk, uzun dönem bağımlılık yapısı, Granger nedensellik analizlerinde yaygın bir şekilde kullanıldığı ancak endeks tahmininde çok fazla çalışma olmadığı görülmektedir. Literatürdeki endeks tahmininde dalgacık dönüşümü kullanılması konusundaki eksikliğin giderilmesine katkıda bulunmak için bu çalışma yapılmaktadır.

## **1.2. DESTEK VEKTÖR MAKİNELERİ KULLANILARAK YAPILAN YÖN TAHMİNİ ÇALIŞMALARI**

DVM yönteminin Borsa İstanbul verilerine uygulandığı uygulamalar çok sınırlıdır. Aşağıda uluslararası ve yerli literatürde DVM kullanılarak yapılan borsa endeksleri yön tahmini uygulamaları yer almaktadır.

K. Kim (2003) tarafından yapılan çalışmada, Kore Borsası (KOSPI) endeksinin Ocak 1989 ile Aralık 1998 tarihleri arasındaki 2928 günlük verisi kullanılarak, endeksin yönü DVM ve yapay sinir ağları (YSA) yöntemleriyle tahmin edilmektedir. Analizde William's %R, Stokastik (%K), Stokastik (%D), Slow (%D), A/D osilatörü, Momentum, ROC, 5 günlük aykırılık endeksi, 10 günlük aykırılık endeksi, OSCP, CCI ve RSI olmak üzere oniki teknik gösterge kullanılmıştır. 2347 gözlem (%80) eğitim için, 581 gözlem (%20) test için kullanılmıştır. Analiz sonucuna göre, DVM yönteminin tahmin performansı %57,83 iken, YSA yönteminin tahmin performansı %54,73 olarak hesaplanmıştır.

Huang, Nakamori ve Wang (2005)'in çalışmalarında, 1 Ocak 1990 ile 31 Aralık 2002 tarih aralığında NIKKEI-225 endeksinin 640 haftalık eğitim ve 36 haftalık test verisi kullanılmış ve haftalık değişim yönünü tahmin etmek için doğrusal diskriminant, kuadratik diskriminant analizi, elman geriyayılım sinir ağları ve DVM yöntemleri karşılaştırılmıştır. NIKKEI borsasının haftalık değişim yönünün tahmini için SP500 endeksi ile Japon Yeni/Amerikan Doları kuru modelde girdi değişkenleri olarak kullanılmıştır. Sonuç olarak DVM yönteminin performansının (%75) diğer yöntemlere göre daha üstün olduğunu ortaya koyulmuştur.

Özdemir, Tolun ve Demirci (2011) tarafından yapılan çalışmada, ikili sınıflandırma kapasitesine sahip olan yöntemlerden, lojistik regresyon ve DVM yöntemleri ile İMKB 100 endeksinin yönünü tahmin etmek için Şubat 1997 ile Aralık 2010 tarihleri arasındaki dönemi kapsayan aylık veriler kullanılmıştır. 167 aylık verinin 138'i eğitim, 29'u ise test verisi olarak değerlendirilmiştir. Çalışmada makroekonomik göstergeler olarak, tüketici fiyatları endeksine dayalı enflasyon oranları, İstanbul Altın borsası endeksi, Türkiye Cumhuriyeti Merkez Bankası (TCMB) döviz rezervleri, reel efektif döviz kuru endeksi, döviz satış kuru (ABD doları), bankalar arası para piyasası gecelik faiz oranları ile ABD Federal Funds Rate farkı, West Texas Intermediate spot petrol fiyatı ve imalat sanayi kapasite kullanım oranı kullanılmıştır. Ayrıca piyasa değeri ağırlıklı hisse senedi piyasa performansını gösteren MSCI European Index (16 Avrupa ülkesi piyasası için), MSCI Emerging Markets Index (21 yükselen piyasa için) ve SP500 endeksleri diğer girdi değişkenlerini oluşturmuştur. Sonuç olarak; her iki yöntem için de test kümesi üzerinde doğru sınıflandırma oranları %86,20 olarak elde edilmiştir.

Kara, Acar Boyacioglu ve Baykan (2011)'in çalışmalarında, 2 Ocak 1997 ile 31 Aralık 2007 tarihleri arası döneme ait 2733 işlem günü verisi kullanılarak İMKB 100 endeksinin yön tahmininde YSA ve DVM yöntemlerinin performansları karşılaştırılmıştır. Analizde bağımsız değişken olarak; momentum, MACD, CCI, RSI, basit ve ağırlıklı hareketli ortalamalar (SMA ve WMA), stokastik (%K), stokastik (%D), A/D osilatörü ve William's %R olmak üzere toplam on teknik gösterge kullanılmıştır. Çalışma elde edilen bulgulara göre; YSA yöntemi ile elde edilen %75,74 tahmin performansının DVM yöntemi ile elde edilen %71,52 tahmin performansından daha iyi olduğu sonucuna ulaşılmıştır.

Yakut (2012) çalışmasında, İMKB'de işlem gören sanayi işletmelerinin 2002 ile 2010 yılları arasındaki verileri kullanılarak, veri madenciliği tekniklerinden C5.0 algoritması, DVM ve YSA yöntemlerini kullanılarak işletmelerin finansal başarısızlıklarını tahmin modelleri oluşturulmuş, bu modellerin dört yıl öncesine kadar ki performansları birbiriyle karşılaştırılarak 2010 yılına ait işletme başarısızlıklarının tahmini gerçekleştirilmiştir. Elde edilen analiz sonuçlarına göre; YSA yönteminin, oluşturulan dört modelin tamamında hem C5.0 algoritması hem de

DVM yönteminin performanslarından daha iyi tahmin performansı gösterdiği ifade edilmektedir.

Emir (2013a)'in makine öğrenmesi yöntemlerinden olan YSA ve DVM yöntemlerinin İMKB 100 Endeksi'nin yön tahmininde kullanılabilirlikleri incelendiği çalışmada, 2005 ile 2011 yılları arasındaki döneme ait 1747 günlük veri seti kullanılmış, sırasıyla teknik gösterge değişkenleri, borsa endeks değişim oranları değişkenleri ve makroekonomik gösterge değişkenlerinden oluşan modellerin yön tahmininde gösterdikleri performanslar karşılaştırılmış, teknik gösterge değişkenlerinin diğer değişken gruplarına göre daha yüksek performans gösterdiği belirlenmiştir. Teknik göstergeler kullanılarak elde edilen modelde DVM yönteminin tahmin performansı %73,70 olarak elde edilirken YSA ile %80,00 tahmin performansı elde edilmiştir. Ayrıca, tüm değişkenlerin kullanıldığı en geniş veri setinde öznitelik seçimi yapılarak minimum sayıdaki değişkenle en iyi yön tahminini sağlayacak model aranmıştır. Çalışmadaki tüm analizlerde YSA yönteminin, DVM yöntemine göre daha iyi sonuçlar verdiği belirtilmektedir.

Tayyar ve Tekin (2013)'in çalışmalarında, İMKB 100 Endeksi hareket yönü tahmini amacıyla sınıflama metodu olarak kullanılan DVM yöntemi kullanılarak, tahmin performansı Lojistik Regresyon yöntemi ile karşılaştırılmıştır. Çalışmada, 3 Nisan 1995 ile 19 Mart 2012 tarihleri arası dönemi kapsayan günlük (4026), haftalık (840) ve aylık (193) veri setleri kullanılmıştır. Çalışmada bağımsız değişkenler 20 farklı teknik göstergeden elde edilmiştir. Günlük modellerde DVM yöntemi ile elde edilen en iyi tahmin performansı %52,42 olurken aylık modellerde en iyi tahmin performansı %63,15 olarak elde edilmiştir. İMKB 100 endeksi yönünü en iyi tahmin eden modelin, 20 ve 200 günlük basit hareketli ortalama, 20 ve 200 günlük üssel hareketli ortalama, 20 ve 200 günlük ağırlıklı hareketli ortalama, Aroon osilatörü, Parabolic SAR, MACD, ROC, RSI, SO, SMI, William's %R, CCI, CMO, BB-HLC, BB-C, ATR, ADX, Chaikin dalgalanma göstergesi, CO, A/D osilatörü, OBV, MFI ve CMF'den oluşan toplam 27 adet bağımsız değişkenin tamamının değişken seçimi uygulanmadan kullanıldığı haftalık DVM modeli olduğu (%70,00) gözlenmiştir.

Yakut ve Elmas (2014), 2005 ile 2012 yılları arasındaki günlük verilerle BIST 100 endeksini tahmin etmek için ileri beslemeli YSA ve DVM yöntemlerini

kullandıkları çalışmalarında, girdi değişkenleri olarak bir, iki ve üç gün öncesine ait endeks değerleri, gecelik faiz oranı, Amerikan dolar kuru ve FTSE, CAC, NIKKEI, DAX ve BOVESPA borsa endeksleri seçilmiştir. Çalışmada, her iki modelde de bir gün önceki veriler ile kurulan modellerin iki ve üç gün önceki verilerin kullanıldığı modellere oranla daha iyi performans gösterdiği, ayrıca YSA ve DVM yöntemlerinin her ikisinin de BIST 100 borsa endeksinin modellenmesinde kullanılabileceği sonucuna ulaşılmıştır.

### **1.3. DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE TAHMİN ÇALIŞMALARI**

Dalgacık dönüşümü ve DVM yönteminin birlikte kullanılarak finansal zaman serileri üzerinde tahmin çalışması yapılan uluslararası literatürde karşılaşılan tüm çalışmalar aşağıda kısaca özetlenmektedir.

Dalgacık dönüşümü ve DVM yönteminin birlikte finansal verilerde kullanıldığı literatürde karşılaşılan ve dalgacık yönteminin önemi ve katkısını ortaya koyan ilk çalışmada, Sui, Hu, Yu, Xie ve Qi (2007), Shanghai Stock Exchange (SSE) bileşik endeksine ait 28 Nisan 1997 ile 12 Eylül 2006 tarihleri arasındaki 2261 günlük verinin 192'si eğitim ve kalan 341'i ise test verisi olarak kullanılmıştır. ALF, RS, RSI, MFI, %BB, V, VB, CHO, MACD, Stokastik (%K), A/D Osc ve Williams %R olmak üzere oniki teknik gösterge kullanılarak endeksin günlük artış ve azalış yönünün tahmin edildiği çalışmada, dalgacık yöntemi ile sıkıştırma yaparak gürültüden arındırma işlemi yapılmış ve sonuç olarak, gürültü içeren veri setinde %54,25, gürültüden arındırılmış veri setinde ise %60,12, oranında daha iyi tahmin performansı elde edildiği belirtilmektedir.

S. Huang ve Wu (2008a)'nın NASDAQ (ABD), NK225 (Japonya), TWSI (Tayvan) ve KOSPI (Güney Kore) endekslerinin Ocak 2003 ile Aralık 2004 tarihleri arasındaki 435 günlük veriyi kullandıkları çalışmada, GARCH ve tekil DVM tahmin modellerine göre dalgacık çekirdekli DVM modelinin daha iyi performans gösterdiği ifade edilmektedir.

S. Huang ve Wu (2008b) NASDAQ (ABD), NK225 (Japonya), TWSI (Tayvan) ve KOSPI (Güney Kore) endekslerinin Ocak 2003 ile Aralık 2004 tarihleri arasındaki 435 günlük verisini ve G7 ülkelerinin CAC40 (Fransa), FTSE100 (UK), DAX30 (Almanya), MIB40 (İtalya), TSX60 (Kanada), SP500 (ABD) ve NK225 (Japonya) endekslerinin Ocak 2004 - Aralık 2005 arasındaki 435 günlük verisini ayrı ayrı inceledikleri çalışmalarında, dalgacık dönüşümü ve genetik algoritma tabanlı DVM yönteminin GARCH, YSA ve tekil DVM yöntemine göre daha iyi tahmin performansı gösterdiği ortaya konulmaktadır.

Tang, Tang ve Sheng (2009)'nin DAXINDEX, FRCAC40, FTSE100, JAPDOWA ve SPCOMP endekslerinin 1 Ocak 1992 ile 31 Aralık 1997 tarihleri arasındaki 1560 günlük verisini kullandıkları çalışmalarında, dalgacık çekirdekli DVM yönteminin Gaussian çekirdekli DVM yöntemine göre istatistiksel olarak anlamlı şekilde daha iyi tahmin performansı gösterdiği ifade edilmektedir.

S. Huang (2011), S. Huang ve Wu (2008b)'deki çalışmasının verilerini kullanarak yaptığı çalışmada dalgacık dönüşümünün kullanılmasının tahmin performansına katkı yaptığını göstermektedir.

Chai, Du, Lai ve Lee (2015), Çin Halk Cumhuriyeti borsasından CSI 300 endeksinin 1 Mayıs 2009 ile 23 Ağustos 2011 tarihleri arasındaki 643 günlük verinin eğitim, 24 Ağustos 2011 ile 20 Ocak 2012 tarihleri arasındaki 100 günlük verinin de test verisi olmak üzere toplam 743 günlük verinin kullanıldığı çalışmalarında, CSI 300, USDX, SHIBOR, REPO, CDS, PE, M2/mkt cap, short-mid not/mkt cap ve New Loan/mkt cap oranları girdi değişkeni olarak seçilmiş, Simplex, GS (grid search), PSO (particle swarm optimization) ve GA (genetic algorithm) optimizasyon metodları parametre seçimi için kullanılmıştır. EMD-LSSVM (Empirical Mode Decomposition Least Square Support Vector Machine) ve WD-LSSVM (Wavelet Denoising Least Square Support Vector Machine) yöntemleri optimizasyon tekniklerinden elde edilen parametreler kullanılarak ayrı ayrı karşılaştırılmış ve EMD-LSSVM (GS), EMD-LSSVM (Simplex), WD-LSSVM (Simplex), WD-LSSVM (PSO), WD-LSSVM (GA), WD-LSSVM (GS) şeklinde bir performans sıralaması elde edilmiştir.



#### 1.4. DİĞER ENDEKS TAHMİN ÇALIŞMALARI (XU100)

Borsa İstanbul'da XU100 kodlu BIST 100 endeksi veya eski adıyla İMKB 100 endeksinin yönünün veya değerinin tahmin edilmeye çalışıldığı literatürde karşılaşılan çalışmalar aşağıda kısaca özetlenmiştir.

Diler (2003) YSA yönteminin ülkemiz piyasalarına uygulanabilirliğini araştırdığı çalışmasında, 1 Ocak 1990'dan 11 Kasım 2003 tarihine kadar olan 3418 günlük veri setinin 2700 günü eğitim 718 günü test için ayrılmıştır. 10 günlük basit hareketli ortalama, 5 ve 10 günlük ağırlıklı hareketli ortalama, 10 günlük momentum, MACD, RSI ve Stokastik (%K) teknik göstergelerinin girdi değişkeni olarak kullanıldığı YSA modeli ile İMKB 100 Endeksi'nin ertesi günü yönü tahmin edilmeye çalışılmış ve test için ayrılan veri kümesi üzerinde %60,81 oranında doğru tahmin performansına ulaşılmıştır.

Egeli, Ozturan ve Badur (2003) İMKB 100 Endeksini tahmin için YSA yönteminin performansını hareketli ortalamalar yaklaşımıyla karşılaştırdıkları çalışmalarında, 1 Temmuz 2001 ile 28 Şubat 2003 tarihleri arasındaki 417 günlük verinin 376 gününü eğitim, geriye kalan 41 gününü test için kullanmışlardır. Bir gün önceki endeks değeri, Amerikan doları kuru değeri, gecelik faiz değeri ve kukla değişkenler olarak haftanın işgünlerini girdi değişkenleri olarak kullandıkları modellerinde YSA ile %82,00 tahmin performansı elde etmişlerdir.

Karaatlı (2003)'nın Ocak 1990 ile Aralık 2002 tarihleri arasında İMKB 100 endeksinin aylık kapanış değerlerinin kullanıldığı çalışmasında, hazine bonusu faiz oranları, TUFİ (enflasyon oranı), cumhuriyet altını fiyatı, tasarruf mevduatı faiz oranı, sanayi üretim endeksi, döviz kuru (Amerikan Doları) ve zaman değişkenleri girdi değişkenleri olarak alınmıştır. YSA yöntemi ile İMKB endeksinin tahmin edildiği çalışmada uzun dönemler için yapılan tahmin performansının yakın dönemlere göre düşük olduğu ifade edilmiştir.

Altay ve Satman (2005)'in çalışmasında, 2 Ocak 1997 ile 28 Şubat 2005 tarihleri arasındaki dönemde İMKB Tüm ve İMKB 30 endekslerine ait toplam 2003 günlük veri kullanılarak YSA ve doğrusal regresyon yöntemlerinin İMKB

endekslerini tahmin performansları karşılaştırılmıştır. Günlük gözlemlerin yanında haftalık (426) ve aylık (100) gözlemlerin de kullanıldığı çalışmada, YSA yönteminin doğrusal regresyon yöntemine göre günlük (%57,80), haftalık (%67,10) ve aylık (%78,30) veri setlerinde daha iyi tahmin performansı gösterdiği ifade edilmiştir.

Yümlü, Gürgen ve Okay (2005) finansal zaman serileri için çeşitli sinir ağı yapıları ve üssel genelleştirilmiş otoregresif koşullu değişen varyans (EGARCH) tekniğinin performanslarını karşılaştırdıkları çalışmalarında, İMKB 100 endeksinin 1990 ile 2002 yılları arasındaki verisinden yararlanmışlardır. Kayan pencere tekniği ile serideki son n gözlemin girdi olarak kullanıldığı çalışmada, YSA modellerinin EGARCH modeline göre daha iyi performans gösterdiği belirtilmektedir.

Karaatlı, Güngör, Demir ve Kalaycı (2005)'nin İMKB 100 endeksinin fiyat hareketlerini YSA yöntemi ile tahmin ettikleri çalışmalarında, Ocak 1990 ile Aralık 2002 tarihleri arasındaki ay sonu verileri kullanılarak 2003 yılının ilk dokuz ayına ait tahminde bulunulmuştur. Hazine bonusu faiz oranları, TUFEE (enflasyon oranı), cumhuriyet altını fiyatı, tasarruf mevduatı faiz oranı, sanayi üretim endeksi, döviz kuru (Amerikan Doları) ve zaman değişkenlerinin girdi değişkenleri olarak seçildiği YSA modeli ile yapılan dokuz aylık tahminde mart ve mayıs aylarında gerçek değerlere daha yakın tahmin değerleri elde edilmiştir.

Yıldız, Yalama ve Coşkun (2008) girdi değişkeni olarak, İMKB 100 endeksinin günlük en yüksek, en düşük ve kapanış değerleri, Amerikan dolar kuru ve faiz oranlarını kullandıkları çalışmalarında, İMKB 100 endeksinin ertesi günkü yönünü tahmin edebilmek için YSA yönteminden yararlanmışlardır. 1805 günlük eğitim veri seti ve 100 günlük test veri setinin kullanıldığı çalışmada YSA yönteminin test veri seti üzerindeki tahmin performansı %74,51 olarak tespit edilmiştir.

Güreşen ve Kayakutlu (2008) endeks değişim hareketlerini YSA ve hibrit modellerle tahmin ettikleri çalışmalarında, İMKB 100 endeksinin Ocak 2003 ile Mart 2008 tarihleri arasındaki verisinin 1132 işgünü eğitim ve kalan 160 günü test veri seti olarak kullanılmıştır. YSA yönteminin endeks değeri tahminindeki eksikliklerini

azaltabilmek için oluşturulan hibrit modellerden YSA için Dinamik Mimari modeli (DAN2), en güçlü sinirsel ağ yapısı olarak elde edilmiştir.

Akel ve Bayramođlu (2008) İMKB 100 endeksinin 04 Ocak 1999 tarihi ile 28 Şubat 2001 tarihleri arasındaki 521 işgününe ait verilerini kullanarak endeksin yönünü YSA yöntemiyle tahmin ettikleri çalışmalarında, girdi değişkenleri olarak dolar kuru, merkez bankası 1 aylık mevduat faiz oranı, Altın borsası kapanış fiyatı, merkez bankası döviz rezervleri ve İMKB işlem hacmi seçilmiştir. Çalışmanın sonuçlarına göre ilgilenilen dönem Şubat 2001 krizi dönemini de içinde bulundurmasına rağmen YSA yöntemi % 73,68 oranında bir tahmin performansı göstermiştir.

Kutlu ve Badur (2009) çalışmalarında, 2 Temmuz 2001 ile 13 Temmuz 2006 tarihleri arasındaki 1270 işgününe ait verilerden 1143 günlük veri setini eğitim, kalan 127 günlük veri setini ise test verisi olarak ayırarak İMKB endeks yönünü tahmin etmeye çalışmışlardır. Bir gün önceki İMKB endeks değeri, Amerikan dolar kuru, gecelik faiz oranı, Almanya, Fransa, İngiltere, DOW JONES, NASDAQ, SP500, Japonya ve Brezilya borsa endeks değerleri ile birlikte kukla değişkenler olarak haftanın işgünleri oluşturulan ileri beslemeli YSA modelinin girdi değişkenleri olarak alınmıştır. Sonuç olarak YSA yönteminin basit hareketli ortalama yöntemine göre biraz daha iyi sonuçlar verdiği ve test veri seti üzerinde %55,10 tahmin performansı ortaya koyduğu belirtilmiştir.

Avacı (2009)'nın çalışmasında, İMKB-30 endeksi içinden seçilmiş hisse senetlerinin günlük getirilerini YSA modellerinin tahmin performansını araştırmak için, seçilen beş hisse senedinin 1 Mayıs 2007 ile 30 Eylül 2008 tarihleri arasındaki günlük kapanış, ağırlıklı ortalama ve hacim verilerinin bir, üç ve beş günlük gecikmeli değerleri kullanılmıştır. Sonuç olarak hisse senedi günlük getirilerinin tahmininde YSA yönteminin etkinliği ortaya konulmaktadır.

Haznedarođlu (2009)'nun kısa vadeli model için Mart 2001 ile Şubat 2008 dönem aralığını, orta vadeli model için Aralık 1997 ile Şubat 2008 dönem aralığını seçtiđi çalışmasında, finansal ve makro ekonomik verilerin İMKB endeksleri üzerindeki etkisini ortaya koyan bir tahmin sistemi geliştirebilmek için günlük

(1750) ve aylık (122) veri setleri kullanılmıştır. Makroekonomik ve finansal piyasalara ait veriler, hisse senedi piyasasına ilişkin temel finansal veriler, teknik analiz göstergeleri ve global piyasalarla etkileşimi gösteren Morgan Stanley bileşik endekslerinin girdi olarak kullanıldığı çalışmada, İMKB endekslerini YSA yöntemi ile tahmin etmede başarılı sonuçlar elde edilemediği ifade edilmektedir.

Ulusoy (2010) İMKB endeks tahmini için ileri beslemeli ağ mimarisine sahip yapay sinir ağı modellemesi yaptığı çalışmasında, İMKB 100 endeksinin 1997 ile 2000 yılları arasında 199 haftalık veri setini kullanmıştır. Bir gün önceki İMKB 100 endeksinin en düşük ve en yüksek, Amerikan dolar kurunun en düşük, en yüksek ve kapanış değerleri, Merkez Bankası gecelik, bir ay vadeli, üç ay vadeli, altı ay vadeli ve oniki ay vadeli faiz oranları, SP500 ve NYSE borsa endekslerinin haftalık kapanış değerleri ve Londra altın borsasının Altın kapanış (ons/\$) değerleri olmak üzere toplam 14 girdi değişkeninin kullanıldığı modellerde İMKB 100 endeksi haftalık veri setinde %33,00'ün altında hata ortalamaları ile tahmin edilmiştir. Araştırmanın bulgularından yola çıkılarak; aşırı artış, azalış olmayan ve siyasi etkilerin olmadığı günlerde daha başarılı tahmin değerleri elde edildiği belirtilmiştir.

Aygören, Sarıtaş ve Moralı (2012)'nin geleneksel zaman serileri, nümerik arama modelleri ve YSA yöntemlerini kullandıkları çalışmada, İMKB 100 endeksi 27 Temmuz 1995 ile 29 Temmuz 2010 tarihleri arasındaki 3779 işgünü verisi kullanılarak, bir gün önceki Altın fiyatları (USD/ONS), faiz oranı, Amerikan doları günlük kapanış değerleri ve bankalar arası çift taraflı işlem miktarlarından (TL) oluşan girdi değişkenleri ile tahmin edilmeye çalışılmıştır. Elde edilen sonuçlara göre, en iyi tahmin performansını YSA geri yayılım algoritması ortaya koyarken, nümerik arama modelleri için kullanılan Newton yöntemi, geleneksel zaman serileri için kullanılan ARMA(1,1) yönteminden daha iyi performans ortaya koyduğu görülmektedir.

Yolcu, Aladag, Egrioglu ve Uslu (2013) yeni bir bulanık zaman serisi yaklaşımı ile zaman serisi öngörüsü yaptıkları çalışmalarında, İMKB 100 endeksinin 3 Ekim 2008 ile 23 Aralık 2010 tarihleri arasındaki verileri üç ayrı döneme ayrılarak kullanılmıştır. Bulanık kümeleme ve YSA modellerinden oluşan hibrid bir bulanık zaman serisi yaklaşımı önerilen çalışmada, literatürdeki diğer çalışmalarla

karşılaştırıldığında önerilen modelin performansının daha iyi olduğu sonucuna varılmaktadır.

Emir (2013b) İMKB 100 endeks getirisini teknik göstergeler ile tahmin etmeye çalıştığı çalışmasında, 4 Ocak 2005 ile 2 Ağustos 2010 tarihleri arasındaki 1397 işgünü gözlem değerlerini eğitim veri seti, 3 Ağustos 2010 ile 30 Aralık 2011 tarihleri arasındaki 349 işgünü gözlem değerlerini test veri seti olarak kullanmıştır. Momentum, CCI, MFI, RSI, Stokastik ve William's %R teknik gösterge değerlerinin girdi değişkeni olarak seçildiği çalışmada, destek vektör regresyonu ve doğrusal regresyon metotları ile İMKB 100 endeksinin getiri tahmini yapılmış, tüm ölçümlerde destek vektör regresyonu metodunun daha iyi tahmin performansı ortaya koyduğu görülmüştür. Çalışmanın bulgularından yola çıkılarak, destek vektör regresyonu metodunun, borsa endeksi getirilerinin tahmininde kullanılabilir olduğu sonucuna varılmaktadır.

Yukarıda verilen çalışmalar incelendiğinde, finansal zaman serileri üzerinde DVM yöntemi ile yapılan tahmin çalışmalarında genellikle tahmin performansları diğer tahmin yöntemleriyle karşılaştırılmıştır. DVM yönteminin performansını artırmak için girdi değişkeni seçimleri, çekirdek seçimi ve parametre seçimi gibi metotlar denenmiştir. Yerli literatürdeki borsa endeks yönü tahmin çalışmalarında da benzer süreçler görülmektedir. Uluslararası literatürde dalgacık dönüşümü ve DVM yöntemlerinin birlikte finansal zaman serileri tahmininde kullanıldığı toplam altı çalışmaya rastlanılmış, bu çalışmaların da üçünün aynı araştırmacılara ait olduğu tespit edilmiştir.

Buradan hareketle finansal zaman serilerinde tahmin konusunda hem uluslararası literatüre katkı yapacak hem de yerel literatürdeki boşluğu giderecek bu çalışmanın gerekliliği ve önemi ortaya çıkmaktadır. Ayrıca bu çalışmada elde edilen modellerin, finans piyasası yatırımcıları için yatırımlarını yönlendirme kararlarına destek olacak sağlıklı tahmin verisi sağlayacağı düşünülmektedir.

## 2. BÖLÜM: DALGACIK DÖNÜŞÜMÜ, DESTEK VEKTÖR MAKİNELERİ VE ÖZNETELİK SEÇİMİNDE KULLANILAN YÖNTEMLER

### 2.1. DALGACIK DÖNÜŞÜMÜ YÖNTEMİ

#### 2.1.1. Dalgacık Analizi

Son zamanlarda özellikle zaman serisi, veri ve görüntü analizi gibi alanlarda kullanılan dalgacık analizi, çeşitli araştırma alanlarında kullanılabilen matematiksel bir araçtır. Zaman serileri, frekans, süre, yoğunluk ve zaman pozisyonu gibi yerel bilgiler ve farklı periyotlar boyunca ortalama durumları gibi global bilgiler ile temsil edilir. Yani bir sinyalin doğru analizi için hem global hem de yerel bilgi önemlidir (Alexandridis ve Zapraniş, 2014, s. 8).

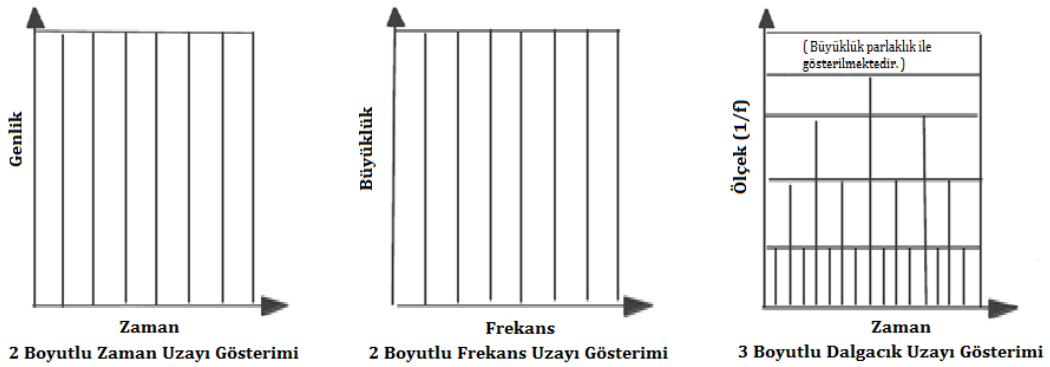
Sinyalin ham halinde bulunmayan bilgileri elde edebilmek için matematiksel dönüşümlerden yararlanılabilir. Zaman serilerinde sinyalin, zaman genlik (amplitude) düzlemi üzerindeki gösterimi elde edilir. Burada genlik, sinyalin zaman eksenindeki her değere karşılık aldığı sayısal değerdir. Ancak bu gösterim, sinyalin analizi için genellikle en iyi gösterim şekli değildir. Çoğu örnekte en önemli bilgi sinyalin frekans bileşeni içinde saklıdır (Polikar, t.y.). Frekans ise bir olayın birim zamanda görülen sıklık miktarı olduğundan, sinyalin frekans genlik düzlemi üzerindeki gösteriminde zaman değil frekans değerleri göz önüne alınır.

Sinyalin frekans bileşenlerine o sinyalin frekans spektrumu denir. Frekans spektrumu ise sinyalin içinde hangi frekansların olduğunu verir. Burada frekans bir değişkenin oranındaki değişimdir. Bu değişim ne kadar hızlı ise frekans o kadar yüksektir ya da değişim yoksa frekans yok yani sıfırdır (Polikar, t.y.).

Jean Baptiste Joseph Fourier 1822'de yayınlanan "Isının Analitik Teorisi" adlı kitabında, birbirinden farklı konu ve değişkenlere ait sinyallerin birbirleriyle karşılaştırılabilecekleri ortak bir dile çevrilebileceğini göstererek, sinyal olarak da

adlandırılan düzensiz periyodik bir fonksiyonun birbirini izleyen düzenli dalgalardan oluşan sinüs ve kosinüs fonksiyonlarının toplamı olarak ifade edilebileceğini ortaya koymuştur. Fourier dönüşümü olarak adlandırılan bu yöntem diğer alanlarda olduğu gibi iktisat ve finans alanlarında da kullanım alanı bulmuştur (Selçuk, 2005, s. 14).

Fourier dönüşümüyle zaman düzlemindeki bir sinyalin, frekans genlik düzlemindeki karşılığı elde edilir. Dönüşümle sinyal zaman uzayından (domain) frekans uzayına aktarılmış olur. Burada hangi frekansın ne kadar olduğu bilgisi elde edilebilirken, kullanılan sinüs ve kosinüs fonksiyonların özelliklerinden dolayı zaman bilgisi kaybolur. Fourier dönüşümü ile ölçeğe göre analiz yapılırken aynı anda zamana göre analiz yapma imkanı bulunmamaktadır. Ancak, temel (baz) fonksiyon olarak sınırlı uzunluğa ve değişken frekanslara sahip dalgacıkların kullanıldığı dalgacık analizi bu sorunun üstesinden gelebilmektedir. Zaman uzayı, frekans uzayı ve dalgacık uzayı karşılaştırılması Şekil 1'de gösterilmektedir.



Şekil 1 : Zaman, Frekans ve Dalgacık Uzayı Gösterimi (Fugal, 2009, s. 10).

Graps (1995, s. 50)'a göre; dalgacık analizinin arkasındaki temel fikir ölçeğe göre analiz etmektir. Veriyi farklı ölçek ve çözünürlüklerde işleyen dalgacık algoritmaları yardımıyla bir sinyale büyük bir pencereden bakıldığında bariz kaba özellikler, küçük bir pencereden bakıldığında ise daha küçük detay özellikler elde edilebilmektedir.

Dalgacık dönüşümü zaman uzayı ve frekans uzayını eş zamanlı analiz eden bir sinyal işleme tekniğidir. Özellikle, dalgacık dönüşümü bir zaman serisini farklı çözünürlük ölçeklerinde alt dizilerine ayrıştırarak, verilen veriyi yüksek ve düşük

frekanslı bileşenler halinde parçalar. Yüksek frekansta, dalgacıklar orijinal verideki süreksizlikleri, kırılmaları ve özellikleri yakalayabilir. Düşük frekansta ise dalgacıklar uzun vadeli trendleri belirlemek için verinin kaba yapısını karakterize eder. Böylece dalgacık analizi, orijinal verilerin gizli ve önemli zamansal özelliklerini ayıklama olanağı sağlar (Lahmiri, 2014, s. 219).

Bir tür matematiksel mikroskop işlevi gören dalgacık analizi ile veri kısa periyotlarda, düşük ölçekte, yüksek çözünürlükte en ince detayına kadar incelenirken aynı zamanda uzun periyotlarda, yüksek ölçekte, düşük çözünürlükte incelenerek veriye kuşbakışı bakma olanağı sağlanmaktadır (Selçuk, 2005, s. 14). Böylece, sinyalin ham halinde veya zaman uzayında görülemeyen özelliklerinin, dalgacık analizi ile elde edilen ölçek zaman düzleminde, düşük frekanslarda karakteristik özellikler ve yüksek frekanslarda ise detay özellikler elde edilerek incelenebilme imkanı ortaya çıkacaktır.

Dalgacık analizinde, dalga şeklindeki bir fonksiyon Fourier analizde olduğu gibi, orijinal fonksiyonu (sinyali) incelemek için kullanılmaktadır. Fourier yönteminden farklı olarak, dalgacık analizinde; incelemede yararlanılan dalga (ana dalgacık), önce sıkıştırılarak çok kısa bir periyoda sığdırılır ve incelenen fonksiyon (sinyal) ile zaman içerisinde kaydırılarak başından sonuna kadar karşılaştırılır. Böylece, incelenen sinyalin o kısa periyotlardaki özellikleri elde edilebilmektedir. Sonrasında, dalga biraz daha uzun bir periyoda yayılarak tekrar baştan aynı işlem tekrarlanır. Bu kez farklı şekilde, bir öncekine göre biraz daha uzun bir periyottaki özellikler elde edilir ve en sonunda çekiştire çekiştire uzatılarak serinin mümkün olan en uzun periyottaki özelliği saptanmaktadır. Yapılan işlem neticesinde incelenen seri değişik ölçeklerde analiz edilmekte, değişik periyodik özelliklerin zamana göre nasıl bir değişim gösterdiği saptanmakta ve serinin farklı çözünürlüklerdeki resimleri elde edilmektedir (Selçuk, 2005, s. 14).

Dalgacıkların tarihi incelendiğinde 1930 öncesi dönem, 1930-1960 arası dönem, 1960-1980 arası dönem ve 1980 sonrası dönem olarak dört döneme ayrılabilir (Graps, 1995, ss. 51–53). Fourier'in “Isının Analitik Teorisi” adlı çalışmasında önerdiği Fourier dönüşümü ile zaman ve ölçek (frekans) bilgisinin aynı anda analiz edilememesi sorununa çözüm geliştirebilmek için; Haar (1910)'ın

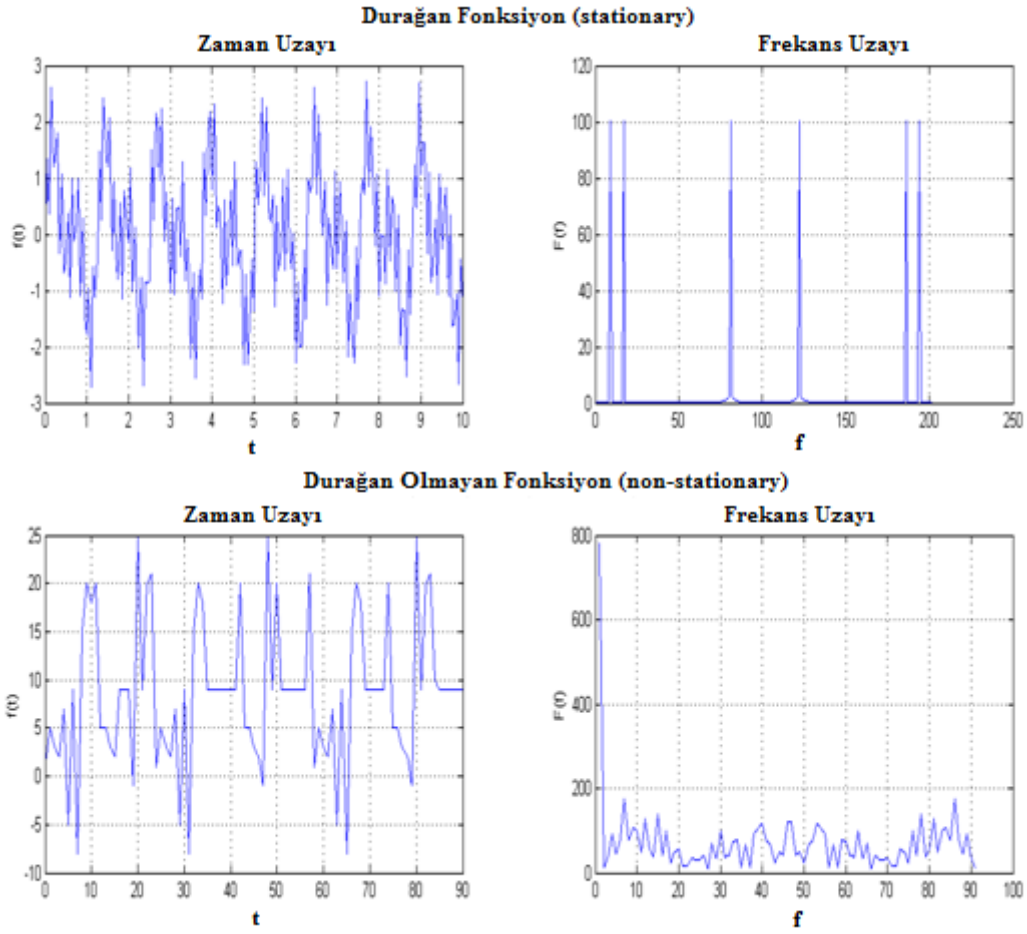


dalgacık terimini bulması ve en ilkel dalgacık türü olan ve kendi adını taşıyan Haar dalgacığını geliştirmesi 1930 öncesi döneminin en önemli olayıdır. 1930'lu yıllarda ise; Littlewood, Paley ve Stein'e ait enerji fonksiyonu ve ölçeği değişen baz fonksiyonları üzerine çalışmalar yapılmıştır (Graps, 1995, s. 52). 1960-1980 dönemindeki en önemli gelişme ise Weiss ve Coifman'ın, fonksiyon uzayının en temel elemanları üzerinde çalışarak atomsal ölçekleme alanında gelişim sağlamalarıdır. Esteban ve Galand'ın 1977'de ortaya attığı yeni bir filtre kavramı ilgi görmüş fakat orjinal sinyalin sentezinde hata oranının oldukça yüksek olduğu görülmüştür (Çifter, 2010, s. 47). 1980 sonrası dönemde dalgacık türleri hızla artmış, dalgacık terimi 1984 yılında kuantum fiziği çalışmalarında Morlet ve Grossman tarafından ilk kez kullanılmıştır. Daha sonra, Daubechies (1988) günümüzdeki birçok uygulamada kullanılan ortogonal dalgacık serileri ailesini ortaya atmıştır. Mallat (1989)'ın dalgacık ve süzgeç grupları arasındaki ilişkiyi ortaya koyan çok çözünürlüklü analizi, Meyer (1993)'in kendi adıyla anılan sürekli türeve sahip dalgacıkları da önemli gelişmeler arasındadır. Türevi sürekli olmayan Haar dalgacıklarının aksine Meyer dalgacıklarının türevi sürekli fakat kompakt desteği yoktur (Graps, 1995, s. 53).

### **2.1.2. Fourier Dönüşümü (FD)**

Sinyal işleme alanında 1807'de Fransız matematikçi Joseph Fourier tarafından ortaya atılan Fourier analizi sayesinde sinyallerin spektrumları incelenerek frekans uzayında özelliklerin ortaya çıkarılabilmesi sağlanmaktadır. Ayrıca, periyodik her fonksiyonun sinüs ve kosinüs trigonometrik fonksiyonlarından yararlanarak trigonometrik seriler halinde ifade edilebileceği ortaya konulmuştur (Çankaya ve Vatansever, 2002, s. 14). Fourier analizi, hem periyodik, yani belli zaman aralıklarıyla kendini tekrar eden, fonksiyonları analiz etmede kullanılan Fourier serilerini, hem de uzayda tanımlı tüm periyodik olmayan fonksiyonların analizi için de geçerli olan genel hali Fourier dönüşümünü (integralini) kapsamaktadır (Goswami ve Chan, 2011, s. 34). Fourier analizinde, sinyal öncelikle frekanslar olarak sinüslere ayrıştırılır ve sinyal durağan ise yani zaman içinde fazla değişim göstermiyorsa, zaman bilgisinin kaybolması önemli görülmemektedir. Doğaları gereği ani geçişler, yapısal kırılmalar, dönemsel dalgalanmaları içermesi nedeniyle finansal zaman serilerinin birçoğu durağan değildir (Çifter, 2010, s. 47). Bu nedenle Fourier analizi

finansal zaman serileri için uygun görülmemektedir. Şekil 2'de durağan ve durağan olmayan fonksiyonların zaman ve frekans uzayı gösterimleri verilmiştir.



Şekil 2 : Durağan ve Durağan Olmayan Fonksiyonların Zaman ve Frekans Uzayı Gösterimi.

Kısaca, sinyali çeşitli frekanslarda sinüsoid bileşenlerinin bir lineer kombinasyonuna ayrıştırması özelliğiyle Fourier dönüşümü, sinyalin frekans-frekans bazında çözünürlüğü olarak ifade edilmektedir (Alexandridis ve Zapraniş, 2014, s. 8). Fourier'in sonuçlarının ortaya koyduğu şekilde  $-\pi \leq x \leq \pi$  aralığında tanımlı  $2\pi$  periyodunda herhangi bir  $f(x)$  fonksiyonunun Fourier serisi, kosinüs ve sinüs fonksiyonlarının sonsuz toplamı olarak ifade edilebilir (Boggess ve Narcowich, 2009, s. 38).

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \quad (2-1)$$

Burada; uygun bir şekilde hesaplanan  $\{ a_0, a_1, b_1, \dots \}$  katsayılarının kümesi karmaşık (kompleks) bir dizidir. Eşitlik (2-1)'de görüldüğü gibi, Fourier serisiyle, herhangi bir sinyal sinüs ve kosinüsün bir fonksiyonu olarak ifade edilebilir (In ve Kim, 2012, s. 2). Eşitlikteki Fourier katsayıları olarak adlandırılan  $a_0, a_k$  ve  $b_k$  katsayıları aşağıdaki gibi ifade edilmektedir (Bogges ve Narcowich, 2009, s. 44):

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx \quad (2-2)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \quad (2-3)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx \quad (2-4)$$

Fourier serisinin genel hali olarak da kabul edilen Fourier dönüşümü üstel olarak ifade edilmektedir. Euler formülünden (2-5) yola çıkarak  $i$  bir karmaşık sayı olmak üzere aşağıdaki denklem (2-6) elde edilir (Alexandridis ve Zapranis, 2014, s. 9):

$$e^{inx} = \cos(nx) + i \sin(nx) \quad (2-5)$$

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \quad (2-6)$$

Burada Fourier katsayıları da aşağıdaki gibidir.

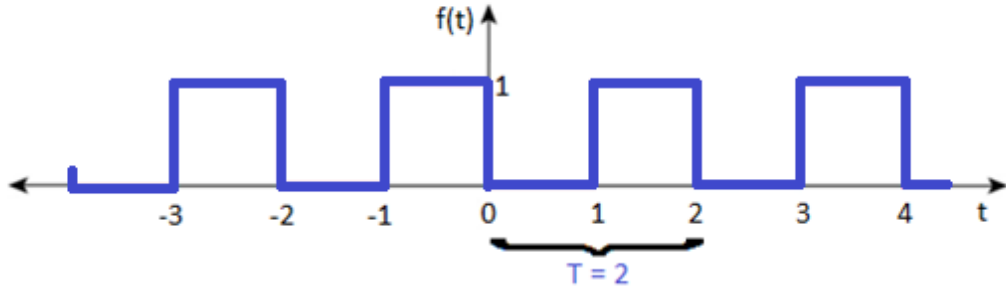
$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx. \quad (2-7)$$

Temel periyodu T olan herhangi bir  $f = x(t)$  fonksiyonu için frekans  $\omega_n = n/T$  ve  $e^{2\pi i \omega_n} = \cos(2\pi \omega_n) + i \sin(2\pi \omega_n)$  olmak üzere Fourier katsayıları ve Fourier dönüşümünün tersi kullanılarak tekrar oluşturulan  $x(t)$  fonksiyonu aşağıdaki gibidir (Alexandridis ve Zapraniş, 2014, s. 9):

$$c_n = \int_{-\infty}^{\infty} e^{-2\pi i \omega_n t} x(t) dt \quad (2-8)$$

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi i \omega t} \hat{x}(\omega) d\omega \quad (2-9)$$

Örneğin, aşağıdaki Şekil 3'te grafik gösterimi verilen fonksiyonun Fourier serisi gösterimi hesaplanıp incelensin.



Şekil 3 : Örnek Fonksiyon Grafiği.

Öncelikle fonksiyonun periyodu ve tanımı belirlenir. Şekilden de anlaşılacağı üzere fonksiyonun periyodu  $T = 2$  'dir. O halde temel frekans  $\omega = 2\pi/T$  'den  $\omega = \pi$  elde edilir. Fonksiyonun tanımı ise:

$$f(t) = \begin{cases} 0, & 0 < x < 1 \\ 1, & 1 < x < 2 \end{cases} \quad f(t + 2) = f(t); \quad (2-10)$$

şeklindedir.  $f(t)$  fonksiyonun Fourier katsayılarını elde etmek için  $a_0, a_n$  ve  $b_n$  katsayıları hesaplanır.

$$a_0 = \frac{2}{T} \int_0^T f(t) dt \text{ ise } a_0 = \frac{2}{2} \int_0^2 f(t) dt = \int_0^1 0. dt + \int_1^2 1. dt \quad (2-11)$$

olacağından (2-11) eşitliğindeki integraller hesaplandığında;  $0 + 1 = 1$  elde edilir.

$$a_n = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt \quad (2-12)$$

$$= \int_0^1 0. dt + \int_1^2 1. \cos(n\pi t) dt = \left[ \frac{\sin(n\pi t)}{n\pi} \right]_1^2 \quad (2-13)$$

$$= \frac{\sin(2n\pi)}{n\pi} - \frac{\sin(n\pi)}{n\pi} \quad (2-14)$$

ise  $n$  tamsayı ve  $\sin(n\pi) = 0$  olduğundan  $a_n = 0$  'dır.

$$b_n = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt \quad (2-15)$$

$$= \int_0^1 0. dt + \int_1^2 1. \sin(n\pi t) dt = \left[ -\frac{\cos(n\pi t)}{n\pi} \right]_1^2 \quad (2-16)$$

$$= -\frac{\cos(2n\pi)}{n\pi} + \frac{\cos(n\pi)}{n\pi} = \frac{-1 + \cos(n\pi)}{n\pi} \quad (2-17)$$

$\cos(2n\pi) = 1$  ve  $\cos(n\pi) = (-1)^n$  olduğundan  $b_n = \frac{-1+(-1)^n}{n\pi} = \begin{cases} -\frac{2}{n\pi}, & n \text{ tek} \\ 0, & n \text{ çift} \end{cases}$  elde

edilir.

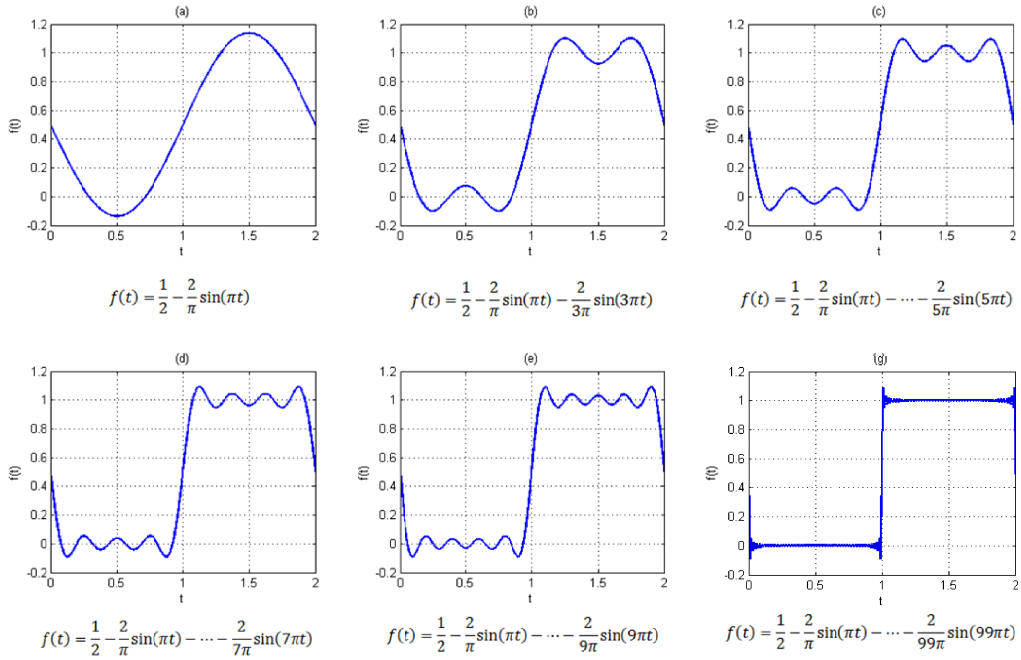
Fourier serisi ise;

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t)) \quad (2-18)$$

$$f(t) = \frac{1}{2} + \sum_{n=1}^{\infty} \left[ \frac{-1 + (-1)^n}{n\pi} \right] \sin(n\pi t) \quad (2-19)$$

$$f(t) = \frac{1}{2} - \frac{2}{\pi} \sin(\pi t) - \frac{2}{3\pi} \sin(3\pi t) - \frac{2}{5\pi} \sin(5\pi t) - \frac{2}{7\pi} \sin(7\pi t) - \dots \quad (2-20)$$

dir. Elde edilen bu toplam (2-20) bize orjinal fonksiyonu verecektir.



Şekil 4 : Fourier Serisi Terimlerinin Toplam Grafiği.

Yukarıdaki Şekil 4'te (a)'da  $n = 1$  için, (g)'de ise  $n = 99$  için Fourier serisi açılımının grafiği görülmektedir. Buradan, Fourier serisi toplamında  $n$  sayısı büyüdükçe elde edilen grafiğin orjinal fonksiyonun grafiğine benzemekte olduğu ve  $n$  sonsuza yaklaştıkça orjinal grafiğe yaklaşacağı görülebilmektedir.

Yukarıdaki fonksiyonun karmaşık Fourier serisi hesaplanırsa;

$$c_0 = \frac{1}{T} \int_0^T f(t) dt = \frac{1}{2} \int_0^2 f(t) dt = \frac{1}{2} \int_0^1 0. dt + \frac{1}{2} \int_1^2 1. dt = \frac{1}{2} \quad (2-21)$$

$$c_n = \frac{1}{T} \int_0^T f(t) e^{-jn\omega t} dt = \frac{1}{2} \int_0^1 0. dt + \frac{1}{2} \int_1^2 e^{-jn\omega t} . dt, \quad \omega = \pi; \quad (2-22)$$

$$c_n = \frac{1}{2} \left[ \frac{e^{-jn\pi t}}{-jn\pi} \right]_1^2 = \frac{e^{-jn\pi} (e^{-jn\pi} - 1)}{-2jn\pi} \quad (2-23)$$

Ancak;

$$e^{-jn\pi} = \cos(n\pi) - j. \sin(n\pi) = \cos(n\pi) = (-1)^n, \quad \sin(n\pi) = 0. \quad (2-24)$$

olduğundan;

$$c_n = \frac{j}{2n\pi} (1 - (-1)^n) = \begin{cases} \frac{j}{n\pi} & , \quad n \text{ tek} \\ 0 & , \quad n \text{ çift} \end{cases} \quad (2-25)$$

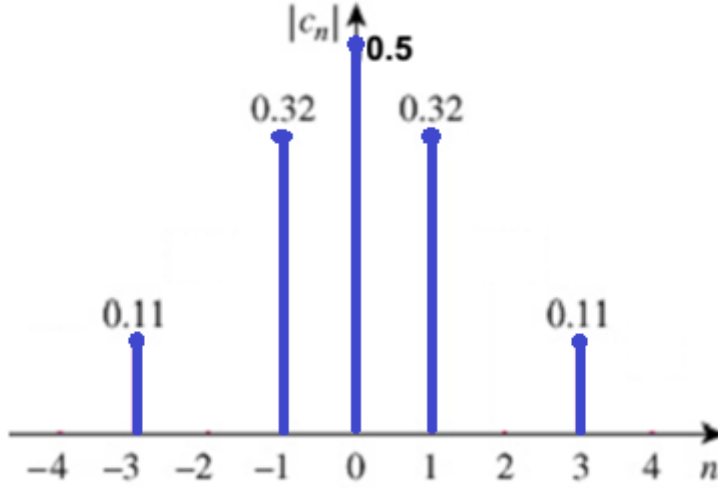
Dolayısıyla,

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega t} = \frac{1}{2} - \sum_{\substack{n=-\infty \\ n \text{ tek}}}^{\infty} \frac{j}{n\pi} e^{jn\pi t} \quad (2-26)$$

elde edilir.

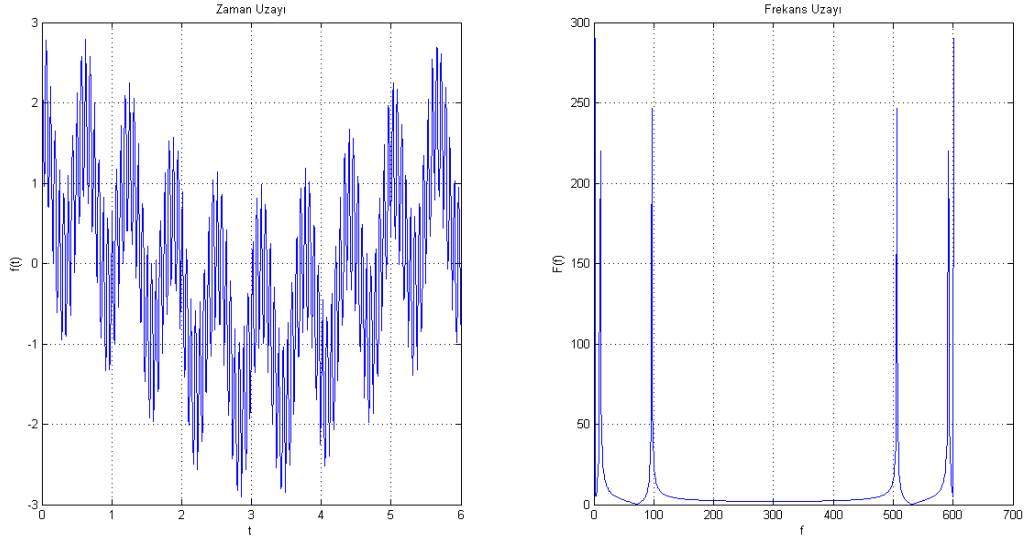
$$c_0 = \frac{1}{2} \text{ ve } |c_n| = \begin{cases} \frac{1}{n\pi} & , \quad n \text{ tek} \\ 0 & , \quad n \text{ çift} \end{cases} \quad (2-27)$$

katsayılarından yararlanarak Şekil 5'teki grafik elde edilir.



Şekil 5 : Fourier Katsayıları Grafiği.

Yine bir diğer örnek olarak  $f(t) = \cos(t) + \cos(10t) + \cos(100t)$  fonksiyonunun grafiği ve Fourier dönüşümü aşağıdaki Şekil 6'da verilmiştir.

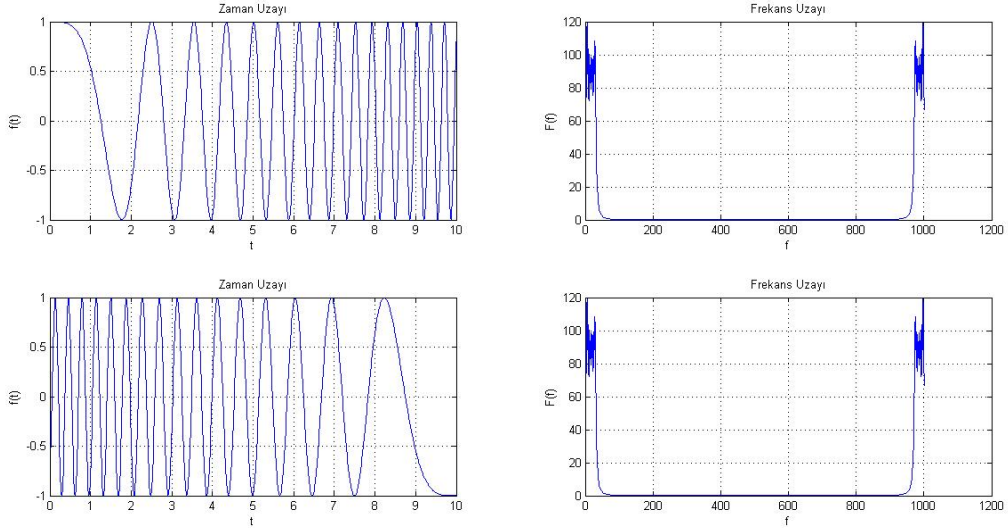


Şekil 6 :  $f(t) = \cos(t) + \cos(10t) + \cos(100t)$  Fonksiyonunun Fourier Dönüşümü.

Şekil 7'deki grafiklerde görüldüğü üzere,  $f_1(t) = \cos(\omega t^2)$  ve  $f_2(t) = -\cos(\omega(10 - t)^2)$  gibi birbirinden farklı iki fonksiyonun Fourier dönüşümleri incelendiğinde; zamana göre fonksiyonların aldığı değerler farklı olmasına rağmen Fourier dönüşümüyle elde edilen frekans uzayındaki değerlerinin



aynı olduğu görülmektedir. Bu örnekte de görüldüğü gibi Fourier dönüşümünün dezavantajı olan zaman bilgisinin kaybolmasının önüne geçebilmek için 1946 yılında Gabor Kısa zaman Fourier dönüşümünü bir diğer adıyla Pencerelemiş Fourier dönüşümünü geliştirmiştir.



Şekil 7 : Birbirinden Farklı İki Fonksiyonun Fourier Dönüşümü.

### 2.1.3. Kısa Zaman Fourier Dönüşümü (KZFD)

Durağan fonksiyonlar (sinyaller) için Fourier dönüşümünün yeterli bilgi sağladığı söylenebilir. Frekans uzayında fonksiyon ile ilgili zaman uzayında elde edilemeyen frekans bileşenleri elde edilebilmektedir.

