

# Borsa Endeksi Hareketlerinin Tahmini: Trend Belirleyici Veri

## Predicting Stock Market Movement: Trend Deterministic Data

Hakan PABUÇCU\*

### ÖZ

Bu çalışma BIST 100 borsa endeksinin negatif ve pozitif yönlü hareketlerinin tahmin edilmesini konu edinmektedir. Yapay sinir ağı, destek vektör makinesi ve naive Bayes algoritmasının tahmin performansları karşılaştırılmaktadır. Analizler iki aşamalı olarak yapılmaktadır. Birinci aşamada tahmin modellerinde girdi olarak kullanılacak dokuz adet teknik gösterge, borsa endeksi açılış, kapanış, en yüksek ve en düşük fiyatlar, kullanılarak hesaplanmakta ve sürekli olan bu teknik göstergeler barındırdıkları trende göre kategorize edilerek yeni bir veri seti oluşturulmaktadır. İkinci aşamada ise, trend belirleyici veri seti girdi olarak kullanılmakta ve seçilen üç makine öğrenme algoritması kullanılarak tahminler yapılmaktadır. BIST 100 veri seti 2009-2018 Aralığını kapsayan günlük kapanış fiyatlarını içermektedir. Analizlerle, destek vektör makineleri algoritmasının en iyi sınıflandırıcı olduğu sonucuna ulaşılmıştır. Ayrıca, daha önceki benzer çalışmalarla karşılaştırmalar yapılarak gerek kullanılan veri seti gerekse tahmin modellerinin etkileri tartışılmaktadır.

### ANAHTAR KELİMELER

BIST100, Yapay Sinir Ağları, Destek Vektör Makineleri, Naive Bayes, Makine Öğrenme

### ABSTRACT

This study focuses on the estimation of negative and positive movements of BIST 100 stock index. The predictive performances of artificial neural network, support vector machine and naive Bayes algorithm are compared. The analyzes are carried out in two stages. In the first stage, nine technical indicators to be used as input in the estimation models are calculated by using the stock index, opening, closing, highest and lowest prices. In the second stage, the trend-setting dataset is used as input and the predictions are made by using three selected machine-learning algorithms. The BIST 100 data set includes the daily closing prices covering the range of 2009-2018. With the analysis, it is concluded that the support vector machines algorithm is the best classifier. In addition, comparisons with previous similar studies and the effects of both the data set used and the prediction models are discussed.

### KEYWORDS

BIST 100, Artificial Neural Network, Support Vector Machine, Naive Bayes Classifier, Machine Learning

Makale Geliş Tarihi / Submission Date	Makale Kabul Tarihi / Date of Acceptance
26.10.2018	28.03.2019
<b>Atf</b>	Pabuçcu, H. (2019). Borsa Endeksi Hareketlerinin Tahmini: Trend Belirleyici Veri. <i>Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi</i> , 22 (1), 246-256.

\* Dr. Öğr. Üyesi, Bayburt Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, hpabuccu@bayburt.edu.tr, ORCID: 0000-0003-2267-5175

## GİRİŞ

Finansal zaman serileri tahminleri, serilerin durağan olmamaları ve belirsizlik barındırmaları sebebiyle oldukça zorlanılan problemlerin başında gelmektedir. Finansal zaman serilerine verilecek en iyi örnek ise hisse senetleri ve borsa endeks verileridir. Borsa endeksleri politik değişimler, ekonominin genel görünümü, yatırımcıların beklenti ve yatırım tercihleri ve diğer endekslerin hareketleri gibi birçok makroekonomik faktörden etkilenmektedir (Tan ve diğ. 2007). Yatırımcıların yatırım kararlarını etkileyen pek çok faktör olmakla birlikte yatırım yapmadan önce faydalandıkları temel ve teknik analiz olmak üzere iki tür analiz vardır. Temel analiz ile yatırımcılar hisse senedinin gerçek değeri, politik iklim, sanayi ve ekonominin performansı gibi göstergeleri dikkate alarak yatırım yapıp yapmayacaklarına karar verirler. Teknik analizde ise, geçmiş değerler ve işlem hacimleri gibi piyasa hareketleri ile oluşturulan istatistiklerin kullanılarak hisse senetlerinin değerlendirilmesi sağlanır. Teknik analizde, bir hisse senedinin gerçek değerini ölçmek yerine, hisse senedinin gelecekte nasıl davranacağını gösterebilecek eğilimleri tanımlamak gerekir ve bu amaçla da hisse senedi grafikleri kullanılır. Malkiel ve Fama (1970) tarafından ileri sürülen etkin piyasalar hipotezi, hisse senedi fiyatlarının önemli ölçüde bilgi barındırdığını ve alım satım verilerine dayanarak hisse senedi fiyatlarının tahmin edilebileceğini ifade etmektedir. Etkin piyasalar hipotezinden hareketle, borsa endeksi hareketleri için başarılı tahminlerin üretilmesi, etkin pazar stratejilerinin geliştirilmesi ve dolayısı ile yatırımcıların piyasa riskleri ve spekülörlere karşı korunma fırsatına sahip olmaları sağlanır (Leung ve diğ. (2000).