Durağan olmayan bir fonksiyonu ayrıştırmak, analiz etmek ve yorumlamak için ve zaman içinde fonksiyonun spektral içeriğindeki değişiklikler hakkında bilgi sağlaması için, bir zaman-frekans gösterimi gereklidir. Bunu elde etmek için, zaman uzayında fonksiyon küçük zaman aralıklarına bölünüp, bu zaman aralıklarında fonksiyonun durağan olduğu kabul edilerek her aralık için ayrı ayrı Fourier dönüşümü alınır. Fonksiyonu dilimleme, uzunluğu fonksiyonun yaklaşık olarak durağan olduğu aralığa göre belirlenmiş pencere fonksiyonu ile yapılır. Kısa zaman Fourier dönüşümü olarak adlandırılan bu dönüşümde (Şekil 8) zaman aralıkları daraltıldığında daha iyi zaman çözünürlüğü ve daha kötü frekans çözünürlüğü elde edilir. Aynı şekilde zaman aralıkları genişletildiğinde daha iyi frekans çözünürlüğü

elde edilirken daha kötü zaman çözünürlüğü elde edilir. Zaman aralığı sınırsız uzunlukta alınması ise, en iyi frekans çözünürlüğün elde edildiği standart Fourier dönüşümüne karşılık gelir (Radunovic, 2010, s. 26).



Şekil 8 : Kısa Zaman Fourier Dönüşümü

(Çifter, 2010, s. 48).

Bu nedenle, KZFD'nde giriş sinyali  $f(x)$  çeşitli bölümlere ayrılır ve her bölüm ayrı ayrı frekans içeriği için analiz edilir. Pencerenin etkisi, sinyali zaman içinde lokalize etmektir. KZFD zaman ile frekans arasında bir çeşit uzlaşmayı temsil ederek, sinyaldeki bir olayla ilgili hem ne zaman hem de hangi frekansların olduğu konusunda bilgi sağlamaktadır. Ancak, sadece pencerenin boyutuna göre belirlenen sınırlı hassasiyetle bu bilgi elde edilebilir. Zaman ve frekans bilgileri arasında KZFD uzlaşması yararlı olabilir ancak zaman penceresi için belirli bir boyut seçildiğinde o pencerenin tüm frekanslar için aynı kalması dezavantaj olarak ortaya çıkmaktadır (In ve Kim, 2012, s. 4). Belirlenen sabit pencere genişliği içinde meydana gelen bir olayın çözümlenemeyecek olması KZFD'nün eksikliği olarak görülmektedir (Gençay, Selçuk ve Whitcher, 2002, s. 99).

Fourier analizi, periyodik sinyallerin analizinde son derece iyi bir performans sergilemektedir. Ancak, sinyalin frekans uzayına dönüştürülmesi sırasında zaman bilgisi kaybolur. Fourier sinyal dönüşümüne bakıldığında, belli bir olayın gerçekleştiği zamanı söylemek mümkün değildir. Sinyal durağan olmayan ya da sürüklenmeler, trendler, ani değişiklikler, ya da olayların başlangıç ve bitişi gibi geçici özellikler içeriyorsa, yani zamana göre sinyalin özellikleri değişiklikler gösteriyorsa, bir zaman serisinin en önemli parçası olan bu özellikleri Fourier dönüşümü ile tespit edilememektedir (Alexandridis ve Zapraniş, 2014, s. 9). Klasik

Fourier dönüşümünün bu tür sorunları aşmak için Gabor, Fourier dönüşümünü pencere olarak adlandırılan kısa zaman aralıklarına uygulamıştır (Mallat, 2008, s. 92).

Frekans ve zaman arasında bir çeşit denge elde edebilmek için Fourier dönüşümü pencerelemiş Fourier ya da kısa zaman Fourier dönüşümü şeklinde genişletilmiştir. Sinyali zaman üzerinde lokalize edebilmek için KZFD'nde  $g(u) = g(-u)$  olacak şekilde simetrik bir pencere fonksiyonu kullanılır.  $t \in \mathbb{R}$ , olacak şekilde (2-28) eşitliğindeki gibi tanımlanır (Alexandridis ve Zapraniş, 2014, ss. 9–10).

$$f_t(u) = \bar{g}(u - t)f(u) \quad (2-28)$$

Burada  $f_t(u)$  fonksiyonu,  $f(u)$  fonksiyonunun  $g(u)$  pencere fonksiyonu ile lokalize edilmiş halidir. Böylece KZFD aşağıdaki gibi ifade edilmektedir (Kaiser, 2011, s. 45).

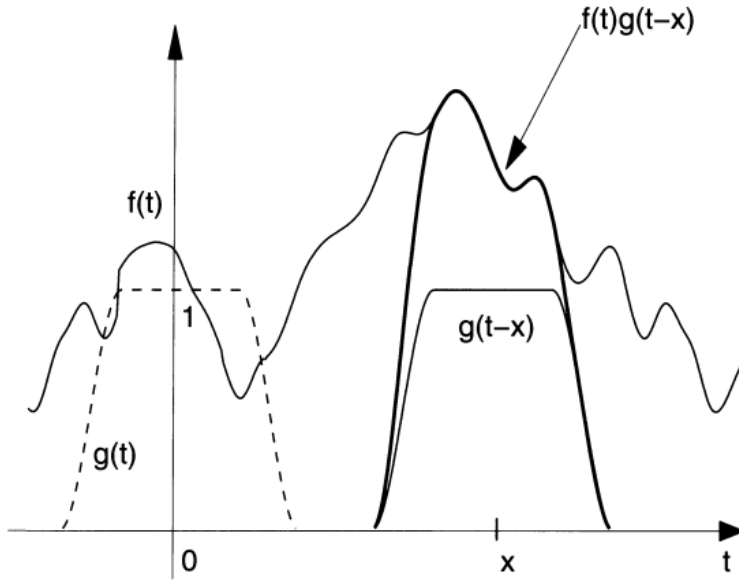
$$KZFD(\omega, x) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t} \hat{g}(t - x)dt \quad (2-29)$$

Yukarıdaki eşitlik (2-29)'da  $\omega$  frekansında  $x$ 'in komşuluğunda  $g(t-x)$  penceresi ile lokalize edilmiş fonksiyonun Fourier dönüşümü ifade edilmektedir. Burada  $g(t)=1$  alınırsa KZFD standart Fourier dönüşümüne indirgenecektir (Mallat, 2008, s. 16). KZFD'nün tersi ise  $C = \|g\|^2$  iken (1-30) eşitliğiyle ifade edilmektedir (Alexandridis ve Zapraniş, 2014, s. 10).

$$f(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} KZFD(\omega, x)e^{2\pi i\omega t} g(t - x)d\omega dx \quad (2-30)$$

$g_{\omega, x}(t) = e^{2\pi i\omega t}g(t - x)$  fonksiyonu bir dalga paketi gibi davranır ve bu dalga paketlerinin her biri birer baz fonksiyon olarak işlev görür. KZFD, zaman

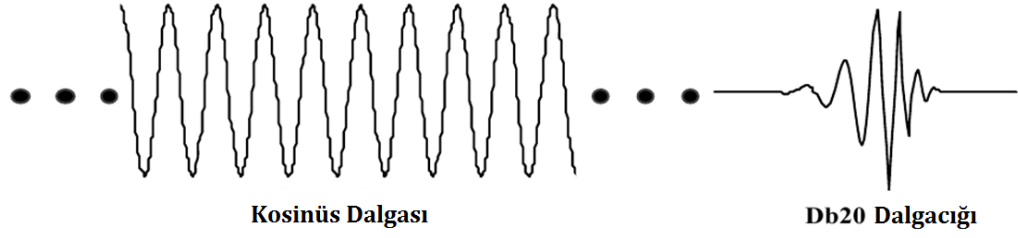
frekans düzleminde  $f(t)$  fonksiyonunun bu dalga paketlerini baz alan bileşenleri olarak yorumlanabilir (Goswami ve Chan, 2011, s. 65). KZFD, aynı zaman ve frekans çözünürlüğüne sahip dalga formları üzerinde sinyalleri ayırır. Ancak bu dönüşüm, sinyalde farklı zaman ve frekans çözünürlükleri ihtiva eden yapılar bulunmadığı zaman yeteri kadar etkili olmaktadır. Dalgacıkların kullanımı zaman ve frekans çözünürlüğünü değiştirerek bu sorunu gidermektedir (Mallat, 2008, s. 17). Şekil 9'da pencere fonksiyonunun sinyali taraması grafik olarak gösterilmektedir.



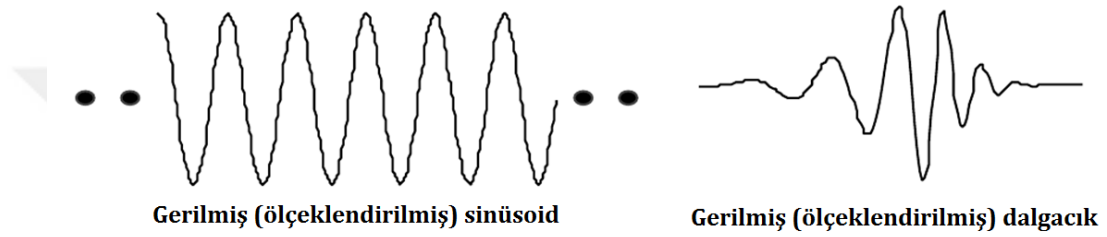
Şekil 9 : Kısa Zaman Fourier Dönüşümünün Sinyali Taraması (Gröchenig, 2001, s. 39).

#### 2.1.4. Dalgacık Dönüşümü (DD)

Bir dalgacık  $x$ -ekseninin altında ve üstünde dalgalanan, değişen frekanslara, sınırlı uzunluğa ve sıfır ortalamaya sahip bir fonksiyondur. Bu fonksiyonlar, FD tarafından kullanılan sonsuz enerjiye sahip sinüsoidlerle farklılık göstermektedir (Fugal, 2009, s. 2). Şekil 10'da sonsuz uzunluğa sahip kosinüs dalgasının bir parçası ile sonlu uzunluğa sahip bir dalgacığın karşılaştırılması ve Şekil 11'de ise, ölçeklendirilmiş formlarının karşılaştırılması gösterilmektedir.



Şekil 10 : Sonsuz Uzunluğa Sahip Kosinüs Dalgasının Bir Parçası ile Sonlu Uzunluğa Sahip Dalgacık  
(Fugal, 2009, s. 2).



Şekil 11 : Gerilmiş (Ölçeklendirilmiş) Sinüsoid ve Dalgacık  
(Fugal, 2009, s. 3).

Dalgacık analizi, ana dalgacık olarak adlandırılan özel bir dalgacık fonksiyonunu temel alır. Bir dalgacık ailesi ise  $\varphi$  ölçekleme fonksiyonunun (baba dalgacık) ve  $\psi$  dalgacık fonksiyonunun (ana dalgacık) ötelenmesi ve ölçeklendirilmesiyle oluşturulan ortogonal bazlı fonksiyonlar kümesidir. Buradaki  $\varphi$  baba dalgacık ile  $\psi$  ana dalgacık aşağıdaki eşitlikleri sağlayacaktır (Alexandridis ve Zapranis, 2014, ss. 11–12).

$$\int \varphi(t)dt = 1 \quad (2-31)$$

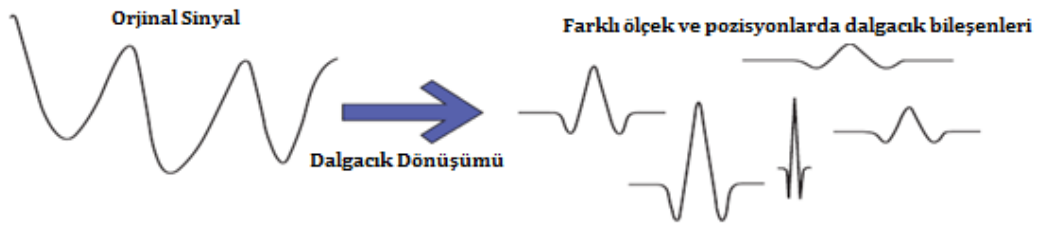
$$\int \psi(t)dt = 0 \quad (2-32)$$

Dalgacık ailesi aynı zamanda ana dalgacığın ötelenmiş ve ölçeklendirilmiş formlarını da kapsamaktadır. Burada, "a" ana dalgacığın ne kadar sıkıştırıldığını ya da geniştirildiğini belirten ölçek parametresi ve "b" de bu dalgacığın ne kadar

kaydırıldığını belirten pozisyon (öteleme) parametresi olmak üzere normalleştirilmiş dalgacık fonksiyonu  $\psi_{a,b}(t)$  (2-33) eşitliğinde verilmektedir (Goswami ve Chan, 2011, s. 72).

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2-33)$$

Mümkün olan her ölçekte dalgacık katsayılarını hesaplamak çok fazla işlem yoğunluğu gerektirdiğinden, çoğu durumda a ve b değerlerinin seçimi ayrık bir küme kullanılarak sınırlandırılır. Frekans analizi, ana dalgacığın gerilmiş, düşük frekanslı versiyonu ile yapılırken, zaman analizi aynı ana dalgacığın sıkıştırılmış, büzülmüş yüksek frekanslı versiyonu ile yapılır (Alexandridis ve Zapranis, 2014, s. 12).



Şekil 12 : Dalgacık Dönüşümü Yardımıyla Sinyalin Ayrıştırılması  
(Alexandridis ve Zapranis, 2014, s. 12).

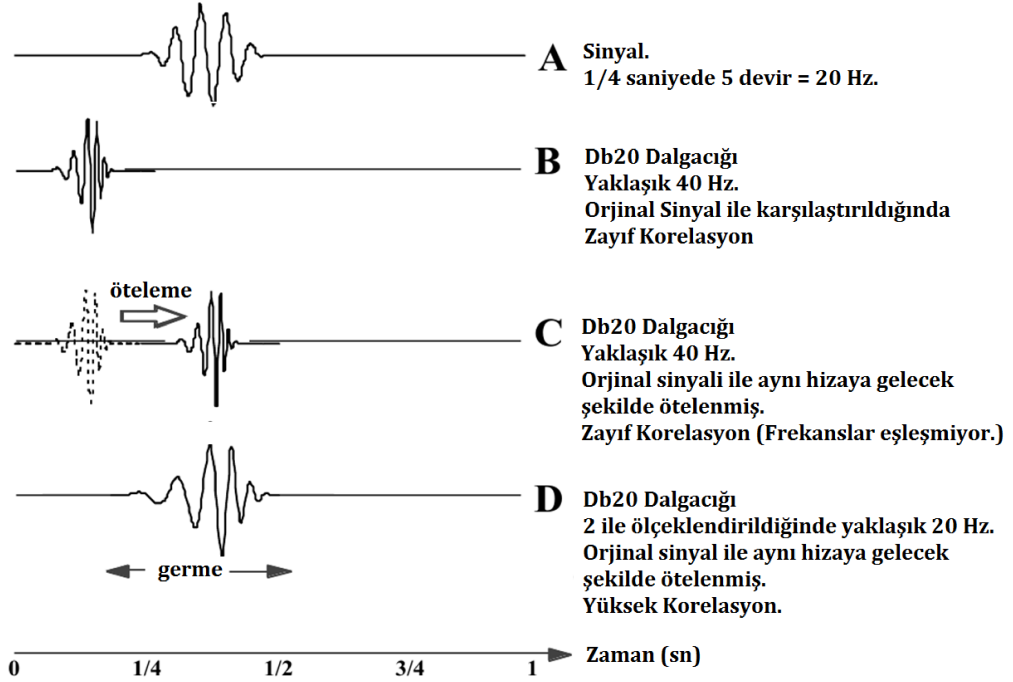
Dalgacık dönüşümü temel olarak ikiye ayrılmaktadır. Bunlardan birincisi belirlenen sınırlar arasındaki her ölçek ve pozisyon için hesaplamaların yapıldığı sürekli dalgacık dönüşümüdür. İkincisi ise; sürekli dalgacık dönüşümünde ortaya çıkan hesaplama yoğunluğunun üstesinden gelebilmek için ölçeklerden kritik örnekleme ile oluşturulan bir alt küme üzerinde yapılan hesaplamalardır ki bu yöntem ayrık dalgacık dönüşümü olarak adlandırılmaktadır.

### 2.1.4.1. Sürekli Dalgacık Dönüşümü (SDD)

Sürekli dalgacık dönüşümü fonksiyon ile ana dalgacığın iç çarpımı olarak tanımlanmaktadır (Radunovic, 2010, s. 29). Sürekli dalgacık dönüşümü sinyal ile ana dalgacığın ölçeklendirilmiş ve kaydırılmış tüm versiyonlarının çarpımının integrali olarak ifade edilebilir (In ve Kim, 2012, s. 11).

$$C(\text{ölçek}, \text{pozisyon}) = \int_{-\infty}^{\infty} x(t) \cdot \psi(\text{ölçek}, \text{pozisyon}, t) \cdot dt \quad (2-34)$$

Sürekli dalgacık dönüşümüyle ölçek ve pozisyonun birer fonksiyonu olan ve işaret ile dalgacığın benzerliğini ortaya koyan birçok C dalgacık katsayıları elde edilmektedir. Burada; ölçek ve pozisyon  $x_t$  zaman serisi boyunca uygun her değeri alabilmektedir (In ve Kim, 2012, s. 11).



Şekil 13 : Orjinal Sinyal ile Dalgacığın SDD Tipi Karşılaştırılması

(Fugal, 2009, s. 14).

Fourier dönüşümünde kullanılan sinüsoidlerin yerine dalgacıklar kullanılmaktadır. Yukarıdaki Şekil 13'te  $t = 0$ 'da başlayan ve 1/8 saniye uzunluğundaki Daubechies 20 dalgacı (B) ile ana sinyal (A) birebir nokta nokta karşılaştırıldığında çok düşük bir korelasyon değeri elde edilir. Dalgacık ötelenip sinyal ile aynı hizaya getirildiğinde de (C) orjinal sinyal ile dalgacığın frekansları eşleşmediği için düşük korelasyon elde edilir. Dalgacık hem ölçeklenerek frekanslar benzeştiğinde hem de öteleme ile orjinal sinyal ile aynı hizaya getirildiğinde (D) yapılan karşılaştırmada yüksek korelasyon elde edilmektedir (Fugal, 2009, s. 14).

Genel olarak ifade edilmek istenirse,  $a$  ölçek parametresi ve  $b$  öteleme parametresi olmak üzere sürekli dalgacık dönüşümü aşağıdaki şekilde yazılabilir (Addison, 2002, s. 12).

$$SDD(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \psi^* \left( \frac{t-b}{a} \right) \cdot dt ; a > 0. \quad (2-35)$$

Eşitlikteki  $a^{-1/2}$  normalizasyon parametresi  $\|\psi_{a,b}\| = \|\psi\|$  eşitliğini gerçekler.  $\psi$  dalgacık fonksiyonunun pencere fonksiyonu olması ve  $f(t)$  fonksiyonunun dalgacık dönüşümünün tersi ile elde edilebilmesi için aşağıdaki eşitliğin sağlanması gerekmektedir (Goswami ve Chan, 2011, s. 73).

$$\int_{-\infty}^{\infty} \psi(t) \cdot dt = 0. \quad (2-36)$$

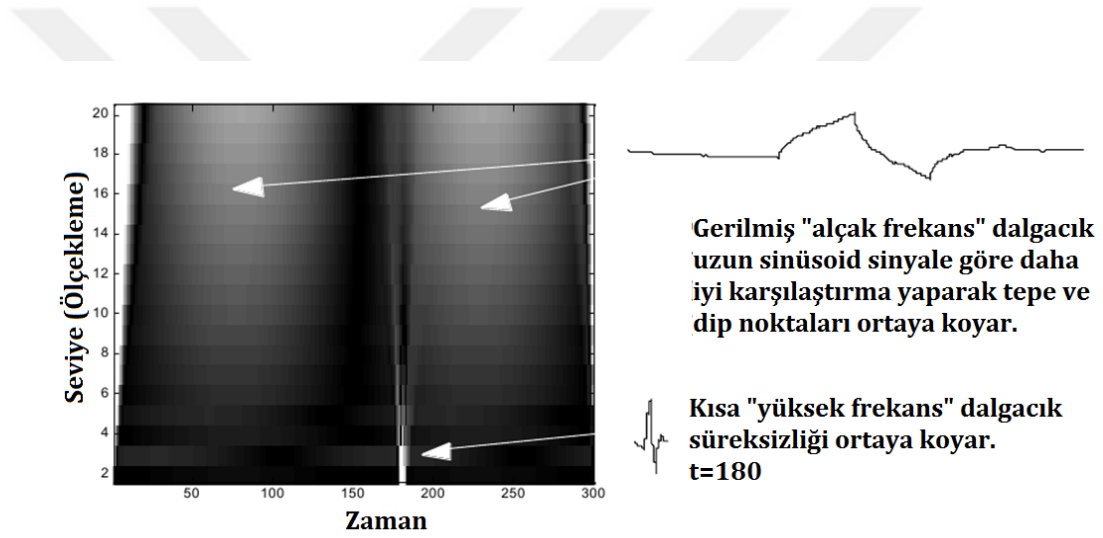
Sürekli dalgacık dönüşümünün tersi ise (2-37) eşitliğindeki çift katlı integral ile hesaplanmaktadır (In ve Kim, 2012, s. 12).

$$x(t) = \iint SDD(a, b) \cdot \psi \left( \frac{t-b}{a} \right) \cdot da \cdot db \quad (2-37)$$



Seçilen ana dalgacık ile incelenen sinyal arasındaki uyum derecesine göre elde edilen sürekli dalgacık dönüşümü katsayıları, dalgacık ile sinyal arasındaki ilişkiyi ortaya koymaktadır. Ana dalgacık gerildiğinde, dalgacığın uzunluğu artarak sinyaldeki düşük frekansları yakalar. Böylece sinyalin kaba yapısı ortaya çıkmaktadır. Tersine, ana dalgacığın sıkıştırılmasıyla yani daha dar bir dalgacıkla sinyal incelendiğinde yüksek frekanslar yani sinyaldeki detay bilgiler elde edilir.

Şekil 14'te SDD'de gerilmiş alçak frekans dalgacığın sinyaldeki tepe ve dip noktaları ortaya koyduğu, kısa yüksek frekans dalgacığın ise süreksizliği ortaya koyduğu görülmektedir (Fugal, 2009, s. 18).



Şekil 14 : SDD'de Sinyal ve Süreksizliğin Dalgacık Gösterimi  
(Fugal, 2009, s. 18).

#### 2.1.4.2. Dalgacıkların Ölçeklenmesi ve Kaydırılması

Bir dalgacığı ölçeklemek dalgacığı basitçe germe veya sıkıştırma anlamına gelir. Örneğin "germe" olarak  $\psi_a(t) = \psi\left(\frac{t}{a}\right)$  olmak üzere  $a$  bir ölçek faktörü alındığında,  $a$  değeri küçüldükçe daha sıkıştırılmış dalgacık elde edilir. Dalgacıkların ölçeği ile frekansı arasındaki ilişki doğaldır. Sıkıştırılmış dalgacık, detayların hızlı bir şekilde değişmesini sağlar. Böylece, küçük ölçekli bir dalgacık yüksek frekans salınımlarını yakalar. Tersine; gerilmiş, büyük ölçeğe sahip bir dalgacık ise

sinyaldeki alçak frekans hareketlerini yakalayabilmektedir. Bir dalgacığın kaydırılması basitçe ileriye ötelenmesidir. Matematiksel olarak  $\psi(t)$  'nin  $b$  kadar ötelenmesi  $\psi(t - b)$  olarak temsil edilir. Bu iki özellik kullanılarak, dalgacık dönüşümü sinyalin geniş bir frekans aralığındaki özelliklerini akıllıca ortaya çıkarmak için kendini adapte etme ve böylece zaman içinde yerel olayları yakalama yeteneğine sahiptir (In ve Kim, 2012, s. 12).

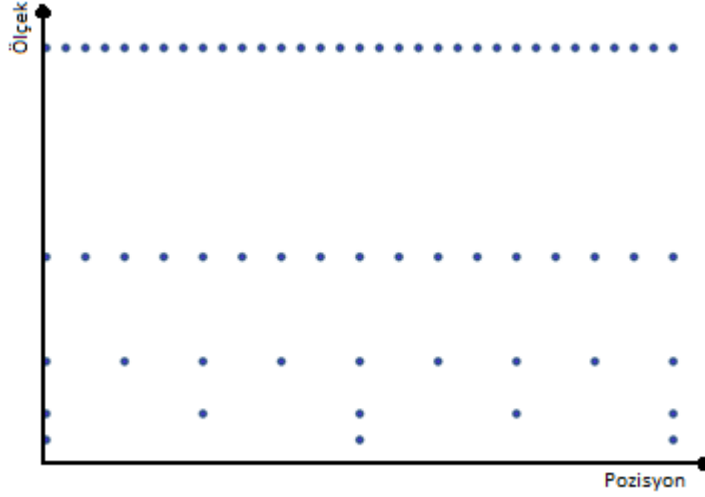
Genel olarak, dalgacıklar ortogonal ve ortogonal olmayan olarak ayrılabilir. Dalgacık fonksiyonu terimi her iki tür için de kullanılmaktadır. Ortogonal dalgacıklar kümesi bir dalgacık bazı olarak adlandırılırken, ortogonal olmayan dalgacıkların kümesi ise dalgacık çerçevesi olarak adlandırılmaktadır. Dalgacık çerçeveleri hem sürekli hem de ayrık dalgacık dönüşümde kullanılabilirken, ortogonal baz, ayrık dalgacık dönüşümünün kullanımını gerektirmektedir (Alexandridis ve Zapranis, 2014, ss. 12–13).

#### **2.1.4.3. Ayrık Dalgacık Dönüşümü (ADD)**

Sürekli dalgacık dönüşümünde dalgacık katsayılarını hesaplama sırasında mümkün olan her ölçek büyüklüğünün kullanılması zaman ve işlem maliyetini oldukça artıracaktır. Eğer ölçek ve öteleme değerlerinin ikinin kuvvetleri şeklinde bir alt kümesi alınrsa, daha verimli bir sonuç elde edilebileceği belirtilmektedir (Alexandridis ve Zapranis, 2014, s. 16).

Sürekli dalgacık dönüşümünün zaman-frekans düzleminde her noktasında değerler alması birçok gereksiz bilgiyi de barındırır. Aksine ADD'nde zaman-frekans düzleminde sadece çok az noktada değer kullanıldığı aşağıdaki Şekil 15'te görülebilmektedir (Ortega ve Khashanah, 2014, s. 60).

Sürekli dalgacık dönüşümünün sinyaldeki gizli bilgiyi adeta bir mikroskop gibi görünür hale getirmesi özelliğinin yanında ADD, KZFD'nün de yaptığı gibi sinyali çeşitli frekans bileşenlerine ayrıştırabilmektedir (Fugal, 2009, s. 19).



Şekil 15 : Zaman-Frekans Düzleminin Kritik Örnekleme  
(Ortega ve Khashanah, 2014, s. 61).

KZFD'nün tüm sinyal genelinde istenmeyen gürültüden kurtulmak için yaygın olarak kullanıldığı düşünülürse, ADD sinyalde belirli zamanlardaki frekans bileşenlerini ortadan kaldırmaya olanak sağlamaktadır. Bu durum sinyaldeki kötü ve gereksiz bilgiyi ortadan kaldırırken esas gerekli bilgiyi tutarak gürültü giderme veya veri sıkıştırma için güçlü bir kapasite ortaya koymaktadır. Ayrıca ADD'nün basitçe tersinin alınabilmesiyle orjinal sinyaldeki gürültü veya gereksiz verinin tespitine ve gürültüden arındırılmış sinyalin tekrar elde edilmesine olanak sağlanmaktadır (Fugal, 2009, s. 19).

ADD'nin matematiksel ifadesini elde etmek için SDD'nin (2-35) eşitliğindeki matematiksel ifadesinde, k ve s birer tamsayı olmak üzere ölçek parametresi a yerine  $2^{-s}$  ve pozisyon parametresi b yerine  $k \cdot 2^{-s}$  alınarak (2-38)'de verilen eşitlik elde edilmektedir. Böylece f(t) orjinal sinyalinin, ayrık (kesikli) hali elde edilmektedir. Örnekleme oranı da bir (1) kabul edilirse, (2-38)'deki eşitliğin integrali (2-39)'daki eşitlikle ifade edilebilmektedir (Goswami ve Chan, 2011, ss. 76–77).

$$ADD(2^{-s}, k2^{-s}) = 2^{\frac{s}{2}} \int_{-\infty}^{\infty} f(t) \psi(2^s t - k) dt \quad (2-38)$$

$$ADD(2^{-s}, k2^{-s}) \cong 2^{\frac{s}{2}} \sum_n f(n) \psi(2^s n - k) \quad (2-39)$$

Bir fonksiyonun zaman ölçek düzleminde bazı noktalarda dalgacık dönüşümünü elde etmek için fonksiyonun zaman ekseninde aldığı tüm değerleri bilinmek zorunda değildir. Burada ihtiyaç duyulan, dalgacığın sıfır olmadığı noktalarda fonksiyonun zaman eksenindeki değerleridir. Sonuç olarak dalgacık analizi zamana göre değişken doğasından dolayı neredeyse gerçek zamanlı yapılabilmektedir. Zaman ekseninde ötelenmiş bir fonksiyonun ADD'ü, orjinal fonksiyonun ADD'nden oldukça farklı olacaktır (Goswami ve Chan, 2011, ss. 76–77).

Temel dalgacık fonksiyonları (2-40) eşitliğindeki genel formda ifade edilebilmektedir. (Alexandridis ve Zapraniş, 2014, s. 16).

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k), \quad j, k \in Z \quad (2-40)$$

Genellikle bir sinyalin en önemli parçası, sinyale temel karakteristiğini veren sinyalin alçak frekans bileşenidir. Sinyalin yüksek frekans bileşeni ise genellikle sinyaldeki gürültünün büyük bir kısmını içeren parçadır. Sinyaldeki yüksek ölçek, alçak frekans bileşenleri yaklaşımlar olarak, düşük ölçek, yüksek frekans bileşenleri ise detaylar olarak adlandırılmaktadır. Her j seviyede, j-seviyede yaklaşım  $a_j$  (veya j. seviyede yaklaşım) ve sapmalar ise j-seviyede detay  $d_j$  (veya j. seviye detay) olarak ifade edilmektedir. Orjinal sinyal sıfırıncı seviye yaklaşım olarak ele alınarak  $a_0$  ile gösterilmekte ve  $a_1$ ,  $a_0$ 'ın alçak frekans bileşeni, aynı şekilde  $d_1$  de  $a_0$ 'ın yüksek frekans bileşeni olarak işlem görmektedir (Alexandridis ve Zapraniş, 2014, s. 17).

Dalgacık terminolojisinde "Ayrıştırma" bir yüksek geçiren ve bir alçak geçiren filtre kullanarak sinyali iki parçaya bölme anlamına gelir. Bu iki parçadan her biri yine aynı şekilde yüksek geçiren ve alçak geçiren filtreler yardımıyla ayrıştırılarak daha fazla sayıda parçalar elde edilebilir. "Yeniden Yapılandırma

(Sentez)" ise filtreleri kullanarak parçaları tekrar birleştirme işlemidir. Oluşturulan yeni sinyal baştaki orjinal halinin aynısı ise buna da "Mükemmel Yeniden Yapılandırma" denir (Fugal, 2009, s. 21).

ADD'nden yararlanarak sinyalin ayrıştırılması için çok çözünürlüklü analiz (1-41) eşitliğindeki gibi tanımlanmaktadır. Burada, her bir  $x_t$  dalgacık detay katsayılarının lineer kombinasyonu olmak üzere;  $N=2^J$  ve  $j=1, \dots, J$  olacak şekilde  $d_j = W_j^T w_j$  sinyalin  $\lambda_j$  ölçeğinde  $j$ .seviye dalgacık detaylarını belirtir.  $W_j^T w_j$  dalgacık analizinin  $\lambda_j$  ölçeğine atfedilen kısmı iken  $w_j = W_j x$  dalgacık katsayıları dalgacık analizinin  $\lambda_j$  ölçeğine dayandırılan kısmını temsil eder. Son dalgacık detay katsayısı  $d_{j+1}$  ise sinyalin ortalamasını verecektir (Gençay vd., 2002, s. 124).

$$x_t = \sum_{j=1}^{J+1} d_{j,t} \quad t = 0, \dots, N - 1 \quad (2-41)$$

Zaman serilerinde verilerin sınırlı uzunluğa sahip olmasından dolayı ölçek ve pozisyonların sınırlı aralıkta olması anlamlıdır. Çok çözünürlüklü analizi açıklamak için ölçek fonksiyonundan (baba dalgacık) başlamak gerekir. Çok çözünürlüklü analiz  $x_t$  zaman serilerini yaklaşık alt uzaylarına başarılı bir şekilde eşlemeyi barındırır. (2-42) eşitliğinde ifade edilen  $g(k)$  ölçekleme fonksiyonu katsayıları (alçak geçiren filtre) ve  $\sqrt{2}$  ise ölçekleme fonksiyonunun normunun ikili ölçeklenmesini sağlamaktadır (In ve Kim, 2012, s. 15).

$$\varphi(t) = \sum_k g(k) \sqrt{2} \varphi(2t - k), \quad k \in Z \quad (2-42)$$

Dalgacık fonksiyonu yukarıdaki denklemde tanımlanan  $\varphi(2t)$  ötelenmiş ölçekleme fonksiyonunun ağırlıklı toplamı ve  $h(k)$  katsayıları (yüksek geçiren filtre) ile ifade edildiğinde (2-43) eşitliği elde edilmektedir (In ve Kim, 2012, s. 17).

$$\psi(t) = \sum_k h(k)\sqrt{2}\varphi(2t - k), \quad k \in Z \quad (2-43)$$

Dalgacıklar analizinde bir diğer alternatif yöntem ise yukarıdaki (2-42) ve (2-43) eşitliklerinde gösterilen alçak geçiren ve yüksek geçiren filtrelerin göz önüne alınarak ayrıştırma yapılmasıdır. Burada önemli olan dalgacık analizinde kullanılacak filtrelerin nasıl üretilceğidir. Alçak geçiren ve yüksek geçiren filtreler baba ve ana dalgacık fonksiyonlarından aşağıdaki ilişkilerle elde edilebilmektedir (In ve Kim, 2012, ss. 17–20):

$$g(k) = \frac{1}{\sqrt{2}} \int \varphi(t)\varphi(2t - k) dt \quad (2-44)$$

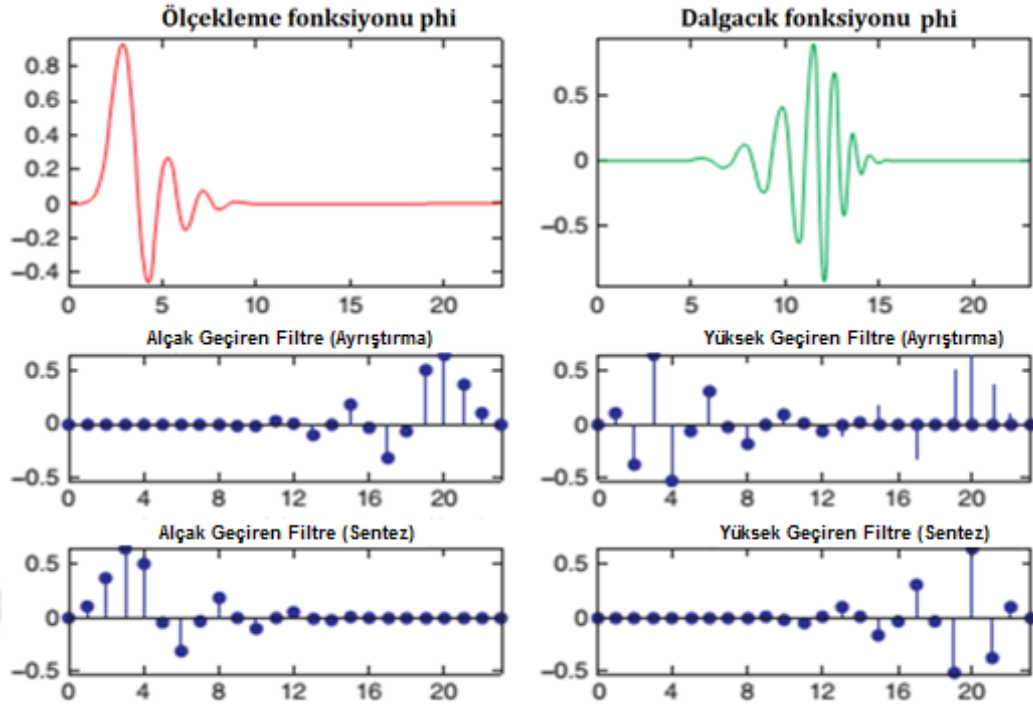
$$h(k) = \frac{1}{\sqrt{2}} \int \psi(t)\psi(2t - k) dt \quad (2-45)$$

veya

$$h(k) = (-1)^k g(k). \quad (2-46)$$

şeklinde yazılabilir.

Şekil 16'da Daubechies 12 dalgacık ve ölçekleme fonksiyonları ile alçak geçiren ve yüksek geçiren filtreleri görülebilmektedir.

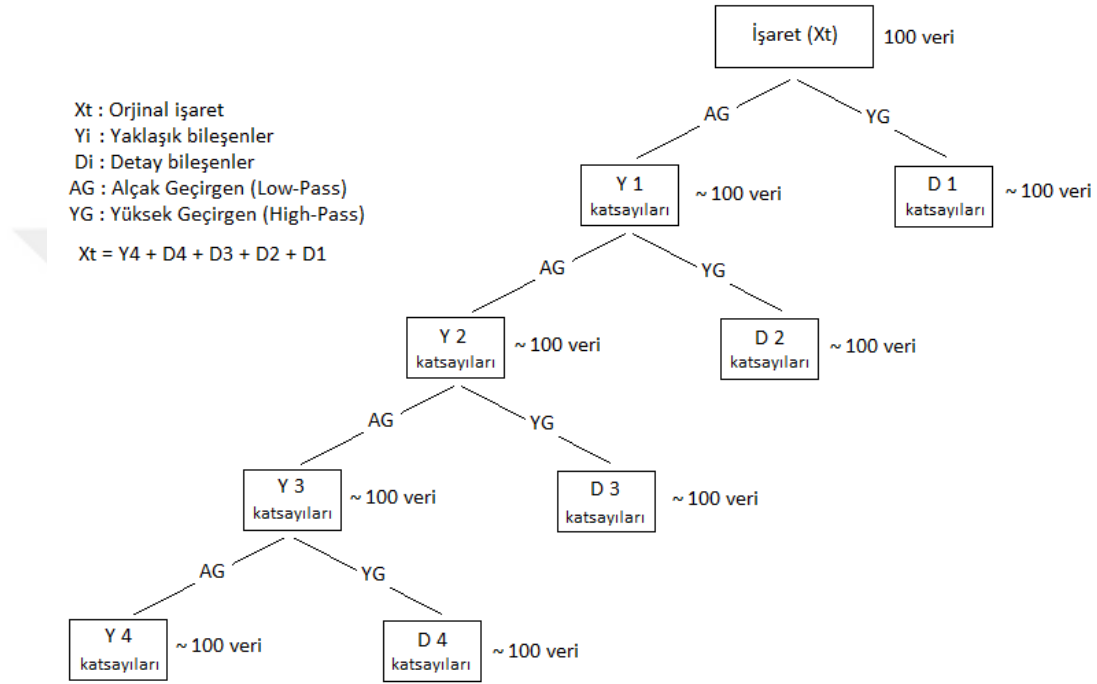


Şekil 16 : Daubechies 12 Dalgacık ve Ölçekleme Fonksiyonları ile Alçak Geçiren ve Yüksek Geçiren Filtreleri  
(Alexandridis ve Zapraniş, 2014, s. 15).

Ayrıştırılma işleminde örnekteki veri sayısının her defasında iki katına çıkmasını önüne geçmek için, her filtrelemede örneklem sayısı genellikle yarıya düşürülür. Dalgacık dönüşümüyle her seviyede alçak geçiren filtreleme ile elde edilen yaklaşım bileşenleri sinyalin daha yumuşatılmış, düzgünleştirilmiş halini verirken yüksek geçiren filtreleme ile elde edilen detay bileşenleri ise sinyaldeki gürültüyü ifade etmektedir (Fugal, 2009, s. 21).

Şekil 17'de  $Y_1$ , sinyale göre daha yumuşatılmış birinci seviyede yaklaşım katsayılarını verirken,  $D_1$  de bu katsayılardan arta kalan detay katsayılarını ifade etmektedir. Burada  $Y_1$  ve  $D_1$  katsayılarının toplamı orjinal sinyali veya işareti verecektir. Aynı şekilde ikinci seviye için elde edilen  $Y_2$  yaklaşım katsayıları birinci seviyede elde edilen  $Y_1$  yaklaşım katsayılarının alçak geçiren filtre ile yumuşatılmış hali olduğu ve  $D_2$  katsayılarının da yüksek geçiren filtre ile elde

edilen detay katsayıları olduğu göze alınırsa, Y2 ve D2 katsayılarının toplamının Y1 katsayılarını vereceği görülecektir. Buradan Y2, D2 ve D1 katsayılarının toplamının da orjinal işareti vereceği ve bu işlemin diğer seviyeler için de uygulanabileceği ifade edilebilir.

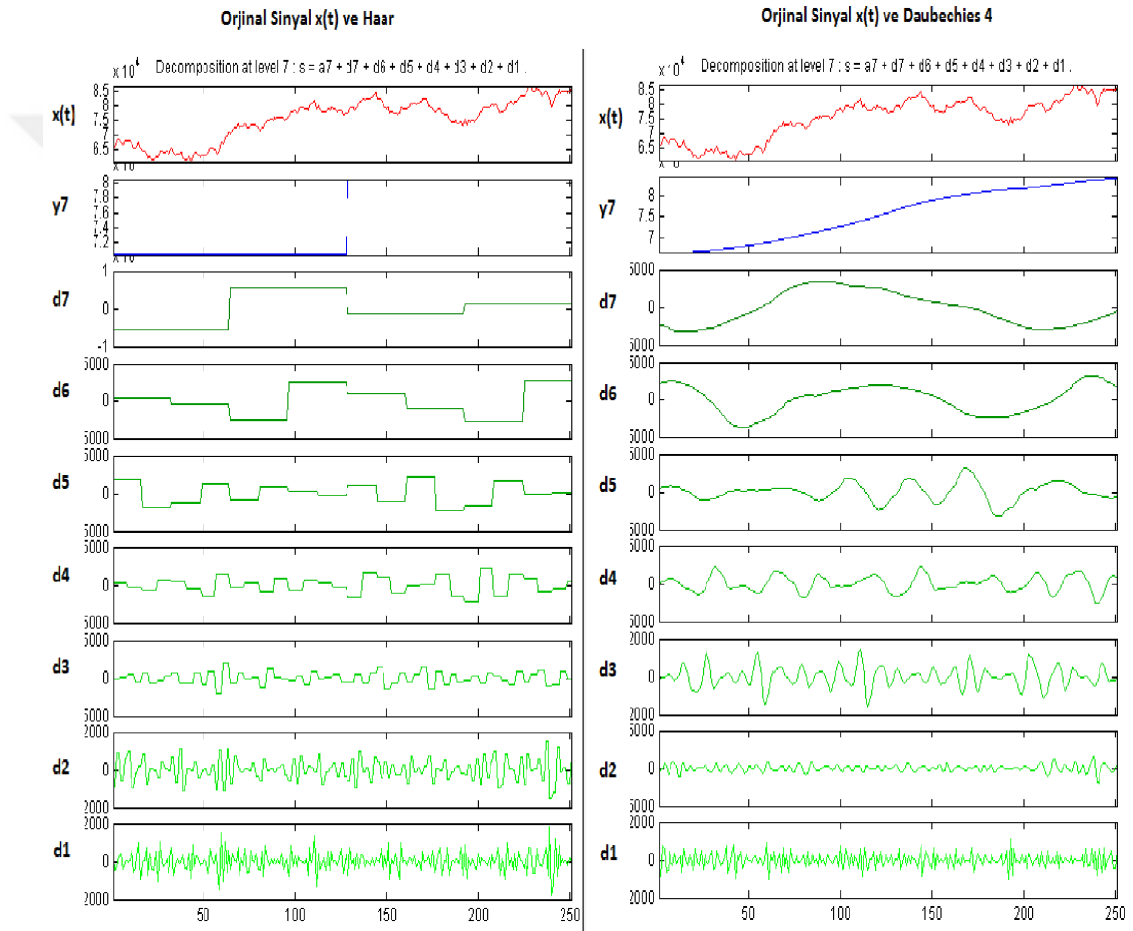


Şekil 17 : Ayrık Dalgacık Dönüşümünün Dört Seviyeli Ayrıştırılması.

Dalgacık dönüşümünde günümüzde ayrık veri ile çalışan dijital bilgisayarlar göz önüne alındığında sürekli dalgacık dönüşümü hakkında neyin sürekli olduğu sorusu düşünülürse ayrık dalgacık dönüşümü ve sürekli dalgacık dönüşümü hesaplanırken her ikisi de teknik olarak ayrık kabul edilir. SDD'nün farkı ölçekleme ve ötelemededir. SDD'ndeki sürekli terimi matematiksel sürekli fonksiyondan ziyade ölçekleme ve ötelemenin mümkün olan tüm tamsayı değerlerini (ör: 2, 3, 4, 5,...) belirtir. Ancak ADD'nde ölçekleme ve öteleme ikinin kuvvetleri (ör: 2, 4, 8,16,...) şeklindedir. Diğer bir farklılık ise SDD tek bir filtre kullanırken, ADD'nde ek olarak üç filtre daha kullanılmasıdır (Fugal, 2009, s. 19).



Dalgacık dönüşümünde Daubechies 4 dalgacık filtresi, Haar dalgacık filtresine göre daha pürüzsüz katsayılar üretmektedir (In ve Kim, 2012, s. 21). Şekil 18'de 2014 yılı BIST 100 endeksi günlük kapanış değerleri solda Haar ve sağda Daubechies 4 dalgacıkları yardımıyla bileşenlerine ayrıştırılmıştır. Katsayılar incelendiğinde Haar dalgacık filtresiyle oluşturulan katsayılara göre Daubechies 4 dalgacık filtresiyle oluşturulan katsayılar arasındaki geçişin daha yumuşak olduğu göze çarpmaktadır.



Şekil 18 : 2014 Yılı BIST 100 Endeksi Günlük Kapanış Değerlerinin Haar ve Daubechies 4 Dalgacıkları Yardımıyla Bileşenlerine Ayrıştırılması.

#### 2.1.4.4. Dalgacık Serileri

Fourier serilerine benzer şekilde dalgacık analizinde de dalgacık serilerinden bahsedilebilir. Fourier serilerinin sadece periyodik fonksiyonlar için var olduğu hatırlanırsa, dalgacık serilerinde biraz daha farklı olarak  $f(t) \in L^2$  olacak şekilde herhangi bir fonksiyon için;

$$\psi_{a,b}(t) = 2^{\frac{a}{2}} \psi(2^a t - b) \quad (2-47)$$

iken dalgacıkların ölçek ve öteleme olmak üzere iki parametresine bağlı çift toplam sembolüyle dalgacık serileri aşağıdaki gibi ifade edilebilir (Goswami ve Chan, 2011, ss. 77–78):

$$f(t) = \sum_a \sum_b w_{a,b} \psi_{a,b}(t). \quad (2-48)$$

Bir sinyalin dalgacık bazlı enerjisi, sinyalin karesinin integrali olarak tanımlanır. Yani  $x(t)$  sinyalinin enerjisi aşağıdaki gibi ifade edilir (Graps, 1995, s. 3):

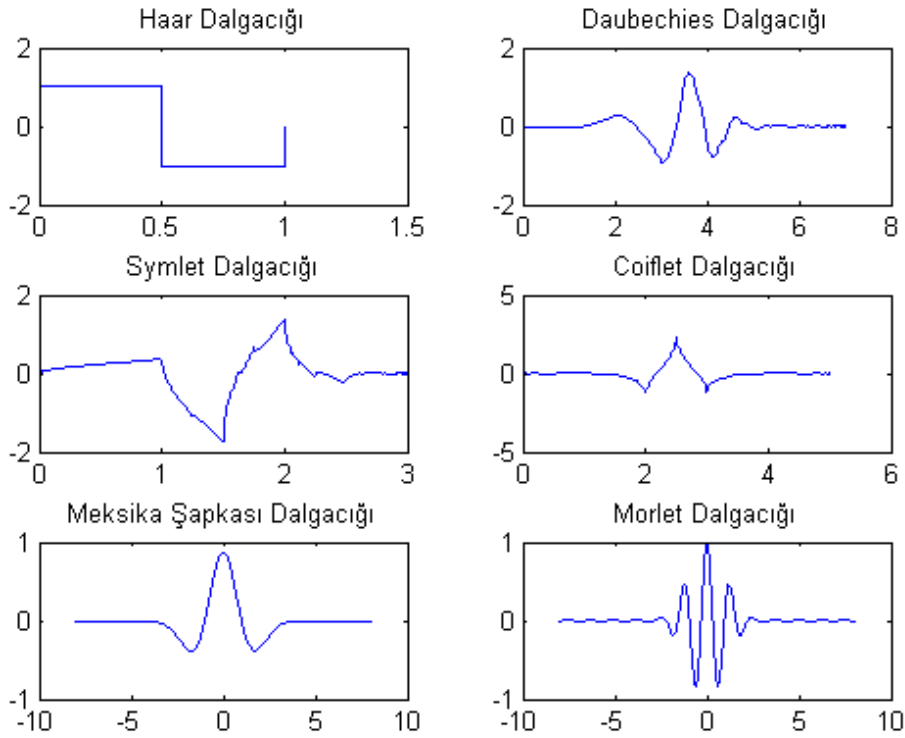
$$E = \int_{-\infty}^{\infty} |x(t)|^2 \cdot dt = \|x(t)\|^2. \quad (2-49)$$

Bir sinyalin dalgacık güç spektrumu ölçek bağımlı enerjiler ile üretilebilir. Dalgacık güç spektrumu ile farklı ölçeklerde elde edilen bileşenlerin toplam enerji seviyelerini belirlemek için,  $a$  ölçek ve  $W_i$  ise  $i$ . seviye dalgacık dönüşümü fonksiyonu olmak üzere aşağıdaki gibi ifade edilebilir (Addison, 2002, ss. 84–85):

$$E = \sum_{i=0}^{N-1} |W_i^{(a)}|^2. \quad (2-50)$$

#### 2.1.4.5. Dalgacık Çeşitleri

Yıllar boyunca literatürde önemli sayıda dalgacık fonksiyonu önerilmiştir. Sadece sürekli ayrıştırmada kullanılacak Gauss, Morlet ve Meksika şapkası dalgacıkları gibi ham dalgacıklar bulunmaktadır. Sonsuz düzenli dalgacıklar olarak Meyer dalgacık ailesinde bulunan dalgacıklar, Daubechies, symlet ve coiflet ailelerinde bulunan ortogonal ve kompakt desteğe sahip dalgacıklar ile kompakt desteğe sahip dalgacık çiftleri olarak B-spline ve ikili ortogonal dalgacık aileleri hem SDD'nde hem de ADD'nde kullanılabilir. Son olarak, karmaşık Gauss, karmaşık Morlet, karmaşık Shannon ve karmaşık frekanslı B-spline dalgacık aileleri gibi karmaşık SDD'nde kullanılan karmaşık dalgacıklar bulunmaktadır (Alexandridis ve Zapraniş, 2014, s. 13). Aşağıda Şekil 19'da yaygın kullanılan bazı dalgacık çeşitleri verilmekte, ardından dalgacıkların temeli olarak kabul edilen Haar dalgacığı ve finansal zaman serileri için literatürde genellikle kullanılan Daubechies dalgacıkları ile ilgili detaylı bilgi verilmektedir.



Şekil 19 : Yaygın Kullanılan Bazı Dalgacık Çeşitleri.

### 2.1.4.5.1. Haar Dalgacıđı

Dalgacıkların atası olarak kabul edilen Haar dalgacıđı, en basit ve kısa dalgacıktır (Fugal, 2009, s. 204). Haar dalgacıđı, parçalı sabit bir fonksiyonun çok çözünlürlüğü ile elde edilir. Ölçekleme fonksiyonu  $\varphi = 1_{[0,1]}$ ,  $h[n]$  filtresinin sıfırdan farklı  $2^{-\frac{1}{2}}$ 'ye eşit iki katsayısı,  $n = 0$  ve  $n = 1$  olmak üzere:

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{\infty} (-1)^{1-n}h[1-n]\varphi(t-n) = \frac{1}{\sqrt{2}}(\varphi(t-1) - \varphi(t)) \quad (2-51)$$

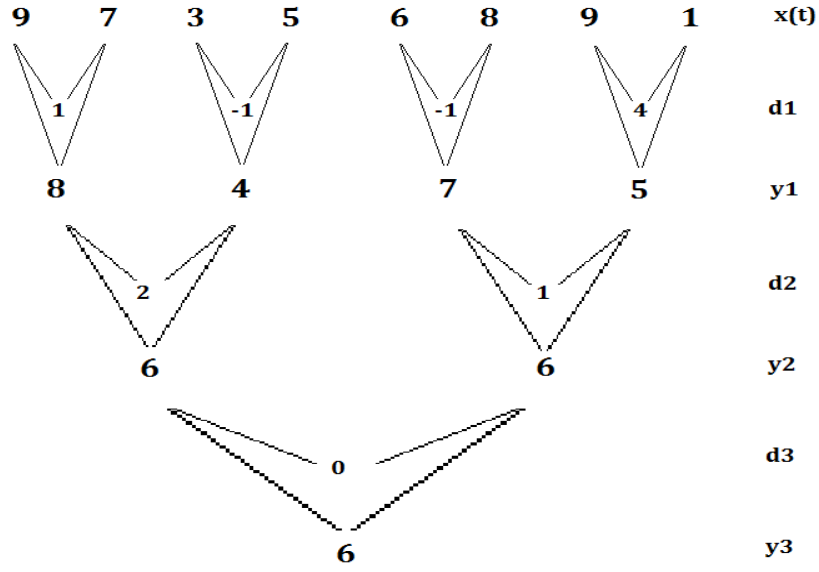
ve böylece;

$$\psi(t) = \begin{cases} -1, & 0 \leq t < \frac{1}{2} \\ 1, & \frac{1}{2} \leq t < 1 \\ 0, & \text{diđer} \end{cases} \quad (2-52)$$

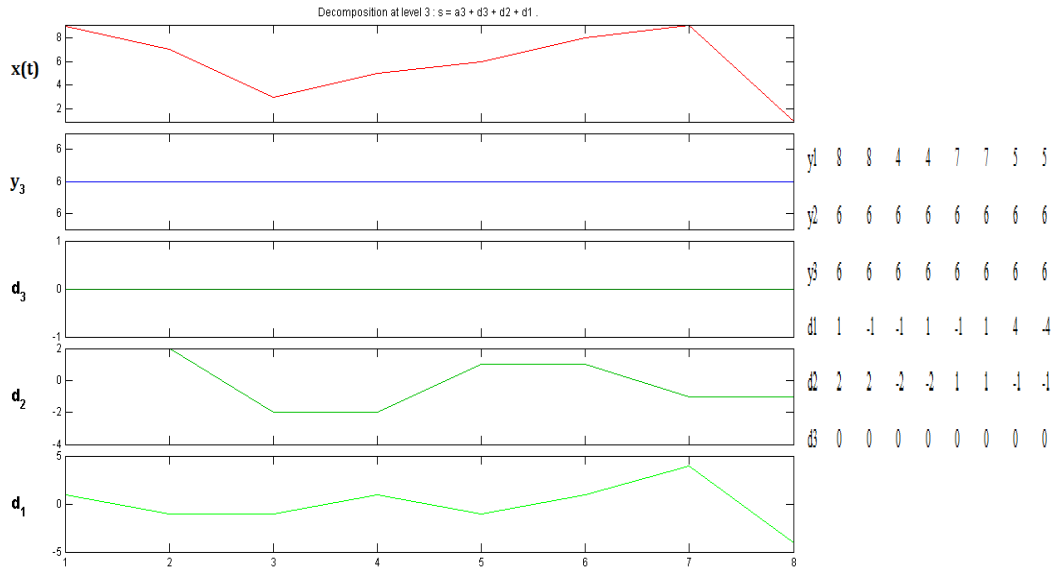
olarak elde edilir (Mallat, 2008, s. 291).

Örnek olarak  $x_t = \{x_1, \dots, x_8\} = \{9, 7, 3, 5, 6, 8, 9, 7\}$  olacak şekildeki  $x_t$  serisinin Haar dalgacıđ dönüşümü, toplam ve fark ortalamaları yardımıyla aşağıdaki Şekil 20'de gösterilmiştir.

Şekilden  $x_t = \{9, 7, 3, 5, 6, 8, 9, 7\}$  serisinin Haar dönüşümü  $\{6, 0, 2, 1, 1, -1, -1, 4\}$  olarak elde edildiđi görülmektedir. Haar dönüşümüyle elde edilen yaklaşım ve detay katsayıları ayrı ayrı hesaplandığında Şekil 21'deki katsayılar elde edilmektedir.



Şekil 20 : Haar Ortalama Dönüşümünün Grafik Gösterimi.



Şekil 21 : Haar Ortalama Dönüşümünün Yaklaşım ve Detay Katsayıları.

### 2.1.4.5.2. Daubechies Dalgacıları

Dalgacık teorisindeki en önemli başarılarından biri Haar dalgacığından daha yumuşak kompakt desteğe sahip ortogonal dalgacıkların bulunmasıdır. Daubechies de çalışmaları sonucunda 1988'de kendi adıyla ifade ettiği ortonormal dalgacık

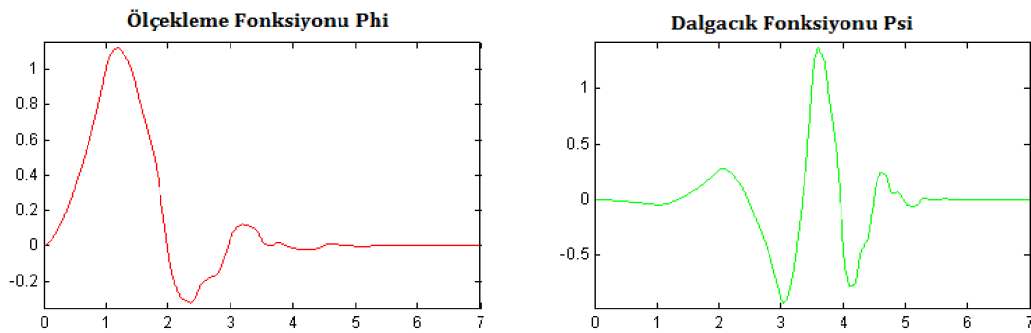
ailesini ortaya atmıştır. Dalgacık ailesinin her bir üyesi N gibi bir sayı ile indekslenir ki burada N,  $h[n]$ 'nin uzunluğunu ve sifıra yaklaşan momentlerin sayısının iki katını ifade eder (Nason, 2010, s. 41). N/2 adet sifıra yaklaşan momentli olan bir fonksiyon  $\psi(t)$ ,  $p=0,1,\dots,N/2-1$  için  $\int t^p \psi(t) dt = 0$  eşitliğini gerçeker (Gençay vd., 2002, s. 112).

Daubechies dalgacık filtreleri Haar dalgacığının frekans alanı özelliklerini geliştiren, komşu ortalamaların genelleştirilmiş farklılıkları olarak yorumlanabilen dalgacıkların bir koleksiyonunu temsil eder. Daubechies bu dalgacıkları kompakt desteğe sahip bir fonksiyon ve maksimum sayıda sifıra yaklaşan momentleri kıstasından üretmiştir. Genel olarak bu sınıf dalgacık filtrelerinin belirgin bir zaman uzayı formülü yoktur. Gerektiğinde Şekil 22'deki gibi filtre katsayıları ile basitçe ifade edilebilmektedir (Gençay vd., 2002, s. 112). Şekil 23'te de Daubechies 4 dalgacığı ve ölçekleme fonksiyonu verilmektedir.

$$h_0 = \frac{1 - \sqrt{3}}{4\sqrt{2}}, h_1 = \frac{-3 + \sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3 + \sqrt{3}}{4\sqrt{2}} \quad \text{ve} \quad h_3 = \frac{-1 - \sqrt{3}}{4\sqrt{2}}.$$

Şekil 22 : Daubechies 4 Dalgacığı için Zaman Uzayında Filtre Katsayılarıyla Gösterimi

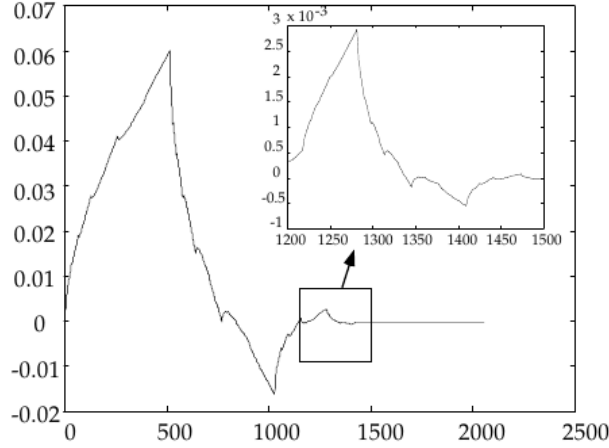
(Gençay vd., 2002, s. 113).



Şekil 23 : Daubechies 4 Dalgacığı ve Ölçekleme Fonksiyonu.

Dalgacık dönüşümleri sonsuz bir dizi oluşturur. Farklı dalgacık aileleri, frekans uzayında lokalize baz fonksiyonlarının kompakt olmaları ve düzgün olmaları

arasında farklı dengeler oluşturur. Bazı dalgacıklar fraktal yapıya sahiptirler. Daubechies dalgacık ailesi Şekil 24'te görüldüğü üzere bu tip dalgacıklara örnektir (Graps, 1995, s. 7).



Şekil 24 : Daubechies Ana Dalgacığının Fraktal Öz Benzerliği  
(Graps, 1995, s. 7).

Daubechies dalgacık ailesi özellikle çok çözünürlüklü analizlerde tercih edildiği görülmektedir. Literatürde özellikle finansal zaman serilerinin analizleri ile ilgili dalgacık dönüşümünün kullanıldığı çalışmalarda Daubechies ailesinden Daubechies 4 dalgacığının seçildiği ve kullanıldığı tespit edilmiştir.

## 2.2. DESTEK VEKTÖR MAKİNELERİ

Destek vektör makineleri (DVM), istatistiksel öğrenme teorisine dayalı bir evrensel yapısal öğrenme sürecidir (Vapnik, 1995, s. 127). Burada evrensel kavramı ile kastedilen DVM yönteminin, sinirsel ağlar (sigmoid aktivasyon), radyal tabanlı fonksiyonlar, splinelar, polinom tahmin ediciler vb. gibi çeşitli gösterimleri öğrenmek için kullanılabilmesidir. Bununla birlikte, daha genel bir anlamda DVM, fonksiyonların parametre olarak yeni bir formunu oluşturma özelliğiyle de tümevarım olmayan öğrenme formülasyonlarına uygulanabilir. Örneğin, destek vektör parametreleştirme, lineer denklemler, bilgisayarlı tomografi, sinyal / görüntü sıkıştırma ve benzeri büyük sistemlerin çözümünde kullanılabilir. DVM parametreleştirme, problemin çok boyutluluğundan bağımsız olarak destek

vektörlerinin sayısı üzerinden fonksiyonun karmaşıklığına anlamlı bir tanımlama sağlar (Cherkassky ve Mulier, 2007, ss. 404–405).

DVM yaklaşımı marjin, çekirdek gösterimi ve dualite gibi bazı temel yaklaşımları birleştirmektedir. 1960'ların başında marjin kavramı Vapnik ve Lerner (1963) ile Vapnik ve Chervonenkis (1964)'in çalışmalarıyla ayrılabilir veriler için geliştirilmiştir. Ancak, DVM yönteminin Boser, Guyon ve Vapnik (1992) ile Cortes ve Vapnik (1995) gibi çalışmalarla çekirdek (kernel) gösterimi kazanması ve ayrılabilir olmayan verileri işleme yeteneğinin ortaya çıkması yaklaşık otuz yıl almıştır.

Aizerman, Braverman ve Rozner (1964)'te çekirdek kullanma fikrini ortaya atarken, Mangasarian (1965)'te matematiksel programlamada, sınıflandırma için DVM benzeri bir doğrusal optimizasyon formülasyonu ileri sürmektedir. Ancak bu gelişmeler, istatistiksel öğrenme kuramının sağladığı sağlam temellerden yoksun olduğundan pratik öğrenme algoritmaları arasında yer bulamamıştır. 90'lı yıllardan beri de DVM, öğrenme problemleri için bir çok uygulamada başarıyla kullanılmıştır (Cherkassky ve Mulier, 2007, s. 408).

DVM, en iyi şekilde temel işlevi olan ikili sınıflandırma ile tanımlanabilmektedir. Birçok gerçek dünya problemi, örneğin, ekonomik verilere bağlı olarak döviz kurunun ya da borsanın aşağı mı yoksa yukarı mı hareket edeceğini tahmin etmek veya bir müşteriye kişisel finansal bilgisinden hareketle kredi verilip verilemeyeceğini tahmin etmek gibi iki sınıf üzerine tahmin içermektedir. DVM eğitim veri kümesini kullanarak öğrenen ve genelleştirmeye çalışan, sonrasında oluşturduğu model yardımıyla yeni veriler üzerinde doğru tahminler yapmaya yarayan soyut öğrenme makinesi olarak tanımlanabilir (Campbell ve Ying, 2011, s. 1).

DVM, çekirdek kaynaklı özellik uzaylarında doğrusal öğrenme makinelerini etkin bir şekilde eğitmek için genelleme teorisini göz önünde bulunduran ve optimizasyon teorisinden yararlanan bir sistem olarak ifade edilmektedir (Stoean, 2014, s. 7).



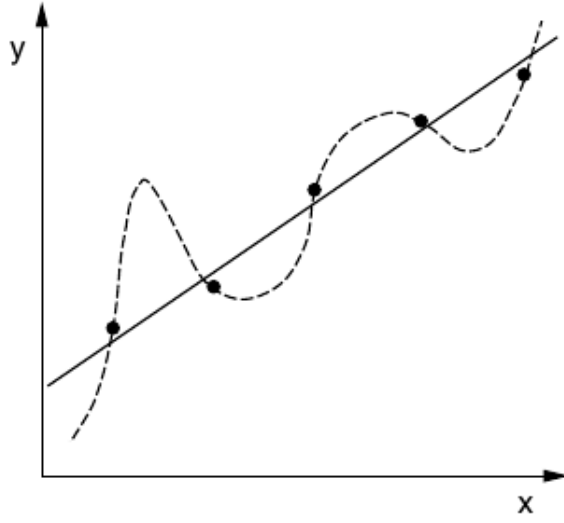
### 2.2.1. İstatistiksel Öğrenme Teorisi

İstatistiksel öğrenme teorisi, 1960'ların sonuna doğru ortaya atılmasına rağmen 1990'lara kadar verilen tam bir veri koleksiyonundan fonksiyon tahmini probleminin teorik analizinden öteye geçmemiştir. 90'lı yılların ortasında istatistiksel öğrenme teorisine dayalı DVM olarak adlandırılan yeni öğrenme algoritması türleri önerildiği görülmektedir. Bu durum, istatistiksel öğrenme teorisini sadece teorik analiz için bir araç değil, aynı zamanda çok boyutlu fonksiyonların tahmini için pratik algoritmalar oluşturabilen bir araç haline getirmektedir (Vapnik, 1999, s. 988).

İstatistiksel öğrenme teorisinin amacı sezgisel argümanları matematiksel çerçeveye yerleştirmektir. Bu amaçla istatistiksel öğrenme teorisi, öğrenme makinesinin matematiksel özelliklerini incelemektedir. Bu özelliklerin, genellikle öğrenme makinesinin uygulayabileceği fonksiyon sınıfının özellikleri olduğu görülmektedir (Schölkopf ve Smola, 2002, ss. 7–8).

İstatistiksel öğrenme teorisinin, sonlu örnekler ile esnek istatistiksel tahmin için halihazırda kullanılacak en iyi teori olduğu ifade edilmektedir. Diğer yaklaşımların ( yapay sinir ağları, cezalandırma çerçevesi ve Bayeşçi çıkarım) aksine istatistiksel öğrenme teorisinde, sonlu örneklerle öğrenme ve tahmin ile önemli sonuçlar elde etmek için titizlikle ilgili tüm kavramlar tanımlanmakta, öğrenme problemi ayarları belirtilmekte ve matematiksel ispatlar sunulmaktadır (Cherkassky ve Mulier, 2007, s. 99).

Şekil 25'te görülen siyah noktalardan oluşan örnek verisi ile fonksiyonel bağımlılığın tahmin edilmek istenildiği varsayılınsın. Karmaşık model tüm veri noktalarına kusursuz bir şekilde uygun iken, düz çizgi veri noktalarına göre artık hata payları ortaya çıkmaktadır. Burada karşılaşılan model seçimi problemi için istatistiksel öğrenme teorisi, model sınıfının karmaşıklığının rolünü formülize etmekte ve olasılıksal teminat sağlamaktadır (Schölkopf ve Smola, 2002, s. 126).



Şekil 25 : Model Seçimi

(Schölkopf ve Smola, 2002, s. 126).

Sistemi tanımlamadan ziyade sistemi taklit hedefini benimseyen istatistiksel öğrenme teorisi tarafından kullanılan kavramsal yaklaşım, klasik istatistikten farklıdır. Bu nedenle, Vapnik ve Chervonenkis (1968) tarafından geliştirilen VC teorik çerçevesi, pratik hedefi bilinmeyen sistemin doğru tanımlanması yerine iyi genellemesi olan birçok uygulama için uygundur. Birçok pratik çok değişkenli problem için çok boyutluluk problemi yüzünden sistemi tanımlamanın gerçekçi bir hedef olmayacağına altı çizilmektedir (Cherkassky ve Mulier, 2007, s. 100).

### 2.2.1.1. VC-Boyutu

VC-boyutu fonksiyon  $\{f(\sigma)\}$  kümesinin bir özelliği olarak fonksiyonun çeşitli sınıfları için tanımlanabilmektedir. Burada parametrelerin bir kümesi olarak her  $\sigma$  ve her  $x$  için  $f(x, \sigma) \in \{-1, 1\}$  olacak şekildeki iki sınıflı fonksiyonlar dikkate alınmaktadır. Verilen  $l$  nokta kümesi mümkün olan  $2^l$  şekilde işaretlenebilir ve her işaretlemeye karşılık  $\{f(\sigma)\}$  kümesinin bir elemanı bulunabiliyorsa, bu nokta kümesi  $\{f(\sigma)\}$  fonksiyon kümesi ile ayrılabilir (shatter) denir.  $\{f(\sigma)\}$  fonksiyon kümesinin VC-boyutu,  $\{f(\sigma)\}$  tarafından ayrılabilen eğitim noktalarının maksimum sayısı olarak tanımlanır. Eğer VC-boyutu  $h$  ise, en az bir  $h$  noktalı küme vardır. Ancak  $h$  noktalı her küme ayrılabilir denilememektedir (Burges, 1998, s. 4). Örneğin, iki boyutlu uzayda en fazla üç nokta doğrusal olarak ayrılabilir olduğundan bu uzayda doğrusal

sınıflandırıcıların VC-boyutunun üç olduğu görülebilmektedir. Ancak doğrusal olmayan sınıflandırıcıların VC-boyutunun daha büyük olabileceği görülmektedir.

VC-boyutu büyüdüğünde, sınıflandırıcının kapasitesi de aynı oranda artmakta, kapasite arttıkça da eğitim noktalarını daha iyi sınıflandırabilmektedir (Hamel, 2011, s. 178).

### 2.2.1.2. Ampirik Risk Minimizasyonu

Öğrenmenin hedefi, verinin temel dağılımı üzerinde yüksek genelleştirme performansı sunan bir model bulunmasıdır. Elde sadece eğitim seti için gözlemler olduğunda, riski tahmin etmek için bu gözlemler kullanılabilir (Hamel, 2011, s. 176).  $X$  bir girdi uzayı,  $Y$  de bir çıktı uzayı olmak üzere  $(X \times Y) \subseteq X \times Y$  ikilileri, ortak dağılımları  $P_{XY}$  olan rasgele değişkenler olsun. Amaç,  $z$  doğru ise  $\delta(z) = 1$  ve doğru değilse  $\delta(z) = 0$  olmak üzere; (2-53) eşitliğinde verilen beklenen riski en aza indiren  $f : X \mapsto Y$  belirleyicisi bulmak olacaktır (Zhang, 2011, s. 312).

$$P(f(X) \neq Y) = E_{(X,Y) \sim P_{XY}}[\delta(f(X) \neq Y)] \quad (2-53)$$

Ancak, uygulamalarda eşdeğer gelen ve  $P_{XY}$  dağılımından bağımsız sadece sonlu sayıda  $n$  çift eğitim örneği  $(X_j, Y_j)$  bulunmaktadır.  $P_{XY}$  dağılımının bilinmesi imkansız olduğundan, genellikle ampirik risk olarak adlandırılan eğitim setindeki risk aşağıdaki dağılım üzerinde beklenen risk yerine kullanılabilir (Zhang, 2011, s. 312).

$$\frac{1}{n} \sum_{j=1}^n \delta(f(X_j) \neq Y_j) \quad (2-54)$$

Ampirik risk minimizasyonu (ARM) ampirik riski minimize edecek bir  $f$  fonksiyonunun bulunması olduğundan bu yöntem, genellikle etkili ve verimli olmasına rağmen ezberleme (overfitting) problemi ile karşı karşıya kaldığı belirtilmektedir (Zhang, 2011, s. 312). Eğitim verisine tam olarak uygun elde edilen bir model, daha önce hiç karşılaşmadığı test verisinde zayıf performans

gösterebilmektedir. Bu da yine ampirik riski minimize etmenin aşırı iyimser olduğunu gösterir (Hamel, 2011, s. 177).

### 2.2.1.3. Yapısal Risk Minimizasyonu

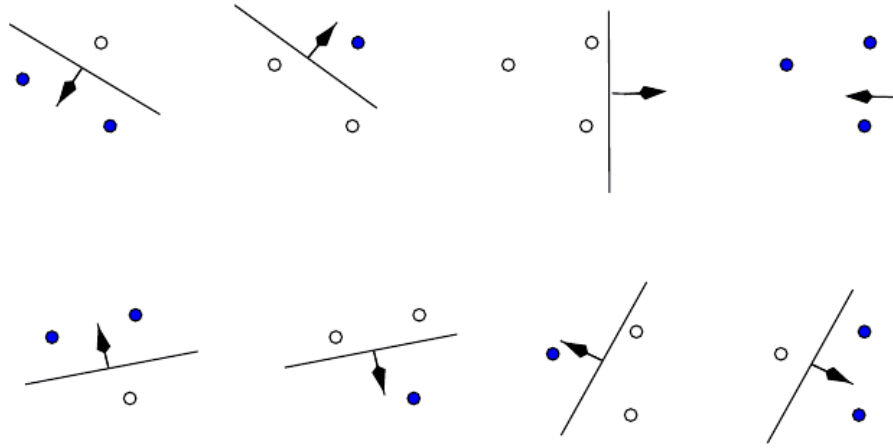
Yapısal risk minimizasyonu (YRM), sınıflandırma problemleri için geliştirilmiş olup, sınırlı gözlemler için optimal model karmaşıklığını seçmeyi amaçlar. YRM, eğer belirli bir eğitim seti üzerindeki doğruluk oranı ile makine kapasitesi arasında diğer bir eğitim seti üzerinde hatasız öğrenmeyi sürdürebilecek bir denge varsa, belirli miktardaki eğitim verisi ile sınıflandırma işlemi için genelleme başarısı elde edileceğini ileri sürmektedir (Stoean, 2014, s. 8).

Bu durumu örneklendiren Burges (1998, s. 2), yüksek kapasiteye sahip bir makineyi, yeni bir ağaçla karşılaştığında onun daha önce gördüğü ağaçların yaprak sayılarına uymadığı için ağaç olduğunu kabul etmeyen fotografik hafızaya sahip bir botanikçiye benzetirken; düşük kapasiteli bir makineyi ise; botanikçinin yeşil olarak gördüğü her şeyin ağaç olduğunu iddia eden kardeşine benzettiği görülmektedir. Dolayısıyla iki durumda da doğru genelleme yapılamamaktadır.

Sınıflandırma denildiğinde ilk akla gelen matematiksel olarak  $y_i \in \{-1, 1\}$  ile işaretlenen ikili sınıflandırma işlemi olmaktadır.  $t$  genel parametre olmak üzere  $f_t$  fonksiyon kümesi (2-55) eşitliğindeki gibi verilsin (Cherkassky ve Mulier, 2007, s. 108).

$$f_t = \mathbb{R}^n \rightarrow \{-1, +1\} \quad (2-55)$$

Verilen  $m$  adet eğitim verisi  $2^m$  farklı şekilde işaretlenebilmektedir. Buradaki her işaretlemeye karşılık bu işaretlemeleri doğru atayan  $\{f_t\}$  fonksiyon kümesinin bir elemanı bulunabiliyorsa, bu veri kümesi  $\{f_t\}$  fonksiyon kümesi tarafından ayrılabilir denir (Stoean, 2014, s. 8). Şekil 26'da  $\mathbb{R}^2$  'de üç noktanın doğru ile ayrılması durumları gösterilmektedir.



Şekil 26 :  $R^2$  'de Üç Noktanın Doğru ile Ayrılması Durumları  
(Burges, 1998, s. 5).

YRM prensibi, sınıflandırma problemleri için risk fonksiyonu ve güven düzeyi ile tanımlanmaktadır. Herhangi bir  $t$  parametresi için  $h$ , VC boyutu ve  $\eta \in [0,1]$  olmak üzere en azından  $1-\eta$  olasılığına sahip  $m > h$  için  $R(t)$  risk fonksiyonu (2-56) eşitliği ve güven düzeyi (2-57) eşitliği ile gösterilmektedir (Stoian, 2014, s. 8).

$$R(t) \leq R_{amp}(t) + G\left(\frac{h}{m}, \frac{\log(\eta)}{m}\right) \quad (2-56)$$

$$G\left(\frac{h}{m}, \frac{\log(\eta)}{m}\right) = \sqrt{\frac{h\left(\log\frac{2m}{h} + 1\right) - \log\frac{\eta}{4}}{m}} \quad (2-57)$$

YRM prensibi, yüksek genelleme yeteneği için hem eğitim hatasının hem de güven teriminin minimal tutulmasını gerektirmekte, sonrasında ise VC-boyutu azaltılarak minimize edilmesinin gerekliliğini ortaya koymaktadır (Stoian, 2014, s. 8).

YRM'ye göre sonlu veriye sahip bir öğrenme probleminin çözümü için yaklaşım fonksiyonları kümesi üzerinde önsel bir yapısal özellik gerekmektedir. Verilen veri seti için optimal bir model tahmini, yapının en uygun karmaşıklığı olan

elemanını seçme ve bu elemandan modeli tahmin etme olmak üzere iki adımda yapılır. Burada birinci adım model seçimini ikinci adım ise istatistiksel modellerde parametre tahminine karşılık gelmektedir. VC sınırlarını minimize etmek için pratik iki strateji bulunmaktadır (Cherkassky ve Mulier, 2007, s. 123):

- 1) Model karmaşıklığını (VC-boyutu) sabit tutup ampirik hata terimini minimize etmek.
- 2) Ampirik hata terimini sabit (küçük) tutup VC-boyutunu indirmek.

DVM yönteminin diğer istatistiksel metotların aksine ikinci stratejiyi uygulaması farkını ortaya koymaktadır.

### **2.2.2. Optimal Ayırıcı Hiperdüzlem**

YRM'nin ikinci stratejisi olan "ampirik risk değerini sabit tutarken VC-boyutunu minimize etme" göz önüne alınarak Vapnik ve Chervonenkis (1974) tarafından yeni bir evrensel öğrenme makinesi tipi ortaya atılmıştır. Burada optimal ayırıcı hiperdüzlem olarak adlandırılan özel bir hiperdüzlem tipi kullanılmaktadır.

Optimal ayırıcı hiperdüzlem öncelikle eğitim verisinin tamamen doğrusal olarak ayrılabilirdiği durumlar için inşa edilmiş olup, eğitim verisinin tamamen doğrusal olarak ayrılamadığı durumlar için de ceza katsayısı kullanılarak genellenebildiği görülmektedir. Bu teknik en genel haliyle bir evrensel öğrenme makinesi tipi olan DVM'ni oluşturmaktadır.

#### **2.2.2.1. Doğrusal Olarak Ayrılabilir Örüntüler için Optimal Hiperdüzlem**

$x \in \mathbb{R}^n$  ve  $y \in \{+1, -1\}$  olmak üzere  $(x_1, y_1), \dots, (x_l, y_l)$  eğitim verisi  $x$  hiperdüzlem üzerindeki herhangi bir nokta,  $w$  hiperdüzlemin normali (ağırlık vektörü),  $|b|/\|w\|$  hiperdüzlemin orjine dik uzaklığı ve  $\|w\|$ ,  $w$ 'nin öklit normu olacak şekilde (2-58) eşitliğinde verilen hiperdüzlem ile ayrılabilirdiği farz edilmektedir (Burges, 1998, s. 8).

$$(w \cdot x) + b = 0. \quad (2-58)$$

Eğer eğitim verisini oluşturan vektörler hatasız bir şekilde ayrılabilir ve hiperdüzleme en yakın vektörün uzaklığı maksimum ise (Şekil 27) bu eğitim verisi optimal hiperdüzlemle ayrılabilir denir (Vapnik, 1995, s. 127).

Diğer bir ifadeyle optimal ayırıcı hiperdüzlem, veriyi marjinin içine herhangi bir gözlem düşmeyecek şekilde maksimum marjinle ayıran hiperdüzlemdir. Bu hiperdüzlem karar sınırı olarak da ifade edilebilir. Sınırlardaki gözlem değerleri de destek vektörleri olarak adlandırılmaktadır.

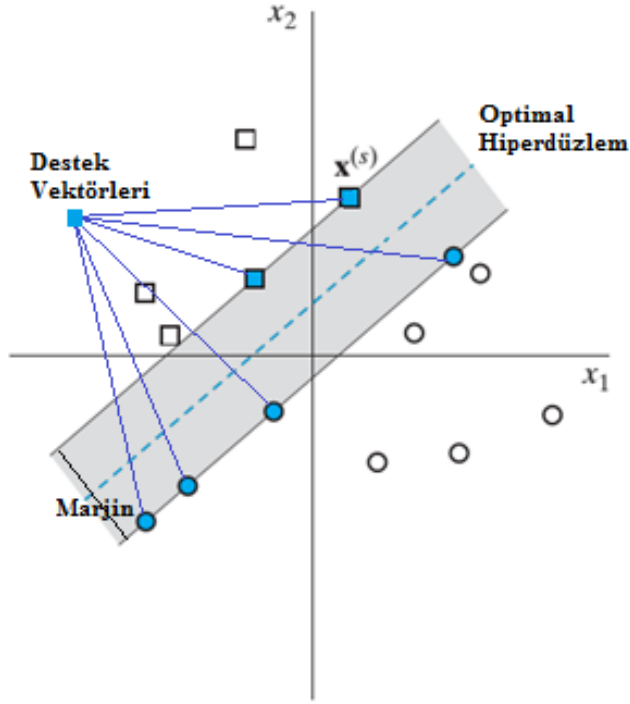
Ayırıcı hiperdüzlemi açıklamak için kanonik form kullanılırsa (Vapnik, 1995, s. 127);

$$\begin{aligned} y_i = +1 \text{ için,} & \quad (w \cdot x) + b \geq +1 \\ y_i = -1 \text{ için,} & \quad (w \cdot x) + b \leq -1 \end{aligned} \quad (2-59)$$

elde edilecektir. Bu eşitsizlikler  $y_i$  değerleri kullanılarak, tek bir eşitsizlik ile ifade edilebilmektedir (Alpaydın, 2010, s. 311).

$$y_i [(w \cdot x) + b] \geq +1 \quad i = 1, \dots, l. \quad (2-60)$$

Böylece uygun  $w$  ve  $b$  değerleri hesaplanarak marjini maksimize edecek şekilde optimum ayırıcı hiperdüzlem bulunabileceği görülmektedir. Şekil 27'de optimal hiperdüzlem ve destek vektörleri görülmektedir.



Şekil 27 : Doğrusal Ayrılabilir Örüntüler için Optimal Hiperdüzlem ve Destek Vektörleri (Haykin, 2009, s. 270).

### 2.2.2.1.1. Kanonik Hiperdüzlemlerin Yapısı

Ayrılcı hiperdüzlemler  $a$  merkezli  $R$  yarıçaplı küre tarafından sınırlanmış  $X^* = x_1, \dots, x_r$  vektör kümesi üzerinde tanımlandığında  $|x_i - a| \leq R$ ,  $x_i \in X^*$  olacağı görülmektedir. Kanonik formdaki  $(w, b)$  çiftiyle tanımlanan ve  $\min_{x_i \in X^*} |(w, x_i) + b| = 1$  şartını sağlayan  $X^*$  üzerinde tanımlı hiperdüzlem kümesi göz önüne alındığında  $\|w\| \leq A$  sınırlılığında  $h \leq \min([R^2 A^2, n]) + 1$  eşitsizliğiyle sınırlanmış  $h$ , VC- boyutu bulunmaktadır (Vapnik, 1995, s. 128).

$n$  boyutlu uzayda hiperdüzlemlerin VC-boyutu  $n+1$  olmasına rağmen  $|w|^2 \leq A^2$  olacak şekilde hiperdüzlemlerin kanonik formdaki altkümesinin VC- boyutu daha az olabilmektedir (Vapnik, 1995, s. 129).

### 2.2.2.1.2. Optimal Hiperdüzlem Oluşturma

Optimal hiperdüzlemi oluşturmak için  $y \in \{+1, -1\}$  olmak üzere  $(x_1, y_1), \dots, (x_n, y_n)$  eğitim verisindeki  $x_i$  vektörlerini ayıracak en küçük norma sahip



hiperdüzlemin katsayıları hesaplanmalıdır. Bu hiperdüzlemin bulunabilmesi için aşağıdaki kvadratik programlama probleminin çözülmesi gerekmektedir (Borges, 1998, ss. 128–129).

$$\text{Minimizasyon: } \Phi(w) = \frac{1}{2}(w \cdot w) \quad (2-61)$$

$$\text{Kısıt: } y_i [(w \cdot x) + b] \geq +1 \quad i = 1, \dots, l. \quad (2-62)$$

Bu optimizasyon problemi konveks ve kısıtları doğrusal olduğu için Lagrange çarpanları kullanılarak çözülebilir. Burada (2-61) eşitliğinde minimize edilecek fonksiyon ve (2-62) eşitliğinde ise çözüm esnasında kullanılacak koşul verilmektedir.

Lagrange (1788) tarafından ortaya konan Lagrange çarpanlarından yararlanarak optimal çözümün elde edilebilmesi için değişken sayısı  $n$ , kısıt sayısı  $m$ 'den küçük olmamalı; yani,  $n \geq m$  olmalı ve amaç ve kısıt fonksiyonları türevlenebilir fonksiyon olması gerekmektedir (Tütek, Gümüšoğlu ve Özdemir, 2012, s. 647).  $x \in R^n, i = 1, 2, \dots, n$  ve  $j = 1, 2, \dots, m$  için problemin primal gösterimi;

$$\text{Amaç fonksiyonu: } \min f(x) \quad (2-63)$$

$$\text{Kısıtlar: } g_j(x) \geq 0 \quad (2-64)$$

olsun. Verilen primalin duali,  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$  negatif olmayan Lagrange çarpanları yardımıyla elde edilen  $L(x, \alpha)$  Lagrange fonksiyonu ile ifade edilmektedir (Campbell ve Ying, 2011, s. 70).

$$L(x, \alpha) = f(x) - \sum_{j=1}^m \alpha_j g_j(x) \quad (2-65)$$

Burada;  $\alpha_j$ 'ler j. kısıta ait Lagrange çarpanlarıdır. Lagrange fonksiyonunun oluşturulmasıyla kısıtlı model kısıtsız bir modele dönüşmektedir. Bu optimizasyon modeli, x primal değişkenine göre minimizasyon iken  $\alpha$  dual değişkenine göre maksimizasyon olmak üzere içiçe ters amaçlı iki optimizasyon operatörü barındırmaktadır.

Eğer  $(x, \alpha) = (x^*, \alpha^*)$ , kısıtsız  $L(x, \alpha)$  fonksiyonunun global minimum ya da maksimum noktası ise  $x^*$  kısıtlı model için de optimal nokta olur. Buradan hareketle  $L(x, \alpha)$  Lagrange fonksiyonunun kısıtsız optimizasyon sorunlarında olduğu gibi kısmi türevi sıfıra eşitlenerek analiz edilmesi gerekmektedir (Tütek vd., 2012, s. 647).

$$\frac{\partial L}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^m \alpha_j \frac{\partial g_j(x)}{\partial x_i} = 0 \quad (2-66)$$

$$\frac{\partial L}{\partial \alpha_j} = - \sum_{j=1}^m g_j(x) = 0 \quad (2-67)$$

Kısmi türevin alınmasıyla, amaç probleminde n değişken ve m kısıt olması nedeniyle  $(n + m)$  bilinmeyen ortaya çıkmaktadır. Yukarıdaki eşitlikler  $(x, \alpha)$  için çözüldüğünde bulunan noktaların fonksiyonun kısmi türevinin sıfıra eşit olduğu muhtemel çözüm noktaları olduğu görülmektedir. Diğer bir ifadeyle yukarıdaki koşullar optimal çözüm için gerek koşulları olduğu ifade edilmektedir (Tütek vd., 2012, s. 648).

Belirli koşullar altında Lagrange dualinin çözümü primalin de çözümünü verir.  $(x^*, \alpha^*)$  değerleri dualinin çözümü iken;  $x^*$  değerinin primal problemin de çözümü olabilmesi için Karush (1939) ve Kuhn ve Tucker (1951) tarafından ortaya konulan KKT koşullarını sağlaması gerekmektedir (Tütek vd., 2012, s. 653):

1.  $\frac{\partial L}{\partial x}(x^*, \alpha^*) = 0$ , ( $x^*$  eğer noktası üzerinde)
2.  $g_j(x^*) \geq 0$ , Primal Kısıt

3.  $\alpha_j^* \geq 0$ , *Dual Kısıt*
4.  $\alpha_j g_j(x^*) = 0$

Maksimizasyon probleminde eğer amaç fonksiyonu konkav, kısıt fonksiyonları konveks ise KKT koşullarını sağlayan bir nokta modelin optimal çözümünü verirken, eğer amaç fonksiyonu konveks ve kısıt fonksiyonları da konveks ise minimizasyon problemlerinde KKT koşullarını sağlayan nokta optimal çözümdür (Tütek vd., 2012, s. 654).

Yukarıda verilen Lagrange çarpanları ve KKT koşulları kullanılarak (2-61) eşitliğindeki optimizasyon problemi Lagrange fonksiyonu  $\alpha_i$ 'ler Lagrange çarpanları olmak üzere (Vapnik, 1995, s. 129):

$$L(w, b, \alpha) = \frac{1}{2}(w \cdot w) - \sum_{i=1}^l \alpha_i \{[(w \cdot x_i) + b]y_i - 1\} \quad (2-68)$$

ile ifade edilmektedir. Lagrange fonksiyonunu,  $w$  ve  $b$  değerlerine göre minimize,  $\alpha_i \geq 0$  değerlerine göre maksimize edecek semer noktası bulunur. Ancak (2-68) eşitliğinde ifade edilen Lagrange fonksiyonunun çözümü oldukça karmaşıktır. Primal çözümün elde edilebilmesi için Karush-Kuhn-Tucker (KKT) şartları kullanılarak problem dualine dönüştürülür. Bu noktada elde edilen  $w_0, b_0$  ve  $\alpha_i^0$  çözümleri aşağıdaki şartları sağlaması gerekmektedir (Alpaydın, 2013, s. 264):

$$\frac{\partial L(w_0, b_0, \alpha_i^0)}{\partial b} = 0 \quad (2-69)$$

$$\frac{\partial L(w_0, b_0, \alpha_i^0)}{\partial w} = 0 \quad (2-70)$$

Bu eşitlikler daha açık bir formda yazılırsa, optimal hiperdüzlemin aşağıdaki özellikleri elde edilir (Vapnik, 1995, s. 130):

- i. Optimal hiperdüzlem için  $\alpha_i^0$  katsayıları aşağıdaki kısıtları sağlamalıdır.

$$\sum_{i=1}^l \alpha_i^0 y_i = 0, \quad \alpha_i^0 \geq 0, \quad i = 1, \dots, l \quad (2-71)$$

- ii. Optimal hiperdüzlem ( $w_0$  vektörü) eğitim seti vektörlerinin lineer kombinasyonudur.

$$w_0 = \sum_{i=1}^l y_i \alpha_i^0 x_i, \quad \alpha_i^0 \geq 0, \quad i = 1, \dots, l \quad (2-72)$$

- iii.  $w_0$  hiperdüzlemin açılımında sadece destek vektörü olarak adlandırılan vektörler sıfırdan farklı  $\alpha_i^0$  katsayılarına sahip olabilir. o halde;

$$w_0 = \sum_{\substack{\text{destek} \\ \text{vektörleri}}} y_i \alpha_i^0 x_i, \quad \alpha_i^0 \geq 0, \quad i = 1, \dots, l. \quad (2-73)$$

Lagrange fonksiyonunda  $w_0$  hiperdüzleminin ifadesi yerine yazıldığında, KKT koşulları kullanılarak elde edilen dual problem aşağıdaki gibi oluşmaktadır (Alpaydin, 2010, s. 313):

$$\underbrace{L_D(\alpha)}_{\max \alpha} = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j). \quad (2-74)$$

Dual problemin kısıtı da (2-75) eşitliğindeki gibi olacaktır.

$$\sum_{i=1}^l \alpha_i y_i = 0, \quad \alpha_i \geq 0, \quad i = 1, \dots, l \quad (2-75)$$

Lagrange çarpanları ve destek vektörlerinin optimal hiperdüzlemi vereceği

görülmektedir. Yani, optimal hiperdüzlemi elde edebilmek için (2-74) eşitliğindeki basit quadratik problemi (2-75) eşitliğindeki kısıt altında çözmek yeterli olacaktır.

Bu quadratik optimizasyon probleminin çözümü  $\alpha_0 = (\alpha_1^0, \dots, \alpha_l^0)$  olsun. Optimal hiperdüzleme karşılık gelen  $w_0$  vektörünün normu (2-76) eşitliğindeki gibi ve optimal hiperdüzleme bağlı ayırma karar fonksiyonu, (2-77) eşitliğindeki gösterge fonksiyonu ile verilir (Vapnik, 1995, s. 131).

$$|w_0|^2 = 2W(\alpha_0) = \sum_{\substack{\text{destek} \\ \text{vektörleri}}} \alpha_i^0 \alpha_j^0 (x_i \cdot x_j) y_i y_j \quad (2-76)$$

$$f(x) = \text{sign} \left( \sum_{\substack{\text{destek} \\ \text{vektörleri}}} y_j \alpha_j^0 (x_i \cdot x) - b_0 \right) \quad (2-77)$$

Burada,  $x_i$ 'ler destek vektörleri,  $\alpha_i^0$  Lagrange katsayıları, eşik (sabit) değer  $b_0$ , her iki sınıfa  $\{-1,1\}$  ait destek vektörlerine bağlı olarak (2-78) eşitliği ile ifade edilmektedir (Vapnik, 1995, s. 131).

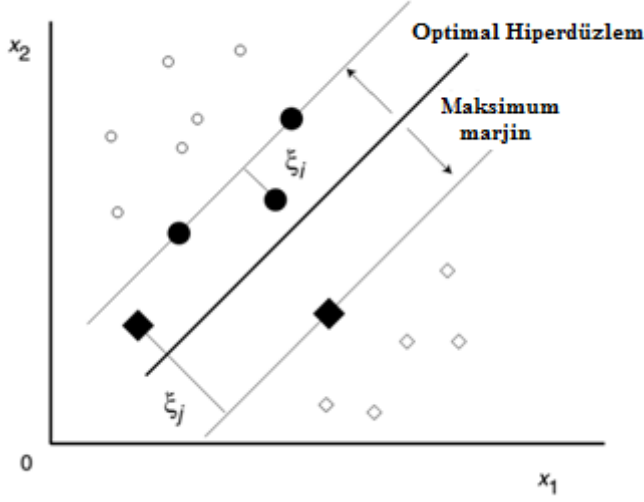
$$b_0 = \frac{1}{2} [(w_0 \cdot x^*(1)) + (w_0 \cdot x^*(-1))] \quad (2-78)$$

Bu şartlarda optimal düzlemin elde edilmesi, katı-marjın destek vektör makinesi olarak da adlandırılmaktadır (Abe, 2005, s. 25).

#### 2.2.2.2. Doğrusal Olarak Ayrılamayan Örüntüler için Optimal Hiperdüzlem

Gerçek dünya problemlerinde verinin doğrusal olarak ayrıldığı durumlarla kıyaslandığında doğrusal olarak ayrılamadığı durumlar çok daha fazladır. Bu durumda veriyi hatasız ayıran bir optimal hiperdüzlem bulunamamaktadır. Optimal hiperdüzlem oluşturmak için sınıflandırma hatasına izin verilerek, en az hata oranı veren hiperdüzlem oluşturulur. Elde edilen marjının içinde kalan ya da yanlış sınıflandırılmış verilerin olması, katı-marjın destek vektör makinelerinin aksine, daha yumuşak ve geçirgen bir yapı ortaya çıkarmaktadır. Burada önemli olan ne kadar

yanlış sınıflandırmaya izin verileceği olacaktır. Şekil 28'de bu durum iki boyutlu uzay için gösterilmektedir (Abe, 2005, s. 28).



Şekil 28 : İki Boyutlu Uzayda Doğrusal Olarak Ayrılamayan Durum (Abe, 2005, s. 28).

Yanlış sınıflandırmaya izin verilebilmesi için hata terimi olarak negatif olmayan  $\xi_i$ ,  $i = 1, \dots, l$  aylak değişkenleri kullanılır.  $\xi_i \geq 0$  ve  $\sigma > 0$  olacak şekilde, (2-79) eşitliğindeki fonksiyon tanımlanıp ve (2-80) eşitliğindeki kısıtlar dahilinde minimize edilecektir (Vapnik, 1995, s. 131).

$$F_{\sigma}(\xi) = \sum_{i=1}^l \xi_i^{\sigma} \quad (2-79)$$

$$\text{Kısıt 1: } y_i [(w \cdot x) + b] \geq 1 - \xi_i, \quad i = 1, \dots, l.$$

$$\text{Kısıt 2: } (w \cdot w) \leq c_n \quad (2-80)$$

Yeterince küçük  $\sigma > 0$  için bu optimizasyon problemi, eğitim hatalarının sayısını minimize eden bir hiperdüzlem tanımlamaktadır. Hesaplama yapılırken,

$\sigma > 0$  şartını yerine getirecek şekilde,  $\sigma = 1$  alındığında (2-81) eşitliğindeki genelleştirilmiş optimal hiperdüzlem elde edilmektedir (Vapnik, 1995, s. 132).

$$w = \frac{1}{C^*} \sum_{i=1}^l \alpha_i y_i x_i \quad (2-81)$$

Burada,  $\alpha_i \geq 0$ ,  $i = 1, \dots, l$  ve  $C^*$  parametrelerinin aşağıdaki konveks (dışbükey) optimizasyon probleminin çözümleri olduğu görülmektedir (Vapnik, 1995, s. 132).

$$\underbrace{W(\alpha, C^*)}_{\text{Maksimizasyon}} = \sum_{i=1}^l \alpha_i - \frac{1}{2C^*} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \frac{c_n C^*}{2} \quad (2-82)$$

$$\text{Kısıtlar: } \sum_{i=1}^l \alpha_i y_i = 0, \quad (2-83)$$

$$0 \leq \alpha_i < 1, \quad i = 1, \dots, l \quad (2-84)$$

$$C^* \geq 0. \quad (2-85)$$

Hesaplama kolaylığı açısından düzenleme parametresi  $C$  ve aylak değişken  $\xi$  ile birlikte doğrusal olarak ayrılamayan veriler için minimizasyon problemi aşağıdaki gibi yazılabilir (Alpaydin, 2010, s. 315):

$$\text{Minimizasyon: } \Phi(w, \xi) = \frac{1}{2} (w \cdot w) + C \sum_{i=1}^l \xi_i \quad (2-86)$$

$$\text{Kısıt: } y_i [(w \cdot x) + b] \geq 1 - \xi_i, \quad i = 1, \dots, l. \quad (2-87)$$

Yukarıdaki optimizasyon probleminin çözümü, doğrusal olarak ayrılabilir veriler için kullanılan teknik ile aynı olacaktır. Buradan yola çıkarak Lagrange fonksiyonu:

$$L(w, b, \alpha, \mu) = \frac{1}{2}(w \cdot w) + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i \{(w \cdot x_i) + b\} y_i - 1 + \xi_i - \sum_{i=1}^l \mu_i \xi_i \quad (2-88)$$

olarak yazılabilir. Burada  $\mu_i$  değişeni  $\xi_i$ 'lerin pozitifliğini sağlamak için kullanılan Lagrange çarpanları olup bu Lagrange fonksiyonunun çözümü için KKT şartları uygulandığında (Borges, 1998, s. 136):

$$\frac{\partial L(w, b, \alpha, \mu)}{\partial w} = w - \sum_{i=1}^l \alpha_i y_i x_i = 0 \quad (2-89)$$

$$\frac{\partial L(w, b, \alpha, \mu)}{\partial b} = - \sum_{i=1}^l \alpha_i y_i = 0 \quad (2-90)$$

$$\frac{\partial L(w, b, \alpha, \mu)}{\partial \xi} = C - \alpha_i - \mu_i = 0 \quad (2-91)$$

eşitlikleri elde edilmektedir. Bu ifadeler Lagrange fonksiyonunda yerlerine yazıldığında, doğrusal olarak ayrılabilir veriler için ortaya çıkan dual problemin aynısı elde edilmektedir.

$$\underbrace{L_D(\alpha)}_{\max \alpha} = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j). \quad (2-92)$$

Ancak dual problemin kısıtları farklılık göstererek aşağıdaki gibi ifade edilmektedir (Alpaydin, 2010, s. 316):



$$\sum_{i=1}^l \alpha_i y_i = 0, \quad 0 \leq \alpha_i < C, \quad i = 1, \dots, l \quad (2-93)$$

Burada, sıfır (0) ve C değerleri arasındaki  $\alpha_i$  Lagrange çarpanlarına karşılık gelen  $x_i$  vektörleri destek vektörleri olarak karşımıza çıkmaktadır. Ayrıca karar fonksiyonunun, doğrusal olarak ayrılabilir veriler için kullanılan fonksiyondan farklı olmadığı ve karar fonksiyonunda aylak değişkenin olmadığı görülmektedir.

### 2.2.3. Doğrusal Olmayan Destek Vektör Makineleri

#### 2.2.3.1. Yüksek Boyutlu Bir Uzaya Eşleme (Çekirdek Hilesi)

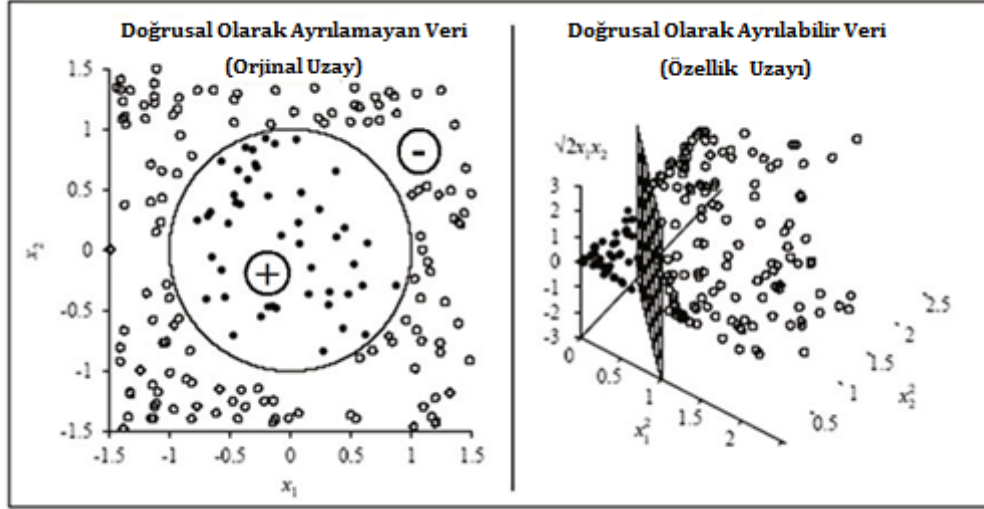
Bir destek vektör makinesinde, genelleme yeteneğini en üst düzeye çıkarmak için optimum hiperdüzlem belirlenmektedir. Eğitim verisinin doğrusal olarak ayrılabilir olmadığı durumlar için ise, hiperdüzlem optimal belirlense de, sınıflandırıcı yüksek genelleme yeteneğine sahip olmayabilir. Bu durumda doğrusal ayrılmayı sağlamak için, orijinal giriş uzayı, özellik uzayı denilen yüksek boyutlu bir iç çarpım uzayına eşleştirilmektedir (Abe, 2005, s. 31).

Destek vektör makinesi, girdi vektörlerini daha yüksek boyutlu bir uzaya doğrusal olmayan önsel seçilmiş bir fonksiyonla eşleyerek, o uzayda optimal ayıran hiperdüzlemi oluşturma temeli üzerine bina edilmektedir (Vapnik, 1995, s. 133).

Orjinal uzayda doğrusal olarak ayrılamayan veriler, doğrusal olmayan bir eşleme fonksiyonu  $\Phi : \mathbb{R}^n \rightarrow H$  aracılığıyla daha yüksek boyutlu H uzayına taşınır. Orjinal uzayda vektörlerin skalar çarpımının kullanılması işlemi bir çekirdek aracılığı ile  $x, y \in \mathbb{R}^n$  olmak üzere (2-94) eşitliğindeki gibi gösterilmektedir (Stoian, 2014, s. 21).

$$K(x, y) = \Phi(x) \cdot \Phi(y) \quad (2-94)$$

Orjinal iki boyutlu uzayda doğrusal olarak ayrılamayan verilerin doğrusal olmayan bir fonksiyon aracılığıyla üç boyutlu bir uzaya eşlenmesi ve o uzayda veriyi doğrusal olarak ayıran optimal hiperdüzlem Şekil 29'da gösterilmektedir.



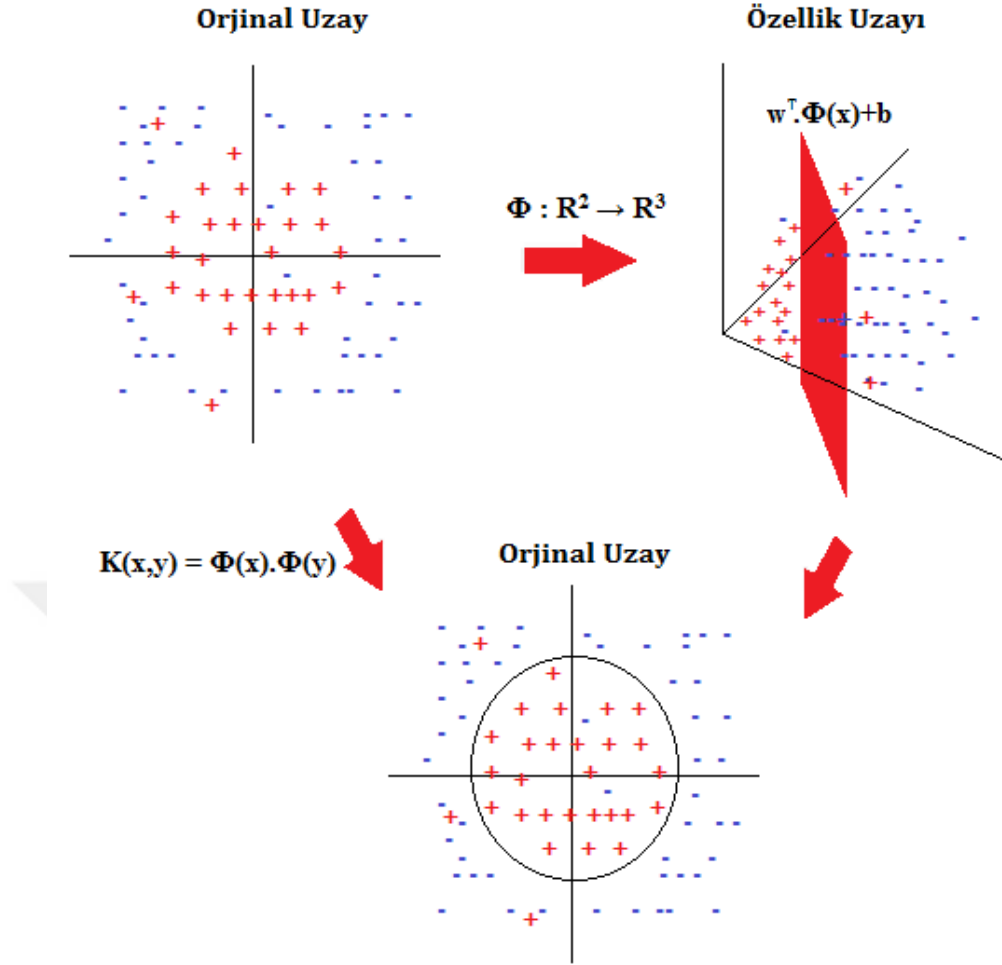
Şekil 29 : Doğrusal Olmayan Eşleme  
(Gorunescu, 2011, s. 234).

Doğrusal olmayan vektör fonksiyonu  $\Phi(x)$  kullanılarak özellik uzayındaki doğrusal karar yüzeyi  $H(x) = w^T \Phi(x) + b$  ile ifade edilebilmektedir. (2-94) eşitliğindeki çekirdek fonksiyon yardımıyla özellik uzayında dual problem aşağıdaki gibi yazılabilmektedir (Abe, 2005, s. 32):

$$\underset{\max \alpha}{L_D(\alpha)} = \sum_{i=1}^M \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (2-95)$$

$$\text{Kısıtlar: } \sum_{i=1}^M \alpha_i y_i = 0, \quad 0 \leq \alpha_i < C, \quad i = 1, \dots, M \quad (2-96)$$

Burada temel problemin çekirdek fonksiyonu ile ilgili olduğuna dikkat edilmesi gerekmektedir. Doğrusal olmayan destek vektör makinelerinde çekirdek fonksiyonu çeşitli şekillerde belirlenebilir ancak bu çekirdek fonksiyonları Mercer koşullarını sağlamak zorundadır.



Şekil 30 : Çekirdek Hilesi.

Şekil 30'da doğrusal olmayan destek vektör makinelerinin işlevi kısaca açıklanmakta olup; iki boyutlu uzayda doğrusal olarak ayrılma imkanı olmayan "+" ve "-" ile işaretlenmiş sınıflandırma verisi,  $\Phi$  doğrusal olmayan eşleme fonksiyonu yardımıyla verinin doğrusal olarak ayrılabilirdiği üç boyutlu özellik uzayına taşındığı görülmektedir. Daha yüksek boyutlu olan bu uzayda veri  $w^T \Phi(x) + b$  hiperdüzlemi ile optimal olarak ayrılabilir. Bu hiperdüzlem orjinal uzayda doğrusal olmayan bir yüzeye karşılık gelmektedir. Ancak Mercer koşullarını sağlayan  $K(x,y)$  çekirdek fonksiyonu, özellik uzayına veriyi taşıyıp orada işlem yapma maliyetinden kurtarmakta ve kısa yoldan orjinal uzayda elde edilen karar yüzeyini sağlamaktadır.

### 2.2.3.2. Mercer Teoremi

Mercer teoremi bir çekirdek fonksiyonunun herhangi bir iç çarpım uzayında tanımlı olması için gerekli koşulları ortaya koymaktadır.

Bir  $K(x, y)$  fonksiyonunun bir çekirdek fonksiyonu olması için şu koşulları sağlaması gereklidir (Özkan, 2008, s. 201):

- Sürekli fonksiyon olmalıdır.
- $K(x_i, x_j) = K(x_j, x_i)$ , yani simetrik olmalıdır.
- Herhangi  $x_1, x_2, \dots, x_n$  değerleri için pozitif tanımlı olmalıdır. Yani

$$K = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & \dots & K(x_1, x_n) \\ \dots & \dots & \dots & \dots \\ K(x_n, x_1) & K(x_n, x_2) & \dots & K(x_n, x_n) \end{bmatrix} \text{ gram matrisi pozitif tanımlı olmalıdır.}$$

### 2.2.3.3. Çekirdek Fonksiyonları

Destek vektör makineleri veriye göre farklı çekirdek seçimi imkanı sunduğu için farklı çekirdekler oluşturarak genelleme performansını artırabilmektedir. Literatürde kullanılan çekirdeklerden biri seçilebileceği gibi özel uygulamalar için literatürden bağımsız Mercer koşulları gibi gerekli koşulları sağlayacak özel çekirdekler oluşturulabilir. Yeni çekirdekler geliştirme devam eden araştırma konuları arasındadır. Literatürde sıkça kullanılan çekirdeklerden aşağıda kısaca bahsedilmektedir.

#### 2.2.3.3.1. Doğrusal Çekirdek Fonksiyonu

Giriş uzayında doğrusal olarak ayrılabilen sınıflandırma problemleri için veriyi daha yüksek boyutlu bir uzaya eşlemeye gerek kalmayacaktır. Bu durumda (2-97) eşitliğindeki doğrusal çekirdek çözüm için yeterli olmaktadır (Abe, 2005, s. 34).

$$K(x_i, x_j) = x_i^T \cdot x_j \quad (2-97)$$

### 2.2.3.3.2. Polinom Çekirdek Fonksiyonu

Polinom çekirdek fonksiyonu,  $d$  parametresi bir doğal sayı olmak üzere aşağıdaki gibi ifade edilmektedir. Eşitlikteki "1" terimi,  $d$  dereceden az veya buna eşit derecedeki çapraz terimlerin dahil edilmesini sağlamaktadır (Abe, 2005, s. 34).

$$K(x_i, x_j) = (x_i^T \cdot x_j + 1)^d \quad (2-98)$$

### 2.2.3.3.3. Radyal Tabanlı Çekirdek Fonksiyonu

Radyal tabanlı, yani dairesel çekirdek fonksiyonu yarıçapı kontrol eden  $\gamma > 0$  parametresiyle aşağıdaki gibi verilmektedir (Hsu, Chang ve Lin, 2003, s. 2):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2). \quad (2-99)$$

### 2.2.3.3.4. Sigmoid Çekirdek Fonksiyonu

Sigmoid çekirdek fonksiyonu, iki katmanlı algılayıcı iç çarpım çekirdeği olarak, Mercer koşullarını sağlamaya yarayan  $\gamma_0$  ve  $\gamma_1$  parametreleri ile tanımlanmaktadır (Gorunescu, 2011, s. 243):

$$K(x_i, x_j) = \tanh(\gamma_0 x_i^T \cdot x_j + \gamma_1). \quad (2-100)$$

## 2.3. DEĞİŞKEN (ÖZNİTELİK) SEÇİMİNDE KULLANILAN YÖNTEMLER

Tahmin çalışmalarında çok değişken kullanımı gerek model kurma gerekse de işlem sayısının fazlalığı ve işlem zamanının artması gibi nedenlerle istenir bir durum değildir. Önemli olan modelin az değişken ve parametre ile tahmin edilmesidir.

Dolayısıyla bu bölümde; literatürde değişken seçme için kullanılan temel yöntemlerden, temel bileşen analizi (principal component analysis), k-en yakın

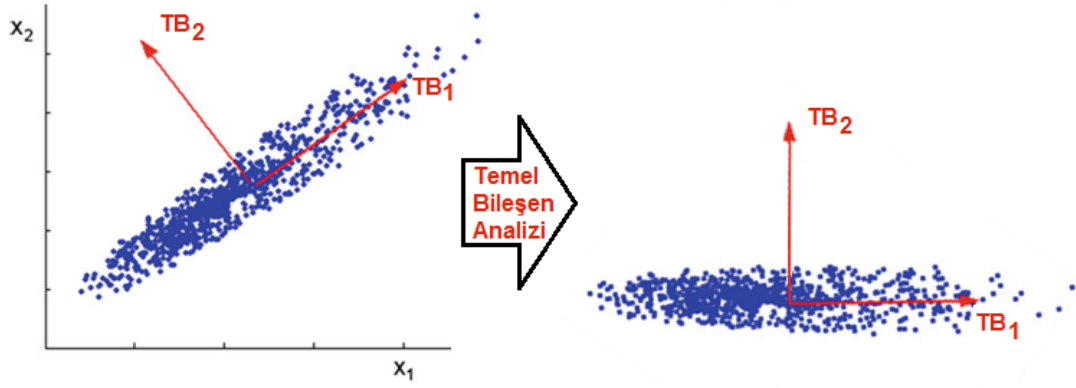
komşuluk (k-nearest neighbors), korelasyon (correlation), lojistik regresyon (logistic regression), kısmi en küçük kareler (partial least squares) ve rastgele orman algoritması (random forest) yöntemleri incelenmektedir. Çalışmada, modeller oluşturulurken bu yöntemler kullanılarak değişken seçimi yapılmış ve elde edilen modellerin tahmin performansları test edilmiştir.

### 2.3.1. Temel Bileşen Analizi (PCA)

Temel Bileşen Analizi (Principal Component Analysis) veriyi daha düşük boyutlu bir uzayda doğru şekilde temsil etmek için kullanılır. Düşük boyutlu uzayda, yüksek boyutlu uzaydaki varyansın büyük oranda korunması, veriyi merkezileştiren ve eksenleri en yüksek varyansa yönelten bir dönüşüm ile mümkün olur. Temel bileşen analizindeki her bir temel bileşen (faktör) yüksek boyutlu uzaydaki orjinal değişkenlerin doğrusal birer kombinasyonu olacak şekilde elde edilir. Olabildiğince yüksek varyansa sahip ilk temel bileşeni yine olabildiğince yüksek varyansa sahip ve ilk temel bileşene ortogonal olan ikinci temel bileşen takip eder (Dougherty, 2013, s. 127). Yani temel bileşenler birbiriyle ilişkisiz olacak şekilde ve en yüksek varyansa göre sırasıyla elde edilir. Varyansın çoğunun temsil edildiği ilk birkaç temel bileşen dışında kalanlar göz ardı edilebilmektedir.

Temel bileşen analizi, verinin geniş bir değişken setinden oluştuğu ve veride aynı yapıyı ölçtüğü düşünülen yüksek korelasyonlu bazı değişkenlerin bulunduğu durumlarda kullanılır. Böylece gözlenen değişkenlerdeki varyansın büyük bölümünü barındıran daha düşük sayıda temel bileşen elde edilmektedir (Verma, 2012, s. 362).

Bileşen analizi, verinin doğru özelliklerini yani özneliklerini bulmak için kullanılan denetimsiz bir yaklaşımdır. Temel bileşen analizinde veriler daha düşük boyutlu bir uzayda temsil edilerek uzay-zaman karışıklığı azaltılmaya çalışılır. Burada temel hedef, veriyi toplam karesel hataya duyarlı şekilde varyansı en iyi temsil eden uzayı elde etmektir (Duda, Hart ve Stork, 2012, s. 655). Şekil 31'de iki boyutlu temel bileşen analizinde eksenlerin en yüksek varyansa sahip TB<sub>1</sub>'e göre döndürüldüğü gösterilmektedir.



Şekil 31 : 2 Boyutta Temel Bileşen Analizi  
(Dougherty, 2013, s. 128).

### 2.3.2. K-En Yakın Komşuluk (KNN)

Cover ve Hart (1967) tarafından önerilen k-en yakın komşuluk modelleri yeni vektörlerin komşu vektörlerin işaretleri referans alınarak sınıflandırılması tekniğine dayanmaktadır. En yakın komşuluktaki manuel seçilen k parametresi, sınıflandırma sürecinde kullanılan en yakın komşu sayısını belirtmektedir. Verilen bir x vektörü için seçilen sınıf, k en yakın komşuluğundaki  $NN(x;k)$  kümesindeki çoğunluk sınıfı olarak tanımlanmaktadır (Guyon, Gunn, Nikraves ve Zadeh, 2006, s. 477).