Finansal zaman serilerinin dinamik, doğrusal olmayan, karmaşık ve kaotik yapısı doğru tahminler üretme noktasında engeller oluşturmaktadır (Abu-Mostafa ve Atiya, 1996). Çok uzun bir süreden beri regresyon analizi, lojistik regresyon analizi, yapay sinir ağları (YSA), destek vektör makineleri (DVM) gibi çok sayıda geleneksel veya yeni yöntemler borsa endeksi ve endeks hareketlerini tahmin etmek için kullanılmaktadır. Son yıllarda ise geleneksel yöntemler yerine yapay zekâ ve makine öğrenme algoritmaları ile yapılan tahminler oldukça popüler olmuş ve başarılı çalışmalar ortaya çıkmıştır. Ayrıca, borsa tahmini çalışmalarında çok çeşitli örneklemeler kullanılmış olmakla birlikte, bunların önemli bir bölümü gelişmiş ülkeler için yapılan çalışmalardan oluşmaktadır.

Borsa endekslerinin durağan olmayan yapısı, makine öğrenme yöntemleri gibi doğrusal olmayan modelleri ön plana çıkarmıştır. Bu yöntemlerin başında YSA ve DVM gelmektedir. YSA, (Hassan ve diğ. (2007); Kara ve diğ. (2011); Kimoto, ve diğ. (1990); Olson ve Mossman (2003) gibi çalışmalarla, DVM ise (Hsu ve diğ. (2009); Huang ve diğ. (2005); Kumar ve Thenmozhi (2005) gibi çalışmalarla borsa tahminlerinde güvenilir sonuçlar üretme noktasında başarısını ispatlamıştır. Bu yöntemlerin kendine özgü yapıları ve öğrenme prensipleri vardır ve makine öğrenme algoritmaları olarak sınıflandırılmaktadırlar.

Chen ve diğ. (2003)'de gelişmekte olan Asya ülkelerinin en hızlı büyüyen finansal borsalarından biri olan Tayvan Borsası'nın piyasa endeksindeki getiri yönünü tahmin etmişlerdir. Fiyat hareketinin yönüne dair tahminlerin yönlendirdiği ticaret stratejilerinin daha etkili olabileceği ve daha yüksek karlara yol açabileceği fikrinden hareketle, olasılıksal sinir ağı tarihsel verilerle eğitildikten sonra endeks hareket yönünü tahmin etmek için kullanılmıştır. Sinir ağı tahminlerinin istatistiksel performansı ölçülmüştür. Thawornwong ve Enke (2004)'te borsa endekslerinin yönü araştırılmıştır. Tahmin modellerine adım adım eklenen değişkenlerin tahminler üzerinde iyileştirme yapıp yapmadıklarını izlemişlerdir. Sonuçlar sinir ağı modellerinin geleneksel tahmin yöntemlerine göre daha düşük riskli ve yüksek getirili tahminler ürettiğini ortaya koymuştur. Enke ve Thawornwong (2005) çok sayıda finansal ve ekonomik değişkenin yordama kapasitelerini değerlendirmek için yapay sinir ağı modellerini kullanmışlardır. Sinir ağı modellerinin sınıflandırma problemi için nasıl sonuçlar ürettiği incelenmiştir. Farklı modellerin genelleme yeteneğini geliştirmek için eğitim aşamasında çapraz doğrulama tekniği kullanılmıştır. Hassan ve diğ. (2007)'de gizli Markov modeli, sinir ağı ve genetik algoritmanın kullanıldığı hibrit bir model önerilmiş ve finansal piyasa davranışının tahmini için uygulanmıştır. Hisse senedi fiyatları YSA ile gizli Markov modeline girdi olacak şekilde dönüştürülmüştür. Kara ve diğ. (2011)'de İstanbul borsası ulusal 100 endeksi verileri kullanılarak YSA ve DVM tahmin performansları karşılaştırılmıştır. Analiz sonuçları YSA performansının %75,7 DVM performansının ise %71,52 olduğunu göstermiştir. Zhong ve Enke (2017)'de 60 ekonomik ve finansal gösterge kullanılarak Standard & Poors 500 endeksinin günlük hareketleri incelenmiştir. Temel bileşenler analizi, bulanık temel bileşenler analizi ve çekirdek tabanlı temel bileşenler analizi yöntemleri ile göstergeler azaltılıp boyutlara indirgenmiş ve YSA ile tahminler yapılmıştır. Farklı yöntem kombinasyonlarının performansları karşılaştırılmıştır.