Makine öğrenmesi yöntemlerinden biri olan en yakın komşuluk sınıflandırıcısında yeni bir gözlem eldeki eğitim gözlemlerine benzerliğine göre sınıflandırılır. Bu benzerliği ölçebilmek için sınıflandırılacak yeni gözlem ile eğitim kümesindeki tüm gözlemler arasındaki Öklid uzaklıkları ölçülür ve en kısa uzaklıktaki gözlemin sınıfına atanır. K-en yakın komşuluk yönteminde ise, sınıflandırma k tane en yakın komşular arasındaki en yaygın sınıfa atama ile yapılmaktadır (Teixeira ve De Oliveira, 2010, s. 6887).

KNN basit bir algoritmayla çalışan örnek tabanlı muhakeme yöntemidir. En yakın komşuların uzaklık ölçüsü veya ayrılık ölçüsü ile hesaplandığı bu yöntemde k genellikle çapraz geçişleme ile hesaplanan 1, 3, 5, 7 vb. gibi tek sayı olarak seçilir (S. Li, 2009, s. 6). Standart k-en yakın komşuluk algoritması, kısaca aşağıdaki gibi ifade edilmektedir (Zheng ve Tropsha, 2000, s. 186):

1. Sınıfı bilinmeyen bir gözlemin (g) eğitim kümesindeki tüm gözlemler ile uzaklıkları hesaplanır.
2. Hesaplanan uzaklıklara göre g gözlemine en yakın k tane gözlem seçilir.
3. Seçilen k tane gözlemin sınıfları incelenerek g gözlemi gruptaki çoğunluğun dahil olduğu sınıfa atanır.

Ölçüm için kullanılan metrik, sınıfa bağımlı seçilebilir ki böylece sınıfa koşullu bir uzaklık elde edilebilir. Böyle bir durumda sınıflandırma süresi artar ancak benzer şekilde sınıflandırma performansı da artacaktır. Bununla beraber, az örneğe sahip sınıflar için belirlenmiş bir değer kullanarak sınıf değeri ile global değer ölçeklendirme parametreleri arasındaki denge sağlanabilecektir (Michie vd., 1994, s. 35).

### 2.3.3. Korelasyon

Çarpım moment korelasyon katsayısı herhangi iki değişken arasındaki doğrusal ilişkinin büyüklüğünü ortaya koyar. n, verideki ikililerin sayısı olmak üzere doğrusal korelasyon (2-101) eşitliği ile hesaplanabilir. Elde edilen R(x,y) değeri -1 ile +1 arasında bir değer almaktadır (Verma, 2012, s. 106).

$$R(x,y) = \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\sqrt{(n \cdot (\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2) \cdot (n \cdot (\sum_{i=1}^n y_i^2) - (\sum_{i=1}^n y_i)^2)}} \quad (2-101)$$

Korelasyon değerinin mutlak değeri değişkenler arasındaki doğrusal ilişkinin gücünü ifade etmektedir. Yüksek değerler iki değişken arasındaki doğrusal ilişkinin yüksek olduğunu göstermektedir (Guyon vd., 2006, s. 348). Pearson korelasyon katsayısı olarak da ifade edilen korelasyon katsayısı -1 değerine yakın ise değişkenler arasında ters yönde yüksek bir doğrusal ilişki olduğu, +1 değerine yakın ise değişkenler arasında aynı yönde yüksek bir doğrusal ilişki olduğu ve korelasyon değeri 0 değerine yakın ise değişkenler arasında doğrusal bir ilişkinin bulunmadığı kabul edilmektedir (Di Lena ve Margara, 2010, s. 680).



Korelasyon katsayısının özellikleri aşağıdaki gibi ifade edilebilir (Verma, 2012, ss. 108–109):

1. Değişkenlere göre simetriktir. Yani,  $R(x, y) = R(y, x)$ .
2. Herhangi iki x ve y değişkenleri için;  $-1 \leq R(x, y) \leq +1$
3. Korelasyon katsayısı iki değişkenin regresyon katsayılarının ( $B_{xy}, B_{yx}$ ) geometrik ortalamasıdır.  $R(x, y) = \sqrt{B_{xy} \cdot B_{yx}}$
4. Değişkenlerden en az birinin sabit olduğu durumda korelasyon hesaplanırken payda sıfır olacağından korelasyon katsayısı hesaplanamaz (Di Lena ve Margara, 2010, s. 680).

#### 2.3.4. Lojistik Regresyon (LR)

Lojistik regresyon doğasında ikili değer (binary) alan olaylar için kullanılan bir tahmin modelidir. Pozitif-negatif, olma-olmama gibi ikili değer alan olaylarda, olma durumu 1 ile olmama durumu ise 0 ile temsil edilmektedir (Verma, 2012, s. 413).

Lojistik regresyon doğrusal regresyona benzer şekilde veriye en uygun doğruyu elde etmeye çalışır. Fakat çıktı değişkeni sadece iki şıklı olduğu için veriye uygun bir doğru optimal elde edilemez. Bu yüzden iki şıklı çıktı değişkenine tam uyan doğruyu elde etmek için lojistik regresyon yönteminde eşitlik (2-102)'de verilen logit veya olasılık dönüşümü kullanılır. Bu dönüşüm çıktının olasılığı (p) ile doğrudan ilgilidir. Olasılıktan farklı olarak logit negatif veya pozitif tüm değerleri alabilir (Kristin L. Sainani, 2014, s. 1157).

$$y = \text{logit}(p) = \ln \frac{p}{1-p} \quad (2-102)$$

Bağımlı ikili değişken y, bağımsız değişken  $x_i$ 'ler ve regresyon katsayıları  $k_i$  olmak üzere lojistik regresyon denklemi eşitlik (2-103)'te verilmektedir (S. Lee, 2005, s. 1482).

$$y = k_0 + \sum_{i=1}^n k_i \cdot x_i \quad (2-103)$$

Eşitlik (2-102) ve (2-103) birlikte çözüldüğünde p gerçekleşme olasılığı ters logaritma yardımıyla aşağıdaki gibi elde edilir (Pagano, 1996, s. 135):

$$p = \frac{e^y}{1 + e^y} \quad (2-104)$$

Lojistik regresyon kullanarak bir olayın ortaya çıkma olasılığı veriye uygun bir lojistik eğri ile tahmin edilir. Ayrıca lojistik regresyon bağımlı değişkenin ikiden daha fazla sınıftan oluştuğu durumlarda da uygulanabilir (Verma, 2012, s. 415).

### 2.3.5. Kısmi En Küçük Kareler (PLS)

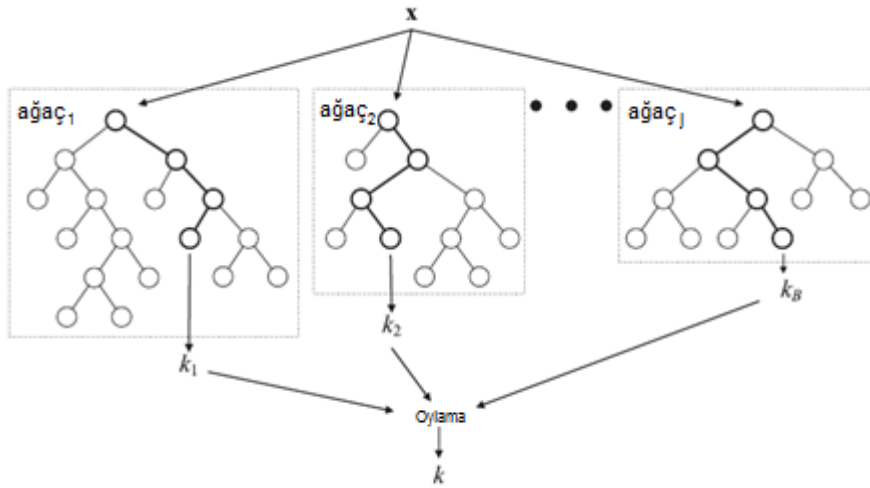
Kısmi en küçük kareler (partial least squares), İsveçli istatistikçi Herman Wold tarafından çok değişkenli zaman serilerinin ekonometrik olarak modellenmesi amacıyla 1975'te ortaya atılmış olup (Guyon vd., 2006, s. 450), yüksek korelasyonlu büyük veri setlerinde seçim metodu olarak da önerilmektedir (Kubinyi, 1996, s. 120). Yüksek ilişkili bağımsız değişkenler, birbirini dengelemek için kullanılan çok sayıda yüksek katsayılı değerler ortaya çıkaracaktır. Bu durumun üstesinden gelmenin bir diğer yolu da katsayı vektörlerinin normunu kontrol altına alabilmektir (Ng, 2013, s. 7).

Kısmi en küçük kareler yöntemi, gizil değişken tabanlı temel bileşen analiziyle ilişkili regresyon metodudur. Gizil değişkenlerin yönleri, değişken vektör blokları arasındaki optimal korelasyonu elde etmek için temel bileşen analiziyle indirgenen boyutlar yardımıyla elde edilir (Kubinyi, 1996, s. 120). Bağımlı ve bağımsız değişkenler arasındaki kovaryansın maksimize edilmesini öngörmektedir (Höskuldsson, 2001, s. 23).

### 2.3.6. Rastgele Orman Algoritması (Randomforest)

Breiman (2001) tarafından geliştirilen rastgele orman algoritması bir sınıflandırma metodu olmasının yanında değişkenlerin önem sırasını da tespit edebilmektedir.

Bir orman rastgele örneklenmiş değişkenlerden seçilmiş örneklerden oluşturulan birçok karar ağacı içermektedir. Burada tahmin karar ağaçlarının çoğunluğuyla elde edilir. Bir değişkenin önem kriterini elde etmek için, önce ikiye ayrılan eğitim kümesinin birinci kısmı ile eğitim, ikinci kısmı ile test yapılarak bu değişkenin doğruluk değeri saptılır.  $j$ . değişkenin önem derecesinin hesaplamak için bu değişkenin değerlerinin sırası eğitim kümesinde değiştirilerek yeni doğruluk değerleri hesaplanır. Hesaplanan iki doğruluk değeri arasındaki fark değişkenin önem değerini vermektedir (Guyon vd., 2006, s. 322).



Şekil 32 : Rastgele Orman Genel Yapısı  
(Verikas vd., 2011, s. 331).

Hem sınıflandırma hem de regresyon problemlerinde kullanılan rastgele orman algoritmasının genel yapısı Şekil 32'de verilmektedir. Yüksek doğruluk oranı için düşük sapma oranı ve düşük korelasyon gereklidir. Maksimum ağaç uzunluğu ile düşük sapma, randomizasyon ile düşük korelasyon elde edilebilir. Dolayısıyla (Verikas, Gelzinis ve Bacauskiene, 2011, s. 331):

1. Her ağaç eğitim kümesinin bir bootstrap örnekleme üzerinde oluşturulur.
2. Bir ağaç oluştururken her düğümde uygun  $N$  değişken arasından rastgele seçilen  $n$  değişken kullanılır.

3. Genellikle  $n = (\log_2 N) + 1$  veya  $n = \sqrt{N}$  olacak şekilde seçilir ve OOB (out-of-bag) hatası minimum olana dek artırıp azaltılır. Her düğümde de, n seçilmiş değişken arasından en iyi bölünmeyi sağlayan bir değişken kullanılır.

Rastgele orman algoritmasında bir değişkenin önemini belirten değer kısaca, gözlem değerlerinin OOB örneklerinde rastgele tekrar sıralandığında ormandaki bir ağacın hata ortalamasındaki artış olarak hesaplanır (Genauer, Poggi ve Tuleau-Malot, 2010, s. 2226).



### 3. BÖLÜM: BORSA ENDEKSLERİ VE TEKNİK GÖSTERGELER

#### 3.1. BORSA İSTANBUL (BIST) VE ENDEKSLER

##### 3.1.1. Borsa İstanbul (BIST)

Ülkemizde 1985 yılında İstanbul Menkul Kıymetler Borsası adıyla açılan ve sermaye piyasalarında faaliyet gösteren şirketlere ve yatırımcılara hizmet veren borsanın adı 5 Nisan 2013 tarihinden itibaren Borsa İstanbul olarak değiştirilmiştir. Borsanın daha önce İMKB olan kısaltması da BIST olarak kullanılmaya başlanılmıştır.

##### 3.1.2. BIST Endeksleri

Hisse senedi endeksleri hisse senedi piyasalarının genel göstergeleri olarak, endeksin kapsadığı hisse senetlerinin fiyatlarını baz alır ve piyasa hakkında anlık durumu yansıtan genel bilgi verir. Dünyada 1884 yılından itibaren kullanılan hisse senedi endeksleri piyasa değeri ağırlıklı, aritmetik ortalama ve geometrik ortalama olmak üzere üç farklı şekilde hesaplanırken, Borsa İstanbul'da piyasa değeri ağırlıklı olarak hesaplanmaktadır. Borsa İstanbul'da Ulusal Pazar için temel endeks olarak XU100 kodlu BIST 100 Endeksi kullanılmaktadır (İMKB, 2010).

Borsa İstanbul'da endekslerin hesaplanması, endeksin kapsadığı hisse senetlerinin toplam piyasa değerleri ile ağırlıklandırılması yöntemiyle yapılmaktadır. Herhangi bir endeksin t zamanındaki değeri (3-1) eşitliğindeki formül aracılığıyla hesaplanabilir (İMKB, 2010).

$$E_t = \frac{\sum_{i=1}^s F_{it} * N_{it} * H_{it}}{B_t} \quad (3-1)$$

- $E_t$  : Endeksin t zamandaki değeri.  
 $s$  : Endekse dahil olan hisse (şirket) sayısı.  
 $F_{it}$  : i. hisse senedinin t zamandaki fiyatı.

- $N_{it}$  : i. hisse senedinin t zamandaki toplam sayısı. (Ödenmiş veya Çıkarılmış sermaye)
- $H_{it}$  : i. hisse senedinin t zamandaki aynen saklamada bulunanlar hariç, MKK saklamasında bulunan miktarının sermayeye göre tamsayıya yuvarlanmış oranı.
- $B_t$  : Bölenin (Düzeltilmiş baz piyasa değeri) t zamandaki değeri.

Çalışmada da ilgilenilen BIST 100 endeksi, Ulusal Pazar'da ve Kurumsal Ürünler Pazarı'nda işlem gören şirket ve yatırım ortaklıkları hisse senetlerinden belirli şartlar dahilinde seçilen 100 hisse senedinden oluşmaktadır. BIST 100 endeksi ilk olarak 3 Ocak 1986 tarihinde hesaplanmıştır. Borsa İstanbul'da hesaplanan diğer endekslerle ilgili bilgi İMKB (2010)'dan yararlanarak hazırlanan Tablo 1'de verilmektedir.

**Tablo 1 : Borsa İstanbul Hisse Senedi Endeksleri**

<b>BIST 100 Endeksi</b>	Ulusal Pazar'da işlem gören şirketlerle, Kurumsal Ürünler Pazarı'nda işlem gören gayrimenkul yatırım ortaklıkları ve girişim sermayesi yatırım ortaklıklarının hisse senetlerinden, belirlenen şartlara göre seçilen 100 hisse senedinden oluşmaktadır. BIST 30 ve BIST 50 Endeksi'nde yer alan hisse senetlerini otomatik olarak kapsamaktadır.
<b>BIST 50 Endeksi</b>	Ulusal Pazar'da işlem gören şirketlerle, Kurumsal Ürünler Pazarı'nda işlem gören gayrimenkul yatırım ortaklıkları ve girişim sermayesi yatırım ortaklıklarının hisse senetlerinden, belirlenen şartlara göre seçilen 50 hisse senedinden oluşmaktadır. BIST 30 Endeksi'nde yer alan hisse senetlerini otomatik olarak kapsamaktadır.
<b>BIST 30 Endeksi</b>	Ulusal Pazar'da işlem gören şirketlerle, Kurumsal Ürünler Pazarı'nda işlem gören gayrimenkul yatırım ortaklıkları ve girişim sermayesi yatırım ortaklıklarının hisse senetlerinden, belirlenen şartlara göre seçilen 30 hisse senedinden oluşmaktadır.
<b>BIST 10 Banka Endeksi</b>	Likiditesi ve piyasa değeri en yüksek olan 10 banka hisse senedinden oluşmaktadır.
<b>BIST 100-30 Endeksi</b>	BIST 100 Endeksi'ne dahil olup da BIST 30 Endeksi'nde yer almayan 70 hisse senedinden oluşmaktadır.
<b>BIST Kurumsal Yönetim Endeksi</b>	BIST pazarlarında (Gözaltı Pazarı ve C Listesi hariç) işlem gören ve Sermaye Piyasasında Derecelendirme Faaliyeti ve Derecelendirme Kuruluşlarına İlişkin Esaslar Tebliği kapsamında, kurumsal yönetim ilkelerine uyuma ilişkin olarak belirlenmiş derecelendirme notu 10 üzerinden en az 6 olan şirketlerin hisse senetlerinden oluşmaktadır.
<b>BIST Tüm-100 Endeksi</b>	BIST Tüm Endeksi'ne dahil olup da BIST 100 Endeksi'nde yer almayan hisse senetlerinden oluşmaktadır. BIST Tüm Endeksi C Listesi'nde işlem görenler hariç olmak üzere, Ulusal Pazar'da işlem gören şirketlerle, Kurumsal Ürünler Pazarı'nda işlem gören gayrimenkul yatırım ortaklıkları ve girişim sermayesi yatırım ortaklıklarının hisse senetlerinden oluşmaktadır.
<b>Sektör Endeksleri ve Alt Sektör Endeksleri</b>	C Listesi'nde işlem görenler hariç olmak üzere, Ulusal Pazar'da işlem gören şirketlerle, Kurumsal Ürünler Pazarı'nda işlem gören gayrimenkul yatırım ortaklıkları ve girişim sermayesi yatırım ortaklıklarının hisse senetlerinin, sektör ve alt sektörler göre gruplandırılması suretiyle oluşmaktadır. Sınai, hizmetler, mali ve teknoloji olmak üzere 4 sektör endeksi ve 20 alt sektör endeksi hesaplanmaktadır.
<b>BIST Menkul Kıymet Yatırım Ortaklıkları Endeksi</b>	Kurumsal Ürünler Pazarı'nda (C Listesi hariç) işlem gören menkul kıymet yatırım ortaklıklarının hisse senetlerinden oluşmaktadır.
<b>BIST İkinci Ulusal Endeksi</b>	İkinci Ulusal Pazar'da (C Listesi hariç) işlem gören hisse senetlerinden oluşmaktadır.
<b>BIST Yeni Ekonomi Endeksi</b>	Yeni Ekonomi Pazarı'nda (C Listesi hariç) işlem gören hisse senetlerinden oluşmaktadır.
<b>BIST Şehir Endeksleri</b>	Ana üretim ya da faaliyet merkezinin bulunduğu şehre göre gruplandırılmış; C Listesi'nde işlem görenler hariç olmak üzere, Ulusal Pazar, İkinci Ulusal Pazar ve Yeni Ekonomi Pazarı'nda işlem gören hisse senetlerinden oluşmaktadır. 30.09.2010 tarihi itibarıyla 9 şehir (Adana, Ankara, Antalya, Bursa, İstanbul, İzmir, Kayseri, Kocaeli ve Tekirdağ) için endeks hesaplanmaktadır

## 3.2. TEKNİK GÖSTERGELER

Teknik analiz sadece piyasada oluşan gerçek fiyat ve hacim bilgilerinin kullanıldığı bir piyasa analiz tekniğidir. Çağımızda bilgisayarların bilgi işleme kapasitelerinin artmasıyla kısa zamanda çok fazla sayıda teknik analiz bilgisinin hesaplanması mümkün olabilmektedir (Wilder, 1987, s. 6). Yani teknik analiz aslında fiyat ve hacim bilgilerini temel alarak fiyat ve hacmin gün içi en yüksek, en düşük ve kapanış değerlerini kullanarak kısa, orta ve uzun vadede çeşitli periyotlar için bu verilerin farklı oranlarıyla elde edilen teknik göstergeler yardımıyla yatırımcılara piyasalarda alış-satış işlemleri için fikir veren bir yöntemdir.

### 3.2.1. Open-High-Low-Close-Volume

Açılış (Open), hisse senedi veya endeksin açılış değerini ifade ederken, gün içi en yüksek değeri (High), gün içi en düşük değeri (Low) ve gün sonu kapanış değeri (Close) değerleri teknik analiz için kullanılan temel bileşenlerdendir. Hacim (Volume) ise hisse senedi veya endeksin o işlem günündeki gördüğü toplam işlem miktarını göstermektedir.

Hacim, piyasalar için önemli bir göstergedir. Trendle birlikte hareket eden hacim fiyatlar ile aynı yönlü hareket eder. Hacim trend ile birlikte fiyat yönünde hareket ediyorsa trendin sağlıklı olduğunu gösterir. Fiyat düşerken hacim artıyorsa bu piyasadaki kaçışı gösterir. Hem düşüş hem de yükseliş trendi için yüksek hacim fiyatın yönünü daha rahat kestirebilmeyi sağlamaktadır (Larson, 2007, ss. 3–4).

### 3.2.2. Aroon Osilatörü (Aroon Oscillator - AO)

Piyasalarda trendler, tıpkı güneş gibi doğar, tepe noktasına yükselir, inişe geçer ve kaybolur. Yeni bir trendin başlangıcına duyarlı bu gösterge için Sanskrit dilinde gündüzü müjdeleyen şafağın ilk ışıkları anlamına gelen "Aroon" adı kullanılmıştır. Bu yeni gösterge fiyat ve zamanı birlikte kullanarak fiyat trendinin gelişiminde trendi izleyen veya anti trend periyotlarını tanımlamak için kullanılabilir (Chande, 1995, s. 370).

Fiyat yönlü trendden sıkışma olarak ifade edilen sabit işlem aralığına geçişi veya ters yönde geçişi hızlı bir şekilde tespit etmeye yarayan Aroon osilatörü

n-periyotluk tanımlanmış bir hareketli zaman penceresi içinde en yakın n-periyot yüksek ve n-periyot düşük fiyatın görülmesinden bu yana geçen periyot sayısını ölçen iki göstergeden oluşur. AroonUp ve AroonDn olarak adlandırılan bu göstergeler aşağıdaki şekilde hesaplanmaktadır (Colby, 2003, s. 102):

$$\text{AroonUp} = 100 \times \left( \frac{n - h}{n} \right) \quad (3-2)$$

$$\text{AroonDn} = 100 \times \left( \frac{n - l}{n} \right) \quad (3-3)$$

Burada "h" en son yüksek n-periyotluk dönemden sonra geçen süreyi, "l" en son düşük n-periyotluk dönemden sonra geçen süreyi ifade ederken "n" ise tanımlanan hareketli zaman penceresinin periyodunu ifade etmektedir.

Chande (1995, s. 371) Aroon osilatörünü matematiksel olarak AroonUp ve AroonDn değerlerinin farkı olarak tanımlamıştır.

$$\text{AroonOsc} = \text{AroonUp} - \text{AroonDn} \quad (3-4)$$

Yukarıdaki (3-2) ve (3-3) eşitliklerinden görüleceği üzere AroonUp ve AroonDn göstergeleri matematiksel olarak 0 ile 100 arasında değerler alabilecektir. Yükselen bir trendde yeni bir n-periyotluk yüksek oluştuğunda h=0 olacağından AroonUp 100 değerini alacaktır ve bu güçlü bir yükseliş trendini gösterecektir. AroonUp 0 değeri aldığı anda ise yükseliş trendinin sonuna gelindiği ifade edilebilir.

AroonUp çizgisinin 70 ile 100 değerleri arasında ve AroonDn çizgisinin 0 ile 30 değerleri arasında ısrarla kalmasıyla güçlü bir yükseliş trendi görülür. Benzer şekilde AroonUp çizgisinin ısrarla 0 ile 30 değerleri arasında ve AroonDn çizgisinin ise 70 ile 100 değerleri arasında kalmasıyla da belirgin bir düşüş trendi görülür. O halde AroonUp çizgisi AroonDn çizgisini yukarı doğru kestiğinde alış, AroonDn çizgisi AroonUp çizgisini yukarı doğru kestiğinde satış kararı verilir. Bu aynı



zamanda Aroon osilatörünün sıfır noktasını kesmesi anlamına gelmektedir (Colby, 2003, s. 102).

### 3.2.3. Bollinger Bantları (Bollinger Bands - BB)

Geliştiricisi John Bollinger'dan adını alan Bollinger bantları, basit hareketli ortalamanın altına ve üstüne iki standart sapma uzaklığında çizilen bantlardan oluşmaktadır (Siegel, Shim, Qureshi ve Brauchler, 2014, s. 22). Fiyat hareketlerini doğrulamak için diğer göstergelerle birlikte kullanılan Bollinger bantları fiyatların yüksek ya da düşük olduğu konusu da fikir vermektedir.

Bollinger bantlarının esnek bir şekilde tüm zaman periyotlarına uygulanabilir olması işlevselliğini artırmaktadır. Bollinger, orta vadeli trend için 20 günlük basit hareketli ortalama ve artı ve eksi iki standart sapma, kısa vadeli trend için 10 günlük basit hareketli ortalama ve artı ve eksi bir buçuk standart sapma ve uzun vadeli trend için 50 günlük basit hareketli ortalama ve artı ve eksi iki buçuk standart sapma değerlerinin kullanılmasının daha sağlıklı sonuçlar üreteceğini ifade etmiştir (Colby, 2003, s. 115).

Kural olarak fiyatların üst banda değdiğinde aşırı alış ve alt banda değdiğinde ise aşırı satışta olduğu göz önünde tutulur (Murphy, 1999, s. 209). Bollinger'e göre bantlardaki hareketler aşağıdaki gibi yorumlanabilir (Siegel vd., 2014, s. 22):

- Bantlar ortalamaya doğru sıkıştığında keskin hamleler meydana gelir.
- Bantlar dışına bir hareket trendin devamı için çağrıda bulunmaktadır.
- Bantlar dışındaki dip ve tepeleri takiben bantlar içinde de dip ve tepelerin oluşması trend değişimine işaret eder.
- Bir banttaki hareket diğer banda gitme eğilimindedir.

Bollinger ayrıca fiyat bantlara değdiğinde veya kestiğinde bantları tek başına mutlak alım veya satım sinyalleri olarak kullanılmamasını, mutlaka bağımsız bir diğer teknik gösterge ile desteklenmesini önermektedir. Bununla birlikte trendlerin sona ermesini belirleyebilmek için Bollinger bantları genişlik indeksi (BBGI) kullanılmaktadır (Colby, 2003, ss. 115–120).

$$BBG\ddot{I} = \frac{\ddot{U}stBB - AltBB}{OrtaBB} \quad (3-5)$$

Burada OrtaBB hareketli ortalamayı ifade ederken, ÜstBB hareketli ortalamasının iki standart sapma üstünü ve AltBB ise hareketli ortalamasının iki standart sapma altını ifade etmektedir.

#### 3.2.4. Chaikin Osilatörü (Chaikin Oscillator - CO)

Chaikin osilatörü, fiyat ağırlıklı hacim hareketlerini inceleyerek toplama/dağıtım endeksi temelli alış veya satış sinyalleri oluşturan bir göstergedir (Colby, 2003, s. 51). Chaikin osilatörünü hesaplamak için toplama/dağıtım endeksinin (Siegel vd., 2014, s. 11) 3-periyotluk üssel hareketli ortalaması, 10-periyotluk üssel hareketli ortalamasından çıkarılır.

#### 3.2.5. Chaikin'in Para Akım Göstergesi (Chaikin's Money Flow - CMF)

Fiyat hareketlerinin altında ne olup bittiğini anlamak için fiyat ve hacim bilgisini kullanan Chaikin'in para akım göstergesi, arz/talep durumunu ortaya koyarak piyasanın toplama veya dağıtım durumunu belirler. Chaikin bu göstergede, bir ayı en iyi 21 günlük periyodun temsil edeceğini, ancak diğer periyotların da kullanılabileceğini belirtmektedir (Narcouzi, 2000, s. 42).

Chaikin'in para akım göstergesini hesaplamak için toplama/dağıtım (A/D) endeksinin (Siegel vd., 2014, s. 11) 21 günlük değerlerinin toplamı yine 21 günlük hacim toplamına bölünür.

$$CMF = \frac{\sum_{i=1}^{21} \left( \frac{(C_i - L_i) - (C_i - H_i)}{(H_i - L_i)} \right) * V_i}{\sum_{i=1}^{21} V_i} \quad (3-6)$$

Eğer Chaikin'in para akım göstergesi değeri 0 ile 0,10 değerleri arasında ise zayıf alış, 0 ile -0,10 arasında ise zayıf satış kabul edilir. Değer 0,10'un üzerinde ise boğa piyasası, -0,10'un altında ise ayı piyasası etkindir. Değerlerin 0,25'in üstünde daha yüksek fiyatlara, -0,25'in altında olması ise daha düşük fiyatlara işaret etmektedir (Narcouzi, 2000, s. 43).

### 3.2.6. Chaikin'in Dalgalanma Göstergesi (Chaikin's Volatility - CV)

Chaikin'in dalgalanma göstergesi, teknik analist Marc Chaikin tarafından geliştirilen göstergelerden biridir. Bu gösterge bir menkul kıymetin yüksek ve düşük fiyatları arasındaki yayılmanın düzeltilmiş hızını ölçmektedir. Önce günlük en yüksek değerden günlük en düşük değer çıkarılarak günlük farklar bulunur. Akabinde günlük farkların 10 günlük üssel hareketli ortalamaları alınır ve bu 10 günlük üssel hareketli ortalamaların değişim oranı hesaplanır. Kısa bir zaman periyodunda Chaikin'in dalgalanma göstergesindeki yükseliş yakın zamanda bir dip olacağını, uzun bir zaman periyodunda göstergede düşüşün olması ise yaklaşan bir tepe noktasını haber vermektedir (Colby, 2003, s. 729).

### 3.2.7. Chande Momentum Osilatörü (Chande Momentum Oscillator - CMO)

Chande Momentum Osilatörü hesaplama yönünden göreceli güç endeksine benzeyen ve eşitlik (4-11)'deki formülle hesaplanabilen bir fiyat momentumu osilatörüdür. Göreceli güç endeksi 0 ile 100 arası değerler alırken Chande momentum osilatörü -100 ile +100 arasında değerler alır. +50 değerinin üstü aşırı alım iken -50 değerinin altı aşırı satım seviyelerini göstermektedir (Chande ve Kroll, 1994, s. 94).

$$CMO = 100 * \frac{Ty - Ta}{Ty + Ta} \quad (3-7)$$

*Ty: Periyottaki yukarı yönlü momentumların toplamı*  
*Ta: Periyottaki aşağı yönlü momentumların toplamı*

### 3.2.8. Denge İşlem Hacmi Göstergesi (On Balance Volume - OBV)

Denge işlem hacmi göstergesi, Joseph E. Granville tarafından 1976 yılında fiyat ve hacim eğilim ölçümü olarak kullanılmıştır. Bugünkü kapanış fiyatı dünün kapanış fiyatından yüksek ise bu günün hacim değeri pozitif, eğer bugünkü kapanış fiyatı dünün kapanış fiyatından düşük ise bu günün hacim değeri negatif olarak işaretlenerek elde edilen pozitif ve negatif hacim değerlerinin kümülatif değerleri ile ilgilenilmektedir (Murphy, 1999, s. 165).

Denge işlem hacmi göstergesi, B en son kapanış değeri, Ö bir önceki kapanış değeri ve V hacim değeri olmak üzere, matematiksel olarak aşağıdaki şekilde ifade edilmektedir (Colby, 2003, s. 768) :

$$OBV = \text{Bir Önceki } OBV + \left( \frac{B - \text{Ö}}{|B - \text{Ö}|} \right) * V \quad (3-8)$$

Denge işlem hacmi göstergesindeki değişim, fiyattaki değişimden önce gelmektedir. Fiyattaki değişim ile denge işlem hacmi göstergesi uyumsuz hareket ediyorsa trend yönünün değişeceği şeklinde yorumlanmaktadır (Siegel vd., 2014, s. 324).

### 3.2.9. Göreceli Güç Endeksi (Relative Strength Index - RSI)

J. Welles Wilder Jr. tarafından 1978'de New Concepts in Technical Trading Systems adlı kitabında tanımlanan göreceli güç endeksi, göreceli güç (RS) olarak ifade edilen n-periyotluk yukarı kapanış hareketli ortalamasının n-periyotluk aşağı kapanış hareketli ortalamasına oranı ile eşitlik (3-9)'daki gibi hesaplanır (Siegel vd., 2014, s. 180).

$$RSI = 100 - \frac{100}{(1 + RS)} \quad (3-9)$$

Göreceli güç endeksi, aşırı alışı veya aşırı satışı ortaya koyan bir momentum osilatörüdür. Seviyelerin ayarlanabilmesine karşın genellikle 70 üzeri aşırı alışı ve 30 altı aşırı satışı gösterir. RSI çizgisi 70'in üzerinde hareket etmesi satış zamanlaması için 30'un altında hareket etmesi ise alışı zamanlaması için kullanılabilir (Thomsett, 2012, s. 16).

### 3.2.10. Hareketli Ortalamalar (Moving Averages - MA)

Hareketli ortalama, incelenen periyotlar için elde edilen kapanış fiyatlarının ortalamasıdır. Genellikle kısa vadeli alım ve satım sinyalleri için 30 günlük periyot kullanılmaktadır. Fiyat 30 günlük hareketli ortalamanın üzerine çıktığında, yükseliş potansiyeli artarken, fiyat 30 günlük hareketli ortalamanın altında hareket ettiğinde aşağı yönlü hareket ihtimali artar. Tüm hareketli ortalamalar içerisinde 200 günlük

hareketli ortalama en güçlü hareketli ortalama olarak önemli bir destek noktası veya direnç olarak adlandırılır (Larson, 2007, ss. 11–12). Kısa vadeli hareketli ortalamanın daha uzun vadeli hareketli ortalamaları yukarı doğru kırması halinde al sinyali, tersi durumda sat sinyali olarak yorumlanmaktadır (Turhan Korkmaz, 2012, s. 345).

### 3.2.10.1. Basit Hareketli Ortalama (Simple Moving Average - SMA)

$C_i$  i. periyodun kapanış fiyatı, n hareketli ortalama hesabında kullanılan periyotların toplam sayısı ve  $SMA_k$  k. periyottaki basit hareketli ortalama olmak üzere matematiksel olarak basit hareketli ortalama aşağıdaki gibi hesaplanmaktadır (Colby, 2003, s. 644):

$$SMA_k = \frac{1}{n} \sum_{k-n+1}^k C_i \quad (3-10)$$

Basit hareketli ortalama n-periyot için son n günlük kapanış fiyatlarının aritmetik ortalaması olarak çoğu teknik analist tarafından kullanılmaktadır. Ancak iki noktada yararı sorgulanmaktadır. İlk eleştiri sadece seçilen periyodun tek dönem olarak ele alınmasıdır. İkincisi ise basit hareketli ortalama her günün fiyatını eşit ağırlıklı olarak değerlendirmesidir. Bazı analistler her günün fiyatını eşit ağırlıklı değil, yakın zamandaki fiyat hareketlerine daha fazla ağırlık verilmesi gerektiğine inanmaktadırlar (Murphy, 1999, s. 199).

### 3.2.10.2. Üssel Hareketli Ortalama (Exponential Moving Average - EMA)

Aşırı yüksek ve düşüklerden kaynaklanan bozulmaların üstesinden gelmek için üssel hareketli ortalama son kapanış fiyatlarını daha önceki kapanış fiyatlarına göre daha fazla ağırlıklandırmaktadır. Çoğu teknik analizci üssel hareketli ortalamayı, basit hareketli ortalamadan daha doğru sonuçlar veren bir gösterge olarak görürler. Üssel hareketli ortalama istenen periyot için öncelikle basit hareketli ortalama hesaplanarak aşağıdaki eşitlik (3-11) aracılığıyla hesaplanır (Siegel vd., 2014, s. 196). Periyot büyüdükçe üssel basit hareketli ortalama yeni veriye daha az duyarlı hale gelmektedir (Colby, 2003, s. 262).

$$EMA = \left[ Son\ EMA * \left( 1 - \frac{2}{n+1} \right) \right] + \left[ SK * \frac{2}{n+1} \right] \quad (3-11)$$

(SK: Son kapanış değeri, n: Hareketli ortalamannın uzunluğu)

### 3.2.10.3. Ağırlıklı Hareketli Ortalama (Weighted Moving Average - WMA)

Ağırlıklı hareketli ortalama hesaplandığı periyotta son günlere daha fazla ağırlık vererek son günlerin ortalamadaki önemini artırır. Örneğin 5 günlük ağırlıklı hareketli ortalama hesaplanırken, beş gün önceki kapanış değerinin ağırlığı bir olarak alınır ve benzer şekilde bir gün önceki kapanışa geleceye birer artırılarak bir gün önceki kapanış değerinin ağırlığı beş olarak elde edilir. Elde edilen ağırlıklı toplam ağırlıkların toplamı olan 15 ile bölünerek ağırlıklı ortalama bulunur (Brooks, 2006, ss. 128–129). Ağırlıklı hareketli ortalama n periyot uzunluğu ve  $C_s$  son kapanış değeri olmak üzere;

$$WMA_n = \frac{n * C_s + (n - 1) * C_{s-1} + (n - 2) * C_{s-2} + \dots + 1 * C_{s-n+1}}{n + (n - 1) + (n - 2) + \dots + 1} \quad (3-12)$$

şeklinde hesaplanabilmektedir.

### 3.2.11. Hareketli Ortalamalar Yakınlaşma - Uzaklaşma Göstergesi (Moving Average Convergence Divergence - MACD)

Hareketli Ortalama Yakınlaşma-Uzaklaşma göstergesi 1970'lerin sonunda ortaya atılmıştır. Kısa, orta ve uzun vade için kullanılabilen bu teknik gösterge aynı zamanda gün içi daha kısa periyotlarda kullanımında da sağlıklı sonuçlar vermektedir. Gösterge hassas giriş ve çıkış sinyalleri üretebilme yeteneğine sahiptir (Appel, 2005, s. 165).

Hareketli ortalama yakınlaşma-uzaklaşma göstergesi 9,12 ve 26 periyotluk üç üssel hareketli ortalama ile hesaplanmaktadır. Hareketli ortalamadaki zaman periyotlarını artırarak daha az alım satım sinyali elde edilebilirken azaltarak daha çok alım satım sinyali elde edilebilir. Herhangi bir fiyat bileşeni için (genellikle kapanış fiyatı) 12 periyotluk EMA'dan 26 periyotluk EMA çıkarılarak hızlı MACD çizgisi üretilir. Hızlı MACD çizgisi için 9 periyotluk EMA kullanılarak yavaş MACD

çizgisi elde edilir. Hızlı çizgi yavaş çizgiyi yukarı doğru keserse alış sinyali, aşağıya doğru keserse satış sinyali olarak yorumlanmaktadır. MACD sinyalleri piyasadaki değişimlere hızlı bir şekilde tepki vermektedirler (Person, 2004, ss. 142–143).

### 3.2.12. Mal Kanal Endeksi (Commodity Channel Index - CCI)

Piyasalarda geçmiş fiyatlardan güncel fiyatlardaki örüntüyü bozmadan yararlanmak amacıyla mal kanal endeksi, üssel hareketli ortalamaları değil basit hareketli ortalamaları kullanmaktadır. Mal kanal endeksi aşağıdaki dört aşamada hesaplanmaktadır (Lambert, 1983, ss. 120–124):

1. Her gün için ortalama fiyat  $OF = (\text{En yüksek} + \text{En düşük} + \text{Kapanış}) / 3$  ile hesaplanır.
2. Seçilen periyot için  $SMA(OF)$  hesaplanır.
3. Ortalama sapma  $OS = \frac{1}{n} * \sum_{i=1}^n |OF_i - SMA(OF)_i|$  hesaplanır.
4.  $OF_1$  son günün ortalama değeri olmak üzere  $CCI = \frac{OF_i - SMA(OF)_i}{OS * 0,015}$  olarak hesaplanır.

Aşırı alış ve aşırı satış osilatörü olan mal kanal endeksi +%100 değerinin üzerinde aşırı alış ve -%100 değerinin altında aşırı satış olarak yorumlanmaktadır (Murphy, 1999, s. 238). Mal kanal endeksi göstergesindeki rastgele dalgalanmaların çoğu yukarıda +%100 ile aşağıda -%100 kanalları arasında olması beklenir. Bu kanalın dışındaki hareketler rastgele olmayan hareketler olup, bu geniş hareketler alış ve satış fırsatları olarak değerlendirilmelidir (Colby, 2003, s. 155).

### 3.2.13. Momentum Göstergesi (Momentum Indicator - MI)

Momentum bugünkü fiyat ile belirli bir periyot öncesindeki fiyat farkı ile hesaplanır. Fiyat değişimini, trendin hızı ve gücünü ortaya koyan momentum göstergesi kısa vadeli değişimlerin izlenmesinde kullanıldığı için seçilecek periyot 30'un altındadır (Turhan Korkmaz, 2012, s. 347).

Momentum göstergesi,  $K$  son kapanış fiyatı ve  $K_n$  n periyot önceki kapanış fiyatı olmak üzere (3-13)'deki eşitlik ile hesaplanabilir. Bu formül 100 değeri civarında değerler alan bir momentum göstergesi üretmektedir. Bazı analizciler

değerden 100 eksilterek sıfır civarında değerler elde ederler. Diğer göstergelerdeki gibi momentum göstergesinde de fiyat ile uyumsuzluklar dikkate alınmalıdır. Gösterge ciddi oranda yükseltilere çıktığında veya aşağıya düştüğünde içinde bulunulan trendin devam edeceği şeklinde yorumlanmaktadır (Colby, 2003, ss. 400–401).

$$MI = \frac{K}{K_n} * 100 \quad (3-13)$$

Momentum çizgisi her zaman fiyat hareketinden bir adım öndedir. Fiyatlardaki azalma veya artışa öncülük eder, sonrasında fiyat trendi devam ettiği halde yatay seyretmeye başlar. Fiyatlar yatay hale gelmeye başladığında momentum göstergesi ters yönde harekete geçmektedir (Murphy, 1999, s. 230).

#### **3.2.14. Ortalama Doğruluk Aralığı (Average True Range - ATR)**

Welles Wilder tarafından 1978'de formülize edilen ortalama doğruluk aralığı piyasadaki değişkenliği ölçmektedir. Geniş aralık daha fazla oynaklığı ifade eder (Siegel vd., 2014, s. 314). Piyasadaki oynaklığı ölçen ortalama doğruluk aralığı, periyot boyunca doğru fiyat aralıklarının ortalaması ile hesaplanır. Doğru aralık (DA), bugünün en yüksek (EY) ile en düşüğü (ED) arasındaki, bir önceki kapanışı (ÖK) ile bugünün en yükseği arasındaki ve bir önceki kapanış ile bugünün en düşüğü arasındaki farklardan en büyüğüdür. Ortalama doğru aralık ise istenen periyot için doğru aralıkların (Eşitlik 3-14) hareketli ortalaması ile elde edilmektedir. Yüksek ATR değerleri genellikle panik satışı takip eden dip fiyatları, düşük ATR değerleri ise tepe fiyatlarda uzayan yatay hareketler sırasında ortaya çıkmaktadır. Yatay ATR hareketleri ise ani fiyat hareketlerinin işareti olarak yorumlanmaktadır (Larson, 2007, s. 85).

$$DA = \max\{|EY - ED|, |\ÖK - EY|, |\ÖK - ED|\} \quad (3-14)$$



$$ATR = MA(DA)$$

(3-15)

### 3.2.15. Para Akışı Endeksi (Money Flow Index - MFI)

Piyasa giriş veya piyasadan çıkış yapan para miktarını belirleyen işlem hacmi göstergesidir. Para akış endeksi aşağıdaki gibi beş adımda hesaplanabilir (Siegel vd., 2014, s. 118):

1. Her gün için ortalama fiyat  $OF = (En\ yüksek + En\ düşük + Kapanış) / 3$  ile hesaplanır.
2. Para akışı  $PA = OF * Hacim$  ile hesaplanır.
3. Belirlenen periyot için toplam pozitif para akışı ve toplam negatif para akışı belirlenir.
4. Para oranı  $PO = Pozitif\ Para\ Akışı / Negatif\ Para\ Akışı$  ile hesaplanır.
5. Son olarak  $MFI = 100 - \frac{100}{1 + PO}$

Fiyattaki aşırı alıŖ veya aşırı satıŖ koŖullarını tespit edebilmesi para akışı endeksini giriŖ veya çıkıŖ zamanlaması için önemli bir gösterge haline getirmektedir. Genellikle periyodun 14 olarak seçildiđi göstergede önemli noktalardan 20 deđerinin altında aşırı satıŖ ve 80 deđerinin üstünde aşırı alıŖ olarak deđerlendirilir (Thomsett, 2012, s. 91). Yani endeksi çizgisi 20 deđerinin altına dođru kestiđinde alıŖ sinyali, 80 deđerinin üstüne dođru kestiđinde satıŖ sinyali olarak yorumlanmaktadır.

### 3.2.16. Parabolik SAR (Parabolic SAR)

Wilder tarafından geliŖtirilen Parabolik sistem, zaman/fiyat yön deđiŖikliđi tespit sistemidir. SAR, İngilizcedeki dur ve geri dönün (Stop And Reverse) kısaltmasıdır (Murphy, 1999, s. 381). Uygun çıkıŖ noktalarını iŖaret eden göstergede fiyatlar SAR noktasının üzerine yükseldiđinde kısa dönemli pozisyon alınmalı, fiyatlar SAR noktasının altına indiđinde ise uzun dönemli pozisyon kapatılmalıdır. Fiyatlardaki artıŖ düz bir dođru Ŗeklinde arttıđı noktaya dođru hız kazandıđında

parabolik bir eğri oluşmakta ve kısa bir zaman aralığında gerçekleşen bu durum tepe fiyatından uzunca bir düşüşle sonuçlanmaktadır (Siegel vd., 2014, s. 209).

SAR hesaplaması bir önceki periyodun SAR değeri kullanılarak iterasyon şeklinde (3-16)'daki eşitlik ile elde edilir. Burada ivme faktörü (İF) başlangıçta 0,02 olarak belirlenen ve sonrasında karşılaşılan her en yüksekte 0,02 artırılarak kullanılan katsayı değeridir. Yeni alış sinyalleri için SAR değeri kısa pozisyonda kaydedilen en düşük değere eşit olmalıdır (Colby, 2003, s. 495).

$$SAR_{uzun} = SAR_{önceki} + İF * (uzun pozisyonda en yüksek fiyat - SAR_{önceki}) \quad (3-16)$$

### 3.2.17. Stokastik Osilatörü (Stochastic Oscillator - SO)

Stokastik şimdiki fiyat ile periyot boyunca görülen fiyat seviyelerini karşılaştırmaktadır. Olası dönüş noktalarını belirlemek için destek ve direnç noktaları en önemli bileşenlerindedir (Thomsett, 2012, s. 106).

George C. Lane tarafından geliştirilen stokastik osilatörü fiyatlar ilerlediğinde kapanış fiyatlarının da periyodun tepe noktasına doğru ilerleme eğilimi olması prensibine dayalı aşırı alış ve satış osilatörüdür. Düşüş trendinde kapanış fiyatları genellikle periyodun dip noktalarına yakın değerlerde görülür. Yapısında 9 ve 14 günlük periyotlar bulunan stokastik değerleri basit hareketli ortalama ile düzgünleştirilir. Aşırı alım ve aşırı satım koşullarını belirtecek zamanlama aracı olarak kullanılırken diğer teknik göstergelerden de yararlanılmalıdır (Siegel vd., 2014, s. 266).

Stokastik osilatöründeki verilen periyot için oluşturulan iki çizgi 0 ile 100 arasında değerler alan stokastik(%K) ve stokastik(%D)'yi gösterir. C periyodun en son kapanış fiyatını, D n-periyot boyunca görülen en düşük değer ve Y n-periyot boyunca görülen en yüksek değer olmak üzere K eşitlik (3-17)'deki gibi hesaplanmaktadır (Person, 2004, s. 145).

$$K = \frac{C - D}{Y - D} * 100 \quad (3-17)$$

Lane, K'nın 3-periyotluk basit hareketli ortalama ile düzgünleştirilmesi ile stokastik (%K)'nin elde edildiğini ve stokastik (%K)'nin 3-periyotluk basit hareketli ortalaması ile stokastik (%D)'nin hesaplandığını belirtmektedir. Burada elde edilen stokastik (%K) ve stokastik (%D) çizgileri yorumlama için kullanılır. Aşağıya doğru fiyat hareketlerinde stokastik (%K) %25'in altına indiğinde al sinyali olarak yorumlanır. %15'in altına indiğinde ise daha anlamlı şekilde aşırı satış bölgesi olduğu ve al sinyali olarak yorumlanır (Colby, 2003, s. 664). %70 seviyesinin üzerinde stokastik (%K) çizgisi stokastik (%D) çizgisini aşağıya keserek birlikte aşağıya doğru yönlendiğinde sat sinyali, %30 seviyesinin altında yukarı doğru kestiğinde ise al sinyali olarak yorumlanmaktadır (Person, 2004, s. 146).

### 3.2.18. Stokastik Momentum Endeksi (Stochastic Momentum Index - SMI)

William Blau tarafından 1991'de önerilen stokastik momentum endeksi, C periyodun en son kapanış fiyatını, D n-periyot boyunca görülen en düşük değer ve Y n-periyot boyunca görülen en yüksek değer olmak üzere (3-18)'deki eşitlikle hesaplanabilmektedir (Blau, 1991, s. 27).

$$SMI = \frac{C - 0,5 * (Y + D)}{0,5 * (Y - D)} * 100 \quad (3-18)$$

Blau (1991, s. 27)'ya göre endeks fiyatları -100 ile + 100 arasında ölçeklemektedir. Gösterge +40'ın üzerindeyken fiyatların yüksek seviyede olduğu, -40'ın altında olduğunda da düşük seviyede olduğu söylenebilir. Stokastik momentum endeksinin üssel hareketli ortalaması ile oluşturulan sinyal çizgisi endeksin altında ise yükseliş trendi, altında ise düşüş trendi olarak yorumlanmaktadır.

### 3.2.19. Williams'ın %R Göstergesi (Williams' %R)

Lary Williams tarafından geliştirilen Williams'ın %R, aşırı alı ve satış bölgelerini gösteren bir momentum göstergesidir. Stokastik osilatöre çok benzer

şekilde hesaplanır ve yorumlanır. Bugünün kapanış değeri seçilen n-periyottaki en yüksek değerinden çıkarılarak elde edilen fark aynı periyottaki toplam fiyat aralığına bölünerek elde edilmektedir (Murphy, 1999, s. 249).

$$Williams'\%R = \frac{\text{periyottaki en yüksek} - \text{bugünün kapanışı}}{\text{periyottaki en yüksek} - \text{periyottaki en düşük}} \cdot (-100) \quad (3-19)$$

Williams'in %R göstergesi -100 ile 0 arasında dalgalanır. Eğer değer -80'in altına inerse piyasanın aşırı satış bölgesinde olduğu, -20'nin üstünde ise aşırı alım bölgesinde olduğunu belirtir. Bu göstergenin en ilginç özelliği fiyatlardaki dönüşü tahmin edebilmesidir. Gösterge hemen her zaman piyasa fiyatının tepe ve dönüş yapmasından bir kaç gün önce bir tepe ve dönüş yapar. Yani Williams'in %R göstergesi piyasadaki fiyat değişimlerinden önce genellikle bir dalga ve dönüş oluşturmaktadır (Siegel vd., 2014, s. 342).

### 3.2.20. Yönsel Hareket Göstergesi (Directional Movement Indicator - DMI)

Yönsel hareket göstergesi Welles Wilder tarafından geliştirilen filtrelenmiş momentum göstergesidir. Kompleks bir trend göstergesi olarak yatay piyasalarda karsılıktan uzak tutmak için piyasaların önemli trend özelliklerini sergilediklerinde girişe izin verir. Piyasa anlamlı bir trend ya da yönsel davranış göstermediğinde yatırımcıları piyasadan uzak tutar. Yönsel hareket (DM) önceki periyodun fiyat aralığı dışındaki şimdiki periyodun fiyat aralığının en geniş bölümü olarak tanımlanmaktadır (Colby, 2003, s. 212).

Yönsel hareket göstergesinde öncelikle yönsel hareketin yönünü belirten +DI ve -DI olmak üzere iki farklı çizgi oluşturulur. +DI pozitif (yukarı) yönlü hareketi ölçerken -DI negatif yönlü hareketi ölçmektedir (Murphy, 1999, s. 386).

Üssel hareketli ortalamalar ve oranları kullanarak en yüksek, en düşük ve kapanış değerlerini 0 ile 100 arasına ölçekleyen yönsel hareket göstergesi, bugünün en yüksek (EY), en düşük (ED), kapanış (BK), bir önceki en yüksek (ÖY), en düşük

(ÖD), kapanışı (ÖK) ve doğru aralık (Eşitlik 3-14) olmak üzere aşağıda verilen aşamalar ile hesaplanır (Colby, 2003, s. 213).

1. Verilen periyot için önce pozitif ve negatif yönsel hareketler hesaplanır ve bu değerlerden küçük olanı sifıra eşitlenir. Eğer negatif değerler varsa onlar da sifıra eşitlenir.

$$\begin{aligned} PDM &= EY - ÖY \\ NDM &= ED - ÖD \end{aligned} \quad (3-20)$$

2. DA (Eşitlik 3-14) hesaplanır.
3. +DI ve -DI değerleri hesaplanır.

$$+DI = \frac{PDM}{DA} \quad \text{ve} \quad -DI = \frac{NDM}{DA}$$

4. Bu değerlere ek olarak Wilder aşağıdaki yönsel hareket endeksi DX'in n-periyotluk (genellikle 14) üssel ortalaması olan ortalama yönsel endeks ADX çizgisi ile bir periyot önceki ADX ile şimdiki ADX'in aritmetik ortalaması olan ADXR çizgilerinin de hesaplanmasını önermektedir.

$$DX = 100 * \frac{(+DI) - (-DI)}{(+DI) + (-DI)} \quad (3-21)$$

ADX çizgisi yükselirken trendin güçlendiği, 40 seviyeden fazla düştüğünde ise trendin zayıfladığı, tekrar 20 seviye üstüne çıkması yeni bir trendin başlangıcı olarak yorumlanır (Murphy, 1999, s. 387). Genel kural olarak, ADXR değerinin 20'nin altında olması trendin kaybolduğunu, 25'ten fazla olması ise trendin oluştuğunu belirtmektedir (Colby, 2003, s. 213).

## **4. BÖLÜM: DALGACIK DÖNÜŞÜMÜ VE DESTEK VEKTÖR MAKİNELERİ İLE TAHMİN: BIST ÜZERİNE BİR UYGULAMA**

### **4.1. ÇALIŞMANIN AMACI VE KAPSAMI**

Temelde sinyal analizinde kullanılan bir yöntem olan dalgacık dönüşümü ile veri madenciliği yöntemlerinden bir yapay öğrenme algoritması olan destek vektör makineleri yöntemlerinin birlikte finansal zaman serileri üzerinde kullanılmasının amaçlandığı bu çalışmada; teknik göstergeler, çeşitli endeksler, döviz ve altın pariteleri ve haftanın günleri gibi bağımsız değişkenlerden uygun olanlarının yardımıyla, borsa endeksinin yönünü ikili sınıflandırma yöntemiyle tahmin edebilecek günlük, haftalık, aylık gibi farklı periyotlarda modeller geliştirilmek istenmektedir. İkinci aşamada ise, geliştirilen modellerin borsa endeksi üzerindeki tahmin performansları ayrı ayrı ölçülerek performanslarına göre sıralanması hedeflenmektedir. Ayrıca ortaya çıkacak bulgulardan hareketle, dalgacık dönüşümünün destek vektör makinelerinin tahmin performansına etkisinin olup olmadığı ve varsa ne tür bir etkisinin olduğunun ortaya çıkarılması da amaçlanmaktadır.

Buradan hareketle çalışmanın amacı; dalgacık dönüşümü ve destek vektör makineleri ile borsa endeksinin yönünü tahmin edebilecek bir model geliştirmek, geliştirilen modelin borsa endeksi üzerindeki tahmin performansını ölçmek ve dalgacık dönüşümünün destek vektör makinelerinin tahmin performansına etkisini incelemektir.

### **4.2. UYGULAMANIN MODELİ**

Borsa endeksini etkileyen çalışmalar incelendiğinde: Brigham ve Houston (2007) faiz oranlarıyla, Fang ve Loo (1996) döviz kurunda meydana gelen değişmelerle, Seler (1996) borsa endeksinin haftanın günleri ile borsa endeksinin ilişkili olduğunu saptamışlardır. Durukan (1999) çalışmasında, borsa endeksi

üzerinde etkili olan değişkenlerin altın fiyatları, ekonomi büyüme hızı, enflasyon, faiz oranı gibi makroekonomik değişkenler olduğunu; Özalp ve Anagün (2003), hisse senedi fiyatlarının; işlem adedi, işlem hacmi, TEFE, TÜFE, Amerikan dolar kuru gibi ekonomik göstergelerin değerlerindeki değişmelerin etkilediğini belirtmektedir. Yüksel (2010) ise, menkul kıymetler borsalarının ülkelerin hem sektörel ekonomiyi hem de genel ekonomiyi ilgilendiren içsel ve dışsal faktörlerden etkilendiğini belirtmiştir. Emir (2013b), Tekin (2013), Tayyar ve Tekin (2013) ve Kara vd. (2011) ise teknik göstergeler ve günlük kapanış fiyatlarının borsa endeksinin tahmininde etkili olduğunu ortaya koymuşlardır. Son olarak, Yakut ve Elmas (2014)'ın çalışmalarında, haftanın günlerinin ve dış borsaların endeks tahmini üzerinde olumlu etkisinin olduğu belirtilmektedir.

Literatürden hareketle bu çalışmada bağımsız değişken olarak teknik gösterge değişkenleri, yurtdışı borsa endeks değişkenleri, makroekonomik gösterge değişkenleri ve gün değişkenleri kullanılmaktadır.

Çalışmada kullanılan teknik gösterge değişkenleri atr.atr, bbands.close.pctB, bbands.HLC.dn, bbands.HLC.ma, bbands.HLC.pctB, bbands.HLC.up, cci, chaikinad, chaikinvolatility, Close, cmf, cmo, dmi.ADX, dmi.DIn, dmi.DIp, dmi.DX, ema.20, ema.200, High, Low, macd.macd, macd.signal, mfi, mom, myAroon, obv, Open, roc, rsi, sar, sma.20, sma.200, smi.signal, smi.smi, stochOSC.fastD, stochOSC.fastK, stochOSC.slowD, Volume, williamsad, wma.20, wma.200 ve wpr değişkenlerinden oluşmaktadır.

Yurtdışı borsa endeks değişkenleri SP500, NASDAQ, DAX ve NIKKEI225 endekslerinin kapanış değerlerinden; makroekonomik gösterge değişkenleri Altın, Dolar, Euro ve Faiz değişkenlerinden ve gün değişkenleri Paz, Sal, Car, Per ve Cum kukla değişkenlerinden oluşmaktadır. Çalışmada kullanılan değişkenlerin istatistik bilgileri ileride değişken gruplarına göre gruplandırılarak verilmektedir.

BIST100 endeksinin bir periyot sonraki artış ve azalışını belirten ikili (binary) değişken ise çalışmanın bağımlı değişkenidir.

Değişken seçme yöntemleri (temel bileşen analizi, k-en yakın komşuluk, korelasyon, lojistik regresyon, kısmi en küçük kareler, rastgele orman algoritması) yardımıyla seçilen değişkenlerden oluşan modeller ile BIST100 endeksinin yönü tahmin edilmektedir.

Çalışmada günlük, haftalık ve aylık periyotlarda değişken seçme yöntemleri yardımıyla çeşitli modeller oluşturulmuş, ayrıca tüm değişkenlerin yer aldığı modeller de oluşturulmuştur. Oluşturulan modeller genel olarak aşağıdaki gibi sembolize edilebilir.

$$\left\{ \bigwedge_{i=1}^{tg} X_t^{TGD_i}, \bigwedge_{j=1}^{yb} X_t^{YBD_j}, \bigwedge_{k=1}^{md} X_t^{MED_k}, \bigwedge_{s=1}^{gd} X_t^{GD_s} \right\} \xrightarrow{\text{sınıflandırıcı}} Y_{t+1} \quad (4-1)$$

$Y_{t+1}$  : Endeksin t+1 zamanındaki yön değeri.

$X_t^{TGD_i}$  : i. teknik gösterge değişkeninin t zamanındaki değeri.

$X_t^{YBD_j}$  : j. yurtdışı borsa endeksi değişkeninin t zamanındaki değeri.

$X_t^{MED_k}$  : k. makroekonomik gösterge değişkeninin t zamanındaki değeri.

$X_t^{GD_s}$  : s. gün değişkeninin t zamanındaki değeri.

$tg$  : Modelde seçilen teknik gösterge değişkeni sayısı.

$yb$  : Modelde seçilen yurtdışı borsa endeksi değişkeni sayısı.

$md$  : Modelde seçilen makroekonomik gösterge değişkeni sayısı.

$gd$  : Modelde seçilen gün değişkeni sayısı.

### 4.3. VERİLERİN TOPLANMASI VE HAZIRLANMASI

Çalışmada kullanılan, BIST 100 Endeksi ile ilgili veriler Borsa İstanbul A.Ş.'den resmi yazı aracılığıyla elde edilmiştir. Yurtdışı borsaların kapanış verileri için ise, finance.yahoo.com internet sitesinden yararlanılmıştır. Kullanılan makroekonomik veriler için Merkez Bankası'nın evds.tcmb.gov.tr adresindeki

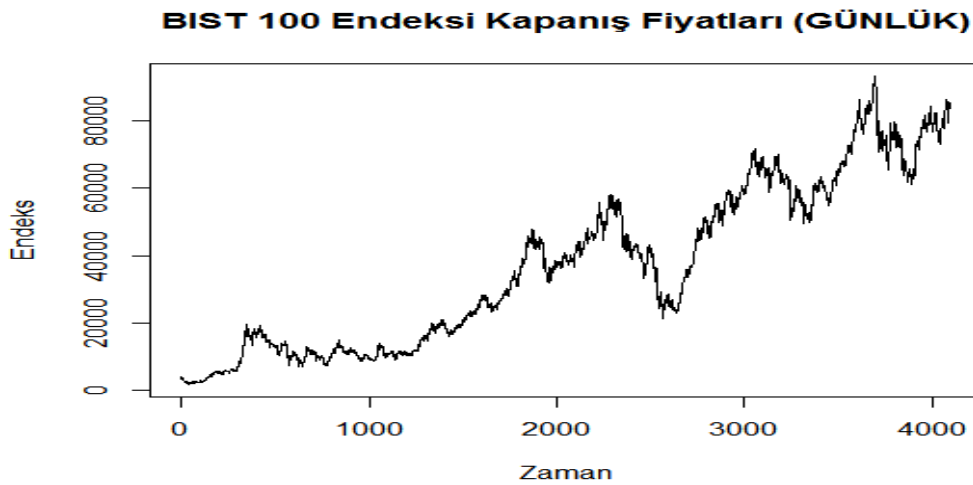


elektronik veri dağıtım sisteminden faydalanılmıştır. Teknik analiz değişkenleri ise, Ulrich (2013) tarafından hazırlanan TTR paketi kullanılarak elde edilmiştir.

Kullanılacak değişkenler için sağlıklı olarak elde edilebilen günümüzden geriye doğru en geniş tarih aralığı alınmaktadır. Burada amaç mümkün olan en geniş veri setini inceleyerek, modeller için kullanılacak periyodu daha sağlıklı belirleyebilmek ve kullanılacak veri setini temsil edecek özellik vektörlerini ortaya koyabilmektir. Böylece, oluşturulan modeller yardımıyla BIST 100 endeksinin yönünün daha hızlı ve tutarlı şekilde tahmin edilebileceği düşünülmektedir.

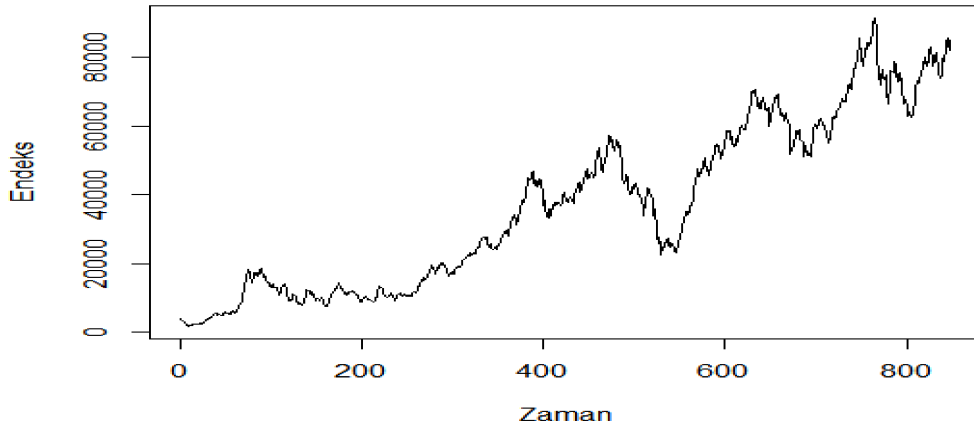
Kullanılan veriler, sağlıklı gözlem değerlerinin elde edilebildiği en geniş tarih aralığı ve ayırık dalgacık dönüşümü için gerekli gözlem sayısı göz önüne alınarak seçilmiştir. Bu doğrultuda 15 Ekim 1997 tarihi ile 30 Aralık 2014 tarihleri arasındaki borsanın açık olduğu işgünlerini kapsayan 4296 günlük gözlem verisi elde edilmiş, bazı teknik gösterge değişkenlerinin hesaplanması için gerekli olan ilk 200 günlük gözlem değerleri yalnızca bu teknik gösterge değerlerinin hesaplanmasında kullanılmıştır.

07 Ağustos 1998 tarihi ile 30 Aralık 2014 tarihleri arasındaki borsanın açık olduğu işgünlerini kapsayan 4096 günlük gözlem verisi ise çalışmasının veri setini teşkil etmektedir. Bu dönem aralığı için BIST 100 endeksinin günlük (Şekil 33), haftalık ve aylık (Şekil 34) kapanış değerleri aşağıdaki grafiklerde görülmektedir.

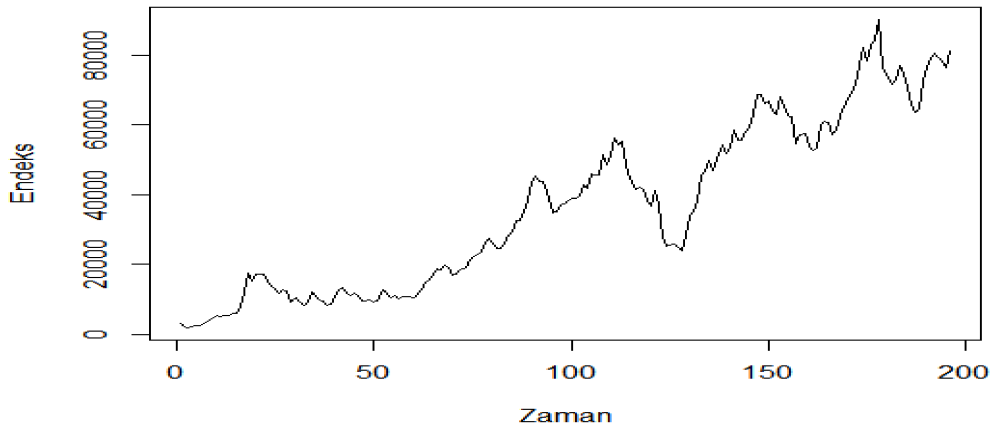


Şekil 33 : BIST 100 Endeksi Günlük Kapanış Grafiği

**BIST 100 Endeksi Kapanış Fiyatları (HAFTALIK)**



**BIST 100 Endeksi Kapanış Fiyatları (AYLIK)**



Şekil 34 : BIST 100 Endeksi Haftalık ve Aylık Kapanış Grafikleri

4096 günlük bu dönem aralığında BIST 100 endeksinin artış yönünde olduğu gün sayısı 2149 iken azalış yönündeki gün sayısı 1947 olduğu Tablo 2'de görülebilmektedir. Yani artış (%52,47) ve azalış (%47,53) yönlü işlem günlerinin oranının birbirine yakın olduğu söylenebilir. Benzer şekilde günlük veriden elde edilen, haftalık ve aylık verilerdeki artış ve azalış yönündeki gözlem sayıları Tablo 2'de verilmektedir.

**Tablo 2 : BIST 100 Endeksi Artış ve Azalış Gözlem Sayıları**

	Azalış	Oran	Artış	Oran	Toplam
Günlük	1947	%47,53	2149	%52,47	4096
Haftalık	369	%43,51	479	%56,49	848
Aylık	84	%42,86	112	%57,14	196

Çalışmada, günlük, haftalık ve aylık veri setlerinin %80'i eğitim veri seti olarak, %20'si ise test veri seti olarak rastgele ikiye ayrılmıştır. Böylece günlük veri setinin 3277'si eğitim veri setini, 819'u ise test veri setini oluşturmuştur. Benzer şekilde, haftalık veri setinin 679'u eğitim veri setini ve 169'u ise test veri setini oluştururken, aylık veri setinin 157'si eğitim veri setini, 39'u ise test veri setini oluşturmuştur.

Eğitim ve test veri setlerine ilişkin artış ve azalış yönlü gözlem sayıları Tablo 3'te verilmektedir. Buna göre tablo incelendiğinde, günlük veri setindeki artış yönündeki oran (%52,47) ile eğitim ve test veri setlerindeki artış yönündeki oranların (sırasıyla %53,13 ve %49,82) birbirine yakın olduğu görülecektir. Aynı şekilde haftalık veri setindeki artış yönündeki oran (%56,49) ile eğitim ve test veri setlerindeki artış yönündeki oranların (sırasıyla %56,85 ve %54,44) ve aylık veri setindeki artış yönündeki oran (%57,14) ile eğitim ve test veri setlerindeki artış yönündeki oranların (sırasıyla %56,69 ve %58,97) birbirine yakın olduğu görülmektedir. Buradan eğitim veri seti ve test veri setlerindeki dağılımların hem birbirleri ile hem de tüm veri seti ile benzer olduğu söylenebilir.

Tablo 3 : Eğitim ve Test Verileri Artış ve Azalış Gözlem Sayıları

		Azalış	Oran	Artış	Oran	Toplam
Günlük	Eğitim	1536	%46,87	1741	%53,13	3277
	Test	411	%50,18	408	%49,82	819
Haftalık	Eğitim	293	%43,15	386	%56,85	679
	Test	77	%45,56	92	%54,44	169
Aylık	Eğitim	68	%43,31	89	%56,69	157
	Test	16	%41,03	23	%58,97	39

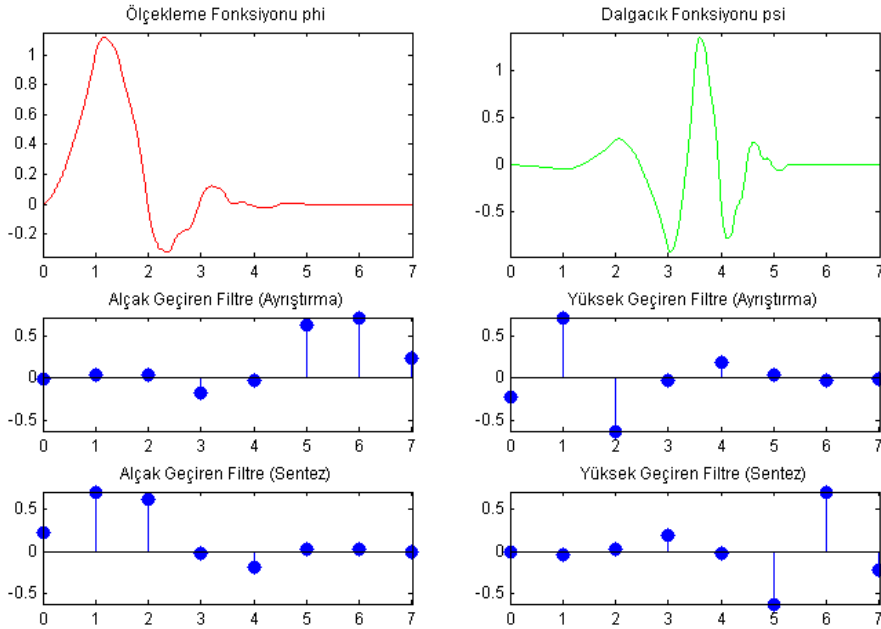
Veri setlerinde bulunan değişkenlerin aldıkları değerlerin birbirinden büyüklük olarak farklılaşmasından dolayı (ör: Tablo 5) veri setinde bulunan dağılım genişliğini korumak için değişkenlerin aldıkları değerlerde z-skor normalleştirilmesi yapılmıştır. Yani, veri ortalama değere göre normalleştirilmekte ve standart sapmaya bölünerek veri setindeki değişim hızı normalleştirilmektedir.

$$X_i^* = \frac{X_i - \mu}{\sigma} \quad (4-2)$$

Ancak burada, normalleştirmenin eğitim ve test veri setlerine ayrı ayrı mı yoksa tüm veri setine aynı anda mı uygulanacağı problemi ortaya çıkmaktadır. Eğitim veri setindeki normalleştirmenin, sınıflandırıcıların eğitim aşamasında görmediği ve modeller oluşturulurken hiçbir şekilde kullanılmayan test veri setindeki gözlem değerlerinden etkilenmemesi için eğitim ve test veri setlerinde ayrı ayrı normalleştirme yapılması gerekmektedir. Fakat bu aşamada, eğitim ve test veri setlerinde ayrı ayrı parametreler elde edileceği için Hsu vd. (2003, s. 3) 'nin de önerdiği gibi eğitim veri setinden elde edilen  $\mu$  ve  $\sigma$  parametreleri test veri setinin normalleştirilmesinde kullanılmıştır.

#### 4.4. UYGULAMA AŞAMALARININ BELİRLENMESİ

Çalışmada kullanılan yöntemler dalgacık dönüşümü ve destek vektör makineleri ile ilgili teorik temeller ve detaylı bilgi önceki bölümlerde verilmişti. Uygulama aşamasında ayrık dalgacık dönüşümü, bağımsız değişkenleri zaman uzayından frekans uzayına taşımak ve orada yaklaşım ve detay katsayı serilerini elde etmek için kullanılmıştır.



Şekil 35 : db4 Dalgacığının Bileşenleri

Dalgacık türü ise finansal zaman serileri için literatürde en sık kullanılan (Kao, Chiu, Lu ve Chang, 2013, s. 1239) daubechies4 dalgacığı (db4) olarak belirlenmiştir. Bağımsız değişkenlerin günlük gözlem değerlerinden oluşan zaman serileri 4096 gözlemden oluştuğı için her seri db4 dalgacığı kullanılarak 12. seviyeden yaklaşım ve detay bileşenlerine ayrıştırılmıştır. Şekil 35'te db4 dalgacığının ölçekleme ve dalgacık fonksiyonları ile birlikte ayrıştırma ve sentez için gerekli alçak ve yüksek geçiren filtreler verilmektedir.

Çalışmada, eğitim ve test kümesi olarak ikiye ayrılan ham veri setinde değişken seçimi eğitim veri seti üzerinde temel bileşen analizi, k-en yakın komşuluk, korelasyon, lojistik regresyon, kısmi en küçük kareler ve rastgele orman algoritması yöntemleri yardımıyla yapılmış ve günlük, haftalık ve aylık modeller oluşturulmuştur.

Hsu vd. (2003, s. 3)'nin önerileri doğrultusunda önce veriler kullanılacak DVM paketine hazır hale getirilerek normalleştirilmiş, ardından endeks yönü tahmininde diğer çekirdeklerden daha iyi performans gösteren (Emir, 2013a, s.202) dairesel tabanlı çekirdek (RBF) seçilerek, bu çekirdek için gerekli cost ve gamma parametreleri çapraz geçirme ile elde edilmiştir. Sonrasında bu parametre değerleri kullanılarak eğitilen modelin test performansı ölçülmüştür.

Modeller oluşturulurken bağımlı değişken, bağımsız değişkenlerin bir periyot önceki gözlem değerleri ile eşleştirilmiştir. Oluşturulan günlük, haftalık ve aylık periyotta veri setleri için kurulan modeller kullanılarak DVM yönteminin önce ham veri seti ile tahmin performansı, daha sonra dalgacık dönüşümü katsayılarından oluşan veri seti ile tahmin performansı incelenmiş ve karşılaştırma yapılmıştır.

#### **4.4.1. Kullanılan Yazılım ve Paketler**

Çalışmada, verileri düzenleme ve ham veri setini oluşturmada Microsoft Excel 2007 Programı kullanılmıştır. Analiz işlemleri için açık kaynak kodlu R programlama dili kullanılmıştır. R istatistiksel hesaplama ve grafik programı Bell laboratuvarlarında geliştirilen S dilinin bir uzantısıdır (Venables, Smith ve Team, 2010, s. 2).

Uygulamanın çeşitli aşamalarında R 3.0.2 versiyonu için yazılmış bazı paketlerden yararlanılmıştır. Bunlardan, Torgo (2010) tarafından geliştirilen "DMwR" ve D. Meyer, Dimitriadou, Hornik, Weingessel ve Leisch (2014) tarafından geliştirilen "e1071" paketleri çapraz geçерleme ve destek vektör makineleri işlemleri için kullanılmış; teknik analiz göstergelerin hesaplanması için Ulrich (2013) tarafından geliştirilen "TTR" paketinden yararlanılmıştır. Değişken seçimi için ise Revelle (2015) tarafından geliştirilen "psych" paketi ile birlikte Kuhn (2008) tarafından geliştirilen "caret" paketi, Mevik, Wehrens ve Liland (2013) tarafından geliştirilen "pls" paketi ve Liaw ve Wiener (2002) tarafından geliştirilen "randomForest" paketlerinden faydalanılmıştır. Ayrıca günlük, haftalık ve aylık verilerin düzenlenmesinde ve zaman serisi olarak kullanılmasında Ryan ve Ulrich (2014) tarafından geliştirilen "xts" paketi ve Zeileis ve Grothendieck (2005) tarafından geliştirilen "zoo" paketlerinin işlevselliğinden yararlanılmıştır. Ayrıca, dalgacık dönüşümü alçak frekans ve yüksek frekans katsayıları Matlab R2013a programı aracılığıyla elde edilmiştir.

#### **4.4.2. Model Parametrelerinin Belirlenmesi ve Çapraz Geçerleme**

Destek vektör makineleri modellerinde kullanıcı tarafından belirlenen model parametreleri olan gamma ve cost değerlerinin belirlenmesinde, sınıflandırmada hata oranlarının tahmini (Michie, Spiegelhalter ve Taylor, 1994, s. 108) için önerilen bootstrap ve çapraz geçerleme yöntemleri kullanılabilir.

Michie vd. (1994, s. 108) deęişkenlięi azaltmak için en iyi yolu temsil eden bootstrap metodunu küçük veri setlerinde kullanmayı önermiştir. Bootstrap metodunda eldeki eğitim veri setinin çeşitli kombinasyonları kullanılarak elde edilen yeni örneklemere bootstrap örneklemeleri denir ve bu örneklemelerden yüksek sayıda elde edilerek bu örneklemelerin hata ortalamaları kullanılır.

Efron (1983) bootstrap ve çapraz geçerleme metotlarını karşılaştırmış ve büyük veri setlerinde çapraz geçerleme metodunun daha sağlıklı sonuçlar verdiğini belirtmiştir. Ayrıca büyük veri setlerinde bootstrap metodunun zaman maliyeti daha yüksektir. Bu nedenle çalışmada kullanılan veri setleri göz önüne alındığında, DVM modelleri için kullanıcı tarafından belirlenen model parametrelerinin belirlenme

aşamasında çapraz geçерleme yönteminin kullanılmasına karar verilmiş, böylece hem modellerde aşırı öğrenme engellenmiş hem de daha geçerli model oluşturabilecek parametreler elde edilmiştir.

Bu amaçla çalışmada DVM model parametrelerini elde etmek için, 10 kat çapraz geçерleme uygulanarak DVM model parametrelerinden gamma ( $\gamma$ ) parametresi için  $10^{-8}$  değeri ile  $10^{-1}$  değeri arasında, cost (c) parametresi içinse  $10^1$  değeri ile  $10^8$  değeri arasında arama yapılmış ve her model için en uygun parametreler elde edilmiştir.

$$10^{-8} \leq \gamma \leq 10^{-1}$$

(4-3)

$$10^1 \leq c \leq 10^8$$

Çapraz geçerlemede (cross-validation) eğitim verisi eşit büyüklükte v tane ayrık alt gruba ayrılır. v-kat geçerleme olarak adlandırılan bu işlemde, bir alt grup test seti olarak seçilir ve kalan diğer v-1 alt grup üzerinde sınıflandırıcının eğitimi yapılır ve test grubunun performansı elde edilir. Bu işlem her alt grup için tekrarlanarak v tane test grubu performansının ortalaması sınıflandırıcının yeni örnekler üzerinde beklenen performansını verir (Michie vd., 1994, s. 108).

Eğitim Verisi (% 80)	Test Verisi (% 20)
1.Altgrup (% 16) Test grubu olarak seçildiğinde 2., 3., 4. ve 5. altgruplar eğitim grubu olarak seçilerek sınıflandırıcı eğitilir.	
2.Altgrup (% 16) Test grubu olarak seçildiğinde 1., 3., 4. ve 5. altgruplar eğitim grubu olarak seçilerek sınıflandırıcı eğitilir.	
3.Altgrup (% 16) Test grubu olarak seçildiğinde 1., 2., 4. ve 5. altgruplar eğitim grubu olarak seçilerek sınıflandırıcı eğitilir.	
4.Altgrup (% 16) Test grubu olarak seçildiğinde 1., 2., 3. ve 5. altgruplar eğitim grubu olarak seçilerek sınıflandırıcı eğitilir.	
5.Altgrup (% 16) Test grubu olarak seçildiğinde 1., 2., 3. ve 4. altgruplar eğitim grubu olarak seçilerek sınıflandırıcı eğitilir.	

Şekil 36 : 5-Kat Çapraz Geçerleme

Örneğin Şekil 36'da görselleştirildiği üzere, %80'i eğitim veri seti olarak ayrılan bir veri setinde 5-kat çapraz geçirme yapılırken, eğitim seti 5 eşit (%16) ayrık alt gruba ayrılarak, her bir alt grup diğer dört alt grup tarafından eğitilen sınıflandırıcı ile test edilip elde edilen test performanslarının ortalaması alınmaktadır.

#### 4.4.3. Sınıflandırıcıların Performans Değerlendirmesi (Sınıflandırma Matrisi)

Sınıflandırma matrisi (Confusion Matrix) n farklı sınıftan oluşan bir sınıflandırma probleminde, sınıflandırıcının gerçek sınıf ve tahmin sınıfı sonuçlarının karşılaştırılmasını sağlayan n x n boyutunda bir matristen oluşur (Visa, Ramsay, Ralescu ve van der Knaap, 2011, s. 121). Sınıflandırma matrisi, aynı zamanda hata matrisi, atama matrisi veya tahmin matrisi olarak da adlandırılmaktadır. Verinin hangi oranda doğru şekilde sınıflandırıldığını ifade eder. Bu oran model tarafından açıklanan bağımlı değişkendeki varyasyon oranını veren  $R^2$  ile bir anlamla benzer işleve sahiptir (Verma, 2012, s. 392). Bu çalışmada da oluşturulan her modelde sınıflandırıcıların performanslarını karşılaştırmak için Sınıflandırma Matrisleri kullanılmıştır.

NG : Negatif Gözlem

PG : Pozitif Gözlem

$N_t$  : Negatif Tahmin

$P_t$  : Pozitif Tahmin

DN : Doğru Negatif Tahmin

YP : Yanlış Pozitif Tahmin

YN : Yanlış Negatif Tahmin

DP : Doğru Pozitif Tahmin

sayılarını göstermek üzere Sınıflandırma Matrisi Tablo 4'teki gibi oluşturulabilir.

Tablo 4 : Sınıflandırma Matrisi

		Tahmin Sınıfı			Doğruluk Oranı
		Sınıf 1	Sınıf 2	Toplam	
Gerçek Sınıf	Sınıf 1	DN	YP	NG	$\frac{DN}{N}$
	Sınıf 2	YN	DP	PG	$\frac{DP}{P}$
	Toplam	$N_t$	$P_t$	NG+PG	$\frac{DN + DP}{NG + PG}$



Tabloda Doğru Negatif Tahmin ve Doğru Pozitif Tahmin değerleri doğru sınıflanan tahmin sayısını verirken, Yanlış Pozitif Tahmin ve Yanlış Negatif Tahmin sayıları ise sınıflandırıcının yaptığı hatalı sınıflandırma sayılarını verecektir. Bu değerlerin sifıra yakın olması tahmin performansının yüksek olduğunu gösterir. Sınıflandırıcının doğru sınıflandırma oranı (accuracy) ise doğru sınıflandırdığı gözlem değerlerinin sayısının toplam gözlem sayısına oranı ile ifade edilir.

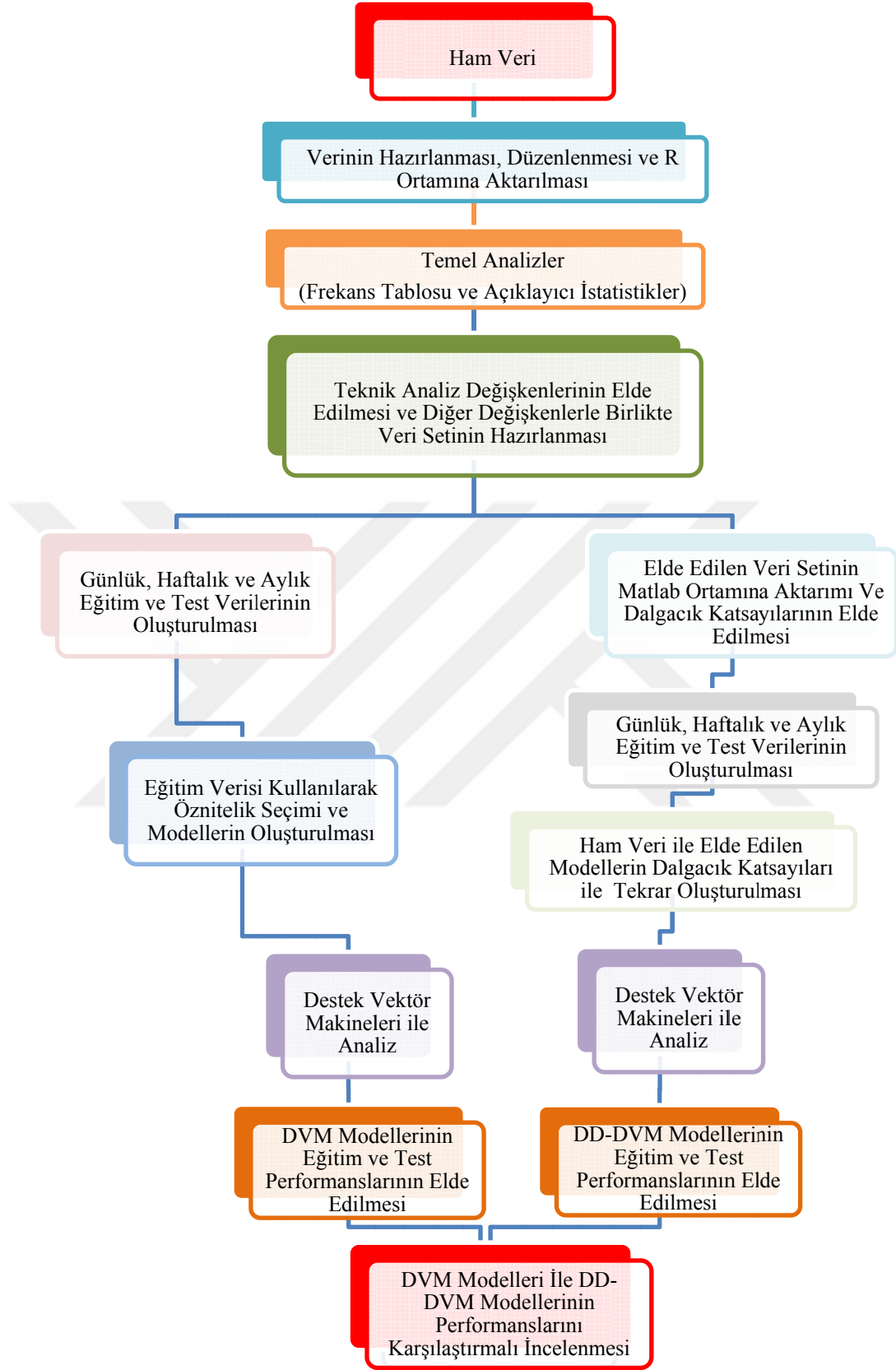
$$\text{Doğruluk Oranı} = \frac{DN + DP}{NG + PG} \quad (4-4)$$

#### 4.4.4. Uygulama Aşamalarının Şema İle Gösterimi

Çalışmada elde edilen veriler kullanılarak oluşturulan ham veri seti Şekil 37'deki aşamalar takip edilerek analize tabi tutulmuştur. Ham veri seti R programına aktarılarak, öncelikle teknik analiz değişkenleri ve açıklayıcı istatistikler elde edilmiştir.

Daha sonra eğitim ve test kümeleri günlük, haftalık ve aylık dönemler için elde edilmiştir. Eğitim veri setleri kullanılarak günlük, haftalık ve aylık veri setlerinde ayrı ayrı değişken seçimi yapılarak modeller oluşturulmuş, elde edilen modellerin destek vektör makineleri kullanılarak eğitim ve test performansları hesaplanmıştır.

Oluşturulan modellerin dalgacık versiyonları Matlab programı aracılığıyla elde edilen dalgacık dönüşümü katsayıları kullanılarak tekrar oluşturulmuş ve destek vektör makineleri kullanılarak eğitim ve test performansları hesaplanmıştır. Son olarak elde edilen performanslar karşılaştırılmıştır.



Şekil 37 : Uygulama Aşamaları

## 4.5. ANALİZDE KULLANILAN DEĞİŞKENLER

### 4.5.1. Makroekonomik Göstergeler

Çalışmada kullanılan makroekonomik değişkenler, literatürde piyasa tahmininde kullanılan çeşitli değişkenler arasından günlük periyotlarla elde edilebilenler arasından seçilmiştir. Bu makro ekonomik değişkenler Altın (ons/\$), dolar, euro ve faiz olarak belirlenmiştir. Aşağıdaki Tablo 5'te kullanılan makroekonomik değişkenlerin ilgili dönem aralığı için temel istatistikleri verilmiştir.

Tablo 5 : Makroekonomik Değişkenler ile İlgili Temel İstatistikler

Değişken	Gözlem Sayısı	Ortalama	Standart Sapma	Minimum	Maksimum
Altın (ons/\$)	4096	781,62	488,41	252,75	1900,20
Dolar	4096	1,40	0,46	0,27	2,37
Euro	4096	1,75	0,67	0,30	3,18
Faiz	4096	31,36	30,26	4,79	276,15

### 4.5.2. Teknik Gösterge Değişkenleri

Çalışmada kullanılan teknik gösterge değişkenleri farklı teknik analiz yöntemlerinden elde edilen ve Tablo 6'da temel istatistiki bilgileri verilen 42 değişkenden oluşmaktadır. Burada değişken sayısının fazla olmasının sebebi, bir teknik analiz yöntemi ile birden fazla gösterge de elde edilebilmesidir. Örneğin, üssel hareketli ortalamalar hesaplanırken 20 günlük üssel hareketli ortalama ile 200 günlük üssel hareketli ortalama birbirinden farklı değişkenler olarak değişken setinde yer almaktadır. Kullanılan değişkenler, Borsa İstanbul A.Ş.'den elde edilen BIST 100 Endeksinin açık olduğu günlerde oluşan açılış, en düşük, en yüksek, kapanış ve hacim verisi ile hesaplandığı için eksik gözlem değeri bulunmamaktadır. Veri seti olarak incelenen dönemin başlangıcından 200 işlem günü öncesine kadar olan 4296 günlük veri kullanılmıştır.

Tablo 6 : Teknik Gösterge Değişkenleri ile İlgili Temel İstatistikler

Teknik Gösterge	Gözlem Sayısı	Ortalama	Standart Sapma	Minimum	Maksimum
atr.atr	4096	863,93	456,12	103,50	2726,94
bbands.HLC.dn	4096	34633,51	23385,75	1726,54	87877,31
bbands.HLC.ma	4096	36818,70	24328,32	2072,08	90463,05
bbands.HLC.pctB	4096	0,55	0,34	-0,42	1,33
bbands.HLC.up	4096	39003,88	25328,74	2322,29	98617,16
bbands.close.pctB	4096	0,55	0,34	-0,43	1,34
cci	4096	15,33	112,16	-376,03	299,47
chaikinad	4096	68064440000	55620900000	308300909	151490900000
chaikinvolatility	4096	0,04	0,29	-0,56	1,52
cmf	4096	0,11	0,15	-0,36	0,62
cmo	4096	7,58	35,48	-88,62	95,11
dmi.ADX	4096	27,44	10,88	8,48	80,14
dmi.DIn	4096	23,26	8,66	1,36	54,58
dmi.DIp	4096	25,59	8,75	4,60	59,38
dmi.DX	4096	27,46	18,67	0,01	94,58
ema.20	4096	36830,23	24301,46	2123,34	89890,89
ema.200	4096	35196,91	23458,29	2803,77	78576,64
macd.macd	4096	0,42	3,07	-13,38	15,07
macd.signal	4096	0,42	2,92	-12,14	14,12
mfi	4096	56,94	16,54	6,82	100,00
mom	4096	19,69	734,37	-9006,35	4805,54
roc	4096	0,00	0,02	-0,20	0,18
myAroon	4096	12,12	65,81	-100,00	100,00
obv	4096	56890040000	49269650000	379452858	152184200000
rsi	4096	53,36	12,99	13,21	89,74
sar	4096	36826,95	24402,48	1818,68	93398,33
sma.20	4096	36830,05	24320,59	2071,36	90446,00
sma.200	4096	35168,38	23575,57	3072,97	79476,74
smi.signal	4096	14,22	33,50	-62,51	79,31
smi.smi	4096	14,24	35,75	-67,60	82,00
stochOSC.fastD	4096	0,57	0,29	0,02	1,00
stochOSC.fastK	4096	0,57	0,31	0,00	1,00
stochOSC.slowD	4096	0,57	0,28	0,02	0,99
williamsad	4096	32342,81	22928,68	169,92	92474,58
wma.20	4096	36891,29	24345,41	2044,37	90934,84
wma.200	4096	35774,76	23798,58	2735,21	81798,03
wpr	4096	0,43	0,31	0,00	1,00
Open	4096	37022,55	24441,19	1852,28	93095,91
High	4096	37404,55	24588,4	1883,11	93398,33
Low	4096	36591,07	24260,64	1818,68	91681,89
Close	4096	37014,83	24412,82	1852,28	93178,87
Volume	4096	369566263	263038939	3618299	1609847000

#### 4.5.3. Yurtdışı Borsa Endeks Verileri

Çalışmada kullanılan yurtdışı borsa endeksleri, literatürde yerel piyasa üzerinde etkisi tespit edilenlerden ilgili dönem aralığı için kapanış verilerine ulaşılabilenler seçilmiştir. Bu endeksler, SP500, NASDAQ, DAX ve NIKKEI225 olarak belirlenmiştir. Aşağıdaki Tablo 7'de kullanılan yurtdışı borsa endekslerinin ilgili dönem aralığı için temel istatistikleri verilmiştir.

Tablo 7 : Endeks Değişkenleri ve İlgili Temel İstatistikler

Ülke	Endeks Adı	Kısaltma	Gözlem Sayısı	Ortalama	Standart Sapma	Minimum	Maksimum
ABD	Standart & Poors 500 Index	SP500	4096	1280,41	260,26	676,53	2090,57
ABD	National Association of Securities Dealers Automated Quotations	NASDAQ	4096	2516,86	810,11	1114,11	5048,62
Almanya	Deutscher Aktien Index	DAX	4096	5987,49	1682,85	2202,96	10087,12
Japonya	Nihon Keizai Shimbun Index 225	NIKKEI225	4096	12615,60	3145,26	7054,98	20833,21

#### 4.5.4. Gün Değişkenleri

Gün değişkenleri, ilgili gün "1", diğer günler "0" olacak şekilde, Pazartesi (Paz), Salı (Sal), Çarşamba (Car), Perşembe (Per) ve Cuma (Cum) kukla değişkenleri olarak oluşturulmuştur. Tablo 8'de günlerin toplam gün sayısı içinde eşit oranlarda (~ %20) dağıldığı, yine eğitim ve test verileri içinde de birbirine yakın oranlarda bir dağılım içerisinde oldukları görülmektedir.

Tablo 8 : Gün Değişkenleri ile İlgili Temel İstatistikler

	Frekans	Oran	Eğitim	Oran	Test	Oran
Pazartesi	819	%20,00	634	%19,35	185	%22,59
Salı	821	%20,04	652	%19,90	169	%20,63
Çarşamba	819	%20,00	668	%20,38	151	%18,44
Perşembe	821	%20,04	674	%20,57	147	%17,95
Cuma	816	%19,92	649	%19,80	167	%20,39
Toplam	4096		3277		819	

#### **4.5.5. Bağımlı Değişken ve Hesaplanması**

Çalışmada BIST 100 Endeksinin yönü ile ilgilenildiği için, bağımlı değişken BIST 100 Endeksinin kapanış değerlerinden yararlanılarak oluşturulmuştur. Endeksin günlük kapanış değeri bir önceki günden yüksek ise bağımlı değişkenin gözlem değeri "1" (bir) olarak, değil ise yani endeksin günlük kapanış değeri bir önceki günden düşük ise bağımlı değişkenin gözlem değeri "0" (sıfır) olarak belirlenmiş ve benzer uygulama haftalık ve aylık bağımlı değişkenler için de uygulanmıştır.

#### **4.6. GÜNLÜK VERİLERDE DEĞİŞKEN SEÇİMİ**

Çalışmada günlük veri seti üzerinde değişken seçme metotları yardımıyla değişken seçimi yapılarak elde edilen modellerin ve değişken seçimi yapmadan tüm değişkenlerin kullanıldığı modelin performansları aşağıda verilmektedir.

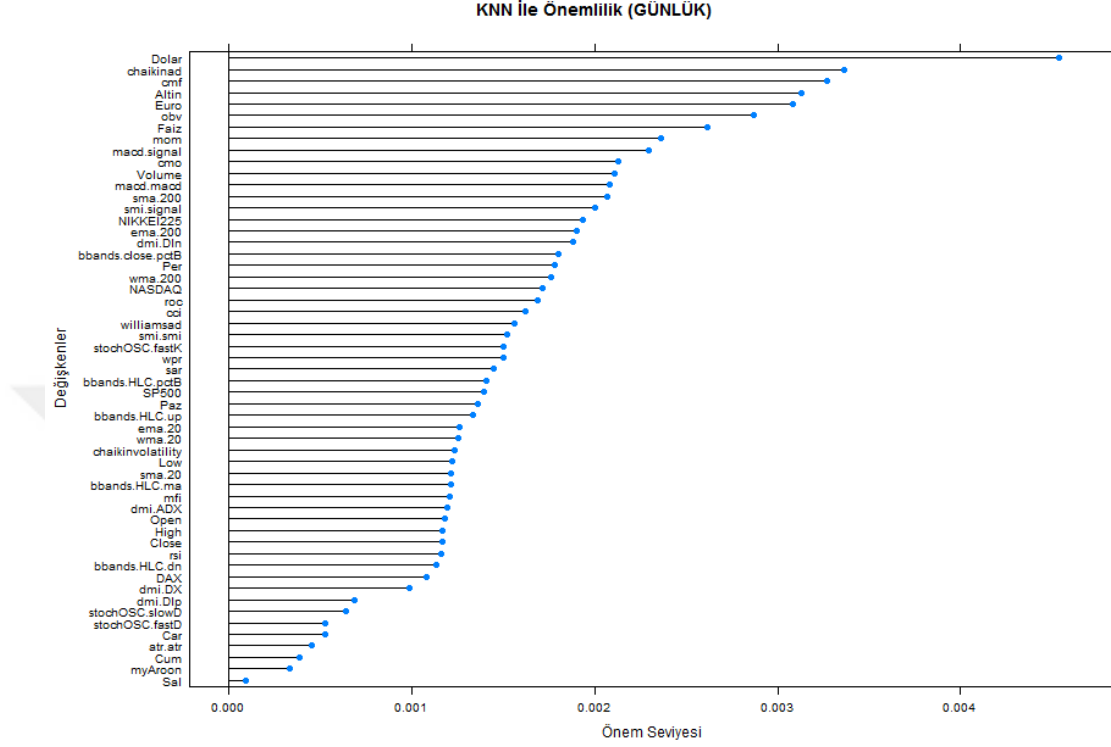
Tablo 9'da çalışmada kullanılan günlük veri setinde bulunan değişkenlere temel bileşen analizi uygulandığında altı temel bileşen elde edildiği ve her bir bileşenin en yüksek yüke sahip değişkenlerinin, bands.HLC.ma, bbands.HLC.pctB, NIKKEI225, macd.signal, dmi.DX ve roc değişkenleri olduğu görülmektedir. Elde edilen bu temel bileşenlerin toplam varyansın 0,81'ini açıkladığı anlaşılmaktadır.

Tablo 9 : Günlük Veri Setinde Temel Bileşen Analizi

Değişkenler	TB1	TB2	TB3	TB4	TB5	TB6
bbands.HLC.ma	0,99	-0,02	0,13	0,02	0,00	0,00
bbands.HLC.up	0,99	-0,02	0,13	0,01	0,02	-0,01
ema.20	0,99	-0,01	0,13	0,01	0,00	-0,01
sma.20	0,99	-0,02	0,13	0,02	0,00	0,00
wma.20	0,99	0,00	0,13	0,03	0,00	-0,01
Open	0,99	0,04	0,13	0,03	-0,01	-0,01
High	0,99	0,04	0,13	0,02	0,00	-0,01
Low	0,99	0,05	0,14	0,02	-0,01	0,00
Close	0,99	0,05	0,13	0,02	-0,01	0,01
bbands.HLC.dn	0,98	-0,02	0,13	0,03	-0,03	0,00
chaikinad	0,98	0,00	-0,07	-0,02	-0,01	0,00
ema.200	0,98	-0,05	0,11	-0,06	-0,03	-0,01
obv	0,98	0,01	0,05	-0,02	-0,01	0,00
sar	0,98	-0,02	0,13	0,05	0,01	-0,01
sma.200	0,98	-0,05	0,11	-0,08	-0,04	-0,01
wma.200	0,98	-0,05	0,12	-0,07	-0,03	-0,01
williamsad	0,95	0,07	0,07	0,05	-0,01	0,01
Altın	0,93	-0,01	-0,09	-0,03	-0,01	0,00
Euro	0,88	-0,01	-0,28	-0,02	-0,07	0,01
atr.atr	0,82	-0,18	0,12	-0,16	0,10	-0,02
Faiz	-0,80	-0,05	0,12	-0,11	0,04	-0,04
Dolar	0,76	-0,02	-0,34	-0,03	-0,07	0,01
Volume	0,75	0,12	-0,27	0,02	0,20	0,05
bbands.HLC.pctB	0,01	0,98	-0,01	-0,05	-0,02	0,04
bbands.close.pctB	0,01	0,97	-0,01	-0,07	-0,01	0,14
cci	0,01	0,97	-0,01	-0,06	-0,02	0,04
stochOSC.fastD	0,03	0,96	0,00	-0,08	-0,05	-0,12
cmo	0,03	0,94	0,00	0,11	0,06	-0,01
rsi	0,03	0,93	0,01	0,26	0,13	0,13
stochOSC.slowD	0,03	0,93	0,00	0,05	-0,05	-0,21
stochOSC.fastK	0,02	0,93	0,00	-0,20	-0,02	0,16
wpr	-0,02	-0,93	0,00	0,20	0,02	-0,16
dmi.DIn	0,06	-0,84	-0,02	-0,36	-0,07	-0,07
dmi.DIp	0,00	0,84	0,05	0,27	0,20	0,04
mfi	-0,03	0,84	-0,01	0,16	0,05	-0,09
myAroon	0,04	0,76	-0,01	0,43	0,03	-0,08
smi.smi	0,08	0,74	0,01	0,61	0,11	-0,05
cmf	-0,35	0,62	-0,14	0,18	0,06	0,04
NIKKEI225	-0,16	0,01	0,90	0,04	0,03	0,02
SP500	0,52	-0,01	0,81	0,01	-0,02	0,02
NASDAQ	0,49	0,00	0,78	0,03	0,02	0,03
DAX	0,66	-0,04	0,70	-0,05	-0,01	0,00
macd.signal	-0,02	0,44	0,10	0,80	0,20	0,07
smi.signal	0,09	0,53	0,01	0,77	0,17	0,01
macd.macd	-0,02	0,62	0,09	0,69	0,15	0,04
dmi.DX	-0,03	0,16	-0,01	-0,04	0,84	-0,15
dmi.ADX	-0,05	0,16	0,01	0,05	0,81	-0,09
chaikinvolatility	0,01	-0,11	0,00	0,18	0,65	0,05
roc	-0,03	0,30	0,01	-0,20	0,10	0,82
mom	0,00	0,29	0,00	-0,22	0,08	0,81
Per	0,00	-0,02	0,00	0,04	0,00	0,18
Cum	0,01	0,00	0,01	-0,01	-0,03	0,11
Car	-0,01	-0,03	-0,02	0,09	-0,06	0,04
Sal	0,00	0,00	0,01	-0,06	0,08	-0,11
Paz	0,01	0,05	0,00	-0,08	0,02	-0,23
Aygen değeri	21,63	12,88	3,14	2,92	2,06	1,67
Açıklanan Yığılmalı Varyans	0,39	0,63	0,68	0,74	0,78	0,81

Çalışmada oluşturulacak modeller için değişken seçimi yapılırken yararlanılan metotlardan bir diğeri olan KNN ile günlük veri setinde elde edilen değişkenlerin önemlilik sırası Şekil 38'de verilmektedir. Günlük veri setinde KNN ile en önemli değişkenin Dolar değişkeni olduğu ve bu değişkeni chaikinad ve cmf değişkenlerinin takip ettiği görülmektedir. Oluşturulan günlük modelde de KNN ile

değişken seçiminde kullanılan Kuhn (2008) tarafından geliştirilen R paketinin önerisi doğrultusunda ilk yirmi değişken ve ayrıca ağırlıklar göz önünde bulundurularak ikinci bir modelde ilk yedi değişken kullanılmıştır.



Şekil 38 : Günlük Veri setinde KNN

Çalışmada değişken seçimi için korelasyondan da yararlanılmış, günlük veri setindeki değişkenlerin Ek'te verilen korelasyon katsayıları incelenerek değişken seçimi yapılmıştır.

Çalışmada değişken seçimi için, lojistik regresyon analizinde değişkenlerin anlamlılık değerlerinden yararlanılmıştır. Günlük eğitim veri setinde anlamlılık değerleri "0,20", "0,10", "0,05" ve "0,01" değerlerinden küçük olan değişkenler ayrı ayrı seçilerek günlük modeller oluşturulmuştur. Tablo 10'da günlük eğitim veri setinde oluşturulan modellerdeki değişkenlerin lojistik regresyon analizindeki anlamlılık değerleri verilmektedir. Tabloda anlamlılık düzeyi 0,01'den küçük dört değişken elde edildiği görülmektedir.



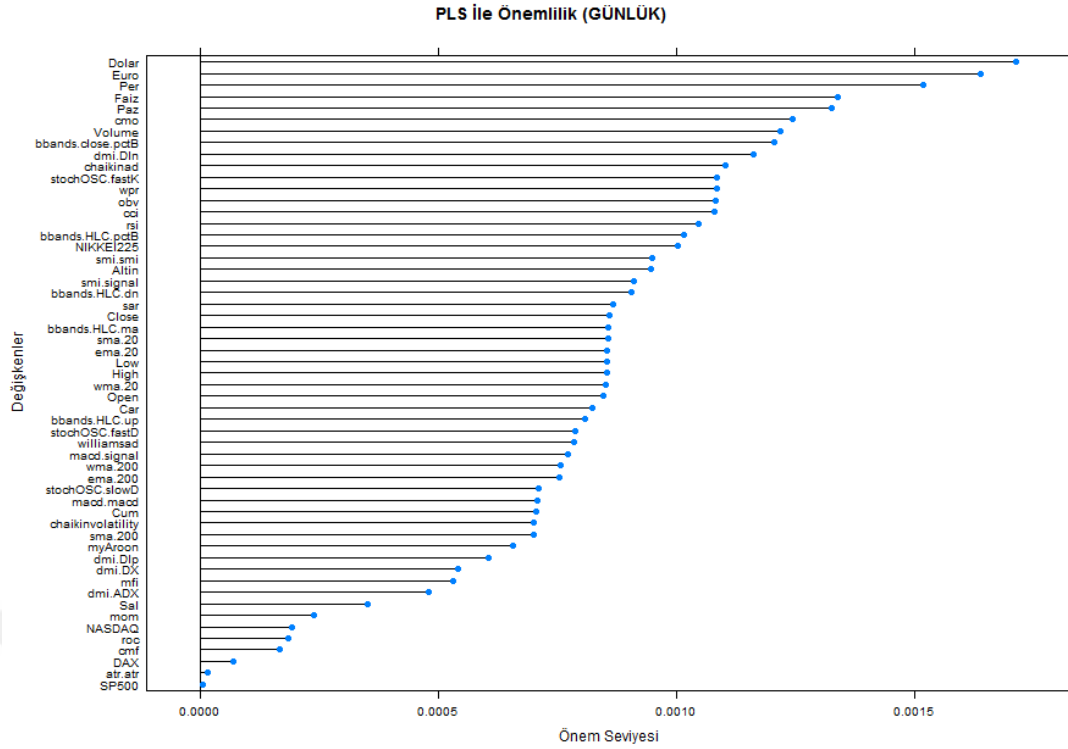
Tablo 10 : LR ile Anlamlılık Düzeyleri (Günlük)

	B	Std. Hata	Wald	Anlamlılık Düzeyi
bbands.HLC.pctB	-2,62917	0,80203	-3,278	0,00104 ****
bbands.close.pctB	1,72828	0,55382	3,121	0,00180 ****
DAX	0,5826	0,19312	3,017	0,00255 ****
wma.20	-21,01075	7,34694	-2,860	0,00424 ****
wma.200	3,79685	1,56821	2,421	0,01547 ***
roc	-0,15469	0,0651	-2,376	0,01749 ***
macd.macd	-0,65136	0,28245	-2,306	0,02111 ***
Per	0,25839	0,11286	2,290	0,02205 ***
ema.200	-4,07253	1,8012	-2,261	0,02376 ***
dmi.DIn	-0,26511	0,12655	-2,095	0,03618 ***
atr.atr	-0,28697	0,13892	-2,066	0,03885 ***
cmo	0,28931	0,14093	2,053	0,04009 ***
chaikinad	1,3134	0,64719	2,029	0,04242 ***
macd.signal	0,53699	0,26761	2,007	0,04479 ***
cci	1,08431	0,55233	1,963	0,04963 ***
chaikinvolatility	0,08803	0,04598	1,915	0,05554 **
sar	1,81002	0,98758	1,833	0,06683 **
High	9,41749	5,14257	1,831	0,06706 **
Car	0,19105	0,11278	1,694	0,09026 **
rsi	-0,44074	0,27925	-1,578	0,11449 *
Euro	0,41296	0,26874	1,537	0,12437 *
ema.20	10,77137	7,19289	1,498	0,13426 *
dmi.DIp	0,16335	0,11576	1,411	0,15819 *
Close	-9,1675	6,64389	-1,380	0,16764 *
dmi.ADX	0,07266	0,05422	1,340	0,18022 *

\* < 0,20 \*\* < 0,10 \*\*\* < 0,05 \*\*\*\* < 0,01

Çalışmada oluşturulacak modeller için değişken seçimi yapılırken yararlanılan metotlardan bir diğeri olan PLS ile günlük veri setinde elde edilen değişkenlerin önemlilik sırası Şekil 39'da verilmektedir.

Günlük veri setinde PLS ile en önemli değişkenin Dolar değişkeni olduğu ve bu değişkeni Euro ve Per değişkenlerinin takip ettiği görülmektedir. Oluşturulan günlük modelde de PLS ile ağırlıklarına göre ilk yirmi değişken ve ayrıca ikinci bir modelde ilk onyedeki değişken kullanılmıştır.



Şekil 39 : Günlük Veri setinde PLS

Çalışmada kullanılan günlük veri seti Liaw ve Wiener (2002) tarafından geliştirilen "randomForest" paketinden yararlanılarak incelenmiş ve günlük veri setinde Dolar, DAX, NASDAQ, Faiz, dmi.DIn ve Cum değişkenlerinden oluşan model elde edilmiştir.

Ayrıca günlük veride yapılan değişken seçiminden elde edilen modellerdeki değişken grupları aynen muhafaza edilerek haftalık ve aylık veri setlerinde de modeller oluşturulmuş ve bu modellerin performansları da incelenmiştir.

#### 4.6.1. Günlük Veriler ile Endeks Yönü Tahmini

Günlük ham eğitim veri seti kullanılarak değişken seçme yöntemleri yardımıyla oluşturulan modellerin DVM yöntemi ile bağımlı değişken olan endeksin yönü tahmin edilmektedir. Aşağıda ham eğitim ve test veri setleri ile dalgacık dönüşümü yardımıyla elde edilen eğitim ve test veri setlerinin performansları model bazında sınıflandırma matrisleri ile verilmektedir.

Temel bileşen analizi ile elde edilen bbands.HLC.ma, bbands.HLC.pctB, NIKKEI225, macd.signal, dmi.DX ve roc değişkenleri ile oluşturulan modelin (Günlük TBA6) DVM sınıflandırıcılarının performansı Tablo 11'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 128 gözlemi (%8,33) doğru sınıflandırırken artış yönündeki 1617 (%92,88) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 36 gözlem (%8,76) doğru sınıflandırılırken artış yönündeki 372 (%91,18) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %53,25'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %49,82'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 11 : Sınıflandırma Matrisi (Günlük TBA6 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	128	1408	1536	% 8,33	36	375	411	% 8,76
	Artış	124	1617	1741	% 92,88	36	372	408	% 91,18
	$\Sigma$	252	3025	3277	<b>% 53,25</b>	72	747	819	<b>% 49,82</b>

\*gamma= 0,00001 cost= 1000000

Temel bileşen analizi ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık TBA6) DVM sınıflandırıcılarının performansı Tablo 12'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1334 gözlemi (%86,85) doğru sınıflandırırken artış yönündeki 1554 (%89,26) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 344 gözlem (%83,70) doğru sınıflandırılırken artış yönündeki 361 (%88,48) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %88,13'ü doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %86,08'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 12 : Sınıflandırma Matrisi (Günlük Dalgacık TBA6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	1334	202	1536	% 86,85	344	67	411	% 83,70
	Artış	187	1554	1741	% 89,26	47	361	408	% 88,48
	Σ	1521	1756	3277	<b>% 88,13</b>	391	428	819	<b>% 86,08</b>

\*gamma= 0,001 cost= 10

Korelasyon analizi ile korelasyonları 0,50'den küçük chaikinvolatility, Paz, Sal, Car, Per ve Cum değişkenleri ile oluşturulan modelin (Günlük KOR<,50) DVM sınıflandırıcılarının performansı Tablo 13'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 196 gözlemi (%12,76) doğru sınıflandırırken artış yönündeki 1546 (%88,80) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 53 gözlem (%12,90) doğru sınıflandırılırken artış yönündeki 360 (%88,24) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %53,16'sı doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %50,43'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 13 : Sınıflandırma Matrisi (Günlük KOR&lt;,50 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	196	1340	1536	% 12,76	53	358	411	% 12,90
	Artış	195	1546	1741	% 88,80	48	360	408	% 88,24
	Σ	391	2886	3277	<b>% 53,16</b>	101	718	819	<b>% 50,43</b>

\*gamma= 0,000001 cost= 1000000

Korelasyon analizi ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık KOR<,50) DVM sınıflandırıcılarının performansı Tablo 14'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^4$ ) 10-kat çapraz

geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 615 gözlemi (%40,04) doğru sınıflandırırken artış yönündeki 1315 (%75,53) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 132 gözlem (%32,12) doğru sınıflandırılırken artış yönündeki 286 (%70,10) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %58,90'ı doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %51,04'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 14 : Sınıflandırma Matrisi (Günlük Dalgacık KOR<,50 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	615	921	1536	% 40,04	132	279	411	% 32,12
	Artış	426	1315	1741	% 75,53	122	286	408	% 70,10
	Σ	1041	2236	3277	<b>% 58,90</b>	254	565	819	<b>% 51,04</b>

\*gamma= 0,00001 cost= 10000

Korelasyon analizi ile korelasyonları 0,80'den küçük atr.atr, chaikinvolatility, cmf, dmi.ADX, dmi.DX, macd.signal, mfi, mom, roc, myAroon, wpr, Volume, Dolar, Altın, Faiz, NIKKEI225, NASDAQ, Paz, Sal, Car, Per, ve Cum değişkenleri ile oluşturulan modelin (Günlük KOR<,80) DVM sınıflandırıcılarının performansı Tablo 15'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 396 gözlemi (%25,78) doğru sınıflandırırken artış yönündeki 1433 (%82,31) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 102 gözlem (%24,82) doğru sınıflandırılırken artış yönündeki 318 (%77,94) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %55,81'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %51,28'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 15 : Sınıflandırma Matrisi (Günlük KOR&lt;,80 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	396	1140	1536	% 25,78	102	309	411	% 24,82
	Artış	308	1433	1741	% 82,31	90	318	408	% 77,94
	Σ	704	2573	3277	<b>% 55,81</b>	192	627	819	<b>% 51,28</b>

\*gamma= 0,000001 cost= 1000000

Korelasyon analizi ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık KOR<,80) DVM sınıflandırıcılarının performansı Tablo 16'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1257 gözlemi (%81,84) doğru sınıflandırırken artış yönündeki 1535 (%88,17) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 325 gözlem (%79,08) doğru sınıflandırılırken artış yönündeki 347 (%85,05) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %85,20'si doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %82,05'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 16 : Sınıflandırma Matrisi (Günlük Dalgacık KOR&lt;,80 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1257	279	1536	% 81,84	325	86	411	% 79,08
	Artış	206	1535	1741	% 88,17	61	347	408	% 85,05
	Σ	1463	1814	3277	<b>% 85,20</b>	386	433	819	<b>% 82,05</b>

\*gamma= 0,0001 cost= 10

Korelasyon analizi ile korelasyonları 0,80'den büyük Close, Low, High, Open, sma.200, wma.200, ema.200, wma.20, ema.20, sar, bbands.HLC.ma, sma.20, bbands.HLC.dn, bbands.HLC.up, obv, williamsad, chaikinad, Euro, DAX, rsi,

dmi.DIn, cmo, SP500, bbands.close.pctB, smi.smi, bbands.HLC.pctB, dmi.DIp, cci, stochOSC.fastD, stochOSC.slowD, stochOSC.fastK, macd.macd ve smi.signal değişkenleri ile oluşturulan modelin (Günlük KOR>,80) DVM sınıflandırıcılarının performansı Tablo 17'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 502 gözlemi (%32,68) doğru sınıflandırırken artış yönündeki 1325 (%76,11) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 136 gözlem (%33,09) doğru sınıflandırılırken artış yönündeki 292 (%71,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %55,75'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %52,26'sının doğru sınıflandırıldığı gözlenmektedir.