Cao ve Tay (2001) DVM'ni finansal zaman serilerinde uygulayan ilk çalışmalardan olması nedeniyle önemlidir. Çalışmada finansal seriler kullanılarak destek vektör makineleri ve geri beslemeli sinir ağları tahminlerinin performansları karşılaştırılmıştır. Sonuçlar DVM'nin YSA'ya göre parametre seçimi ve tahmin hata oranları bakımından daha uygulanabilir olduğunu ve daha iyi sonuçlar ürettiği ortaya koymuştur. Tay ve

Cao (2001)'de ise DVM tahminlerinin geliştirilebilmesi için öz düzenleyici haritalar yardımıyla iki aşamalı hibrit bir yöntem önerilmiş ve DVM tahminlerine oranla daha iyi sonuçlar elde edilmiştir. (Kim, 2003)'de ise DVM, YSA ve durum temelli muhakeme (case-based reasoning) yöntemlerinin finansal zaman serileri tahminleri karşılaştırılmış ve DVM en iyi sınıflandırıcı olarak belirlenmiştir. Patel ve diğ.(2015a)'da DVM, YSA, NB ve random forest algoritmaları kullanılarak Hindistan borsası hareketlerinin yönü araştırılmıştır. Veri seti barındırdığı trende göre yeniden kodlanarak kategorik hale getirilmiştir. Random forest algoritması en iyi sınıflandırma performansını göstermiştir. Yine, Hint borsalarına ait iki endeksin on yıllık veri seti kullanılarak (Patel ve diğ. (2015b)'de ise iki aşamalı bir endeks tahmin yöntemi önerilmiştir. İlk aşamada destek vektör regresyonu ikinci aşamada ise YSA ve random forest algoritmaları kullanılmış ve sonuçlar tartışılmıştır. DVM ve veri zarflama analizi ile Çin'de faaliyet gösteren bankalara ait finansal veriler kullanılarak banka etkinliklerinin değerlendirildiği Chen ve diğ. (2018)'de bankaların önemli bir kısmının istenilen etkinlik seviyesine ulaşamadığı belirlenmiştir. Literatürde çok sayıda çalışma makine öğrenme yöntemlerinin finansal zaman serileri tahminlerinde kullanıldığı görülmektedir. Bu çalışmalarla ilgili kapsamlı incelemeler Atsalakis ve Valavanis (2009a), (2009b); Chun ve Kim (2004); Vanstone ve Finnie (2009) çalışmalarında yer almaktadır.

Yapılan kapsamlı literatür araştırması sonucundan hareketle, bu araştırmanın temel amacı BIST 100 endeksinin hareketlerinin yönünü YSA, DVM ve NB makine öğrenme algoritmalarını kullanarak tahmin etmektir. Ayrıca, araştırmanın ilgili alana en önemli katkılarından biri de endeks hareketlerinin yönünün tahmininde ANN, DVM ve NB algoritmalarının kullanılabilirliğini göstermek, doğrulamak ve bu yöntemlerin performanslarını karşılaştırmaktır.

Çalışmanın devamı üç bölümden oluşmaktadır. İkinci bölümde veri seti ve kullanılan yöntem sunulmaktadır. Üçüncü bölüm araştırmanın bulgularını dördüncü bölüm ise sonuç ve önerileri kapsamaktadır.