Tablo 17 : Sınıflandırma Matrisi (Günlük KOR>,80 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
	Azalış	Artış				Azalış	Artış		
Gerçek Sınıf	Azalış	502	1034	1536	% 32,68	136	275	411	% 33,09
	Artış	416	1325	1741	% 76,11	116	292	408	% 71,57
	$\Sigma$	918	2359	3277	<b>% 55,75</b>	252	567	819	<b>% 52,26</b>

\*gamma= 0,000001 cost= 1000000

Korelasyon analizi ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık KOR>,80) DVM sınıflandırıcılarının performansı Tablo 18'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^2$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1446 gözlemi (%94,14) doğru sınıflandırırken artış yönündeki 1657 (%95,18) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 371 gözlem (%90,27) doğru sınıflandırılırken artış yönündeki 372 (%91,18) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %94,69'u doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %90,72'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 18 : Sınıflandırma Matrisi (Günlük Dalgacık KOR&gt;,80 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1446	90	1536	% 94,14	371	40	411	% 90,27
	Artış	84	1657	1741	% 95,18	36	372	408	% 91,18
	Σ	1530	1747	3277	<b>% 94,69</b>	407	412	819	<b>% 90,72</b>

\*gamma= 0,0001 cost= 100

KNN ile elde edilen değişkenlerin önemlilik sırasındaki ilk yedi değişken; Dolar, chaikinad, cmf, Altın, Euro, obv ve Faiz değişkenleri ile oluşturulan modelin (Günlük KNN7) DVM sınıflandırıcılarının performansı Tablo 19'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 278 gözlemi (%18,10) doğru sınıflandırırken artış yönündeki 1485 (%85,30) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 66 gözlem (%16,06) doğru sınıflandırılırken artış yönündeki 331 (%81,13) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %53,80'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %48,47'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 19 : Sınıflandırma Matrisi (Günlük KNN7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	278	1258	1536	% 18,10	66	345	411	% 16,06
	Artış	256	1485	1741	% 85,30	77	331	408	% 81,13
	Σ	534	2743	3277	<b>% 53,80</b>	143	676	819	<b>% 48,47</b>

\*gamma= 0,000001 cost= 1000000

KNN ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık KNN7) DVM sınıflandırıcılarının performansı Tablo 20'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM



sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1199 gözlemi (%78,06) doğru sınıflandırırken artış yönündeki 1435 (%82,42) gözlemi doğru sınıflandırdığı görölmektedir. Test veri setinde ise azalış yönündeki 295 gözlem (%71,78) doğru sınıflandırılırken artış yönündeki 331 (%81,13) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %80,38'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %76,43'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 20 : Sınıflandırma Matrisi (Günlük Dalgacık KNN7 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Azalış		1199	337	1536	% 78,06	295	116	411	% 71,78
Artış		306	1435	1741	% 82,42	77	331	408	% 81,13
$\Sigma$		1505	1772	3277	<b>% 80,38</b>	372	447	819	<b>% 76,43</b>

\*gamma= 0,001 cost= 10

KNN ile elde edilen deęişkenlerin önemlilik sırasındaki ilk yirmi deęişken; Dolar, chaikinad, cmf, Altin, Euro, obv, Faiz, mom, macd.signal, cmo, Volume, macd.macd, sma.200, smi.signal, NIKKEI225, ema.200, dmi.DIn, bbands.close.pctB, Per ve wma.200 deęişkenleri ile oluşturulan modelin (Günlük KNN20) DVM sınıflandırıcılarının performansı Tablo 21'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 403 gözlemi (%26,24) doğru sınıflandırırken artış yönündeki 1402 (%80,53) gözlemi doğru sınıflandırdığı görölmektedir. Test veri setinde ise azalış yönündeki 94 gözlem (%22,87) doğru sınıflandırılırken artış yönündeki 323 (%79,17) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %55,08'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %50,92'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 21 : Sınıflandırma Matrisi (Günlük KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	403	1133	1536	% 26,24	94	317	411	% 22,87
	Artış	339	1402	1741	% 80,53	85	323	408	% 79,17
	Σ	742	2535	3277	<b>% 55,08</b>	179	640	819	<b>% 50,92</b>

\*gamma= 0,000001 cost= 1000000

KNN ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık KNN20) DVM sınıflandırıcılarının performansı Tablo 22'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1481 gözlemi (%96,42) doğru sınıflandırırken artış yönündeki 1684 (%96,73) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 365 gözlem (%88,81) doğru sınıflandırılırken artış yönündeki 373 (%91,42) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %96,58'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %90,11'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 22 : Sınıflandırma Matrisi (Günlük Dalgacık KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1481	55	1536	% 96,42	365	46	411	% 88,81
	Artış	57	1684	1741	% 96,73	35	373	408	% 91,42
	Σ	1538	1739	3277	<b>% 96,58</b>	400	419	819	<b>% 90,11</b>

\*gamma= 0,001 cost= 10

PLS ile elde edilen değişkenlerin önemlilik sırasındaki ilk onyedeki değişken; Dolar, Euro, Per, Faiz, Paz, cmo, Volume, bbands.close.pctB, dmi.DIn, chaikinad, stochOSC.fastK, wpr, obv, cci, rsi, bbands.HLC.pctB ve NIKKEI225 değişkenleri ile

oluşturulan modelin (Günlük PLS17) DVM sınıflandırıcılarının performansı Tablo 23'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçişle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 466 gözlemi (%30,34) doğru sınıflandırırken artış yönündeki 1369 (%78,63) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 124 gözlem (%30,17) doğru sınıflandırılırken artış yönündeki 298 (%73,04) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %56,00'ü doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %51,53'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 23 : Sınıflandırma Matrisi (Günlük PLS17 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	466	1070	1536	% 30,34	124	287	411	% 30,17
	Artış	372	1369	1741	% 78,63	110	298	408	% 73,04
	$\Sigma$	838	2439	3277	<b>% 56,00</b>	234	585	819	<b>% 51,53</b>

\*gamma= 0,000001 cost= 1000000

PLS ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık PLS17) DVM sınıflandırıcılarının performansı Tablo 24'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^4$ ) 10-kat çapraz geçişle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1321 gözlemi (%86,00) doğru sınıflandırırken artış yönündeki 1551 (%89,09) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 322 gözlem (%78,35) doğru sınıflandırılırken artış yönündeki 354 (%86,76) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %87,64'ü doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %82,54'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 24 : Sınıflandırma Matrisi (Günlük Dalgacık PLS17 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1321	215	1536	% 86,00	322	89	411	% 78,35
	Artış	190	1551	1741	% 89,09	54	354	408	% 86,76
	Σ	1511	1766	3277	<b>% 87,64</b>	376	443	819	<b>% 82,54</b>

\*gamma= 0,00001 cost= 10000

PLS ile elde edilen değişkenlerin önemlilik sırasındaki ilk yirmi değişken; Dolar, Euro, Per, Faiz, Paz, cmo, Volume, bbands.close.pctB, dmi.DIn, chaikinad, stochOSC.fastK, wpr, obv, cci, rsi, bbands.HLC.pctB, NIKKEI225, smi.smi, Altın ve smi.signal değişkenleri ile oluşturulan modelin (Günlük PLS20) DVM sınıflandırıcılarının performansı Tablo 25'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 434 gözlemi (%28,26) doğru sınıflandırırken artış yönündeki 1373 (%78,86) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 107 gözlem (%26,03) doğru sınıflandırılırken artış yönündeki 304 (%74,51) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %55,14'ü doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %50,18'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 25 : Sınıflandırma Matrisi (Günlük PLS20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	434	1102	1536	% 28,26	107	304	411	% 26,03
	Artış	368	1373	1741	% 78,86	104	304	408	% 74,51
	Σ	802	2475	3277	<b>% 55,14</b>	211	608	819	<b>% 50,18</b>

\*gamma= 0,00001 cost= 100000

PLS ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık PLS20) DVM sınıflandırıcılarının performansı Tablo 26'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^2$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1388 gözlemi (%90,36) doğru sınıflandırırken artış yönündeki 1591 (%91,38) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 354 gözlem (%86,13) doğru sınıflandırılırken artış yönündeki 366 (%89,71) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %90,91'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %87,91'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 26 : Sınıflandırma Matrisi (Günlük Dalgacık PLS20 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	$\Sigma$		Azalış	Artış	$\Sigma$		
Gerçek Sınıf	Azalış	1388	148	1536	% 90,36	354	57	411	% 86,13
	Artış	150	1591	1741	% 91,38	42	366	408	% 89,71
	$\Sigma$	1538	1739	3277	<b>% 90,91</b>	396	423	819	<b>% 87,91</b>

\*gamma= 0,0001 cost= 100

Randomforest ile elde edilen değişkenlerin önemlilik sırasındaki ilk altı değişken; Dolar, DAX, NASDAQ, Faiz, dmi.DIn ve Cum değişkenleri ile oluşturulan modelin (Günlük RF6) DVM sınıflandırıcılarının performansı Tablo 27'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 301 gözlemi (%19,60) doğru sınıflandırırken artış yönündeki 1489 (%85,53) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 66 gözlem (%16,06) doğru sınıflandırılırken artış yönündeki 337 (%82,60) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %54,62'si doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %49,21'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 27 : Sınıflandırma Matrisi (Günlük RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	301	1235	1536	% 19,60	66	345	411	% 16,06
	Artış	252	1489	1741	% 85,53	71	337	408	% 82,60
	Σ	553	2724	3277	<b>% 54,62</b>	137	682	819	<b>% 49,21</b>

\*gamma= 0,00001 cost= 1000000

Randomforest ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık RF6) DVM sınıflandırıcılarının performansı Tablo 28'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 895 gözlemi (%58,27) doğru sınıflandırırken artış yönündeki 1441 (%82,77) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 223 gözlem (%54,26) doğru sınıflandırılırken artış yönündeki 337 (%82,60) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %71,28'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %68,38'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 28 : Sınıflandırma Matrisi (Günlük Dalgacık RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	895	641	1536	% 58,27	223	188	411	% 54,26
	Artış	300	1441	1741	% 82,77	71	337	408	% 82,60
	Σ	1195	2082	3277	<b>% 71,28</b>	294	525	819	<b>% 68,38</b>

\*gamma= 0,0001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,20'nin altında olan; bbands.HLC.pctB, bbands.close.pctB, DAX, wma.20, wma.200, roc, macd.macd, Per, ema.200, dmi.DIn, atr.atr, cmo, chaikinad, macd.signal, cci, chaikinvolatility,

sar, High, Car, rsi, Euro, ema.20, dmi.Dİp, Close ve dmi.ADX deęişkenleri ile oluşturulan modelin (Günlük LR<,20) DVM sınıflandırıcılarının performansı Tablo 29'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 579 gözlemi (%37,70) doğru sınıflandırırken artış yönündeki 1295 (%74,38) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 160 gözlem (%38,93) doğru sınıflandırılırken artış yönündeki 293 (%71,81) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %57,19'u doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %55,31'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 29 : Sınıflandırma Matrisi (Günlük LR<,20 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	579	957	1536	% 37,70	160	251	411	% 38,93
	Artış	446	1295	1741	% 74,38	115	293	408	% 71,81
	$\Sigma$	1025	2252	3277	<b>% 57,19</b>	275	544	819	<b>% 55,31</b>

\*gamma= 0,000001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,20'nin altında olan yukarıdaki deęişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık LR<,20) DVM sınıflandırıcılarının performansı Tablo 30'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^2$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1398 gözlemi (%91,02) doğru sınıflandırırken artış yönündeki 1594 (%91,56) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 342 gözlem (%83,21) doğru sınıflandırılırken artış yönündeki 368 (%90,20) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %91,30'u doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %86,69'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 30 : Sınıflandırma Matrisi (Günlük Dalgacık LR&lt;,20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1398	138	1536	% 91,02	342	69	411	% 83,21
	Artış	147	1594	1741	% 91,56	40	368	408	% 90,20
	Σ	1545	1732	3277	<b>% 91,30</b>	382	437	819	<b>% 86,69</b>

\*gamma= 0,0001 cost= 100

Lojistik regresyon analizinde anlamlılık düzeyi 0,10'un altında olan; bbands.HLC.pctB, bbands.close.pctB, DAX, wma.20, wma.200, roc, macd.macd, Per, ema.200, dmi.DIn, atr.atr, cmo, chaikinad, macd.signal, cci, chaikinvolatility, sar, High ve Car değişkenleri ile oluşturulan modelin (Günlük LR<,10) DVM sınıflandırıcılarının performansı Tablo 31'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 462 gözlemi (%30,08) doğru sınıflandırırken artış yönündeki 1376 (%79,04) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 111 gözlem (%27,01) doğru sınıflandırılırken artış yönündeki 318 (%77,94) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %56,09'u doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %52,38'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 31 : Sınıflandırma Matrisi (Günlük LR&lt;,10 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	462	1074	1536	% 30,08	111	300	411	% 27,01
	Artış	365	1376	1741	% 79,04	90	318	408	% 77,94
	Σ	827	2450	3277	<b>% 56,09</b>	201	618	819	<b>% 52,38</b>

\*gamma= 0,000001 cost= 1000000



Lojistik regresyon analizinde anlamlılık düzeyi 0,10'un altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık LR<,10) DVM sınıflandırıcılarının performansı Tablo 32'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1328 gözlemi (%86,46) doğru sınıflandırırken artış yönündeki 1555 (%89,32) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 343 gözlem (%83,45) doğru sınıflandırılırken artış yönündeki 369 (%90,44) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %87,98'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %86,94'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 32 : Sınıflandırma Matrisi (Günlük Dalgacık LR<,10 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Azalış		1328	208	1536	% 86,46	343	68	411	% 83,45
Artış		186	1555	1741	% 89,32	39	369	408	% 90,44
Σ		1514	1763	3277	<b>% 87,98</b>	382	437	819	<b>% 86,94</b>

\*gamma= 0,0001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,05'in altında olan; bbands.HLC.pctB, bbands.close.pctB, DAX, wma.20, wma.200, roc, macd.macd, Per, ema.200, dmi.DIn, atr.atr, cmo, chaikinad, macd.signal ve cci değişkenleri ile oluşturulan modelin (Günlük LR<,05) DVM sınıflandırıcılarının performansı Tablo 33'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 570 gözlemi (%37,11) doğru sınıflandırırken artış yönündeki 1279 (%73,46) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 138 gözlem (%33,58) doğru sınıflandırılırken artış yönündeki 282 (%69,12) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %56,42'si doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %51,28'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 33 : Sınıflandırma Matrisi (Günlük LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	570	966	1536	% 37,11	138	273	411	% 33,58
	Artış	462	1279	1741	% 73,46	126	282	408	% 69,12
	Σ	1032	2245	3277	<b>% 56,42</b>	264	555	819	<b>% 51,28</b>

\*gamma= 0,00001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,05'in altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık LR<,05) DVM sınıflandırıcılarının performansı Tablo 34'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1324 gözlemi (%86,20) doğru sınıflandırırken artış yönündeki 1542 (%88,57) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 348 gözlem (%84,67) doğru sınıflandırılırken artış yönündeki 364 (%89,22) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %87,46'sı doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %86,94'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 34 : Sınıflandırma Matrisi (Günlük Dalgacık LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1324	212	1536	% 86,20	348	63	411	% 84,67
	Artış	199	1542	1741	% 88,57	44	364	408	% 89,22
	Σ	1523	1754	3277	<b>% 87,46</b>	392	427	819	<b>% 86,94</b>

\*gamma= 0,0001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,01'in altında olan; bbands.HLC.pctB, bbands.close.pctB, DAX ve wma.20 değişkenleri ile oluşturulan modelin (Günlük LR<,01) DVM sınıflandırıcılarının performansı Tablo 35'te

verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 284 gözlemi (%18,49) doğru sınıflandırırken artış yönündeki 1462 (%83,97) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 72 gözlem (%17,52) doğru sınıflandırılırken artış yönündeki 337 (%82,60) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %53,28'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %49,94'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 35 : Sınıflandırma Matrisi (Günlük LR<,01 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	284	1252	1536	% 18,49	72	339	411	% 17,52
	Artış	279	1462	1741	% 83,97	71	337	408	% 82,60
	Σ	563	2714	3277	<b>% 53,28</b>	143	676	819	<b>% 49,94</b>

\*gamma= 0,000001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,01'in altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık LR<,01) DVM sınıflandırıcılarının performansı Tablo 36'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^4$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1287 gözlemi (%83,79) doğru sınıflandırırken artış yönündeki 1492 (%85,70) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 341 gözlem (%82,97) doğru sınıflandırılırken artış yönündeki 367 (%89,95) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %84,80'i doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %86,45'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 36 : Sınıflandırma Matrisi (Günlük Dalgacık LR&lt;,01 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1287	249	1536	% 83,79	341	70	411	% 82,97
	Artış	249	1492	1741	% 85,70	41	367	408	% 89,95
	Σ	1536	1741	3277	<b>% 84,80</b>	382	437	819	<b>% 86,45</b>

\*gamma= 0,00001 cost= 10000

Son olarak günlük veri setindeki tüm değişkenler ile oluşturulan modelin (Günlük TUM55) DVM sınıflandırıcılarının performansı Tablo 37'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 621 gözlemi (%40,43) doğru sınıflandırırken artış yönündeki 1269 (%72,89) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 171 gözlem (%41,61) doğru sınıflandırılırken artış yönündeki 295 (%72,30) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3277 gözlemin %57,67'si doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %56,90'ının doğru sınıflandırıldığı gözlenmektedir.

Tablo 37 : Sınıflandırma Matrisi (Günlük TUM55 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	621	915	1536	% 40,43	171	240	411	% 41,61
	Artış	472	1269	1741	% 72,89	113	295	408	% 72,30
	Σ	1093	2184	3277	<b>% 57,67</b>	284	535	819	<b>% 56,90</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setindeki yukarıdaki tüm değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Günlük Dalgacık TUM55) DVM sınıflandırıcılarının performansı Tablo 38'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde

edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 1453 gözlemi (%94,60) doğru sınıflandırırken artış yönündeki 1670 (%95,92) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 373 gözlem (%90,75) doğru sınıflandırılırken artış yönündeki 371 (%90,93) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 3227 gözlemin %95,30'u doğru sınıflandırılırken test veri setindeki toplam 819 gözlemin %90,84'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 38 : Sınıflandırma Matrisi (Günlük Dalgacık TUM55 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	1453	83	1536	% 94,60	373	38	411	% 90,75
	Artış	71	1670	1741	% 95,92	37	371	408	% 90,93
	Σ	1524	1753	3277	<b>% 95,30</b>	410	409	819	<b>% 90,84</b>

\*gamma= 0,00001 cost= 1000

#### 4.6.2. Haftalık Veriler ile Endeks Yönü Tahmini

Günlük ham eğitim veri seti kullanılarak değişken seçme yöntemleri yardımıyla oluşturulan modellerdeki değişkenler (gün değişkenleri modellerden çıkarılarak) kullanılarak haftalık veri setinde DVM yöntemi ile bağımlı değişken olan endeksin yönü tahmin edilmektedir. Aşağıda ham eğitim ve test veri setleri ile dalgacık dönüşümü yardımıyla elde edilen eğitim ve test veri setlerinin performansları model bazında sınıflandırma matrisleri ile verilmektedir.

Günlük veri setinde Günlük TBA6 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-TBA6) DVM sınıflandırıcısının performansı Tablo 39'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçişle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 12 gözlemi (%4,10) doğru sınıflandırırken artış yönündeki 374 (%96,86) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki hiçbir gözlem (%0,00) doğru sınıflandırılmazken artış yönündeki 90 (%97,83) gözlem doğru sınıflandırılmıştır.

Eđitim veri setindeki toplam 679 gözlemin %56,85'i dođru sınıflandırılırken test veri setindeki toplam 169 gözlemin %53,25'inin dođru sınıflandırıldığı gözlenmektedir.

Tablo 39 : Sınıflandırma Matrisi (H-TBA6 Modeli)

		Eđitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Dođruluk Oranı	$\Sigma$	Tahmin Sınıfı		Dođruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	12	281	293	% 4,10	0	77	77	% 0,00
	Artış	12	374	386	% 96,89	2	90	92	% 97,83
	$\Sigma$	24	655	679	<b>% 56,85</b>	2	167	169	<b>% 53,25</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük TBA6 modelinde yer alan deđişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-TBA6) DVM sınıflandırıcılarının performansı Tablo 40'ta verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eđitim veri setinde, azalış yönündeki 240 gözlemi (%81,91) dođru sınıflandırırken artış yönündeki 353 (%91,45) gözlemi dođru sınıflandırdığı görölmektedir. Test veri setinde ise azalış yönündeki 62 gözlem (%80,52) dođru sınıflandırılırken artış yönündeki 79 (%85,87) gözlem dođru sınıflandırılmıştır. Eđitim veri setindeki toplam 679 gözlemin %87,33'ü dođru sınıflandırılırken test veri setindeki toplam 169 gözlemin %83,43'ünün dođru sınıflandırıldığı gözlenmektedir.

Tablo 40 : Sınıflandırma Matrisi (Dalgacık H-TBA6 Modeli)

		Eđitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Dođruluk Oranı	$\Sigma$	Tahmin Sınıfı		Dođruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	240	53	293	% 81,91	62	15	77	% 80,52
	Artış	33	353	386	% 91,45	13	79	92	% 85,87
	$\Sigma$	273	406	679	<b>% 87,33</b>	75	94	169	<b>% 83,43</b>

\*gamma= 0.001 cost= 10

Günlük veri setinde Günlük KOR>,80 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H- KOR>,80) DVM sınıflandırıcılarının performansı Tablo 41'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 76 gözlemi (%25,94) doğru sınıflandırırken artış yönündeki 331 (%85,71) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 17 gözlem (%22,08) doğru sınıflandırılırken artış yönündeki 70 (%76,09) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %59,94'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %51,48'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 41 : Sınıflandırma Matrisi (H-KOR>,80 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	76	217	293	% 25,94	17	60	77	% 22,08
	Artış	55	331	386	% 85,75	22	70	92	% 76,09
	$\Sigma$	131	548	679	<b>% 59,94</b>	39	130	169	<b>% 51,48</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük KOR>,80 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-KOR>,80) DVM sınıflandırıcılarının performansı Tablo 42'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 225 gözlemi (%76,79) doğru sınıflandırırken artış yönündeki 347 (%89,90) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 59 gözlem (%76,62) doğru sınıflandırılırken artış yönündeki 78 (%84,78) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %84,24'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %81,07'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 42 : Sınıflandırma Matrisi (Dalgacık H-KOR&gt;,80 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	225	68	293	% 76,79	59	18	77	% 76,62
	Artış	39	347	386	% 89,90	14	78	92	% 84,78
	Σ	264	415	679	<b>% 84,24</b>	73	96	169	<b>% 81,07</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük KNN7 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H- KNN7) DVM sınıflandırıcılarının performansı Tablo 43'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 31 gözlemi (%10,58) doğru sınıflandırırken artış yönündeki 371 (%96,11) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 5 gözlem (%6,49) doğru sınıflandırılırken artış yönündeki 86 (%93,48) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %59,20'si doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 43 : Sınıflandırma Matrisi (H-KNN7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	31	262	293	% 10,58	5	72	77	% 6,49
	Artış	15	371	386	% 96,11	6	86	92	% 93,48
	Σ	46	633	679	<b>% 59,20</b>	11	158	169	<b>% 53,85</b>

\*gamma= 0,0001 cost= 1000000

Günlük veri setinde Günlük KNN7 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H- KNN7) DVM sınıflandırıcılarının performansı Tablo 44'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM



sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 153 gözlemi (%52,22) doğru sınıflandırırken artış yönündeki 328 (%84,87) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 45 gözlem (%58,44) doğru sınıflandırılırken artış yönündeki 75 (%81,52) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %70,84'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %71,01'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 44 : Sınıflandırma Matrisi (Dalgacık H-KNN7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	153	140	293	% 52,22	45	32	77	% 58,44
	Artış	58	328	386	% 84,97	17	75	92	% 81,52
	$\Sigma$	211	468	679	<b>% 70,84</b>	62	107	169	<b>% 71,01</b>

\*gamma= 0,0001 cost= 10

Günlük veri setinde Günlük KNN20 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H- KNN20) DVM sınıflandırıcılarının performansı Tablo 45'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 45 gözlemi (%15,36) doğru sınıflandırırken artış yönündeki 366 (%94,82) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 7 gözlem (%9,09) doğru sınıflandırılırken artış yönündeki 85 (%92,39) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %60,53'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %54,44'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 45 : Sınıflandırma Matrisi (H-KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	45	248	293	% 15,36	7	70	77	% 9,09
	Artış	20	366	386	% 94,82	7	85	92	% 92,39
	Σ	65	614	679	<b>% 60,53</b>	14	155	169	<b>% 54,44</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük KNN20 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-KNN20) DVM sınıflandırıcılarının performansı Tablo 46'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçişlemlerle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 230 gözlemi (%78,50) doğru sınıflandırırken artış yönündeki 346 (%89,64) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 58 gözlem (%75,32) doğru sınıflandırılırken artış yönündeki 80 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %84,83'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %81,66'sının doğru sınıflandırıldığı gözlenmektedir.

Tablo 46 : Sınıflandırma Matrisi (Dalgacık H-KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	230	63	293	% 78,50	58	19	77	% 75,32
	Artış	40	346	386	% 89,64	12	80	92	% 86,96
	Σ	270	409	679	<b>% 84,83</b>	70	99	169	<b>% 81,66</b>

\*gamma= 0,0001 cost= 10

Günlük veri setinde Günlük PLS17 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H- PLS17) DVM sınıflandırıcılarının performansı Tablo 47'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM

sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 67 gözlemi (%22,87) doğru sınıflandırırken artış yönündeki 343 (%88,86) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 12 gözlem (%15,58) doğru sınıflandırılırken artış yönündeki 74 (%80,43) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %60,38'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %50,89'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 47 : Sınıflandırma Matrisi (H-PLS17 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Azalış	67	226	293	% 22,87	12	65	77	% 15,58	
Artış	43	343	386	% 88,86	18	74	92	% 80,43	
$\Sigma$	110	569	679	<b>% 60,38</b>	30	139	169	<b>% 50,89</b>	

\*gamma= 0,00001 cost= 1000000

Günlük veri setinde Günlük PLS17 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H- PLS17) DVM sınıflandırıcılarının performansı Tablo 48'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 188 gözlemi (%64,16) doğru sınıflandırırken artış yönündeki 334 (%86,53) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 52 gözlem (%67,53) doğru sınıflandırılırken artış yönündeki 77 (%83,70) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %76,88'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %76,33'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 48 : Sınıflandırma Matrisi (Dalgacık H-PLS17 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	188	105	293	% 64,16	52	25	77	% 67,53
	Artış	52	334	386	% 86,53	15	77	92	% 83,70
	Σ	240	439	679	<b>% 76,88</b>	67	102	169	<b>% 76,33</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük PLS20 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-PLS20) DVM sınıflandırıcılarının performansı Tablo 49'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 78 gözlemi (%26,62) doğru sınıflandırırken artış yönündeki 336 (%87,05) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 16 gözlem (%20,78) doğru sınıflandırılırken artış yönündeki 70 (%76,09) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %60,97'si doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %50,89'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 49 : Sınıflandırma Matrisi (H-PLS20 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	78	215	293	% 26,62	16	61	77	% 20,78
	Artış	50	336	386	% 87,05	22	70	92	% 76,09
	Σ	128	551	679	<b>% 60,97</b>	38	131	169	<b>% 50,89</b>

\*gamma= 0,00001 cost= 1000000

Günlük veri setinde Günlük PLS20 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-PLS20) DVM sınıflandırıcılarının performansı Tablo 50'de verilmektedir. Dairesel tabanlı çekirdek

kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 216 gözlemi (%73,72) doğru sınıflandırırken artış yönündeki 345 (%89,38) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 50 gözlem (%64,94) doğru sınıflandırılırken artış yönündeki 78 (%84,78) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %82,62'si doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %75,74'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 50 : Sınıflandırma Matrisi (Dalgacık H-PLS20 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
	Azalış	Artış				Azalış	Artış		
Gerçek Sınıf	Azalış	216	77	293	% 73,72	50	27	77	% 64,94
	Artış	41	345	386	% 89,38	14	78	92	% 84,78
	$\Sigma$	257	422	679	<b>% 82,62</b>	64	105	169	<b>% 75,74</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük RF6 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-RF6) DVM sınıflandırıcılarının performansı Tablo 51'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 2 gözlemi (%0,68) doğru sınıflandırırken artış yönündeki 386 (%100,00) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 1 gözlem (%1,30) doğru sınıflandırılırken artış yönündeki 91 (%98,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %57,14'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %54,44'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 51 : Sınıflandırma Matrisi (H-RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	2	291	293	% 0,68	1	76	77	% 1,30
	Artış	0	386	386	% 100,00	1	91	92	% 98,91
	Σ	2	677	679	<b>% 57,14</b>	2	167	169	<b>% 54,44</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük RF6 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-RF6) DVM sınıflandırıcılarının performansı Tablo 52'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 123 gözlemi (%41,98) doğru sınıflandırırken artış yönündeki 345 (%89,38) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 32 gözlem (%41,56) doğru sınıflandırılırken artış yönündeki 77 (%83,70) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %68,92'si doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %64,50'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 52 : Sınıflandırma Matrisi (Dalgacık H-RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	123	170	293	% 41,98	32	45	77	% 41,56
	Artış	41	345	386	% 89,38	15	77	92	% 83,70
	Σ	164	515	679	<b>% 68,92</b>	47	122	169	<b>% 64,50</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük LR<,20 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-LR<,20) DVM sınıflandırıcılarının performansı Tablo 53'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM

sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 54 gözlemi (%18,43) doğru sınıflandırırken artış yönündeki 357 (%92,49) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 11 gözlem (%14,29) doğru sınıflandırılırken artış yönündeki 77 (%83,70) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %60,53'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %52,07'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 53 : Sınıflandırma Matrisi (H-LR<,20 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Azalış	54	239	293	% 18,43	11	66	77	% 14,29	
Artış	29	357	386	% 92,49	15	77	92	% 83,70	
$\Sigma$	83	596	679	<b>% 60,53</b>	26	143	169	<b>% 52,07</b>	

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük LR<,20 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-LR<,20) DVM sınıflandırıcılarının performansı Tablo 54'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 229 gözlemi (%78,16) doğru sınıflandırırken artış yönündeki 349 (%90,41) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 63 gözlem (%81,82) doğru sınıflandırılırken artış yönündeki 80 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %85,13'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %84,62'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 54 : Sınıflandırma Matrisi (Dalgacık H-LR&lt;,20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	229	64	293	% 78,16	63	14	77	% 81,82
	Artış	37	349	386	% 90,41	12	80	92	% 86,96
	Σ	266	413	679	<b>% 85,13</b>	75	94	169	<b>% 84,62</b>

\*gamma= 0,0001 cost= 10

Günlük veri setinde Günlük LR<,10 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-LR<,10) DVM sınıflandırıcılarının performansı Tablo 55'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 30 gözlemi (%10,24) doğru sınıflandırırken artış yönündeki 361 (%93,52) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 6 gözlem (%7,79) doğru sınıflandırılırken artış yönündeki 80 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %57,58'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %50,89'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 55 : Sınıflandırma Matrisi (H-LR&lt;,10 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	30	263	293	% 10,24	6	71	77	% 7,79
	Artış	25	361	386	% 93,52	12	80	92	% 86,96
	Σ	55	624	679	<b>% 57,58</b>	18	151	169	<b>% 50,89</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük LR<,10 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-LR<,10) DVM sınıflandırıcılarının performansı Tablo 56'da verilmektedir. Dairesel tabanlı çekirdek



kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 220 gözlemi (%75,09) doğru sınıflandırırken artış yönündeki 346 (%89,64) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 60 gözlem (%77,92) doğru sınıflandırılırken artış yönündeki 80 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %83,36'sı doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %82,84'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 56 : Sınıflandırma Matrisi (Dalgacık H-LR<,10 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Azalış	220	73	293	% 75,09	60	17	77	% 77,92	
Artış	40	346	386	% 89,64	12	80	92	% 86,96	
$\Sigma$	260	419	679	<b>% 83,36</b>	72	97	169	<b>% 82,84</b>	

\*gamma= 0,0001 cost= 10

Günlük veri setinde Günlük LR<,05 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-LR<,05) DVM sınıflandırıcılarının performansı Tablo 57'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 63 gözlemi (%21,50) doğru sınıflandırırken artış yönündeki 344 (%89,12) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%11,69) doğru sınıflandırılırken artış yönündeki 75 (%81,52) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %59,94'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %49,70'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 57 : Sınıflandırma Matrisi (H-LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	63	230	293	% 21,50	9	68	77	% 11,69
	Artış	42	344	386	% 89,12	17	75	92	% 81,52
	Σ	105	574	679	<b>% 59,94</b>	26	143	169	<b>% 49,70</b>

\*gamma= 0,00001 cost= 1000000

Günlük veri setinde Günlük LR<,05 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-LR<,05) DVM sınıflandırıcılarının performansı Tablo 58'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 220 gözlemi (%75,09) doğru sınıflandırırken artış yönündeki 345 (%89,38) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 62 gözlem (%80,52) doğru sınıflandırılırken artış yönündeki 81 (%88,04) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %83,21'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %84,62'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 58 : Sınıflandırma Matrisi (Dalgacık H-LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	220	73	293	% 75,09	62	15	77	% 80,52
	Artış	41	345	386	% 89,38	11	81	92	% 88,04
	Σ	261	418	679	<b>% 83,21</b>	73	96	169	<b>% 84,62</b>

\*gamma= 0,0001 cost= 10

Günlük veri setinde Günlük LR<,01 modelinde yer alan değişkenler ile haftalık veri setinde oluşturulan modelin (H-LR<,01) DVM sınıflandırıcılarının performansı Tablo 59'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM

sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 28 gözlemi (%9,56) doğru sınıflandırırken artış yönündeki 377 (%97,67) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 6 gözlem (%7,79) doğru sınıflandırılırken artış yönündeki 85 (%92,39) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %59,65'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 59 : Sınıflandırma Matrisi (H-LR<,01 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	28	265	293	% 9,56	6	71	77	% 7,79
	Artış	9	377	386	% 97,67	7	85	92	% 92,39
	$\Sigma$	37	642	679	% 59,65	13	156	169	% 53,85

\*gamma= 0,0001 cost= 1000000

Günlük veri setinde Günlük LR<,01 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-LR<,01) DVM sınıflandırıcılarının performansı Tablo 60'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^5$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 215 gözlemi (%73,38) doğru sınıflandırırken artış yönündeki 326 (%84,46) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 57 gözlem (%74,03) doğru sınıflandırılırken artış yönündeki 77 (%83,70) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %79,68'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %79,29'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 60 : Sınıflandırma Matrisi (Dalgacık H-LR&lt;,01 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	215	78	293	% 73,38	57	20	77	% 74,03
	Artış	60	326	386	% 84,46	15	77	92	% 83,70
	Σ	275	404	679	<b>% 79,68</b>	72	97	169	<b>% 79,29</b>

\*gamma= 0,000001 cost= 100000

Günlük veri setinde Günlük TÜR55 modelinde yer alan değişkenler (gün değişkenleri dışında) ile haftalık veri setinde oluşturulan modelin (H-TÜR55) DVM sınıflandırıcılarının performansı Tablo 61'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-1}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 292 gözlemi (%99,66) doğru sınıflandırırken artış yönündeki 385 (%99,74) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 28 gözlem (%36,36) doğru sınıflandırılırken artış yönündeki 60 (%65,22) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %99,71'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %52,07'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 61 : Sınıflandırma Matrisi (H-TÜR55 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	292	1	293	% 99,66	28	49	77	% 36,36
	Artış	1	385	386	% 99,74	32	60	92	% 65,22
	Σ	293	386	679	<b>% 99,71</b>	60	109	169	<b>% 52,07</b>

\*gamma= 0.1 cost= 10

Günlük veri setinde Günlük TÜR55 modelinde yer alan değişkenlerin (gün değişkenleri dışında) dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık H-TÜR55) DVM sınıflandırıcılarının performansı Tablo 62'de verilmektedir. Dairesel

tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^1$ ) 10-kat çapraz geçişle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 200 gözlemi (%68,26) doğru sınıflandırırken artış yönündeki 353 (%91,45) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 51 gözlem (%66,23) doğru sınıflandırılırken artış yönündeki 83 (%90,22) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %81,44'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %79,29'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 62 : Sınıflandırma Matrisi (Dalgacık H-TÜM55 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	200	93	293	% 68,26	51	26	77	% 66,23
	Artış	33	353	386	% 91,45	9	83	92	% 90,22
	$\Sigma$	233	446	679	<b>% 81,44</b>	60	109	169	<b>% 79,29</b>

\*gamma= 0,00001 cost= 10

#### 4.6.3. Aylık Veriler ile Endeks Yönü Tahmini

Günlük ham eğitim veri seti kullanılarak değişken seçme yöntemleri yardımıyla oluşturulan modellerdeki değişkenler (gün değişkenleri modellerden çıkarılarak) kullanılarak aylık veri setinde DVM yöntemi ile bağımlı değişken olan endeksin yönü tahmin edilmektedir. Aşağıda ham eğitim ve test veri setleri ile dalgacık dönüşümü yardımıyla elde edilen eğitim ve test veri setlerinin performansları model bazında sınıflandırma matrisleri ile verilmektedir.

Günlük veri setinde Günlük TBA6 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-TBA6) DVM sınıflandırıcılarının performansı Tablo 63'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçişle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 31 gözlemi (%45,59) doğru sınıflandırırken artış yönündeki 64 (%71,91) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 3 gözlem (%18,75) doğru

sınıflandırılırken artış yönündeki 16 (%69,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %60,51'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %448,72'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 63 : Sınıflandırma Matrisi (A-TBA6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	31	37	68	% 45,59	3	13	16	% 18,75
	Artış	25	64	89	% 71,91	7	16	23	% 69,57
	Σ	56	101	157	<b>% 60,51</b>	10	29	39	<b>% 48,72</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük TBA6 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-TBA6) DVM sınıflandırıcılarının performansı Tablo 64'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^4$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 64 gözlemi (%94,12) doğru sınıflandırırken artış yönündeki 84 (%94,38) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 7 gözlem (%43,75) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %94,27'si doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %61,54'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 64 : Sınıflandırma Matrisi (Dalgacık A-TBA6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	64	4	68	% 94,12	7	9	16	% 43,75
	Artış	5	84	89	% 94,38	6	17	23	% 73,91
	Σ	69	88	157	<b>% 94,27</b>	13	26	39	<b>% 61,54</b>

\*gamma= 0,000001 cost= 10000

Günlük veri setinde Günlük  $KOR > ,80$  modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A- $KOR > ,80$ ) DVM sınıflandırıcılarının performansı Tablo 65'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 36 gözlemi (%52,94) doğru sınıflandırırken artış yönündeki 71 (%79,78) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 4 gözlem (%25,00) doğru sınıflandırılırken artış yönündeki 14 (%60,87) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %68,15'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %46,15'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 65 : Sınıflandırma Matrisi (A- $KOR > ,80$  Modeli)

	Eğitim Veri Seti					Test Veri Seti			
		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	36	32	68	% 52,94	4	12	16	% 25,00
	Artış	18	71	89	% 79,78	9	14	23	% 60,87
	$\Sigma$	54	103	157	<b>% 68,15</b>	13	26	39	<b>% 46,15</b>

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük  $KOR > ,80$  modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A- $KOR > ,80$ ) DVM sınıflandırıcılarının performansı Tablo 66'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^2$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 55 gözlemi (%80,88) doğru sınıflandırırken artış yönündeki 83 (%93,26) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 18 (%78,26) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %87,90'ı doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %69,23'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 66 : Sınıflandırma Matrisi (Dalgacık A-KOR&gt;,80 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	55	13	68	% 80,88	9	7	16	% 56,25
	Artış	6	83	89	% 93,26	5	18	23	% 78,26
	Σ	61	96	157	<b>% 87,90</b>	14	25	39	<b>% 69,23</b>

\*gamma= 0,00001 cost= 100

Günlük veri setinde Günlük KNN7 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-KNN7) DVM sınıflandırıcılarının performansı Tablo 67'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-2}$ ,  $c = 10^2$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 22 gözlemi (%32,35) doğru sınıflandırırken artış yönündeki 78 (%87,64) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %63,69'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 67 : Sınıflandırma Matrisi (A-KNN7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	22	46	68	% 32,35	2	14	16	% 12,50
	Artış	11	78	89	% 87,64	4	19	23	% 82,61
	Σ	33	124	157	<b>% 63,69</b>	6	33	39	<b>% 53,85</b>

\*gamma= 0.01 cost= 100

Günlük veri setinde Günlük KNN7 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-KNN7) DVM sınıflandırıcılarının performansı Tablo 68'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 59 gözlemi



(%86,76) doğru sınıflandırırken artış yönündeki 83 (%93,26) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %90,45'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %66,67'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 68 : Sınıflandırma Matrisi (Dalgacık A-KNN7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	59	9	68	% 86,76	9	7	16	% 56,25
	Artış	6	83	89	% 93,26	6	17	23	% 73,91
	Σ	65	92	157	<b>% 90,45</b>	15	24	39	<b>% 66,67</b>

\*gamma= 0.001 cost= 10

Günlük veri setinde Günlük KNN20 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-KNN20) DVM sınıflandırıcılarının performansı Tablo 69'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-2}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 33 gözlemi (%48,53) doğru sınıflandırırken artış yönündeki 78 (%87,64) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 1 gözlem (%6,25) doğru sınıflandırılırken artış yönündeki 18 (%78,26) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %70,70'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %48,72'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 69 : Sınıflandırma Matrisi (A-KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	33	35	68	% 48,53	1	15	16	% 6,25
	Artış	11	78	89	% 87,64	5	18	23	% 78,26
	Σ	44	113	157	<b>% 70,70</b>	6	33	39	<b>% 48,72</b>

\*gamma= 0.01 cost= 10

Günlük veri setinde Günlük KNN20 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-KNN20) DVM sınıflandırıcılarının performansı Tablo 70'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 59 gözlemi (%86,76) doğru sınıflandırırken artış yönündeki 85 (%95,51) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %91,72'si doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %74,36'sının doğru sınıflandırıldığı gözlenmektedir.

Tablo 70 : Sınıflandırma Matrisi (Dalgacık A-KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	59	9	68	% 86,76	9	7	16	% 56,25
	Artış	4	85	89	% 95,51	3	20	23	% 86,96
	Σ	63	94	157	<b>% 91,72</b>	12	27	39	<b>% 74,36</b>

\*gamma= 0,0001 cost= 10

Günlük veri setinde Günlük PLS17 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-PLS17) DVM sınıflandırıcılarının performansı Tablo 71'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı

için parametreler ( $\gamma = 10^{-2}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 29 gözlemi (%42,65) doğru sınıflandırırken artış yönündeki 71 (%79,78) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 16 (%69,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %63,69'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %46,15'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 71 : Sınıflandırma Matrisi (A-PLS17 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Azalış		29	39	68	% 42,65	2	14	16	% 12,50
Artış		18	71	89	% 79,78	7	16	23	% 69,57
$\Sigma$		47	110	157	<b>% 63,69</b>	9	30	39	<b>% 46,15</b>

\*gamma= 0.01 cost= 10

Günlük veri setinde Günlük PLS17 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-PLS17) DVM sınıflandırıcılarının performansı Tablo 72'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 64 gözlemi (%94,12) doğru sınıflandırırken artış yönündeki 88 (%98,88) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 6 gözlem (%37,50) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %96,82'si doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %66,67'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 72 : Sınıflandırma Matrisi (Dalgacık A-PLS17 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	64	4	68	% 94,12	6	10	16	% 37,50
	Artış	1	88	89	% 98,88	3	20	23	% 86,96
	Σ	65	92	157	<b>% 96,82</b>	9	30	39	<b>% 66,67</b>

\*gamma= 0.001 cost= 10

Günlük veri setinde Günlük PLS20 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-PLS20) DVM sınıflandırıcılarının performansı Tablo 73'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-2}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 29 gözlemi (%42,65) doğru sınıflandırırken artış yönündeki 76 (%85,39) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 16 (%69,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %66,88'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %46,15'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 73 : Sınıflandırma Matrisi (A-PLS20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	29	39	68	% 42,65	2	14	16	% 12,50
	Artış	13	76	89	% 85,39	7	16	23	% 69,57
	Σ	42	115	157	<b>% 66,88</b>	9	30	39	<b>% 46,15</b>

\*gamma= 0.01 cost= 10

Günlük veri setinde Günlük PLS20 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-PLS20) DVM sınıflandırıcılarının performansı Tablo 74'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz

geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 67 gözlemi (%98,53) doğru sınıflandırırken artış yönündeki 89 (%100,00) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 18 (%78,26) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %99,36'sı doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %69,23'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 74 : Sınıflandırma Matrisi (Dalgacık A-PLS20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	67	1	68	% 98,53	9	7	16	% 56,25
	Artış	0	89	89	% 100,00	5	18	23	% 78,26
	Σ	67	90	157	<b>% 99,36</b>	14	25	39	<b>% 69,23</b>

\*gamma= 0.001 cost= 10

Günlük veri setinde Günlük RF6 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-RF6) DVM sınıflandırıcılarının performansı Tablo 75'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 25 gözlemi (%36,76) doğru sınıflandırırken artış yönündeki 76 (%85,39) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %64,33'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 75 : Sınıflandırma Matrisi (A-RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	25	43	68	% 36,76	2	14	16	% 12,50
	Artış	13	76	89	% 85,39	4	19	23	% 82,61
	Σ	38	119	157	<b>% 64,33</b>	6	33	39	<b>% 53,85</b>

\*gamma= 0,0001 cost= 1000000

Günlük veri setinde Günlük RF6 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-RF6) DVM sınıflandırıcılarının performansı Tablo 76'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-1}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 68 gözlemi (%100,00) doğru sınıflandırırken artış yönündeki 89 (%100,00) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki hiçbir gözlem doğru sınıflandırılmazken artış yönündeki 23 (%100,00) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %100,00'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %58,97'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 76 : Sınıflandırma Matrisi (Dalgacık A-RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	68	0	68	% 100,00	0	16	16	% 0,00
	Artış	0	89	89	% 100,00	0	23	23	% 100,00
	Σ	68	89	157	<b>% 100,00</b>	0	39	39	<b>% 58,97</b>

\*gamma= 0.1 cost= 10

Günlük veri setinde Günlük LR<,20 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-LR<,20) DVM sınıflandırıcılarının performansı Tablo 77'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı

için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^2$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 29 gözlemi (%42,65) doğru sınıflandırırken artış yönündeki 74 (%83,15) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 16 (%69,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %65,61'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %46,15'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 77 : Sınıflandırma Matrisi (A-LR<,20 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Azalış	29	39	68	% 42,65	2	14	16	% 12,50	
Artış	15	74	89	% 83,15	7	16	23	% 69,57	
Σ	44	113	157	<b>% 65,61</b>	9	30	39	<b>% 46,15</b>	

\*gamma= 0.001 cost= 100

Günlük veri setinde Günlük LR<,20 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-LR<,20) DVM sınıflandırıcılarının performansı Tablo 78'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 54 gözlemi (%79,41) doğru sınıflandırırken artış yönündeki 85 (%95,51) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 8 gözlem (%50,00) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %88,54'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %69,23'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 78 : Sınıflandırma Matrisi (Dalgacık A-LR&lt;,20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	54	14	68	% 79,41	8	8	16	% 50,00
	Artış	4	85	89	% 95,51	4	19	23	% 82,61
	Σ	58	99	157	<b>% 88,54</b>	12	27	39	<b>% 69,23</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük LR<,10 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-LR<,10) DVM sınıflandırıcılarının performansı Tablo 79'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^2$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 11 gözlemi (%16,18) doğru sınıflandırırken artış yönündeki 88 (%98,88) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 1 gözlem (%6,25) doğru sınıflandırılırken artış yönündeki 21 (%91,30) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %63,06'sı doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %56,41'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 79 : Sınıflandırma Matrisi (A-LR&lt;,10 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	11	57	68	% 16,18	1	15	16	% 6,25
	Artış	1	88	89	% 98,88	2	21	23	% 91,30
	Σ	12	145	157	<b>% 63,06</b>	3	36	39	<b>% 56,41</b>

\*gamma= 0.001 cost= 100

Günlük veri setinde Günlük LR<,10 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-LR<,10) DVM sınıflandırıcılarının performansı Tablo 80'de verilmektedir. Dairesel tabanlı çekirdek



kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 53 gözlemi (%77,94) doğru sınıflandırırken artış yönündeki 82 (%92,13) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 10 gözlem (%62,50) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %85,99'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %74,36'sının doğru sınıflandırıldığı gözlenmektedir.

Tablo 80 : Sınıflandırma Matrisi (Dalgacık A-LR<,10 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	53	15	68	% 77,94	10	6	16	% 62,50
	Artış	7	82	89	% 92,13	4	19	23	% 82,61
	Σ	60	97	157	<b>% 85,99</b>	14	25	39	<b>% 74,36</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük LR<,05 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-LR<,05) DVM sınıflandırıcılarının performansı Tablo 81'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-2}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 24 gözlemi (%35,29) doğru sınıflandırırken artış yönündeki 81 (%91,01) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %66,88'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %56,41'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 81 : Sınıflandırma Matrisi (A-LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	24	44	68	% 35,29	2	14	16	% 12,50
	Artış	8	81	89	% 91,01	3	20	23	% 86,96
	Σ	32	125	157	<b>% 66,88</b>	5	34	39	<b>% 56,41</b>

\*gamma= 0.01 cost= 10

Günlük veri setinde Günlük LR<,05 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-LR<,05) DVM sınıflandırıcılarının performansı Tablo 82'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 52 gözlemi (%76,47) doğru sınıflandırırken artış yönündeki 83 (%93,26) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %85,99'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %71,79'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 82 : Sınıflandırma Matrisi (Dalgacık A-LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	52	16	68	% 76,47	9	7	16	% 56,25
	Artış	6	83	89	% 93,26	4	19	23	% 82,61
	Σ	58	99	157	<b>% 85,99</b>	13	26	39	<b>% 71,79</b>

\*gamma= 0,000001 cost= 1000

Günlük veri setinde Günlük LR<,01 modelinde yer alan değişkenler ile aylık veri setinde oluşturulan modelin (A-LR<,01) DVM sınıflandırıcılarının performansı Tablo 83'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı

için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^7$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 25 gözlemi (%36,76) doğru sınıflandırırken artış yönündeki 64 (%71,91) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 6 gözlem (%37,50) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %56,69'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %64,10'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 83 : Sınıflandırma Matrisi (A-LR<,01 Modeli)

	Eğitim Veri Seti					Test Veri Seti				
	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	
	Azalış	Artış				Azalış	Artış			
Gerçek Sınıf	Azalış	25	43	68	% 36,76	6	10	16	% 37,50	
	Artış	25	64	89	% 71,91	4	19	23	% 82,61	
	$\Sigma$	50	107	157	<b>% 56,69</b>	10	29	39	<b>% 64,10</b>	

\*gamma= 0,000001 cost= 1000000

Günlük veri setinde Günlük LR<,01 modelinde yer alan değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-LR<,01) DVM sınıflandırıcılarının performansı Tablo 84'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 54 gözlemi (%79,41) doğru sınıflandırırken artış yönündeki 83 (%93,26) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 8 gözlem (%50,00) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %87,26'sı doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %71,79'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 84 : Sınıflandırma Matrisi (Dalgacık A-LR&lt;,01 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	54	14	68	% 79,41	8	8	16	% 50,00
	Artış	6	83	89	% 93,26	3	20	23	% 86,96
	Σ	60	97	157	<b>% 87,26</b>	11	28	39	<b>% 71,79</b>

\*gamma= 0.001 cost= 10

Günlük veri setinde Günlük TUM55 modelinde yer alan değişkenler (gün değişkenleri dışında) ile aylık veri setinde oluşturulan modelin (A-TUM55) DVM sınıflandırıcılarının performansı Tablo 85'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-1}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 68 gözlemi (%100,00) doğru sınıflandırırken artış yönündeki 89 (%100,00) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 5 gözlem (%31,25) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %100,00'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %56,41'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 85 : Sınıflandırma Matrisi (A-TUM55 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	68	0	68	% 100,00	5	11	16	% 31,25
	Artış	0	89	89	% 100,00	6	17	23	% 73,91
	Σ	68	89	157	<b>% 100,00</b>	11	28	39	<b>% 56,41</b>

\*gamma= 0.1 cost= 10

Günlük veri setinde Günlük TUM55 modelinde yer alan değişkenlerin (gün değişkenleri dışında) dalgacık dönüşümleri ile oluşturulan modelin (Dalgacık A-TUM55) DVM sınıflandırıcılarının performansı Tablo 86'da verilmektedir. Dairesel

tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 63 gözlemi (%92,65) doğru sınıflandırırken artış yönündeki 86 (%96,63) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 7 gözlem (%43,75) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %94,90'ı doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %66,67'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 86 : Sınıflandırma Matrisi (Dalgacık A-TÜM55 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	63	5	68	% 92,65	7	9	16	% 43,75
	Artış	3	86	89	% 96,63	4	19	23	% 82,61
	$\Sigma$	66	91	157	<b>% 94,90</b>	11	28	39	<b>% 66,67</b>

\*gamma= 0,0001 cost= 10

#### 4.7. HAFTALIK VERİLERDE DEĞİŞKEN SEÇİMİ VE ENDEKS YÖNÜ TAHMİNİ

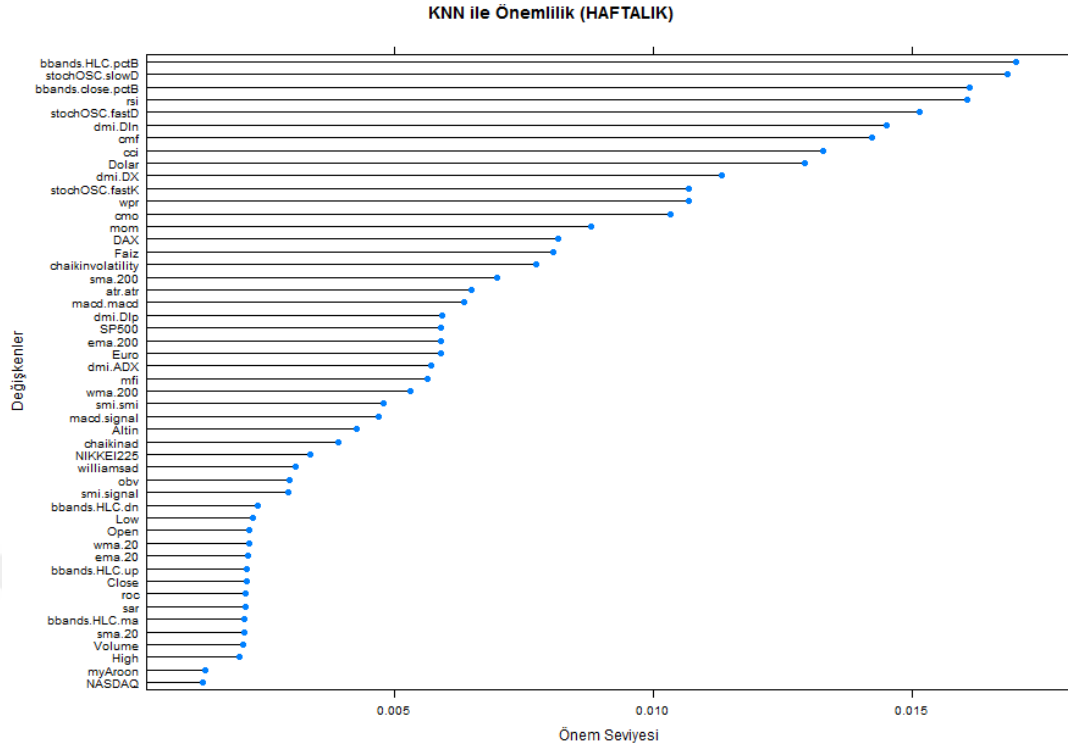
Çalışmada haftalık veri seti üzerinde değişken seçme metotları yardımıyla değişken seçimi yapılarak elde edilen modellerin performansları aşağıda verilmektedir.

Tablo 87'de haftalık veri setinde bulunan değişkenlere temel bileşen analizi uygulandığında altı temel bileşen elde edildiği ve her bir bileşenin en yüksek yüke sahip değişkenlerinin, bbands.HLC.ma, stochOSC.fastD, NIKKEI225, macd.signal, dmi.DX ve roc değişkenleri olduğu görülmektedir. Elde edilen bu temel bileşenlerin toplam varyansın 0,90'ını açıkladığı görülmektedir.

Tablo 87 : Haftalık Veri Setinde Temel Bileşen Analizi

Değişkenler	TB1	TB2	TB3	TB4	TB5	TB6
bbands.HLC.ma	0,99	-0,01	0,14	0,02	-0,02	-0,01
bbands.HLC.up	0,99	-0,02	0,14	0,01	0,01	-0,02
ema.20	0,99	0,00	0,14	0,02	-0,02	-0,01
sma.20	0,99	-0,01	0,14	0,02	-0,02	-0,01
wma.20	0,99	0,00	0,14	0,03	-0,02	-0,02
bbands.HLC.dn	0,98	-0,01	0,14	0,03	-0,04	-0,01
chaikinad	0,98	0,00	-0,05	-0,02	-0,03	0,02
ema.200	0,98	-0,04	0,12	-0,07	-0,04	-0,01
obv	0,98	0,01	0,07	-0,01	-0,02	0,02
sar	0,98	-0,01	0,14	0,05	-0,01	-0,03
sma.200	0,98	-0,05	0,12	-0,08	-0,04	-0,01
wma.200	0,98	-0,04	0,13	-0,07	-0,04	-0,01
Open	0,98	0,04	0,14	0,03	-0,02	-0,01
High	0,98	0,04	0,14	0,03	-0,02	-0,01
Low	0,98	0,05	0,14	0,03	-0,02	0,00
Close	0,98	0,05	0,14	0,03	-0,02	0,00
williamsad	0,95	0,07	0,06	0,06	-0,03	-0,01
Altın	0,93	-0,01	-0,07	-0,02	-0,03	0,03
Euro	0,88	-0,02	-0,26	-0,05	-0,05	0,04
atr.atr	0,83	-0,17	0,13	-0,15	0,10	-0,03
Faiz	-0,82	-0,04	0,11	-0,07	0,00	-0,02
Volume	0,80	0,12	-0,23	-0,01	0,19	0,03
Dolar	0,77	-0,03	-0,34	-0,06	-0,05	0,05
stochOSC.fastD	0,03	0,98	0,00	-0,06	-0,02	0,04
stochOSC.slowD	0,03	0,98	0,01	0,01	-0,01	-0,10
bbands.HLC.pctB	0,02	0,97	-0,01	0,02	-0,01	0,16
bbands.close.pctB	0,02	0,96	-0,02	0,01	-0,01	0,20
cci	0,03	0,96	-0,01	0,01	-0,01	0,18
cmo	0,02	0,95	0,00	0,15	0,08	0,02
stochOSC.fastK	0,03	0,94	-0,01	-0,11	-0,03	0,22
rsi	0,03	0,90	0,01	0,34	0,12	0,15
mfi	-0,05	0,86	0,00	0,15	0,06	-0,02
dmi.Dip	0,00	0,83	0,07	0,33	0,21	0,12
myAroon	0,04	0,80	0,00	0,39	0,04	-0,13
smi.smi	0,07	0,75	0,01	0,59	0,09	-0,13
cmf	-0,37	0,62	-0,17	0,17	0,07	-0,03
NIKKEI225	-0,18	0,01	0,90	0,05	0,03	-0,04
SP500	0,51	-0,02	0,81	0,04	-0,04	0,00
NASDAQ	0,48	-0,02	0,79	0,08	0,01	0,03
DAX	0,66	-0,04	0,71	-0,01	-0,02	-0,02
macd.signal	-0,05	0,42	0,12	0,83	0,20	0,00
smi.signal	0,08	0,54	0,00	0,77	0,13	-0,13
macd.macd	-0,06	0,60	0,11	0,71	0,17	0,02
dmi.DX	-0,03	0,19	-0,02	-0,03	0,87	-0,05
dmi.ADX	-0,05	0,17	-0,02	0,10	0,82	0,03
chaikinvolatility	-0,01	-0,10	0,04	0,22	0,67	0,03
roc	-0,04	0,40	-0,01	-0,01	0,03	0,84
mom	0,02	0,37	-0,03	-0,07	-0,01	0,84
dmi.DIn	0,07	-0,82	-0,01	-0,41	-0,07	-0,12
wpr	-0,03	-0,94	0,01	0,11	0,03	-0,22
Aygen değeri	21,74	13,17	3,19	2,92	2,12	1,74
Açıklanan Yıgımlı Varyans	0,43	0,70	0,76	0,82	0,86	0,90

KNN ile haftalık veri setinde elde edilen değişkenlerin önemlilik sırası Şekil 40'da verilmektedir. Haftalık veri setinde KNN ile en önemli değişkenin bbands.HLC.pctB değişkeni olduğu ve bu değişkeni stochOSC.slowD değişkeninin takip ettiği görülmektedir. Günlük modellerin oluşturulmasına benzer şekilde haftalık veri setinde de ilk yirmi değişken ve ilk dokuz değişken kullanılarak iki ayrı model oluşturulmuştur.



Şekil 40 : Haftalık Veri setinde KNN

Çalışmada değişken seçimi için korelasyondan da yararlanılmış, haftalık veri setindeki değişkenlerin Ek'te verilen korelasyon katsayıları incelenerek değişken seçimi yapılmıştır.

Çalışmada değişken seçimi için lojistik regresyon analizinde değişkenlerin anlamlılık değerlerinden yararlanılmıştır. Haftalık eğitim veri setinde anlamlılık değerleri "0,20", "0,10", "0,05" ve "0,01" değerlerinden küçük olan değişkenler ayrı ayrı seçilerek haftalık modeller oluşturulmuştur.

Tablo 88'de haftalık veri setine uygulanan LR analizinde anlamlılık değerlerine göre modellere alınan değişkenlerin anlamlılık düzeyleri verilmektedir. Tabloda anlamlılık düzeyi 0,01'den küçük hiç değişken olmadığı, anlamlılık düzeyi 0,05'den küçük üç değişken elde edildiği görülmektedir.

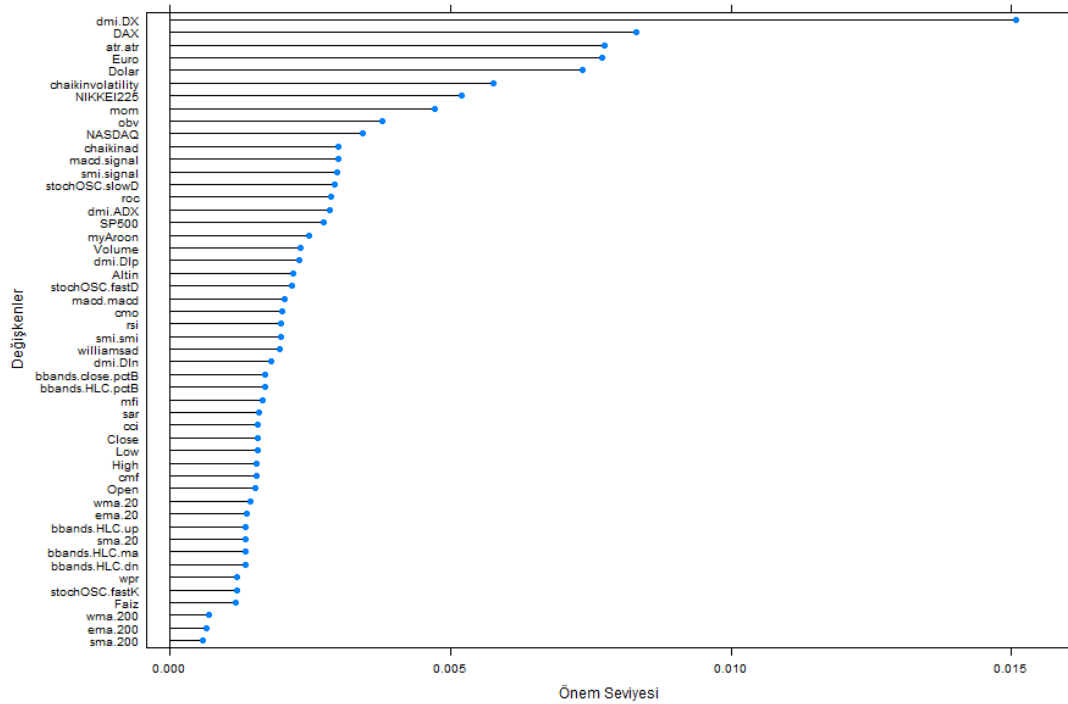
Tablo 88 : LR ile Anlamlılık Düzeyleri (Haftalık)

	B	Std. Hata	Wald	Anlamlılık Düzeyi
rsi	0,152	0,06061	2,508	0,0122 ***
atr.atr	-0,001853	0,0008344	-2,22	0,0264 ***
wma.200	0,0003566	0,0001634	2,182	0,0291 ***
NIKKEI225	0,0001566	8,157E-05	1,92	0,0548 **
chaikinad	4,231E-11	2,838E-11	1,491	0,1360 *
dmi.Dlp	-0,04544	0,03274	-1,388	0,1652 *
Euro	1,326	0,9583	1,383	0,1666 *
sar	-0,0001585	0,0001155	-1,372	0,1701 *
Volume	-1,288E-09	9,804E-10	-1,314	0,1890 *

\* < 0,20 \*\* < 0,10 \*\*\* < 0,05 \*\*\*\* < 0,01

PLS ile haftalık veri setinde elde edilen değişkenlerin önemlilik sırası Şekil 41'de verilmektedir. Haftalık veri setinde PLS ile en önemli değişkenin dmi.DX değişkeni olduğu ve bu değişkeni DAX değişkeninin takip ettiği görülmektedir. Haftalık veri setinde de ilk yirmi değişken kullanılarak PLS yardımıyla bir model oluşturulmuştur.

PLS ile Önemlilik (HAFTALIK)



Şekil 41 : Haftalık Veri setinde PLS



Çalışmada kullanılan haftalık veri seti Liaw ve Wiener (2002) tarafından geliştirilen "randomForest" paketinden yararlanılarak incelenmiş ve haftalık veri setinde Dolar, stochOSC.slowD, stochOSC.fastD, bbands.HLC.pctB, dmi.DX, mom değişkenlerinden oluşan model elde edilmiştir.

Haftalık ham eğitim veri seti kullanılarak değişken seçme yöntemleri yardımıyla oluşturulan modellerin DVM yöntemi ile bağımlı değişken olan endeksin yönünü tahmin konusunda eğitim ve test veri setlerinin ve dalgacık dönüşümü yardımıyla elde edilen eğitim ve test veri setlerinin performansları model bazında sınıflandırma matrisleri ile aşağıda verilmektedir.

Temel bileşen analizi ile elde edilen bbands.HLC.ma, stochOSC.fastD, NIKKEI225, macd.signal, dmi.DX ve roc değişkenleri ile oluşturulan modelin (Haftalık TBA6) DVM sınıflandırıcılarının performansı Tablo 89'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki hiçbir gözlem (%0,00) doğru sınıflandırılmazken artış yönündeki 383 (%99,22) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki hiçbir gözlem (%0,00) doğru sınıflandırılmazken artış yönündeki 91 (%98,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %56,41'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %53,82'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 89 : Sınıflandırma Matrisi (Haftalık TBA6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	0	293	293	% 0,00	0	77	77	% 0,00
	Artış	3	383	386	% 99,22	1	91	92	% 98,91
	Σ	3	676	679	<b>% 56,41</b>	1	168	169	<b>% 53,85</b>

\*gamma= 0,000001 cost= 1000000

Temel bileşen analizi ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık TBA6) DVM

sınıflandırıcılarının performansı Tablo 90'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^2$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 233 gözlemi (%79,52) doğru sınıflandırırken artış yönündeki 343 (%88,86) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 58 gözlem (%75,32) doğru sınıflandırılırken artış yönündeki 81 (%88,04) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %84,83'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %82,25'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 90 : Sınıflandırma Matrisi (Haftalık Dalgacık TBA6 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	233	60	293	% 79,52	58	19	77	% 75,32
	Artış	43	343	386	% 88,86	11	81	92	% 88,04
	$\Sigma$	276	403	679	<b>% 84,83</b>	69	100	169	<b>% 82,25</b>

\*gamma=0,0001 cost=100

KNN ile elde edilen değişkenlerin önemlilik sırasındaki ilk yirmi değişken; bbands.HLC.pctB, stochOSC.slowD, bbands.close.pctB, rsi, stochOSC.fastD, dmi.DIn, cmf, cci, Dolar, dmi.DX, wpr, stochOSC.fastK, cmo, mom, DAX, Faiz, chaikinvolatility, sma.200, atr.atr ve macd.macd değişkenleri ile oluşturulan modelin (Haftalık KNN20) DVM sınıflandırıcılarının performansı Tablo 91'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 44 gözlemi (%15,02) doğru sınıflandırırken artış yönündeki 361 (%93,52) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 8 gözlem (%10,39) doğru sınıflandırılırken artış yönündeki 79 (%85,87) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %59,65'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %51,48'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 91 : Sınıflandırma Matrisi (Haftalık KNN20 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	44	249	293	% 15,02	8	69	77	% 10,39
	Artış	25	361	386	% 93,52	13	79	92	% 85,87
	Σ	69	610	679	<b>% 59,65</b>	21	148	169	<b>% 51,48</b>

\*gamma= 0,000001 cost= 1000000

KNN ile elde edilen değişkenlerin önemlilik sırasındaki ilk yirmi değişkenin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık KNN20) DVM sınıflandırıcılarının performansı Tablo 92'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^2$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 217 gözlemi (%74,06) doğru sınıflandırırken artış yönündeki 348 (%90,16) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 57 gözlem (%74,03) doğru sınıflandırılırken artış yönündeki 82 (%89,13) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %83,21'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %82,25'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 92 : Sınıflandırma Matrisi (Haftalık Dalgacık KNN20 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	Σ		Azalış	Artış	Σ		
Gerçek Sınıf	Azalış	217	76	293	% 74,06	57	20	77	% 74,03
	Artış	38	348	386	% 90,16	10	82	92	% 89,13
	Σ	255	424	679	<b>% 83,21</b>	67	102	169	<b>% 82,25</b>

\*gamma= 0,00001 cost= 100

PLS ile elde edilen değişkenlerin önemlilik sırasındaki ilk yirmi değişken; dmi.DX, DAX, atr.atr, Euro, Dolar, chaikinvolatility, NIKKEI225, mom, obv, NASDAQ, chaikinad, macd.signal, smi.signal, stochOSC.slowD, roc, dmi.ADX,

SP500, myAroon, Volume ve dmi.Dİp değişkenleri ile oluşturulan modelin (Haftalık PLS20) DVM sınıflandırıcılarının performansı Tablo 93'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 91 gözlemi (%31,06) doğru sınıflandırırken artış yönündeki 333 (%86,27) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 15 gözlem (%19,48) doğru sınıflandırılırken artış yönündeki 72 (%78,26) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %62,44'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %51,48'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 93 : Sınıflandırma Matrisi (Haftalık PLS20 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı
		Azalış	Artış	$\Sigma$		Azalış	Artış	$\Sigma$	
Azalış	91	202	293	% 31,06	15	62	77	% 19,48	
Artış	53	333	386	% 86,27	20	72	92	% 78,26	
$\Sigma$	144	535	679	<b>% 62,44</b>	35	134	169	<b>% 51,48</b>	

\*gamma= 0,00001 cost= 1000000

PLS ile elde edilen değişkenlerin önemlilik sırasındaki ilk yirmi değişkenin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık PLS20) DVM sınıflandırıcılarının performansı Tablo 94'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 225 gözlemi (%76,79) doğru sınıflandırırken artış yönündeki 354 (%91,71) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 52 gözlem (%67,53) doğru sınıflandırılırken artış yönündeki 80 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %85,27'si doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %78,11'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 94 : Sınıflandırma Matrisi (Haftalık Dalgacık PLS20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	225	68	293	% 76,79	52	25	77	% 67,53
	Artış	32	354	386	% 91,71	12	80	92	% 86,96
	Σ	257	422	679	<b>% 85,27</b>	64	105	169	<b>% 78,11</b>

\*gamma= 0,0001 cost= 10

Randomforest ile elde edilen değişkenlerin önemlilik sırasındaki ilk altı değişken; Dolar, stochOSC.slowD, stochOSC.fastD, bbands.HLC.pctB, dmi.DX ve mom değişkenleri ile oluşturulan modelin (Haftalık RF6) DVM sınıflandırıcılarının performansı Tablo 95'te verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 17 gözlemi (%5,80) doğru sınıflandırırken artış yönündeki 368 (%95,34) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%2,60) doğru sınıflandırılırken artış yönündeki 90 (%97,83) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %56,70'i doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %54,44'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 95 : Sınıflandırma Matrisi (Haftalık RF6 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	17	276	293	% 5,80	2	75	77	% 2,60
	Artış	18	368	386	% 95,34	2	90	92	% 97,83
	Σ	35	644	679	<b>% 56,70</b>	4	165	169	<b>% 54,44</b>

\*gamma= 0,00001 cost= 1000000

Randomforest ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık RF6) DVM sınıflandırıcılarının performansı Tablo 96'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM

sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^4$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 221 gözlemi (%75,43) doğru sınıflandırırken artış yönündeki 343 (%88,86) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 54 gözlem (%70,13) doğru sınıflandırılırken artış yönündeki 76 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %83,06'sı doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %76,92'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 96 : Sınıflandırma Matrisi (Haftalık Dalgacık RF6 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı			Doğruluk Oranı	Tahmin Sınıfı			Doğruluk Oranı	
	Azalış	Artış	$\Sigma$		Azalış	Artış	$\Sigma$		
Gerçek Sınıf	Azalış	221	72	293	% 75,43	54	23	77	% 70,13
	Artış	43	343	386	% 88,86	16	76	92	% 82,61
	$\Sigma$	264	415	679	<b>% 83,06</b>	70	99	169	<b>% 76,92</b>

\*gamma= 0,000001 cost= 10000

Lojistik regresyon analizinde anlamlılık düzeyi 0,20'nin altında olan; rsi, atr.atr, wma.200, NIKKEI225, chaikinad, dmi.Dİp, Euro, sar ve Volume değişkenleri ile oluşturulan modelin (Haftalık LR<,20) DVM sınıflandırıcılarının performansı Tablo 97'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 75 gözlemi (%25,60) doğru sınıflandırırken artış yönündeki 333 (%86,27) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%11,69) doğru sınıflandırılırken artış yönündeki 76 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %60,09'u doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %50,30'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 97 : Sınıflandırma Matrisi (Haftalık LR&lt;,20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	75	218	293	% 25,60	9	68	77	% 11,69
	Artış	53	333	386	% 86,27	16	76	92	% 82,61
	Σ	128	551	679	<b>% 60,09</b>	25	144	169	<b>% 50,30</b>

\*gamma= 0,000001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,20'nin altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık LR<,20) DVM sınıflandırıcılarının performansı Tablo 98'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 186 gözlemi (%63,48) doğru sınıflandırırken artış yönündeki 339 (%87,82) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 53 gözlem (%68,83) doğru sınıflandırılırken artış yönündeki 80 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %77,32'si doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %78,70'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 98 : Sınıflandırma Matrisi (Haftalık Dalgacık LR&lt;,20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	186	107	293	% 63,48	53	24	77	% 68,83
	Artış	47	339	386	% 87,82	12	80	92	% 86,96
	Σ	233	446	679	<b>% 77,32</b>	65	104	169	<b>% 78,70</b>

\*gamma= 0,000001 cost= 1000

Lojistik regresyon analizinde anlamlılık düzeyi 0,10'un altında olan; rsi, atr.atr, wma.200 ve NIKKEI225 değişkenleri ile oluşturulan modelin (Haftalık LR<,10) DVM sınıflandırıcılarının performansı Tablo 99'da verilmektedir. Dairesel

tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 17 gözlemi (%5,80) doğru sınıflandırırken artış yönündeki 372 (%96,37) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 4 gözlem (%5,19) doğru sınıflandırılırken artış yönündeki 85 (%92,39) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %57,29'u doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %52,66'sının doğru sınıflandırıldığı gözlenmektedir.

Tablo 99 : Sınıflandırma Matrisi (Haftalık LR<,10 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		Doğruluk			Tahmin Sınıfı		Doğruluk	
	Azalış	Artış	$\Sigma$	Oranı	Azalış	Artış	$\Sigma$	Oranı	
Gerçek Sınıf	Azalış	17	276	293	% 5,80	4	73	77	% 5,19
	Artış	14	372	386	% 96,37	7	85	92	% 92,39
	$\Sigma$	31	648	679	<b>% 57,29</b>	11	158	169	<b>% 52,66</b>

\*gamma= 0,000001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,10'un altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık LR<,10) DVM sınıflandırıcılarının performansı Tablo 100'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 218 gözlemi (%74,40) doğru sınıflandırırken artış yönündeki 337 (%87,31) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 54 gözlem (%70,13) doğru sınıflandırılırken artış yönündeki 78 (%84,78) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %81,74'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %78,11'inin doğru sınıflandırıldığı gözlenmektedir.



Tablo 100 : Sınıflandırma Matrisi (Haftalık Dalgacık LR<,10 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	218	75	293	% 74,40	54	23	77	% 70,13
	Artış	49	337	386	% 87,31	14	78	92	% 84,78
	Σ	267	412	679	<b>% 81,74</b>	68	101	169	<b>% 78,11</b>

\*gamma= 0.001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,05'in altında olan; rsi, atr.atr ve wma.200 değişkenleri ile oluşturulan modelin (Haftalık LR<,05) DVM sınıflandırıcılarının performansı Tablo 101'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 12 gözlemi (%4,10) doğru sınıflandırırken artış yönündeki 380 (%98,45) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 3 gözlem (%3,90) doğru sınıflandırılırken artış yönündeki 87 (%94,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %57,73'ü doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %53,25'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 101 : Sınıflandırma Matrisi (Haftalık LR<,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	12	281	293	% 4,10	3	74	77	% 3,90
	Artış	6	380	386	% 98,45	5	87	92	% 94,57
	Σ	18	661	679	<b>% 57,73</b>	8	161	169	<b>% 53,25</b>

\*gamma= 0,00001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,05'in altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Haftalık Dalgacık LR<,05) DVM sınıflandırıcılarının performansı Tablo 102'de

verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 211 gözlemi (%72,01) doğru sınıflandırırken artış yönündeki 336 (%87,05) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 54 gözlem (%70,13) doğru sınıflandırılırken artış yönündeki 77 (%83,70) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 679 gözlemin %80,56'sı doğru sınıflandırılırken test veri setindeki toplam 169 gözlemin %77,51'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 102 : Sınıflandırma Matrisi (Haftalık Dalgacık LR<,05 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	211	82	293	% 72,01	54	23	77	% 70,13
	Artış	50	336	386	% 87,05	15	77	92	% 83,70
	Σ	261	418	679	<b>% 80,56</b>	69	100	169	<b>% 77,51</b>

\*gamma=0.001 cost= 10

Haftalık veri setinde tüm değişkenler ile oluşturulan modeldeki (Tablo 61) ve korelasyonları 0,80'in üzerindeki değişkenler ile oluşturulan modeldeki (Tablo 41) değişkenler günlük veri setinden seçilen değişkenler ile aynı olduğu için DVM sınıflandırıcısı ile tahmin performanslarına burada tekrar yer verilmemiştir.

#### 4.8. AYLIK VERİLERDE DEĞİŞKEN SEÇİMİ VE ENDEKS YÖNÜ TAHMİNİ

Çalışmada aylık veri seti üzerinde değişken seçme metodları yardımıyla değişken seçimi yapılarak elde edilen modellerin performansları aşağıda verilmektedir.

Tablo 103'te aylık veri setinde bulunan değişkenlere temel bileşen analizi uygulandığında beş temel bileşen elde edildiği ve her bir bileşenin en yüksek yüke sahip değişkenlerinin, bbands.HLC.ma, bbands.HLC.pctB, NIKKEI225, dmi.DX ve

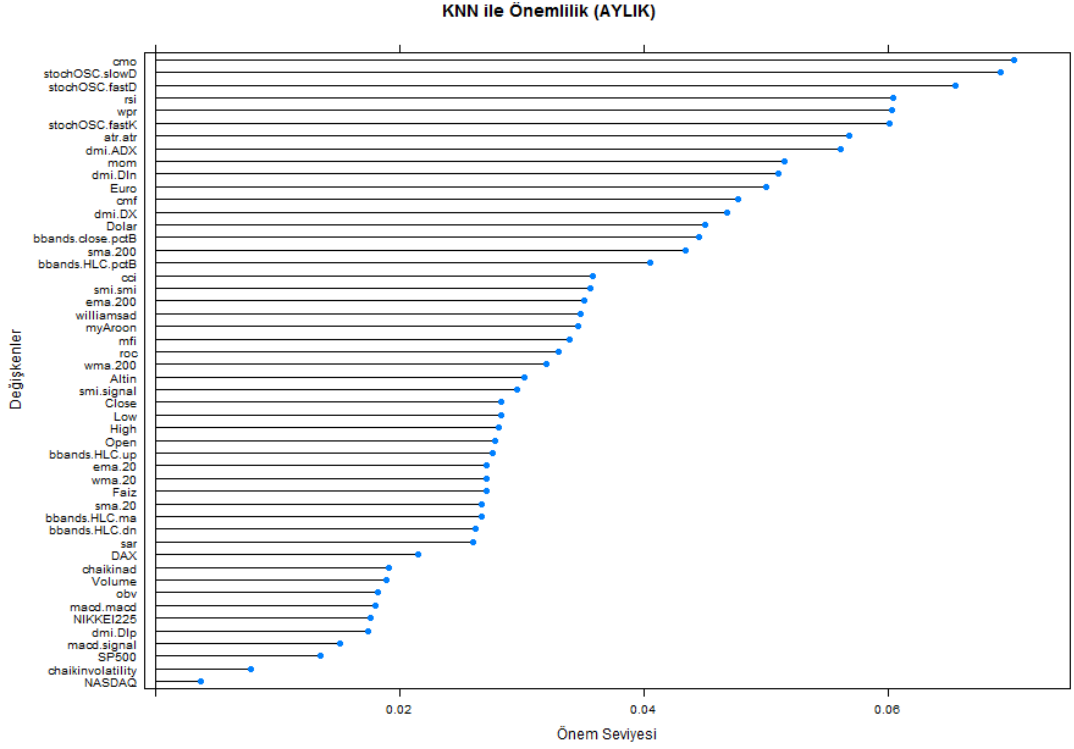
macd.signal değişkenleri olduğu görülmektedir. Elde edilen bu beş temel bileşenin toplam varyansın 0,90'ını açıkladığı anlaşılmaktadır.

Tablo 103 : Aylık Veri Setinde Temel Bileşen Analizi

Değişkenler	TB1	TB2	TB3	TB4	TB5
bbands.HLC.ma	0,99	-0,06	0,14	0,00	0,01
ema.20	0,99	-0,05	0,14	0,00	0,01
sar	0,99	-0,04	0,14	0,01	0,03
sma.20	0,99	-0,06	0,14	0,00	0,01
wma.20	0,99	-0,04	0,14	0,00	0,02
Open	0,99	-0,01	0,14	0,00	0,02
High	0,99	-0,01	0,14	0,00	0,02
Low	0,99	0,00	0,14	-0,01	0,02
Close	0,99	0,00	0,14	0,00	0,02
bbands.HLC.dn	0,98	-0,05	0,13	-0,02	0,02
bbands.HLC.up	0,98	-0,06	0,14	0,02	0,00
chaikinad	0,98	-0,02	-0,05	-0,02	-0,02
ema.200	0,98	-0,08	0,12	-0,04	-0,07
obv	0,98	-0,01	0,06	-0,01	-0,03
wma.200	0,98	-0,08	0,13	-0,04	-0,07
sma.200	0,97	-0,09	0,11	-0,05	-0,09
williamsad	0,95	0,03	0,08	-0,01	0,05
Altın	0,93	-0,03	-0,07	0,00	-0,04
Euro	0,87	-0,03	-0,30	-0,08	-0,08
atr.atr	0,84	-0,22	0,17	0,10	-0,12
Volume	0,83	0,10	-0,24	0,12	0,05
Dolar	0,73	-0,03	-0,39	-0,08	-0,13
bbands.HLC.pctB	-0,02	0,99	-0,02	0,04	0,06
bbands.close.pctB	-0,02	0,99	-0,02	0,04	0,05
cci	-0,02	0,99	-0,02	0,03	0,06
stochOSC.fastD	0,01	0,98	-0,01	0,01	-0,01
stochOSC.slowD	0,01	0,98	-0,02	0,01	0,07
stochOSC.fastK	0,00	0,98	-0,01	0,01	-0,07
cmo	0,00	0,97	-0,02	0,11	0,13
rsi	-0,01	0,90	-0,01	0,21	0,36
myAroon	0,01	0,89	-0,02	0,04	0,37
mfi	-0,08	0,89	-0,01	0,08	0,17
mom	-0,01	0,83	0,03	-0,08	-0,20
dmi.DIp	-0,05	0,81	0,05	0,28	0,35
roc	-0,14	0,79	0,08	0,15	-0,22
smi.smi	0,04	0,78	-0,02	0,16	0,57
cmf	-0,38	0,66	-0,17	0,06	0,12
macd.macd	-0,09	0,63	0,08	0,31	0,62
NIKKEI225	-0,17	0,02	0,89	0,02	0,07
SP500	0,47	-0,03	0,83	0,00	-0,02
NASDAQ	0,43	-0,02	0,80	0,08	0,01
DAX	0,63	-0,08	0,72	0,00	-0,07
dmi.DX	0,00	0,25	0,01	0,89	0,07
dmi.ADX	-0,02	0,25	0,03	0,85	0,05
chaikinvolatility	0,02	0,06	0,04	0,74	0,24
macd.signal	-0,08	0,48	0,09	0,34	0,73
smi.signal	0,05	0,61	-0,02	0,21	0,72
wpr	0,00	-0,98	0,01	-0,01	0,07
Faiz	-0,81	-0,03	0,11	0,01	-0,11
dmi.DIn	0,13	-0,82	0,02	-0,19	-0,37
Aygen değeri	21,65	15,01	3,30	2,64	2,61
Açıklanan Yıgmalı Varyans	0,43	0,73	0,80	0,85	0,90

KNN ile aylık veri setinde elde edilen değişkenlerin önemlilik sırası ise Şekil 42'de verilmektedir. Aylık veri setinde KNN ile en önemli değişkenin cmo olduğu ve bu değişkeni stochOSC.slowD değişkeninin takip ettiği görülmektedir. Günlük

modellerin oluşturulmasına benzer şekilde aylık veri setinde de ilk yirmi değişken ve ilk sekiz değişken kullanılarak iki ayrı model oluşturulmuştur.



Şekil 42 : Aylık Veri setinde KNN

Çalışmada değişken seçimi için korelasyondan da yararlanılmış, aylık veri setindeki değişkenlerin Ek'te verilen korelasyon katsayıları incelenerek değişken seçimi yapılmıştır.

Çalışmada değişken seçimi için lojistik regresyon analizinde değişkenlerin anlamlılık değerlerinden yararlanılmıştır. Aylık eğitim veri setinde anlamlılık değerleri "0,20", "0,10", "0,05" ve "0,01" değerlerinden küçük olan değişkenler ayrı ayrı seçilerek aylık modeller oluşturulmuştur.

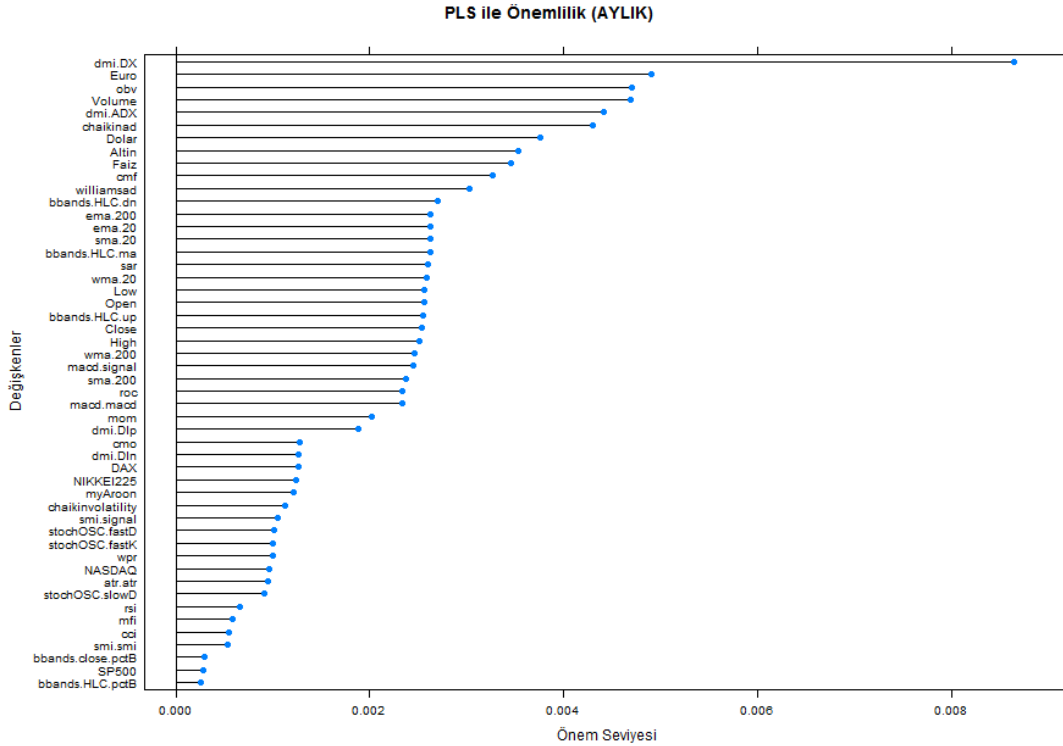
Tablo 104'te aylık veri setine uygulanan LR analizinde anlamlılık değerlerine göre modellere alınan değişkenlerin anlamlılık düzeyleri verilmektedir. Tabloda anlamlılık düzeyi 0,01'den küçük iki değişken elde edildiği görülmektedir.

Tablo 104 : LR ile Anlamlılık Düzeyleri (Aylık)

	B	Std. Hata	Wald	Anlamlılık Düzeyi
High	-0,02597	0,008874	-2,926	0,00343 ****
dmi.DX	0,1483	0,0556	2,668	0,00762 ****
Open	0,02241	0,009244	2,425	0,01532 ***
wma.20	-0,01674	0,009861	-1,697	0,08962 **
dmi.DIp	0,228	0,1344	1,697	0,08971 **
obv	1,6E-10	9,558E-11	1,674	0,09419 **
Altın	-0,009196	0,005997	-1,533	0,12520 *
williamsad	0,0001281	8,386E-05	1,527	0,12669 *
NIKKEI225	0,0004117	0,000271	1,519	0,12869 *
stochOSC.slowD	105,4	71,47	1,475	0,14014 *
atr.atr	0,007133	0,004942	1,443	0,14897 *
stochOSC.fastD	-152,8	114,7	-1,333	0,18268 *
sma.20	0,04185	0,03263	1,283	0,19964 *

\* < 0,20 \*\* < 0,10 \*\*\* < 0,05 \*\*\*\* < 0,01

PLS ile aylık veri setinde elde edilen değişkenlerin önemlilik sırası Şekil 43'te verilmektedir. Aylık veri setinde PLS ile en önemli değişkenin dmi.DX olduğu ve bu değişkeni Euro değişkeninin takip ettiği görülmektedir. Günlük modellerin oluşturulmasına benzer şekilde aylık veri setinde de ilk yirmi değişken kullanılarak PLS yardımıyla bir model elde edilmiştir.



Şekil 43 : Aylık Veri setinde PLS

Çalışmada kullanılan aylık veri seti Liaw ve Wiener (2002) tarafından geliştirilen "randomForest" paketinden yararlanılarak incelenmiş ve aylık veri setinde stochOSC.slowD, Dolar, dmi.ADX, stochOSC.fastD, wpr, cmf, stochOSC.fastK değişkenlerinden oluşan model elde edilmiştir.

Aylık ham eğitim veri seti kullanılarak değişken seçme yöntemleri yardımıyla oluşturulan modellerin DVM yöntemi ile bağımlı değişken olan endeksin yönünü tahmin konusunda eğitim ve test veri setlerinin ve dalgacık dönüşümü yardımıyla elde edilen eğitim ve test veri setlerinin performansları model bazında sınıflandırma matrisleri ile aşağıda verilmektedir.

Temel bileşen analizi ile elde edilen bbands.HLC.ma, bbands.HLC.pctB, NIKKEI225, dmi.DX ve macd.signal değişkenleri ile oluşturulan modelin (Aylık TBA5) DVM sınıflandırıcılarının performansı Tablo 105'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^6$ ) 10-kat çapraz geçişle elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 27 gözlemi (%39,71) doğru sınıflandırırken artış yönündeki 76 (%85,39) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 3 gözlem (%18,75) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %65,61'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %51,28'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 105 : Sınıflandırma Matrisi (Aylık TBA5 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	27	41	68	% 39,71	3	13	16	% 18,75
	Artış	13	76	89	% 85,39	6	17	23	% 73,91
	Σ	40	117	157	<b>% 65,61</b>	9	30	39	<b>% 51,28</b>

\*gamma= 0,0001 cost= 1000000

Temel bileşen analizi ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık TBA5) DVM

sınıflandırıcılarının performansı Tablo 106'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 62 gözlemi (%91,18) doğru sınıflandırırken artış yönündeki 86 (%96,63) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 4 gözlem (%25,00) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %94,27'si doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %58,97'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 106 : Sınıflandırma Matrisi (Aylık Dalgacık TBA5 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	62	6	68	% 91,18	4	12	16	% 25,00
	Artış	3	86	89	% 96,63	4	19	23	% 82,61
	$\Sigma$	65	92	157	<b>% 94,27</b>	8	31	39	<b>% 58,97</b>

\*gamma= 0.001 cost= 10

Korelasyon analizi ile korelasyonları 0,80'den büyük bbands.sma.200, wma.200, ema.200, bbands.HLC.up, bbands.HLC.ma, sma.20, bbands.HLC.dn, ema.20, atr.atr, sar, wma.20, Open, High, Low, Close, obv, chaikinad, williamsad, Euro, dmi.DIn, DAX, rsi, mfi, bbands.close.pctB, bbands.HLC.pctB, cmo, cci, dmi.DIp, myAroon, stochOSC.slowD, stochOSC.fastD, macd.macd, wpr, smi.smi, smi.signal, SP500 ve dmi.ADX değişkenleri ile oluşturulan modelin (Aylık KOR>,80) DVM sınıflandırıcılarının performansı Tablo 107'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^3$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 26 gözlemi (%38,24) doğru sınıflandırırken artış yönündeki 81 (%91,01) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 3 gözlem (%18,75) doğru sınıflandırılırken artış yönündeki 18 (%78,26) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %68,15'i

doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 107 : Sınıflandırma Matrisi (Aylık KOR>,80 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Azalış		26	42	68	% 38,24	3	13	16	% 18,75
Artış		8	81	89	% 91,01	5	18	23	% 78,26
Σ		34	123	157	<b>% 68,15</b>	8	31	39	<b>% 53,85</b>

\*gamma= 0,0001 cost= 1000

Korelasyon analizi ile korelasyonları 0,80'den büyük yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık KOR>,80) DVM sınıflandırıcılarının performansı Tablo 108'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 57 gözlemi (%83,82) doğru sınıflandırırken artış yönündeki 85 (%95,51) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 8 gözlem (%50,00) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %90,45'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %69,23'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 108 : Sınıflandırma Matrisi (Aylık Dalgacık KOR>,80 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı
		Azalış	Artış			Azalış	Artış		
Azalış		57	11	68	% 83,82	8	8	16	% 50,00
Artış		4	85	89	% 95,51	4	19	23	% 82,61
Σ		61	96	157	<b>% 90,45</b>	12	27	39	<b>% 69,23</b>

\*gamma= 0,0001 cost= 10



KNN ile elde edilen deęişkenlerin önemlilik sırasındaki ilk yirmi deęişken; cmo, stochOSC.slowD, stochOSC.fastD, rsi, wpr, stochOSC.fastK, atr.atr, dmi.ADX, mom, dmi.DIn, Euro, cmf, dmi.DX, Dolar, bbands.close.pctB, sma.200, bbands.HLC.pctB, cci, smi.smi ve ema.200 deęişkenleri ile oluşturulan modelin (Aylık KNN20) DVM sınıflandırıcılarının performansı Tablo 109'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 33 gözlemi (%48,53) doğru sınıflandırırken artış yönündeki 76 (%85,39) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 7 gözlem (%43,75) doğru sınıflandırılırken artış yönündeki 14 (%60,87) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %69,43'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 109 : Sınıflandırma Matrisi (Aylık KNN20 Modeli)

		Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$
		Azalış	Artış			Azalış	Artış		
Gerçek Sınıf	Azalış	33	35	68	% 48,53	7	9	16	% 43,75
	Artış	13	76	89	% 85,39	9	14	23	% 60,87
	$\Sigma$	46	111	157	<b>% 69,43</b>	16	23	39	<b>% 53,85</b>

\*gamma= 0,000001 cost= 1000000

KNN ile elde edilen yukarıdaki deęişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık KNN20) DVM sınıflandırıcılarının performansı Tablo 110'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 54 gözlemi (%79,41) doğru sınıflandırırken artış yönündeki 88 (%98,88) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 8 gözlem (%50,00) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %90,45'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %71,79'unun doğru sınıflandırıldığı gözlenmektedir.

Tablo 110 : Sınıflandırma Matrisi (Aylık Dalgacık KNN20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	54	14	68	% 79,41	8	8	16	% 50,00
	Artış	1	88	89	% 98,88	3	20	23	% 86,96
	Σ	55	102	157	<b>% 90,45</b>	11	28	39	<b>% 71,79</b>

\*gamma= 0,0001 cost= 10

PLS ile elde edilen değişkenlerin önemlilik sırasındaki ilk yirmi değişken; dmi.DX, Euro, obv, Volume, dmi.ADX, chaikinad, Dolar, Altın, Faiz, cmf, williamsad, bbands.HLC.dn, ema.200, ema.20, sma.20, bbands.HLC.ma, sar, wma.20, Low ve Open değişkenleri ile oluşturulan modelin (Aylık PLS20) DVM sınıflandırıcılarının performansı Tablo 111'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^5$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 32 gözlemi (%47,06) doğru sınıflandırırken artış yönündeki 72 (%80,90) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 19 (%82,61) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %66,24'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 111 : Sınıflandırma Matrisi (Aylık PLS20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	32	36	68	% 47,06	2	14	16	% 12,50
	Artış	17	72	89	% 80,90	4	19	23	% 82,61
	Σ	49	108	157	<b>% 66,24</b>	6	33	39	<b>% 53,85</b>

\*gamma= 0,00001 cost= 100000

PLS ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık PLS20) DVM sınıflandırıcılarının performansı Tablo 112'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 55 gözlemi (%80,88) doğru sınıflandırırken artış yönündeki 80 (%89,89) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 7 gözlem (%43,75) doğru sınıflandırılırken artış yönündeki 16 (%69,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %85,99'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %58,97'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 112 : Sınıflandırma Matrisi (Aylık Dalgacık PLS20 Modeli)

	Gerçek Sınıf	Eğitim Veri Seti				Test Veri Seti			
		Tahmin Sınıfı		Doğruluk Oranı	Tahmin Sınıfı		Doğruluk Oranı		
		Azalış	Artış		Azalış	Artış			
	Azalış	55	13	68	% 80,88	7	9	16	% 43,75
	Artış	9	80	89	% 89,89	7	16	23	% 69,57
	$\Sigma$	64	93	157	<b>% 85,99</b>	14	25	39	<b>% 58,97</b>

\*gamma= 0,000001 cost= 1000

Randomforest ile elde edilen değişkenlerin önemlilik sırasındaki ilk yedi değişken; stochOSC.slowD, Dolar, dmi.ADX, stochOSC.fastD, wpr, cmf ve stochOSC.fastK değişkenleri ile oluşturulan modelin (Aylık RF7) DVM sınıflandırıcılarının performansı Tablo 113'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-1}$ ,  $c = 10^1$ ) 10-kat çapraz geçерleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 38 gözlemi (%55,88) doğru sınıflandırırken artış yönündeki 81 (%91,01) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 5 gözlem (%31,25) doğru sınıflandırılırken artış yönündeki 16 (%69,57) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %75,80'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 113 : Sınıflandırma Matrisi (Aylık RF7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	38	30	68	% 55,88	5	11	16	% 31,25
	Artış	8	81	89	% 91,01	7	16	23	% 69,57
	Σ	46	111	157	<b>% 75,80</b>	12	27	39	<b>% 53,85</b>

\*gamma= 0.1 cost= 10

Randomforest ile elde edilen yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık RF7) DVM sınıflandırıcılarının performansı Tablo 114'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^1$ ) 10-kat çapraz geçirme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 43 gözlemi (%63,24) doğru sınıflandırırken artış yönündeki 79 (%88,76) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %77,71'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %74,36'sının doğru sınıflandırıldığı gözlenmektedir.

Tablo 114 : Sınıflandırma Matrisi (Aylık Dalgacık RF7 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	43	25	68	% 63,24	9	7	16	% 56,25
	Artış	10	79	89	% 88,76	3	20	23	% 86,96
	Σ	53	104	157	<b>% 77,71</b>	12	27	39	<b>% 74,36</b>

\*gamma= 0,0001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,20'nin altında olan; High, dmi.DX, Open, wma.20, dmi.DIp, obv, Altin, williamsad, NIKKEI225, stochOSC.slowD, atr.atr, stochOSC.fastD ve sma.20 değişkenleri ile oluşturulan modelin (Aylık LR<,20) DVM sınıflandırıcılarının performansı Tablo 115'de

verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 28 gözlemi (%41,18) doğru sınıflandırırken artış yönündeki 76 (%85,39) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 2 gözlem (%12,50) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %66,24'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %48,72'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 115 : Sınıflandırma Matrisi (Aylık LR<,20 Modeli)

	Eğitim Veri Seti					Test Veri Seti				
	Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$		Tahmin Sınıfı		Doğruluk Oranı	$\Sigma$	
	Azalış	Artış				Azalış	Artış			
Gerçek Sınıf	Azalış	28	40	68	% 41,18	2	14	16	% 12,50	
	Artış	13	76	89	% 85,39	6	17	23	% 73,91	
	$\Sigma$	41	116	157	<b>% 66,24</b>	8	31	39	<b>% 48,72</b>	

\*gamma= 0,000001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,20'nin altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık LR<,20) DVM sınıflandırıcılarının performansı Tablo 116'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-6}$ ,  $c = 10^3$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 51 gözlemi (%75,00) doğru sınıflandırırken artış yönündeki 82 (%84,71) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 9 gözlem (%56,25) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %84,71'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %66,67'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 116 : Sınıflandırma Matrisi (Aylık Dalgacık LR&lt;,20 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	51	17	68	% 75,00	9	7	16	% 56,25
	Artış	7	82	89	% 92,13	6	17	23	% 73,91
	Σ	58	99	157	<b>% 84,71</b>	15	24	39	<b>% 66,67</b>

\*gamma= 0,000001 cost= 1000

Lojistik regresyon analizinde anlamlılık düzeyi 0,10'un altında olan High, dmi.DX, Open, wma.20, dmi.Dip ve obv değişkenleri ile oluşturulan modelin (Aylık LR<,10) DVM sınıflandırıcılarının performansı Tablo 117'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-5}$ ,  $c = 10^6$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 29 gözlemi (%42,65) doğru sınıflandırırken artış yönündeki 61 (%68,54) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 3 gözlem (%18,75) doğru sınıflandırılırken artış yönündeki 15 (%65,22) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %57,32'si doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %46,15'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 117 : Sınıflandırma Matrisi (Aylık LR&lt;,10 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	29	39	68	% 42,65	3	13	16	% 18,75
	Artış	28	61	89	% 68,54	8	15	23	% 65,22
	Σ	57	100	157	<b>% 57,32</b>	11	28	39	<b>% 46,15</b>

\*gamma= 0,00001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,10'un altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık LR<,10) DVM sınıflandırıcılarının performansı Tablo 118'de

verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 59 gözlemi (%86,76) doğru sınıflandırırken artış yönündeki 87 (%97,75) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 6 gözlem (%37,50) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %92,99'u doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %58,97'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 118 : Sınıflandırma Matrisi (Aylık Dalgacık LR<,10 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı		Tahmin Sınıfı		$\Sigma$	Doğruluk Oranı
	Azalış	Artış				Azalış	Artış		
Gerçek Sınıf	Azalış	59	9	68	% 86,76	6	10	16	% 37,50
	Artış	2	87	89	% 97,75	6	17	23	% 73,91
	$\Sigma$	61	96	157	<b>% 92,99</b>	12	27	39	<b>% 58,97</b>

\*gamma= 0.001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,05'in altında olan High, dmi.DX ve Open değişkenleri ile oluşturulan modelin (Aylık LR<,05) DVM sınıflandırıcılarının performansı Tablo 119'da verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 24 gözlemi (%35,29) doğru sınıflandırırken artış yönündeki 73 (%82,02) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 1 gözlem (%6,25) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %61,78'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 119 : Sınıflandırma Matrisi (Aylık LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	24	44	68	% 35,29	1	15	16	% 6,25
	Artış	16	73	89	% 82,02	3	20	23	% 86,96
	Σ	40	117	157	<b>% 61,78</b>	4	35	39	<b>% 53,85</b>

\*gamma= 0.001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,05'in altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık LR<,05) DVM sınıflandırıcılarının performansı Tablo 120'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçişleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 49 gözlemi (%72,06) doğru sınıflandırırken artış yönündeki 81 (%91,01) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 6 gözlem (%37,50) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %82,80'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %58,97'sinin doğru sınıflandırıldığı gözlenmektedir.

Tablo 120 : Sınıflandırma Matrisi (Aylık Dalgacık LR&lt;,05 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	49	19	68	% 72,06	6	10	16	% 37,50
	Artış	8	81	89	% 91,01	6	17	23	% 73,91
	Σ	57	100	157	<b>% 82,80</b>	12	27	39	<b>% 58,97</b>

\*gamma= 0.001 cost= 10

Lojistik regresyon analizinde anlamlılık düzeyi 0,01'in altında olan High ve dmi.DX değişkenleri ile oluşturulan modelin (Aylık LR<,01) DVM sınıflandırıcılarının performansı Tablo 121'de verilmektedir. Dairesel tabanlı



çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-4}$ ,  $c = 10^6$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 21 gözlemi (%30,88) doğru sınıflandırırken artış yönündeki 75 (%84,27) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 1 gözlem (%6,25) doğru sınıflandırılırken artış yönündeki 20 (%86,96) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %61,15'i doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %53,85'inin doğru sınıflandırıldığı gözlenmektedir.

Tablo 121 : Sınıflandırma Matrisi (Aylık LR<,01 Modeli)

	Eğitim Veri Seti					Test Veri Seti			
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	21	47	68	% 30,88	1	15	16	% 6,25
	Artış	14	75	89	% 84,27	3	20	23	% 86,96
	Σ	35	122	157	<b>% 61,15</b>	4	35	39	<b>% 53,85</b>

\*gamma= 0,0001 cost= 1000000

Lojistik regresyon analizinde anlamlılık düzeyi 0,01'in altında olan yukarıdaki değişkenlerin dalgacık dönüşümleri ile oluşturulan modelin (Aylık Dalgacık LR<,01) DVM sınıflandırıcılarının performansı Tablo 122'de verilmektedir. Dairesel tabanlı çekirdek kullanılan DVM sınıflandırıcısı için parametreler ( $\gamma = 10^{-3}$ ,  $c = 10^1$ ) 10-kat çapraz geçerleme ile elde edilmiştir. Sınıflandırıcının eğitim veri setinde, azalış yönündeki 49 gözlemi (%72,06) doğru sınıflandırırken artış yönündeki 79 (%88,76) gözlemi doğru sınıflandırdığı görülmektedir. Test veri setinde ise azalış yönündeki 7 gözlem (%43,75) doğru sınıflandırılırken artış yönündeki 17 (%73,91) gözlem doğru sınıflandırılmıştır. Eğitim veri setindeki toplam 157 gözlemin %81,53'ü doğru sınıflandırılırken test veri setindeki toplam 39 gözlemin %61,54'ünün doğru sınıflandırıldığı gözlenmektedir.

Tablo 122 : Sınıflandırma Matrisi (Aylık Dalgacık LR<,01 Modeli)

	Eğitim Veri Seti				Test Veri Seti				
	Tahmin Sınıfı		Σ	Doğruluk Oranı	Tahmin Sınıfı		Σ	Doğruluk Oranı	
	Azalış	Artış			Azalış	Artış			
Gerçek Sınıf	Azalış	49	19	68	% 72,06	7	9	16	% 43,75
	Artış	10	79	89	% 88,76	6	17	23	% 73,91
	Σ	59	98	157	<b>% 81,53</b>	13	26	39	<b>% 61,54</b>

\*gamma= 0.001 cost= 10

Aylık veri setinde tüm değişkenler ile oluşturulan modeldeki (Tablo 85) değişkenler günlük veri setinden seçilen tüm değişkenler ile aynı olduğu için DVM sınıflandırıcısı ile tahmin performanslarına burada tekrar yer verilmemiştir.

## 5. BÖLÜM: TARTIŞMALAR VE YORUMLAR

Çalışmada değişken seçimi yöntemleri yardımıyla oluşturulan günlük, haftalık ve aylık modellerin endeksin artış ve azalış yönlerini doğru sınıflandırma oranları toplu olarak Tablo 123'te verilmektedir.

Tablo 123 : Günlük Modeller Artış ve Azalış Sınıflandırma Oranları

Modeller	Ham Veriseti				Dalgacık Veriseti			
	Eğitim		Test		Eğitim		Test	
	Azalış	Artış	Azalış	Artış	Azalış	Artış	Azalış	Artış
TBA6	% 8,33	% 92,88	% 8,76	% 91,18	% 86,85	% 89,26	% 83,70	% 88,48
KOR<,50	% 12,76	% 88,80	% 12,90	% 88,24	% 40,04	% 75,53	% 32,12	% 70,10
KOR<,80	% 25,78	% 82,31	% 24,82	% 77,94	% 81,84	% 88,17	% 79,08	% 85,05
KOR>,80	% 32,68	% 76,11	% 33,09	% 71,57	% 94,14	% 95,18	% 90,27	% 91,18
KNN7	% 18,10	% 85,30	% 16,06	% 81,13	% 78,06	% 82,42	% 71,78	% 81,13
KNN20	% 26,24	% 80,53	% 22,87	% 79,17	% 96,42	% 96,73	% 88,81	% 91,42
PLS17	% 30,34	% 78,63	% 30,17	% 73,04	% 86,00	% 89,09	% 78,35	% 86,76
PLS20	% 28,26	% 78,86	% 26,03	% 74,51	% 90,36	% 91,38	% 86,13	% 89,71
RF6	% 19,60	% 85,53	% 16,06	% 82,60	% 58,27	% 82,77	% 54,26	% 82,60
LR<,20	% 37,70	% 74,38	% 38,93	% 71,81	% 91,02	% 91,56	% 83,21	% 90,20
LR<,10	% 30,08	% 79,04	% 27,01	% 77,94	% 86,46	% 89,32	% 83,45	% 90,44
LR<,05	% 37,11	% 73,46	% 33,58	% 69,12	% 86,20	% 88,57	% 84,67	% 89,22
LR<,01	% 18,49	% 83,97	% 17,52	% 82,60	% 83,79	% 85,70	% 82,97	% 89,95
TÜM55	% 40,43	% 72,89	% 41,61	% 72,30	% 94,60	% 95,92	% 90,75	% 90,93

Tablo 123 incelediğinde günlük modellerde DVM sınıflandırıcıları ham eğitim veri setinde azalışları en yüksek TÜM55 modelinde (%40,43) doğru

sınıflandırılırken artış yönü en yüksek TBA6 modelinde (%92,88) sınıflandırılmaktadır. Benzer şekilde ham test veri setinde azalış yönü en yüksek TÜR55 modelinde (%41,61) doğru sınıflandırılırken artış yönü en yüksek TBA6 modelinde (%91,18) sınıflandırıldığı görülmektedir. Yani günlük tüm modellerde azalış yönünü tahmin etmede düşük performans elde edilmektedir. Dalgacık eğitim veri setinde ise azalış yönü en yüksek TÜR55 modelinde (%94,60) doğru sınıflandırılırken artış yönü en yüksek KNN20 modelinde (%96,73) doğru sınıflandırılmıştır. Dalgacık test veri setinde, azalış yönü en yüksek TÜR55 modelinde (%90,75) doğru sınıflandırılırken artış yönü en yüksek KNN20 modelinde (%91,42) sınıflandırıldığı görülmektedir. Yani dalgacık veri setlerinde modellerin hem artış hem de azalış yönünü doğru tahmin etmede yüksek performans gösterdikleri görülmektedir.

Tablo 124 : Haftalık Modeller Artış ve Azalış Sınıflandırma Oranları

Modeller	Ham Veriseti				Dalgacık Veriseti			
	Eğitim		Test		Eğitim		Test	
	Azalış	Artış	Azalış	Artış	Azalış	Artış	Azalış	Artış
TBA6	% 0,00	% 99,22	% 0,00	% 98,91	% 79,52	% 88,86	% 75,32	% 88,04
KOR>,80	% 25,94	% 85,75	% 22,08	% 76,09	% 76,79	% 89,90	% 76,62	% 84,78
KNN20	% 15,02	% 93,52	% 10,39	% 85,87	% 74,06	% 90,16	% 74,03	% 89,13
PLS20	% 31,06	% 86,27	% 19,48	% 78,26	% 76,79	% 91,71	% 67,53	% 86,96
RF6	% 5,80	% 95,34	% 2,60	% 97,83	% 75,43	% 88,86	% 70,13	% 82,61
LR<,20	% 25,60	% 86,27	% 11,69	% 82,61	% 63,48	% 87,82	% 68,83	% 86,96
LR<,10	% 5,80	% 96,37	% 5,19	% 92,39	% 74,40	% 87,31	% 70,13	% 84,78
LR<,05	% 4,10	% 98,45	% 3,90	% 94,57	% 72,01	% 87,05	% 70,13	% 83,70
TÜR50	% 99,66	% 99,74	% 36,36	% 65,22	% 68,26	% 91,45	% 66,23	% 90,22

Tablo 124 incelediğinde haftalık modellerde DVM sınıflandırıcıları ham eğitim veri setinde azalışları en yüksek TÜR50 modelinde (%99,66) doğru sınıflandırılırken artış yönü en yüksek TÜR50 modelinde (%99,74) sınıflandırılmaktadır. Benzer şekilde ham test veri setinde azalış yönü en yüksek TÜR50 modelinde (%36,36) doğru sınıflandırılırken artış yönü en yüksek TBA6 modelinde (%98,91) sınıflandırıldığı görülmektedir. Yani haftalık test veri setinde

tüm modellerde azalış yönünü tahmin etmede düşük performans görülmektedir. Dalgacık eğitim veri setinde ise azalış yönü en yüksek TBA6 modelinde (%79,52) doğru sınıflandırılırken artış yönü en yüksek PLS20 modelinde (%91,71) doğru sınıflandırılmıştır. Dalgacık test veri setinde, azalış yönü en yüksek KOR>,80 modelinde (%76,62) doğru sınıflandırılırken artış yönü en yüksek TÜM50 modelinde (%90,22) sınıflandırıldığı görülmektedir. Günlük dalgacık modellerde olduğu gibi haftalık dalgacık veri setlerinde modellerin hem artış hem de azalış yönünü doğru tahmin etmede yüksek performanslar gösterdiği görülmektedir.

Tablo 125 : Aylık Modeller Artış ve Azalış Sınıflandırma Oranları

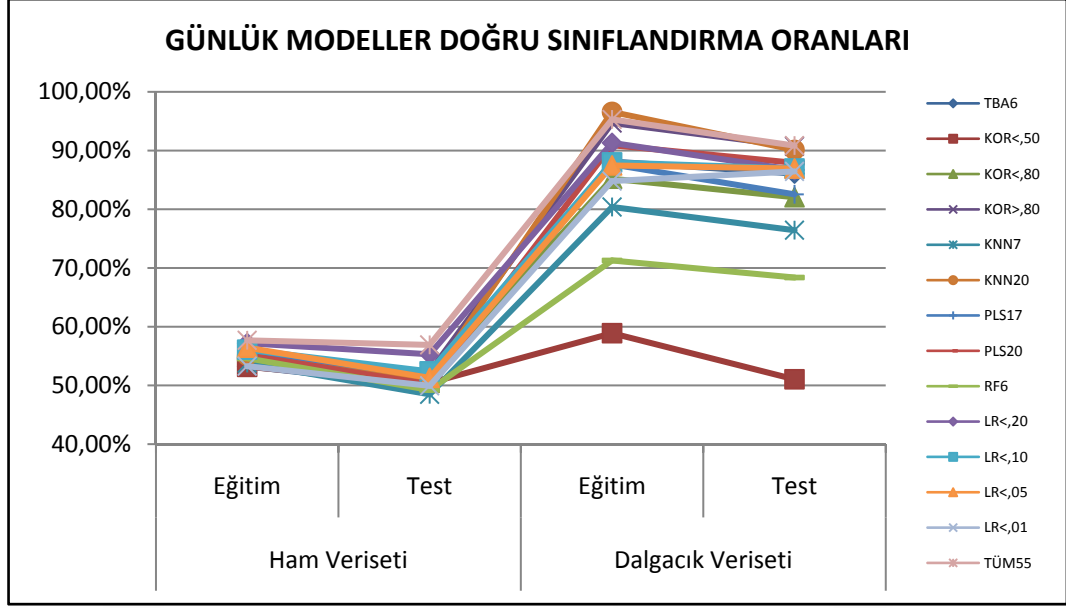
Modeller	Ham Veriseti				Dalgacık Veriseti			
	Eğitim		Test		Eğitim		Test	
	Azalış	Artış	Azalış	Artış	Azalış	Artış	Azalış	Artış
TBA5	% 39,71	% 85,39	% 18,75	% 73,91	% 91,18	% 96,63	% 25,00	% 82,61
KOR>,80	% 38,24	% 91,01	% 18,75	% 78,26	% 83,82	% 95,51	% 50,00	% 82,61
KNN20	% 48,53	% 85,39	% 43,75	% 60,87	% 79,41	% 98,88	% 50,00	% 86,96
PLS20	% 47,06	% 80,90	% 12,50	% 82,61	% 80,88	% 89,89	% 43,75	% 69,57
RF7	% 55,88	% 91,01	% 31,25	% 69,57	% 63,24	% 88,76	% 56,25	% 86,96
LR<,20	% 41,18	% 85,39	% 12,50	% 73,91	% 75,00	% 92,13	% 56,25	% 73,91
LR<,10	% 42,65	% 68,54	% 18,75	% 65,22	% 86,76	% 97,75	% 37,50	% 73,91
LR<,05	% 35,29	% 82,02	% 6,25	% 86,96	% 72,06	% 91,01	% 37,50	% 73,91
LR<,01	% 30,88	% 84,27	% 6,25	% 86,96	% 72,06	% 88,76	% 43,75	% 73,91
TÜM50	% 100,00	% 100,00	% 31,25	% 73,91	% 92,65	% 96,63	% 43,75	% 82,61

Tablo 125 incelediğinde aylık modellerde DVM sınıflandırıcıları ham eğitim veri setinde azalışları en yüksek TÜM50 modelinde (%100,00) doğru sınıflandırılırken artış yönü en yüksek TÜM50 modelinde (%100,00) sınıflandırılmaktadır. Benzer şekilde ham test veri setinde azalış yönü en yüksek KNN20 modelinde (%43,75) doğru sınıflandırılırken artış yönü en yüksek LR<,0,5 ve LR<,01 modellerinde (%86,96) doğru sınıflandırıldığı görülmektedir. Aylık test veri setinde tüm modellerde azalış yönünü tahmin etmede düşük performans görülmektedir. Dalgacık eğitim veri setinde ise azalış yönü en yüksek TÜM50 modelinde (%92,65) doğru sınıflandırılırken artış yönü en yüksek KNN20 modelinde

(%98,88) doğru sınıflandırılmıştır. Dalgacık test veri setinde, azalış yönü en yüksek RF7 ve LR>,20 modellerinde (%56,25) doğru sınıflandırılırken artış yönü en yüksek RF7 ve KNN20 modellerinde (%86,96) sınıflandırıldığı görülmektedir. Günlük ve haftalık dalgacık modellerde olduğu kadar yüksek olmasa da aylık dalgacık veri setlerinde modellerin hem artış hem de azalış yönünü doğru tahmin etmede ham veri setine göre daha yüksek performanslar gösterdiği görülmektedir.