## 1. VERİ SETİ VE YÖNTEM

Çalışmada BIST 100 endeksinin negatif ve pozitif yönlü hareketlerinin tahmin edilebilmesi için iki aşamalı bir yol izlenmektedir. İlk aşamada. Kara ve diğ. (2011)'deki teknik göstergeler endeks kapanış, açılış, en yüksek ve en düşük fiyatları kullanılarak ayrı ayrı hesaplanmıştır. Hesaplanan teknik göstergelerin barındırdıkları trende göre belli kurallar dahilinde kategorize edilerek *trend belirleyici* veri seti oluşturulmuştur. Yani sürekli değerlerden oluşan veri seti  $\{-1,1\}$  değerleri ile yeniden kodlanarak yeni bir veri seti oluşturulmuştur. İkinci aşamada ise, seçilen makine öğrenme algoritmaları kullanılarak borsa endeksi hareketlerinin yönünü tahmin edebilecek en iyi modeller araştırılmıştır. Borsa endeksi verisi ise 2009-2018 Aralığı kapsayan günlük kapanış fiyatlarından oluşmaktadır. Veri seti T.C. Merkez Bankasından alınarak analize hazır hale getirmek üzere düzenlenmiştir. Veri analizinde algoritmaların eğitimi için veri setini iki veya üç parçaya bölmek yerine, 10 kat çapraz doğrulama (Ten-fold cross validation) yöntemi kullanılmıştır. Yöntemlerin daha iyi tahminler üretebilmesi için eğitim sürecinde her bir yöntemin kendine özgü parametreleri değiştirilerek en iyi tahmin modeli araştırılmıştır. Kullanılan teknik göstergeler (Tablo 1) ve her bir göstergenin kategorik hale dönüştürülme aşamaları (Patel et al., 2015a) için şu tanımlamaları yapmak mümkündür.

Hareketli ortalamalar (MA-WMA): Hareketli ortalama ve ağırlıklı hareketli ortalama bir ürünün belli zaman aralığındaki ortalama fiyat değişimidir. Hareketli ortalamalar eğilimin genel yönünü gösterir. Teknik analizde çok sık kullanılan en basit göstergedir. Bu çalışmada 14 günlük hareketli ortalamalar kullanılmıştır. Genel bir kural olarak, fiyat hareketli ortalamasının üzerindeyse eğilim artış yönünde, fiyat hareketli ortalamasının altındaysa eğilim azalış yönündedir <<https://stockcharts.com>>. Çalışmada, hareketli ortalamaların fiyattan yüksek olduğu günler için +1, düşük olduğu günler için de -1 değeri atanmıştır. Yani +1 eğilimin artış yönünde -1 ise eğilimin azalış yönünde olduğunu göstermektedir.

Momentum: Bu gösterge endeksin fiyat artış ve azalış oranlarını ölçer. Pozitif değer eğilimin artış yönünde olduğunu gösterir ve +1 ile kodlanır. Negatif değer ise eğilimin azalış yönünde olduğunu anlamına gelir ve -1 ile kodlanır.

Stokastik %K-%D ve Larry Williams %R: Bu göstergeler trendden arındırılmış stokastik göstergelerdir. Stokastik göstergeler artıyor ise endeks fiyatının da aynı şekilde artış göstermesi, azalıyorsa da yine aynı şekilde azalış göstermesi beklenir. Bu nedenle, stokastik göstergenin  $t$  zamanındaki değeri  $t-1$  zamanına göre daha yüksek ise trendin artış yönlü olduğu daha düşük ise trendin azalış yönlü olduğu ifade edilir ve sırasıyla +1 ve -1 ile kodlanır.

Göreceli Güç Endeksi (RSI): Bu gösterge genel olarak aşırı alım-satım noktalarının belirlenmesi için kullanılır <<https://stockcharts.com>>. 0 ile 100 arasında değerler alabilir. 70'in üzerinde olması aşırı alım, 30'un altında olması ise aşırı satım anlamına gelir. Sırasıyla -1 ve +1 ile kodlanır. 30 ile 70 arasında olduğu durumlar

için; eğer göstergenin  $t$  zamanındaki değeri  $t-1$  zamanına göre daha yüksek ise  $+1$  ile daha düşük ise  $-1$  ile kodlanır.