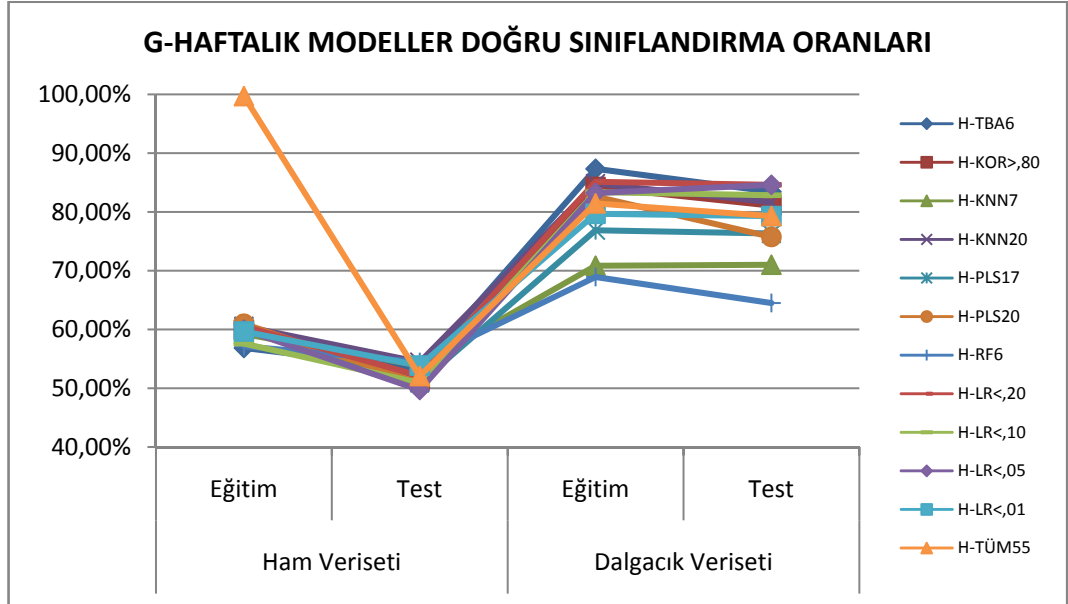
Emir (2013a)'in bulgularından hareketle, hem YSA hem de DVM yöntemlerinin azalışları tahmin etmede başarısız olduğunu belirtmektedir. Benzer şekilde Tayyar ve Tekin (2013)'in çalışmalarındaki tüm modellerde hem LR hem de DVM yöntemleri ile azalışların tahmininde artışların tahmininde elde edilen başarıya ulaşamadığı gösterilmektedir. Yakut (2012)'un çalışmasında ise YSA, DVM yöntemleri ve CS5.0 algoritması ile finansal başarısızlıkların finansal başarı tahminindeki kadar iyi tahmin edilemediği ve arada bariz performans farkı olduğu görülmektedir. Bu çalışmaların bulguları yapılan bu çalışma ile birlikte değerlendirildiğinde, hem eğitim hem de test veri setlerinde dalgacık dönüşümünün DVM sınıflandırıcılarının performansını artırdığı, azalışları doğru sınıflandırma konusunda artışları sınıflandırmadaki başarıyı yakalamasını sağladığı görülebilmektedir.

Modellerin doğru sınıflandırma performansları toplu olarak incelendiğinde, hem eğitim veri setlerinde hem de test veri setlerinde dalgacık modellerin performanslarının aynı modelin ham veri setleri üzerindeki performanslardan daha yüksek olduğu net bir şekilde görülebilmektedir. Şekil 44'deki günlük modellerin doğru sınıflandırma oranları incelendiğinde; ham test veri setindeki en yüksek doğru sınıflandırma oranına sahip TÜM55 modelinin (%56,90) dalgacık test veri setindeki doğru sınıflandırma oranı (%90,84) yaklaşık %60 artmıştır. Benzer şekilde tüm modellerin dalgacık veri setlerinde hem eğitim hem de test performanslarının arttığı en yüksek artışın da ham test performansı %50,92 iken dalgacık test performansı %77 artarak %90,11 olduğu KNN20 modelinde olduğu gözlenmiştir.



Şekil 44 : Günlük Modeller Doğru Sınıflandırma Oranları

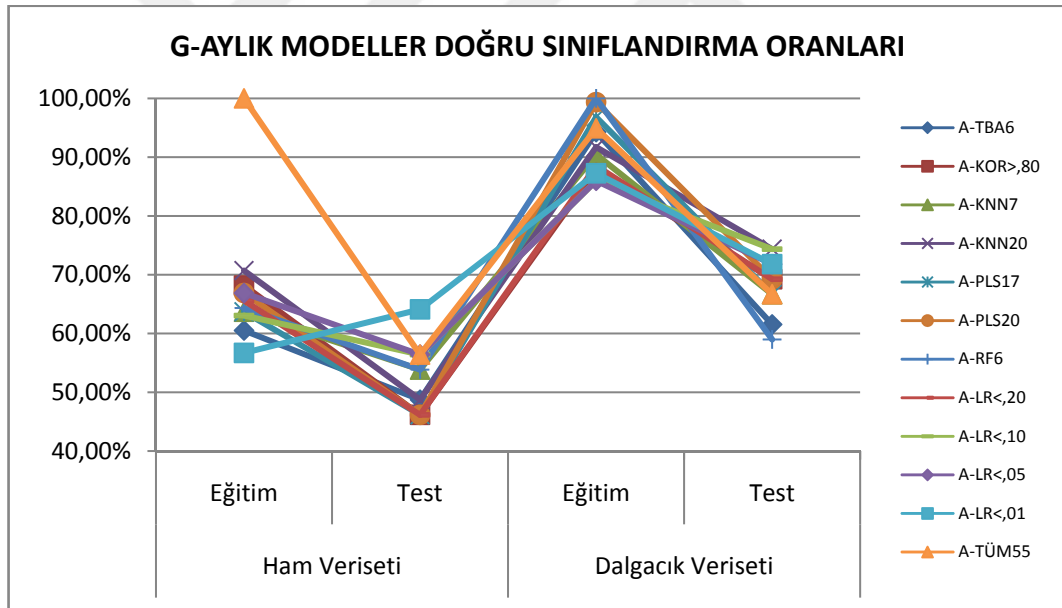
Şekil 45'teki günlük modellerde kullanılan değişkenler ile oluşturulan haftalık modellerin doğru sınıflandırma oranları incelendiğinde; ham test veri setindeki en yüksek doğru sınıflandırma oranına sahip H-KNN20 modelinin (%54,44) dalgacık test veri setindeki doğru sınıflandırma oranı (%81,66) yaklaşık %50 artmıştır.



Şekil 45 : G-Haftalık Modeller Doğru Sınıflandırma Oranları

Benzer şekilde tüm modellerin dalgacık veri setlerinde hem eğitim hem de test performanslarının arttığı, en yüksek artışın da ham test performansı %49,70 iken dalgacık test performansı %70 artarak %84,62 olan H-LR<,05 modelinde olduğu gözlenmiştir.

Şekil 46'daki günlük modellerde kullanılan değişkenler ile oluşturulan aylık modellerin doğru sınıflandırma oranları incelendiğinde; ham test veri setindeki en yüksek doğru sınıflandırma oranına sahip A-LR<,01 modelinin (%64,10) dalgacık test veri setindeki doğru sınıflandırma oranı (%71,79) yaklaşık %50 artmıştır. Benzer şekilde tüm modellerin dalgacık veri setlerinde hem eğitim hem de test performanslarının arttığı, en yüksek artışın da ham test performansı %48,72 iken dalgacık test performansı %53 artarak %74,36 olan A-KNN20 modelinde olduğu gözlenmiştir.

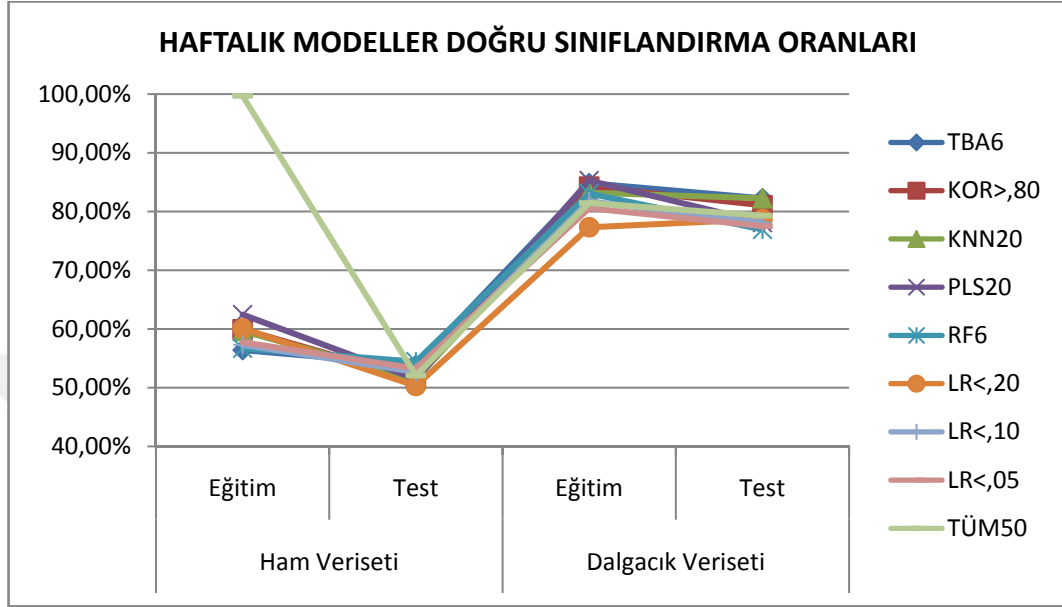


Şekil 46 : G-Aylık Modeller Doğru Sınıflandırma Oranları

Şekil 47'deki haftalık modellerin doğru sınıflandırma oranları incelendiğinde; ham test veri setindeki en yüksek doğru sınıflandırma oranına sahip RF6 modelinin (%54,44) dalgacık test veri setindeki doğru sınıflandırma oranı (%76,92) yaklaşık %41 artmıştır.



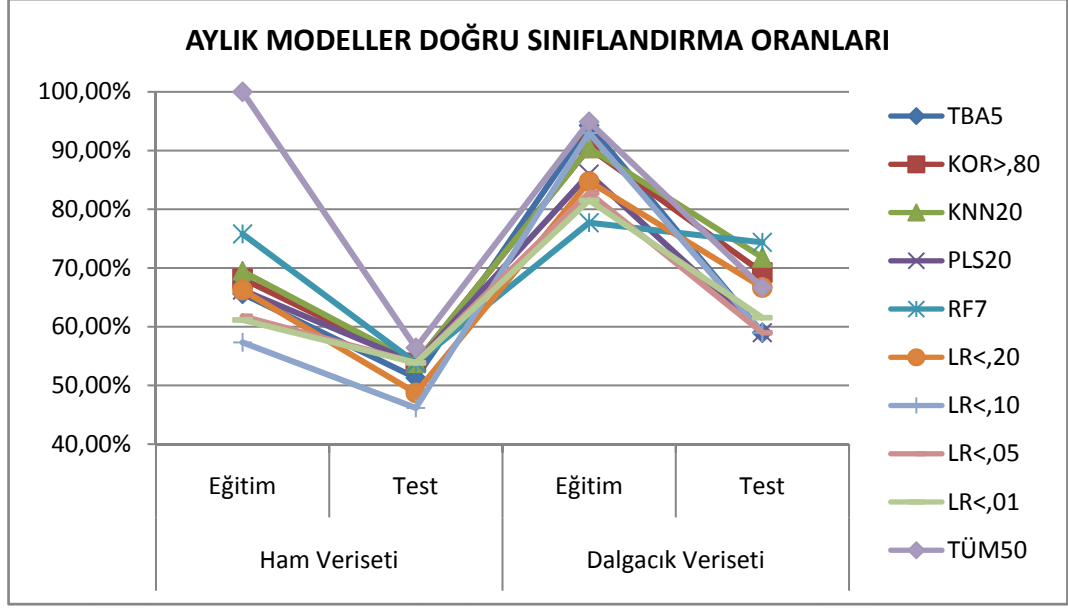
Benzer şekilde tüm modellerin dalgacık veri setlerinde hem eğitim hem de test performanslarının arttığı, en yüksek artışın da ham test performansı %51,48 iken dalgacık test performansı %60 artarak %82,25 olan KNN20 modelinde olduğu gözlenmiştir.



Şekil 47 : Haftalık Modeller Doğru Sınıflandırma Oranları

Şekil 48'deki aylık modellerin doğru sınıflandırma oranları incelendiğinde; ham test veri setindeki en yüksek doğru sınıflandırma oranına sahip TUM50 modelinin (%56,41) dalgacık test veri setindeki doğru sınıflandırma oranı (%66,67) yaklaşık %18 artmıştır.

Benzer şekilde tüm modellerin dalgacık veri setlerinde hem eğitim hem de test performanslarının arttığı, en yüksek artışın da ham test performansı %53,85 iken dalgacık test performansı %38 artarak %74,36 olan RF7 modelinde olduğu gözlenmiştir. Diğer periyotlarda oluşturulan modellere göre dalgacık tahmin kümesinde en düşük artış aylık modellerde gerçekleşmiştir.



Şekil 48 : Aylık Modeller Doğru Sınıflandırma Oranları

Günlük veri setinde değişken seçimi yöntemleriyle elde edilen değişkenler kullanılarak oluşturulan haftalık ve aylık modellerin performansları, haftalık ve aylık veri setlerinde değişken seçimi yöntemleriyle elde edilen değişkenlerden oluşan modellerin performansları ile karşılaştırıldığında; haftalık dalgacık test veri setinde en yüksek performans %82,25 (TBA6 ve KNN20) iken günlük modellerdeki değişkenlerden oluşan haftalık dalgacık test veri setinde en yüksek performans %84,62 (H-LR<,20 ve H-LR<,05) olduğu görülmektedir. Benzer şekilde aylık dalgacık test veri setinde en yüksek performans %74,36 (RF7) iken günlük modellerdeki değişkenlerden oluşan aylık dalgacık test veri setinde en yüksek performans %74,36 (A-KNN20 ve A-LR<,10) olduğu gözlenmiştir. Bu durum günlük veri setinde değişken seçme ile elde edilen değişkenler kullanılarak haftalık ve aylık modellerin oluşturulabileceği, haftalık ve aylık modeller için ayrıca değişken seçimine gerek olmadığı şeklinde yorumlanabilir.

Değişken seçme yöntemleri ile seçilen değişkenlerin oluşturulan günlük, haftalık ve aylık modellerde yer alma sıklıkları Tablo 126'da verilmektedir.

Tablo 126 : Değişkenlerin Modellerde Yer Alma Sıklıkları

Günlük Modeller		Haftalık Modeller		Aylık Modeller	
Değişken	f	Değişken	f	Değişken	f
Per	9	rsi	6	dmi.DX	8
dmi.DIn	9	atr.atr	6	Open	6
chaikinad	9	wma.200	5	High	6
bbands.HLC.pctB	9	stochOSC.slowD	5	wma.20	5
bbands.close.pctB	9	stochOSC.fastD	5	stochOSC.slowD	5
cmo	8	NIKKEI225	5	stochOSC.fastD	5
macd.signal	7	dmi.DX	5	obv	5
Faiz	7	mom	4	dmi.ADX	5
Euro	7	Euro	4	wpr	4
Dolar	7	Dolar	4	williamsad	4
DAX	7	dmi.DIp	4	sma.20	4
cci	7	DAX	4	Euro	4
wma.200	6	chaikinad	4	ema.200	4
wma.20	6	bbands.HLC.pctB	4	Dolar	4
roc	6	Volume	3	dmi.DIp	4
obv	6	stochOSC.fastK	3	cmf	4
NIKKEI225	6	SP500	3	bbands.HLC.pctB	4
macd.macd	6	smi.signal	3	bbands.HLC.ma	4
ema.200	6	sma.200	3	atr.atr	4
Volume	5	sar	3	stochOSC.fastK	3
rsi	5	roc	3	smi.smi	3
Paz	5	obv	3	sma.200	3
chaikinvolatility	5	macd.signal	3	sar	3
Car	5	macd.macd	3	rsi	3
atr.atr	5	dmi.DIn	3	NIKKEI225	3
Altın	5	cmo	3	Low	3
wpr	4	chaikinvolatility	3	ema.20	3
stochOSC.fastK	4	cci	3	dmi.DIn	3
smi.signal	4	bbands.HLC.ma	3	cmo	3
sar	4	bbands.close.pctB	3	chaikinad	3
High	4	wpr	2	cci	3
Cum	4	wma.20	2	bbands.HLC.dn	3
cmf	4	williamsad	2	bbands.close.pctB	3
smi.smi	3	smi.smi	2	Altın	3
sma.200	3	sma.20	2	wma.200	2
Sal	3	Open	2	Volume	2
NASDAQ	3	NASDAQ	2	SP500	2
mom	3	myAroon	2	smi.signal	2
ema.20	3	Low	2	myAroon	2
dmi.DX	3	High	2	mom	2
dmi.DIp	3	Faiz	2	mfi	2
dmi.ADX	3	ema.200	2	macd.signal	2
Close	3	ema.20	2	macd.macd	2
bbands.HLC.ma	3	dmi.ADX	2	Faiz	2
williamsad	2	cmf	2	DAX	2
stochOSC.slowD	2	Close	2	Close	2
stochOSC.fastD	2	bbands.HLC.up	2	bbands.HLC.up	2
SP500	2	bbands.HLC.dn	2	roc	1
sma.20	2	mfi	1	NASDAQ	1
Open	2	Altın	1	chaikinvolatility	1
myAroon	2	Sal	0	Sal	0
mfi	2	Per	0	Per	0
Low	2	Paz	0	Paz	0
bbands.HLC.up	2	Cum	0	Cum	0
bbands.HLC.dn	2	Car	0	Car	0

Tablo 126 incelendiğinde günlük modellerde değişken seçme yöntemleri tarafından en çok seçilen yani günlük modeller içinde en çok yer bulan değişkenler; gün değişkenlerinden Per ve teknik analiz değişkenlerinden dmi.DIn, chaikinad, bbands.HLC.pctB ve bbands.close.pctB olduğu görülmektedir. Burada özellikle dikkat çeken Aktaş ve Kozanoğlu (2007)'nin sonuçlarını destekler şekilde, Perşembe gün değişkeninin günlük veri setinde endeksin yön tahminindeki önemidir. Haftalık modeller içinde en çok yer bulan değişkenlerin ise teknik analiz değişkenlerinden rsi ve atr.atr olduğu görülmektedir. Aylık modeller içinde ise en çok yer bulan değişken yine teknik analiz değişkeni olan dmi.DX olarak ortaya çıkmıştır.

Günlük değişkenlerde hiçbir değişken seçme metodu ile seçilemeyen değişken yani tüm yöntemleri için de önemsiz görülen değişken olmamıştır. Bu durum çalışmada kullanılan değişkenlerin hepsinin endeksin yönünü tahmin konusunda önemi olduğunu göstermektedir. Haftalık modeller incelendiğinde mfi ve Altın değişkenleri değişken seçimi yöntemleri tarafından hiç bir modelde seçilmemiştir. Aylık veri setinde yapılan değişken seçimlerinde ise roc, chaikinvolatility ve NASDAQ değişkenleri tüm değişkenlerin kullanıldığı model dışında hiçbir modelde yer almamaktadırlar.

Tablo 127'de verilen modellerdeki değişkenlerin türleri ve sayıları incelendiğinde içerisinde teknik analiz değişkeni yer almayan hiçbir modelin olmadığı görülecektir. Bu durum toplam değişken sayısı içindeki teknik analiz değişkenlerinin oranından (~%76) kaynaklanabileceği gibi, endeksin yön tahmininde teknik analiz değişkenlerinin önemini de göstermektedir.

Değişken seçiminin amacının en az değişkenle en iyi sonucu elde edebilmek olduğu göz önüne alındığında günlük modeller içindeki  $LR < ,01$  modelinin sadece dört değişken (bbands.HLC.pctB, bbands.close.pctB, DAX ve wma.20) ile endeksin yönünü dalgacık test kümesinde %86,45 oranında doğru tahmin başarısı göstermesi dikkat çekmektedir. Haftalık modellerde sadece üç değişken (rsi, atr.atr ve wma.200) ile en az değişkenden oluşan  $LR < ,05$  modelinin %77,51 tahmin başarısı gösterdiği, aylık modellerde ise en az değişkenden oluşan  $LR < ,01$  modeli sadece iki değişken (High ve dmi.DX) ile dalgacık test veri setinde %61,54 doğru sınıflandırma performansına ulaştığı tespit edilmiştir.

Tablo 127 : Modellerdeki Değişken Sayıları

	Model	Teknik Analiz	Makro Ekonomik	Yurtdışı Borsa	Gün	Toplam	Tahmin Başarısı
Günlük	TBA6	5	-	1	-	6	% 86,08
	KOR<,50	1	-	-	5	6	% 51,04
	KOR<,80	12	3	2	5	22	% 82,05
	KOR>,80	30	1	2	-	33	% 90,72
	KNN7	3	4	-	-	7	% 76,43
	KNN20	14	4	1	1	20	% 90,11
	PLS17	11	3	1	2	17	% 82,54
	PLS20	13	4	1	2	20	% 87,91
	RF6	2	2	1	1	6	% 68,38
	LR<,20	21	1	1	2	25	% 86,69
	LR<,10	16	-	1	2	19	% 86,94
	LR<,05	13	-	1	1	15	% 86,94
	LR<,01	3	-	1	-	4	% 86,45
	TÜM55	42	4	4	5	55	% 90,84
Haftalık	TBA6	5	-	1	-	6	% 82,25
	KOR>,80	30	1	2	-	33	% 81,07
	KNN20	17	2	1	-	20	% 82,25
	PLS20	14	2	4	-	20	% 78,11
	RF6	5	1	-	-	6	% 76,92
	LR<,20	7	1	1	-	9	% 78,70
	LR<,10	3	-	1	-	4	% 78,11
	LR<,05	3	-	-	-	3	% 77,51
	TÜM50	42	4	4	-	50	% 79,29
	TBA5	4	-	1	-	5	% 58,97
	KOR>,80	34	1	2	-	37	% 69,23
	KNN20	18	2	-	-	20	% 71,79
	PLS20	16	4	-	-	20	% 58,97
	RF7	6	1	-	-	7	% 74,36
Aylık	LR<,20	11	1	1	-	13	% 66,67
	LR<,10	6	-	-	-	6	% 58,97
	LR<,05	3	-	-	-	3	% 58,97
	LR<,01	2	-	-	-	2	% 61,54
	TÜM50	42	4	4	-	50	% 66,67

Ayrıca Tablo 127'deki modellerin tahmin başarıları incelendiğinde; tüm periyotlardaki değişken seçimi ile oluşturulan modeller içindeki en yüksek tahmin performansı günlük modellerde, tüm değişkenlerin kullanıldığı TÜM55 modelinde

(%90,84) elde edilmiştir. Haftalık modellerde, altı değişkenden oluşan TBA6 ve yirmi değişkenden oluşan KNN20 modelleri en iyi tahmin performansı (%82,25) göstermişlerdir. Aylık modellerde ise tahmin performansları biraz daha düşük olsa da en yüksek tahmin performansı yedi değişkenden oluşan RF7 modelinde (%74,36) elde edilmiştir. Toplu olarak değerlendirildiğinde günlük modellerin haftalık ve aylık modellere göre daha başarılı olduğu, endeksin yön tahmininin günlük verilerle yapılmasının daha sağlıklı sonuçlar vereceği şeklinde saptama yapılabilir.

Çalışmada elde edilen günlük, haftalık ve aylık modellerin endeksin yönünü tahmin performanslarının literatürdeki XU100 endeksi yön tahmin çalışmaları ile kıyaslama yapılabilmesi için literatürde karşılaşılan çalışmalar Tablo 128'de verilmektedir.

Tablo 128 : XU100 Endeksi Yön Tahmini Çalışmaları

Çalışma	Aralık	Periyot	Değişkenler	Yöntem	Tahmin Oranı
Diler (2003)	1990-2003	Günlük	Teknik Gösterge (7)	YSA	% 60,81
Egeli vd. (2003)	Haz 2001-Şub 2003	Günlük	Önceki Gün Endeks Değerleri (1), Makroekonomik Göstergeler (2), Haftanın günleri (5)	YSA	% 82,00
Altay ve Satman (2005)	1997-2005	Günlük			% 57,80
		Haftalık	Önceki Gün Endeks Değerleri (7)	YSA	% 67,10
		Aylık			% 78,30
Yıldız vd. (2008)	1905 gün	Günlük	Önceki Gün Endeks Değerleri (3), Makroekonomik Göstergeler (2)	YSA	% 74,51
Akel ve Bayramoğlu (2008)	Ocak 1999 - Şubat 2001	Günlük	Makroekonomik Gösterge (5)	YSA	% 73,68
Kutlu ve Badur (2009)	2001-2006	Günlük	Önceki Gün endeks Değeri (1), Yurtdışı Endeksler (8), Makroekonomik Göstergeler (2), Haftanın günleri (5)	YSA	% 55,10
Kara vd. (2011)	1997-2007	Günlük	Teknik Gösterge (10)	YSA DVM	% 75,74 % 71,52
Özdemir vd. (2011)	1997-2010	Aylık	Makroekonomik Gösterge (8), Yurtdışı Endeks (3)	LR DVM	% 86,20
Emir (2013a)	2005-2011	Günlük	Teknik Göstergeler (11), Yurtdışı Endeksler (18), Makroekonomik Göstergeler (5)	YSA ve DVM	% 80,00 % 73,70
Tayyar ve Tekin (2013)	1995-2012	Günlük			% 52,42
		Haftalık	Teknik Gösterge (27)	LR	% 70,00
		Aylık		DVM	% 63,15

Günlük periyotta TUM55 modelinde elde edilen %90,84 tahmin performansı Tablo 128'deki en yüksek tahmin oranı Egeli vd. (2003)'nin YSA yöntemiyle elde ettikleri %82,00 ve Emir (2013a)'in DVM ile elde ettiği %73,70 tahmin performanslarından daha iyi olduğu; haftalık periyotta TBA6 ve KNN20 modelleri ile elde edilen %82,25 tahmin performansı tablodaki haftalık en yüksek tahmin oranı olan Tayyar ve Tekin (2013)'in %70,00 tahmin performansından daha iyi olduğu açıkça görülmektedir. Aylık periyotta ise RF7 modelinde erişilen %74,36 tahmin performansı tablodaki aylık en yüksek tahmin oranı olan Özdemir vd. (2011)'nin %86,20 olan aylık tahmin performansından biraz geri kalmaktadır. Bu durum dalgacık dönüşümü ile oluşturulan günlük ve haftalık modellerin tahmin gücünde bariz bir performans farkı olduğunu göstermektedir.

Tüm bu bulgular ışığında çalışmada elde edilen tüm modeller, modellerdeki değişkenler, modellerin ham veri seti üzerindeki tahmin performansları ve dalgacık veri setindeki tahmin performansları toplu olarak değerlendirilirse; dalgacık dönüşümünün DVM yönteminin tahmin performansı üzerinde pozitif yönde açık ve net bir etkisi olduğu görülmektedir. Ayrıca dalgacık dönüşümünün, DVM yönteminin endeksin azalış yönlü gözlemlerinde artış yönlü gözlemlerdeki tahmin performansını ortaya koyamama sorununu çözdüğü ve hem azalışlardaki hem de artışlardaki yapıyı benzer oranda öğrenmesini sağladığı tespit edilmiştir.

## 6. BÖLÜM: SONUÇ VE ÖNERİLER

Günümüzde, karar vermede kullanılan temel veri kaynağı sahip olunan bilgi ve veri setidir. Geleceği öngörebilmek ve en doğru kararı alabilmek için girdi ve çıktıları bilinen eldeki veriler detaylıca incelenmeli ve bu verilerin öznelikleri ortaya çıkarılmalıdır. Gelecekte ortaya çıkacak benzer girdi bilgisiyle, nasıl bir çıktı bilgisi elde edilebileceği konusunda makine öğrenmesi yöntemleri başarılı sonuçlar ortaya koyabilmekte, bu başarıyı birden fazla yöntemin birlikte kullanıldığı hibrit yöntemlerin daha da artırdığı literatürde görülmektedir.

Finansal piyasalarda da finansal varlıkların fiyatlarında belli bir periyot sonra nasıl bir değişimin gerçekleşeceğini, fiyatların yönünün ne olacağını öngörebilmek ve bu doğrultuda doğru pozisyon almak yatırımcıların pozitif getiri elde etmesi için önemlidir. Burada üzerine odaklanılması gereken konu tahminlerin doğruluk oranını artıracak model ve yöntemler olmaktadır.

Temelde sinyal analizinde kullanılan bir yöntem olan dalgacık dönüşümü ile veri madenciliği yöntemlerinden bir yapay öğrenme algoritması olan destek vektör makineleri yöntemlerinin birlikte finansal zaman serileri üzerinde kullanıldığı bu çalışmada, BIST 100 endeksinin yönü tahmin edebilecek günlük, haftalık ve aylık periyotlar için ham ve dalgacık tahmin modelleri oluşturularak bu modellerin tahmin performansları incelenmiştir.

Çalışmanın bağımsız değişkenleri ise, teknik göstergeler, yurtdışı borsa endeksleri, makroekonomik göstergeler ve haftanın günlerinden oluşmaktadır. Modeller belirlenirken değişken seçme yöntemlerinden (temel bileşen analizi, k-en yakın komşuluk, korelasyon, lojistik regresyon, kısmi en küçük kareler, rastgele orman algoritması) yararlanılmış, ayrıca tüm değişkenlerin yer aldığı modeller de kullanılmıştır. Çalışmada veri seti olarak, 15 Ekim 1997 tarihi ile 30 Aralık 2014



tarihleri arasındaki borsanın açık olduğu işgünlerini kapsayan 4296 günlük gözlem kullanılmış, bu veriler ile günlük, haftalık ve aylık veri setleri ayrı ayrı oluşturulmuştur.

Çalışmada elde edilen bulguların değerlendirilmesi ve literatürdeki diğer çalışmalarla karşılaştırılması bir önceki bölümde yapılmış olup, bu çalışmanın literatüre katkıları ve sonuçları aşağıdaki şekilde ifade edilebilir.

- Dalgacık dönüşümünün incelenen tüm modellerin hem eğitim hem de test veri setlerinde, destek vektör makinelerin tahmin performansını artırdığı tespit edilmiştir.
- Dalgacık dönüşümünün yön tahmininde ham veri setinde karşılaşılan azalış yönündeki sınıflandırma hatalarını azaltarak, artış ve azalışların birbirine yakın oranlarda sınıflandırılmasını sağladığı sonucuna ulaşılmıştır.
- Dalgacık dönüşümü ve destek vektör makineleri yöntemlerinin birlikte BIST 100 endeksinin hareket yönü tahmininde başarılı olduğu sonucuna ulaşılmıştır.
- Günlük veri setinde değişken seçme yöntemleri aracılığıyla seçilen değişkenlerden oluşan modellerin tahmin performanslarının haftalık ve aylık veri setlerinde değişken seçme yöntemleriyle seçilen değişkenlerden oluşan modellerin tahmin performanslarından daha iyi olduğu görülmüştür. Buradan yola çıkılarak, günlük veri setinde değişken seçimi ile elde edilen değişkenler kullanılarak haftalık ve aylık modellerin oluşturulabileceği, haftalık ve aylık veri setlerinde ayrıca değişken seçimi yapılmasına gerek olmadığı sonucu literatüre önemli bir katkı olarak değerlendirilebilir.
- Değişken seçmek için, temel bileşen analizi, k-en yakın komşuluk, korelasyon, lojistik regresyon, kısmi en küçük kareler ve rastgele orman algoritması yöntemlerinden yararlanılmış ve değişken seçme konusunda birbirlerine yakın performans gösterdikleri görülmüştür.

- BIST 100 endeksinin yönünü en iyi tahmin eden modelin, günlük veri setindeki tüm teknik gösterge, yurtdışı borsa endeksleri, makroekonomik göstergeler ve gün değişkenlerinden oluşan Dalgacık TUM55 modeli (%90,84) olduğu, bu modelin ham veri setindeki (%56,90) tahmin oranını yaklaşık %60 artırdığı sonucuna ulaşılmıştır.
- BIST 100 endeksinin yönünü en iyi tahmin eden haftalık modellerin, lojistik regresyon ile seçilen H-LR<,20 ve H-LR<,05 dalgacık modellerinin (%84,62) olduğu sonucuna ulaşılmıştır.
- BIST 100 endeksinin yönünü en iyi tahmin eden aylık modellerin, KNN ile seçilen A-KNN20 ve LR ile seçilen A-LR<,10 dalgacık modellerinin (%74,36) olduğu sonucuna ulaşılmıştır.
- BIST 100 endeksinin yönünü en az değişkenle tahmin eden günlük modelin, LR analizinde anlamlılık düzeyi 0,01'in altında olan dört değişken (bbands.HLC.pctB, bbands.close.pctB, DAX ve wma.20) ile oluşturulan dalgacık LR<,01 modeli (%86,45) olduğu görülmüştür.
- BIST 100 endeksinin yönünü en az değişkenle tahmin eden haftalık modelin, LR analizinde anlamlılık düzeyi 0,05'in altında olan üç değişken (rsi, atr.atr ve wma.200) ile oluşturulan dalgacık LR<,05 modeli (%77,51) olduğu görülmüştür.
- BIST 100 endeksinin yönünü en az değişkenle tahmin eden aylık modelin, LR analizinde anlamlılık düzeyi 0,01'in altında olan iki değişken (High ve dmi.DX) ile oluşturulan dalgacık LR<,01 modeli (%61,54) olduğu görülmüştür.
- Günlük değişkenlerde hiçbir değişken seçme yöntemi tarafından seçilemeyen değişken, diğer bir ifadeyle tüm yöntemler için de önemsiz görülen değişken olmamıştır. Bu durum çalışmada kullanılan değişkenlerin hepsinin BIST 100 endeksinin yön tahmininde önemli olduğu sonucunu ortaya koymuştur.
- BIST 100 endeksinin yön tahmininde daha önemli yani Günlük modeller içinde en çok yer bulan değişkenler; gün değişkenlerinden Per ve teknik

analiz deęişkenlerinden dmi.DIn, chaikinad, bbands.HLC.pctB ve bbands.close.pctB olduęu sonucu ortaya çıkmıştır. Buradan hareketle, Perşembe gününün BIST 100 endeksin yön tahminindeki önemli yer tuttuęu sonucu elde edilmiştir.

- BIST 100 endeksinin yön tahmininde haftalık modeller içinde en çok yer bulan deęişkenlerin ise teknik analiz deęişkenlerinden rsi ve atr.atr olduęu görölmektedir. Aylık modeller içinde ise en çok yer bulan deęişken yine teknik analiz deęişkeni olan dmi.DX olarak ortaya çıkmıştır.
- Deęişken seçme yöntemleri yardımıyla oluşturulan tüm modellerde teknik analiz deęişkenlerinin ağırlıklı olarak yer alması, teknik analiz deęişkenlerinin BIST 100 endeksinin yön tahmininde çok önemli olduęunu göstermektedir.

Bu çalışmada elde edilen modellerin ve çalışmanın sonuçlarının, finans piyasası yatırımcıları için yatırımlarını yönlendirme kararlarına destek olacak sağlıklı tahmin verisi sağlayacağı düşünülmektedir. Ayrıca, çalışmada ulaşılan en temel sonuçlardan biri olan, dalgacık dönüşümünün destek vektör makinelere tahmin performansına pozitif etkisi araştırmacılar için yol gösterici olacaktır.

Daha sonra yapılacak çalışmalarda; dalgacık dönüşümü dięer makine öğrenmesi yöntemleriyle birlikte kullanılıp, o yöntemlerin tahmin performanslarını nasıl etkileyeceęi incelenerek bu çalışmadaki dalgacık dönüşümünün etkisinin spesifik mi yoksa global mi olduęu incelenebilir.

Ayrıca, bu çalışmadan farklı olarak, farklı ana dalgacık kullanma, farklı seviyelerden ayırıştırma, daha kısa zaman periyodu kullanma, haftalık ve aylık veri setleri oluştururken haftanın veya ayın son işgünü gözlem deęerlerinin kullanılması, eğitim ve test veri setleri oluştururken rastgele seçim yapmadan belirli periyotları eğitim ve test için ayırma gibi deęişik yaklaşımlar farklı sonuçlar ortaya koyabilecektir. Yine hafta ve ay deęişkenleri oluşturularak, hangi haftaların veya hangi ayların öne çıktığı incelenebilecektir.

Modelleri oluřturan deęiřkenlerin seęimi dalgacık veri setinde de yapılarak, elde edilen modellerde sınıflandırıcıların tahmin performansının ölçölmesiyle dalgacık dönüřümünün etkisinin daha net görölceęi düşünölmektedir.













## KAYNAKÇA

- Abe, S. (2005). *Support vector machines for pattern classification*. London: Springer.
- Abramovich, F., Bailey, T. C. ve Sapatinas, T. (2000). Wavelet analysis and its statistical applications. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(1), 1–29.
- Addison, P. S. (2002). *The illustrated wavelet transform handbook: Introductory theory and applications in science, Engineering, Medicine and Finance*. Institute of Physics Publishing.
- Aizerman, A., Braverman, E. M. ve Rozoner, L. (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and remote control*, 25, 821–837.
- Akel, V. ve Bayramođlu, M. F. (2008). Kriz dönemlerinde yapay sinir ađları ile finansal öngörüde bulunma: İMKB 100 endeksi örneđi. *International Symposium on International Capital Flows and Emerging Markets* içinde (ss. 24–28).
- Aktan, B., Mabrouk, A. B., Ozturk, M. ve Rhaiem, N. (2009). Wavelet-Based systematic risk estimation an application on istanbul stock exchange. *International Research Journal of Finance and Economics*, 23(3), 33–45.
- Aktaş, H. ve Kozanođlu, M. (2007). Haftanın günleri etkisinin İstanbul menkul kıymetler borsası'nda GARCH modeli ile test edilmesi. *Finansal Politik & Ekonomik Yorumlar Dergisi*, 44(514), 37–45.
- Alexandridis, A. K. ve Zapranis, A. D. (2014). *Wavelet neural networks: With applications in financial engineering, chaos, and classification*. New Jersey: John Wiley & Sons.
- Alpaydın, E. (2013). *Yapay öğrenme*. İstanbul: Boğaziçi Üniversitesi Yayınları.
- Alpaydın, E. (2010). *Introduction To Machine Learning*. London: MIT press.
- Altay, E. ve Satman, M. H. (2005). Stock market forecasting: Artificial neural network and linear regression comparison in an emerging market. *Journal of Financial Management & Analysis*, 18(2), 18–33.
- Appel, G. (2005). *Technical analysis: Power tools for active investors*. New Jersey: FT Press.
- Avcı, E. (2009). Yapay sinir ađları modelleri ile hisse senedi getiri tahminleri. *Marmara Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 26(1), 443–461.



- Aygören, H., Sarıtaş, H. ve Moralı, T. (2012). IMKB 100 endeksinin yapay sinir ağları ve newton nümerik arama modelleri ile tahmini. *Alanya İisletme Fakültesi Dergisi*, 4(1), 73–88.
- Blau, W. (1991). Stochastic momentum. *Technical Analysis of Stocks and Commodities*, 11(1), 26–32.
- Boggess, A. ve Narcowich, F. J. (2009). *A first course in wavelets with Fourier analysis*. New Jersey: John Wiley & Sons.
- Boser, B. E., Guyon, I. M. ve Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory* içinde (ss. 144–152), USA: ACM.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Brigham, E. ve Houston, J. (2007). *Fundamentals of financial management*. USA: Thomson Corporation.
- Brooks, J. C. (2006). *Mastering Technical Analysis Using the Tools of Technical Analysis for Profitable Trading*. USA: McGraw-Hill.
- Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121–167.
- Campbell, C. ve Ying, Y. (2011). Learning with support vector machines. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 5(1), 1–95.
- Chai, J., Du, J., Lai, K. K. ve Lee, Y. P. (2015). A hybrid least square support vector machine model with parameters optimization for stock forecasting. *Mathematical Problems in Engineering*, 2015, 1–7.
- Chande, T. S. (1995). A time price oscillator. *Technical Analysis of Stocks & Commodities*, 13, 369–374.
- Chande, T. S. ve Kroll, S. (1994). *The new technical trader: Boost your profit by plugging into the latest indicators*. New Jersey: John Wiley & Sons Inc.
- Cherkassky, V. ve Mulier, F. M. (2007). *Learning from data: Concepts, theory, and methods*. New Jersey: John Wiley & Sons.
- Colby, R. W. (2003). *The Encyclopedia of Technical Market Indicators*. USA: McGraw-Hill.
- Cortes, C. ve Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273–297.
- Cover, T. M. ve Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27.

- Çankaya, I. ve Vatansever, F. (2002). Fourier ve dalgacık dönüşümünün karşılaştırılması. *Fen Bilimleri Enstitüsü Dergisi*, 6(3), 14–24.
- Çifter, A. (2010). *Dalgacık bazlı uç değer teorisi ile parametrik olmayan volatilité modellenmesi*. Yayınlanmamış doktora tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on pure and applied mathematics*, 41(7), 909–996.
- Di Lena, P. ve Margara, L. (2010). Optimal global alignment of signals by maximization of Pearson correlation. *Information Processing Letters*, 110(16), 679–686.
- Diler, A. İ. (2003). İMKB Ulusal-100 endeksinin yönünün yapay sinir ağıları hata geriye yayma yöntemi ile tahmin edilmesi. *İMKB Dergisi*, 7(25-26), 65–81.
- DiSario, R., Saraoglu, H., McCarthy, J. ve Li, H. (2008). Long memory in the volatility of an emerging equity market: The case of Turkey. *Journal of International Financial Markets, Institutions and Money*, 18(4), 305–312.
- Dougherty, G. (2013). *Pattern Recognition And Classification: An Introduction*. New York: Springer Science & Business Media.
- Duda, R. O., Hart, P. E. ve Stork, D. G. (2012). *Pattern Classification*. New Jersey: John Wiley & Sons.
- Durukan, M. B. (1999). İstanbul menkul kıymetler borsasında makroekonomik değişkenlerin hisse senedi fiyatlarına etkisi. *İMKB Dergisi*, 3(11), 20–41.
- Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association*, 78(382), 316–331.
- Egeli, B., Ozturan, and M. ve Badur, B. (2003). Stock market prediction using artificial neural networks. *Decision Support Systems*, 22, 171–185.
- Emir, Ş. (2013a). *Yapay sinir ağıları ve destek vektör makineleri yöntemlerinin sınıflandırma performanslarının karşılaştırılması: Borsa Endeks yönünün tahmini üzerine bir uygulama*. Yayınlanmamış doktora tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Emir, Ş. (2013b). Predicting the Istanbul stock exchange index return using technical indicators: A comparative study. *International Journal of Finance & Banking Studies*, 2(3), 111–117.
- Erdemir, A. (2012). *Analyzing The Forecast Performance of SP500 Index Options Implied Volatility*. Yayınlanmamış yüksek lisans tezi, Bilkent University, Economics and Social Sciences Institute, Ankara.

- Fang, H. ve Loo, J. C. (1996). Foreign exchange risk and common stock returns: a note on international evidence. *Journal of Business Finance & Accounting*, 23(3), 473–480.
- Fugal, D. L. (2009). *Conceptual wavelets in digital signal processing: an in-depth, practical approach for the non-mathematician*. USA: Space & Signals Technical Pub.
- Gallegati, M. (2005). A wavelet analysis of MENA stock markets. *Department of economics*. [https://www.researchgate.net/profile/Marco\\_Gallegati/publication/23744118\\_A\\_Wavelet\\_Analysis\\_of\\_MENA\\_Stock\\_Markets/links/09e4151122459e2fab000000.pdf](https://www.researchgate.net/profile/Marco_Gallegati/publication/23744118_A_Wavelet_Analysis_of_MENA_Stock_Markets/links/09e4151122459e2fab000000.pdf) adresinden erişilmiştir.
- Gençay, R., Selçuk, F. ve Whitcher, B. (2001). Scaling properties of foreign exchange volatility. *Physica A: Statistical Mechanics and its Applications*, 289(1), 249–266.
- Gençay, R., Selçuk, F. ve Whitcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance*, 24, 55–70.
- Gençay, R., Selçuk, F. ve Whitcher, B. J. (2002). *An introduction to wavelets and other filtering methods in finance and economics*. USA: Academic press.
- Genuer, R., Poggi, J.-M. ve Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236.
- Gorunescu, F. (2011). *Data Mining: Concepts, Models and Techniques*. New York: Springer Science & Business Media.
- Goswami, J. C. ve Chan, A. K. (2011). *Fundamentals of wavelets: Theory, algorithms, and applications* (C. 233). New Jersey: John Wiley & Sons.
- Graps, A. (1995). An introduction to wavelets. *Computational Science & Engineering, IEEE*, 2(2), 50–61.
- Gröchenig, K. (2001). *Foundations of time-frequency analysis*. USA: Springer.
- Guyon, I., Gunn, S., Nikravesh, M. ve Zadeh, L. (2006). *Feature Extraction, Foundations and Applications*. USA: Springer.
- Güreşen, E. ve Kayakutlu, G. (2008). Forecasting Stock Exchange Movements Using Artificial Neural Network Models And Hybrid Models. *Intelligent Information Processing IV* içinde (ss. 129–137). USA: Springer.
- Gürsakal, S. (2009). *Finansal zaman serileri analizine frekans boyutu yaklaşımı: Dalgacıklar yöntemi ve İMKB’de bir uygulama*. Yayınlanmamış doktora tezi, Uludağ Üniversitesi, Sosyal Bilimler Enstitüsü, Bursa.
- Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3), 331–371.

- Hamel, L. H. (2011). *Knowledge discovery with support vector machines* (C. 3). New Jersey: John Wiley & Sons.
- Haykin, S. (2009). *Neural networks and learning machines* (3. edition). New Jersey: Pearson Education Inc.
- Haznedaroğlu, F. (2009). *Makro ekonomik ve finansal verilerin İMKB endeksleri üzerinde etkisini belirleyen bir tahmin sistemi geliştirilmesi*. Yayınlanmamış doktora tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Höskuldsson, A. (2001). Variable and subset selection in PLS regression. *Chemometrics and intelligent laboratory systems*, 55(1), 23–38.
- Hsu, C.-W., Chang, C.-C. ve Lin, C.-J. (2003). A practical guide to support vector classification. <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> adresinden erişilmiştir.
- Huang, S.-C. (2011). Forecasting stock indices with wavelet domain kernel partial least square regressions. *Applied Soft Computing*, 11(8), 5433–5443.
- Huang, S.-C. ve Wu, T.-K. (2008a). Combining wavelet-based feature extractions with relevance vector machines for stock index forecasting. *Expert Systems*, 25(2), 133–149.
- Huang, S.-C. ve Wu, T.-K. (2008b). Integrating GA-based time-scale feature extractions with SVMs for stock index forecasting. *Expert Systems with Applications*, 35(4), 2080–2088.
- Huang, W., Nakamori, Y. ve Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10), 2513–2522.
- İMKB. (2010). *İMKB Endeksleri*. [www.borsaistanbul.com/data/kilavuzlar/İMKB%20Endeksleri.pdf](http://www.borsaistanbul.com/data/kilavuzlar/İMKB%20Endeksleri.pdf) adresinden erişilmiştir.
- In, F. ve Kim, S. (2012). *An introduction to wavelet theory in finance: A wavelet multiscale approach*. Singapore: World Scientific.
- Kaiser, G. (2011). *A Friendly Guide to Wavelets*. USA: Springer.
- Kao, L.-J., Chiu, C.-C., Lu, C.-J. ve Chang, C.-H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems*, 54(3), 1228–1244.
- Kara, Y., Acar Boyacioglu, M. ve Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul stock exchange. *Expert systems with Applications*, 38(5), 5311–5319.

- Karaatlı, M. A. (2003). *Yapay Sinir Ağları Yöntemi İle Tahmin (İMKB’de Bir Uygulama)*. Yayınlanmamış yüksek lisans tezi, Süleyman Demirel Üniversitesi, Sosyal Bilimler Enstitüsü, Isparta.
- Karaatlı, M., Güngör, İ., Demir, Y. ve Kalaycı, Ş. (2005). Hisse senedi fiyat hareketlerinin yapay sinir ağları yöntemi ile tahmin edilmesi. *Balıkesir Üniversitesi İİBF Dergisi*, 2(1), 22–48.
- Kim, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1), 307–319.
- Kim, S. ve In, F. (2007). On the relationship between changes in stock prices and bond yields in the G7 countries: Wavelet analysis. *Journal of International Financial Markets, Institutions and Money*, 17(2), 167–179.
- Kim, S. ve In, F. H. (2003). The relationship between financial variables and real economic activity: evidence from spectral and wavelet analyses. *Studies in Nonlinear Dynamics & Econometrics*, 7(4), 1–16.
- Kristin L. Sainani. (2014). Logistic Regression. *PM&R*, 6(12), 1157–1162.
- Kubinyi, H. (1996). Evolutionary variable selection in regression and PLS analyses. *Journal of Chemometrics*, 10(2), 119–133.
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1–26.
- Kutlu, B. ve Badur, B. (2009). Yapay sinir ağları ile borsa endeksi tahmini. *Yönetim*, 20(63), 25–40.
- Lahmiri, S. (2014). Wavelet low-and high-frequency components as features for predicting stock prices with backpropagation neural networks. *Journal of King Saud University-Computer and Information Sciences*, 26(2), 218–227.
- Lambert, D. R. (1983). Commodity channel index: Tool for trading cyclic trends. *Technical Analysis of Stocks & Commodities*, 1(5), 120–122.
- Larson, M. (2007). *12 Simple Technical Indicators that Really Work*. USA: Marketplace Books Inc.
- Lee, H. S. (2002). International transmission of stock market movements: A wavelet analysis on MENA stock market. *Economic Research Forum, ERF Eighth Annual Conference, Cairo, Egypt* içinde .
- Lee, S. (2005). Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing*, 26(7), 1477–1491.

- Li, J., Shi, Z. ve Li, X. (2006). Genetic programming with wavelet-based indicators for financial forecasting. *Transactions of the Institute of Measurement and Control*, 28(3), 285–297.
- Li, S. (2009). *Random KNN modeling and variable selection for high dimensional data*. Yayınlanmamış doktora tezi, West Virginia University, USA.
- Liaw, A. ve Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18–22.
- Mallat, S. (2008). *A wavelet tour of signal processing: The sparse way*. USA: Academic press.
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.
- Mangasarian, O. L. (1965). Linear and nonlinear separation of patterns by linear programming. *Operations research*, 13(3), 444–452.
- Mevik, B.-H., Wehrens, R. ve Liland, K. H. (2013). *pls: Partial least squares and principal component regression*. <http://CRAN.R-project.org/package=pls> adresinden erişilmiştir.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. ve Leisch, F. (2014). *Misc Functions of the Department of Statistics (e1071)*, TU Wien. <http://CRAN.R-project.org/package=e1071> adresinden erişilmiştir.
- Meyer, Y. (1993). *Wavelets: Algorithms and applications*. Philadelphia : SIAM.
- Michie, D., Spiegelhalter, D. J. ve Taylor, C. C. (1994). *Machine learning, neural and statistical classification*. USA: Overseas Press.
- Morettin, P. A. (1996). From Fourier to wavelet analysis of time series. *Computational Statistics* içinde (ss. 111–122), Physica-Verlag HD.
- Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. USA: NYIF.
- Narcouzi, C. (2000). Chaikin's Money Flow. *Technical Analysis Of Stocks And Commodities*, 18(8), 42–47.
- Nason, G. P. (2010). *Wavelet methods in statistics with R*. USA: Springer.
- Nason, G. P. ve Von Sachs, R. (1999). Wavelets in time-series analysis. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760), 2511–2526.

- Ng, K. S. (2013). *A simple explanation of partial least squares*. <http://users.rsise.anu.edu.au/~kee/pls.pdf> adresinden erişilmiştir.
- Norsworthy, J. R., Li, D. ve Gorener, R. (2000). Wavelet-based analysis of time series: an export from engineering to finance. *Engineering Management Society, Proceedings of the 2000 IEEE* içinde (ss. 126–132).
- Ortega, L. ve Khashanah, K. (2014). A Neuro-wavelet model for the short-term forecasting of high-frequency time series of stock returns. *Journal of Forecasting*, 33(2), 134–146.
- Ozun, A. ve Cifter, A. (2007). Multi-scale causality between energy consumption and GNP in emerging markets: evidence from Turkey. *Investment management and financial innovations*, 4(2), 61–70.
- Özalp, A. ve Anagün, S. (2003). Yapay sinir ağı performansına etki eden faktörlerin analizinde Taguchi yöntemi: Hisse senedi fiyat tahmini uygulaması. *İstatistik Araştırma Dergisi*, 2(1), 29–45.
- Özdemir, A. K., Tolun, S. ve Demirci, E. (2011). Endeks getirisi yönünün ikili sınıflandırma yöntemiyle tahmin edilmesi: İMKB-100 endeksi örneği. *Niğde Üniversitesi İİBF Dergisi*, 4(2), 45–59.
- Özkan, Y. (2008). *Veri Madenciliği Yöntemleri*. İstanbul: Papatya Yayıncılık Eğitim.
- Özün, A. ve Çifter, A. (2006). Bankaların hisse senedi getirilerinde faiz oranı riski: dalgalıklar analizi ile Türk bankacılık sektörü üzerine bir uygulama. *Bankacılar Dergisi*, 59, 3–15.
- Pagano, M. (1996). Logistic Regression. *Nutrition*, 12(2), 135.
- Person, J. L. (2004). *A complete guide to technical trading tactics*. New Jersey: John Wiley & Sons.
- Polikar, R. (t.y.). The Wavelet Tutorial Part I. 7–12 2014 tarihinde <http://users.rowan.edu/~polikar/WAVELETS/WTpart1.html> adresinden erişilmiştir.
- Priestley, M. (1996). Wavelets And Time-Dependent Spectral Analysis. *Journal of Time Series Analysis*, 17(1), 85–103.
- Radunovic, D. P. (2010). *Wavelets: from math to practice*. USA: Springer.
- Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760), 2593–2606.
- Ramsey, J. B. ve Lampart, C. (1998). Decomposition of economic relationships by timescale using wavelets. *Macroeconomic Dynamics*, 2(1), 49–71.

- Ramsey, J. B. ve Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance*, 4(4), 341–372.
- Revelle, W. (2015). *psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois. <http://CRAN.R-project.org/package=psych> adresinden erişilmiştir.
- Ryan, J. A. ve Ulrich, J. M. (2014). *xts: eXtensible Time Series*. <http://CRAN.R-project.org/package=xts> adresinden erişilmiştir.
- Schlüter, S. ve Deuschle, C. (2010). *Using wavelets for time series forecasting: Does it pay off?*, IWQW discussion paper series, No. 04/2010.
- Schölkopf, B. ve Smola, A. J. (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. London: MIT press.
- Selçuk, F. (2005). Dalgacıklar yeni bir analiz yöntemi . *Bilkent Dergisi*, 3, 12–15.
- Seler, İ. T. (1996). Haftanın günleri: İMKB'ye etkileri üzerine bir inceleme. *Sermaye Piyasası ve İMKB Üzerine Çalışmalar, İşletme ve Finans Yayınları*, (4), 147–168.
- Siegel, J. G., Shim, J. K., Qureshi, A. A. ve Brauchler, J. (2014). *International encyclopedia of technical analysis*. USA: Glenlake Publishing Company.
- Stoean, C. (2014). *Support vector machines and evolutionary algorithms for classification: Single or together?* (C. 69). USA: Springer.
- Sui, X., Hu, Q., Yu, D., Xie, Z. ve Qi, Z. (2007). A Hybrid Method for Forecasting Stock Market Trend Using Soft-Thresholding De-noise Model and SVM. *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing* içinde (ss. 387–394). USA: Springer.
- Tang, L.-B., Tang, L.-X. ve Sheng, H.-Y. (2009). Forecasting volatility based on wavelet support vector machine. *Expert Systems with Applications*, 36(2), 2901–2909.
- Tayyar, N. ve Tekin, S. (2013). İMKB-100 endeksinin destek vektör makineleri ile günlük, haftalık ve aylık veriler kullanarak tahmin edilmesi. *Abant İzzet Baysal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 13(1), 189–217.
- Teixeira, L. A. ve De Oliveira, A. L. I. (2010). A method for automatic stock trading combining technical analysis and nearest neighbor classification. *Expert systems with applications*, 37(10), 6885–6890.
- Tekin, S. (2013). *Destek vektör makineleri yöntemi ile İMKB 100 endeksi hareket yönü tahmini*. Yayınlanmamış yüksek lisans tezi, Uşak Üniversitesi, Sosyal Bilimler Enstitüsü, Uşak.
- Thomsett, M. (2012). *Technical analysis of stock trends explained: An easy-to-understand system for successful trading*. USA: Ethan Hathaway Co Ltd.



- Torgo, L. (2010). *Data mining with R, learning with case studies*. Chapman and Hall/CRC. <http://www.dcc.fc.up.pt/~ltorgo/DataMiningWithR> adresinden erişilmiştir.
- Turhan Korkmaz, A. C. (2012). *Sermaye piyasası ve menkul değer analizi*. Bursa: Ekin Yayınları.
- Tütek, H., Gümüsoğlu, Ş. ve Özdemir, A. (2012). *Sayısal Yöntemler: Yönetmel Yaklaşım*. İstanbul: Beta.
- Ulrich, J. (2013). *TTR: Technical trading rules*. <http://CRAN.R-project.org/package=TTR> adresinden erişilmiştir.
- Ulusoy, T. (2010). İMKB Endeks öngörüsü için ileri beslemeli ağ mimarisine sahip yapay sinir ağı modellenmesi. *International Journal of Economic and Administrative Studies*, 1(5), 21–40.
- Vapnik, V. ve Chervonenkis, A. (1964). A note on one class of perceptrons. *Automation and remote control*, 25(1), 821-837.
- Vapnik, V. ve Lerner, A. (1963). Generalized portrait method for pattern recognition. *Automation and Remote Control*, 24(6), 774–780.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.
- Vapnik, V. N. (1999). An overview of statistical learning theory. *Neural Networks, IEEE Transactions on*, 10(5), 988–999.
- Vapnik, V. N. ve Chervonenkis, A. J. (1974). *Theory of pattern recognition*. USSR: Nauka.
- Venables, W., Smith, D. ve Team, R. C. (2010). *An introduction to R*. R Development Core Team, <https://cran.r-project.org/doc/manuals/R-intro.pdf> adresinden erişildi.
- Verikas, A., Gelzinis, A. ve Bacauskiene, M. (2011). Mining data with random forests: A survey and results of new tests. *Pattern Recognition*, 44(2), 330–349.
- Verma, J. P. (2012). *Data analysis in management with SPSS software*. New York: Springer Science & Business Media.
- Visa, S., Ramsay, B., Ralescu, A. L. ve van der Knaap, E. (2011). Confusion Matrix-based Feature Selection. *MAICS* içinde (ss. 120–127).
- Wilder, J. W. (1987). *The Adam Theory of Markets or What Matters is Profit*. USA: Cavida.

- Yakut, E. (2012). *Veri madenciliği tekniklerinden C5.0 algoritması ve destek vektör makineleri ile yapay sinir ağlarının sınıflandırma başarılarının karşılaştırılması*. Yayınlanmamış doktora tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, Erzurum.
- Yakut, E. ve Elmas, B. Y. S. (2014). Yapay sinir ağları ve destek vektör makineleri yöntemleriyle borsa endeksi tahmini. *Suleyman Demirel University Journal of Faculty of Economics & Administrative Sciences*, 19(2), 139-157.
- Yamada, H. (2005). Wavelet-based beta estimation and Japanese industrial stock prices. *Applied Economics Letters*, 12(2), 85–88.
- Yıldız, B., Yalama, A. ve Coşkun, M. (2008). Forecasting the Istanbul stock exchange national 100 index using an artificial neural network. *World Academy of Science, Engineering and Technology*, 46, 36–39.
- Yolcu, U., Aladag, C. H., Egrioglu, E. ve Uslu, V. R. (2013). Time-series forecasting with a novel fuzzy time-series approach: An example for Istanbul stock market. *Journal of Statistical Computation and Simulation*, 83(4), 599–612.
- Yüksel, E. G. G. (2010). İstanbul menkul kıymetler borsası'nda eşhareketlilik ve asimetrik ayarlama. *Maliye Finans Yazıları*, 24(88), 97–113.
- Yümlü, S., Gürgen, F. S. ve Okay, N. (2005). A comparison of global, recurrent and smoothed-piecewise neural models for Istanbul stock exchange (ISE) prediction. *Pattern Recognition Letters*, 26(13), 2093–2103.
- Zeileis, A. ve Grothendieck, G. (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software*, 14(6), 1–27.
- Zhang, X. (2011). *Encyclopedia of machine learning*. New York: Springer Science & Business Media.
- Zheng, W. ve Tropsha, A. (2000). Novel variable selection quantitative structure-property relationship approach based on the k-nearest-neighbor principle. *Journal of chemical information and computer sciences*, 40(1), 185–194.