**Tablo 1: Seçilen Teknik Göstergeler ve Formülleri**

Göstergeler	Formül
Basit 14 günlük hareketli ortalama (MA)	$C_t + C_{t-1} + \dots + C_{t-14}/14$
Ağırlıklı 14 günlük hareketli ortalama (WMA)	$\frac{((n) * C_t + (n - 1) * C_{t-1} + \dots + C_{t-14})}{(n + (n - 1) + \dots + 1)}$
Momentum (Mom)	$C_t - C_{t-n}$
Stokastik K% (K%)	$\frac{C_t - LL_{t-n}}{HH_{t-n} - LL_{t-n}} * 100$
Stokastik D% (D%)	$\frac{\sum_{i=0}^{n-1} K_{t-i}\%}{n}$
Göreceli güç indeksi (RSI)	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} Up_{t-i}/n) / (\sum_{i=0}^{n-1} Dw_{t-i}/n)}$
MACD(Hareketli Ortalamaların Uyumu/Uyumsuzluğu)	$MACD(n)_{t-1} + \frac{2}{n+1} * (DIFF_t - MACD(n)_{t-1})$
Larry William's R% (LW)	$\frac{H_n - C_t}{H_n - L_n} * 100$
A/D (Birikim/Dağılım göstergesi)	$\frac{H_t - C_{t-1}}{H_t - L_t}$

**Kaynak: Kara vd. (2011).**

\* $C_t$  kapanış fiyatı,  $L_t$  en düşük fiyat ve  $H_t$  en yüksek fiyat.  $DIFF_t$ :  $EMA(12)_t - EMA(26)_t$ .  $EMA$  Üstel hareketli ortalama,  $EMA(k)_t$ :  $EMA(k)_{t-1} + \alpha * (C_t - EMA(k)_{t-1})$ ,  $\alpha$  düzeltme faktörü.  $LL_t$  son  $t$  gün için minimumların minimumu,  $HH_t$  son  $t$  gün için maksimumların maksimum değeri.  $M_t = (H_t + L_t + C_t)/3$ ,  $SM_t = (\sum_{i=0}^n M_{t-i+1}/n)$ ,  $D_t = (\sum_{i=1}^n |M_{t-i+1} - SM_t|/n)$ ,  $Up_t$  ve  $Dw_t$  sırasıyla  $t$  zamanındaki yukarı ve aşağı yönlü fiyat değişimlerini ifade etmektedir.

Hareketli Ortalamaların Uyumu/Uyumsuzluğu (MACD): İki farklı hareketli ortalama arasındaki ilişkiyi gösteren ve piyasa katılımcılarının sık kullandığı orta vadeli göstergedir. Eğer MACD yükselirse endeks fiyatı da yükselir, düşerse endeks fiyatı da düşer. Göstergenin  $t$  zamanındaki değeri  $t-1$  zamanına göre daha yüksek ise trendin artış yönlü olduğu daha düşük ise trendin azalış yönlü olduğu ifade edilir ve sırasıyla  $+1$  ve  $-1$  ile kodlanır.

Birikim/Dağılım göstergesi (A/D endeksi): Bu gösterge de bir trendin var olup olmadığını ve var olan trendin sürekliliğini araştırır. Diğer göstergelere benzer şekilde  $t$  zamanındaki değeri  $t-1$  zamanına göre daha yüksek ise trendin artış yönlü olduğu daha düşük ise trendin azalış yönlü olduğu ifade edilir ve sırasıyla  $+1$  ve  $-1$  ile kodlanır.

Seçilen teknik göstergelerin her biri tek başına endeks hareketlerinin tahmininde kullanılan göstergelerdir. Dolayısı ile kendi içlerinde önemli düzeyde bilgi barındırmaktadırlar. Bu çalışmanın amacı, göstergelerin gerçek sürekli değerlerinin yerine barındırdıkları trend bilgisi hesaplanarak doğrudan tahmin modeline girdi olarak sunulmasının nasıl tahminler üreteceğinin araştırılmasıdır.

Tablo 2'de modeller için girdi setlerini oluşturan teknik göstergelere ait bazı tanımlayıcı istatistikler sunulmaktadır.

**Tablo 2: Göstergelere ilişkin özet istatistikler**

Gösterge	Minimum	Maksimum	Ortalama	Standart sapma
<i>MA</i>	34706.677	118348.507	74514.625	17339.498
<i>WMA</i>	34482.403	119403.421	74601.188	17314.628
<i>Mom</i>	-16792.68	10920.66	318.826	3671.642
<i>K%</i>	14.08	80.193	51.864	14.501
<i>D%</i>	4.023	74.37	51.742	10.755
<i>RSI</i>	18.016	88.366	53.443	12.617
<i>MACD</i>	-3586.069	2491.367	184.149	1137.383
<i>LW</i>	-85.92	-19.807	-48.136	14.501
<i>A/D</i>	-0.67	1.166	0.503	0.144

### 1.1. Yapay Sinir Ağları

Yapay sinir ağları biyolojik sinir hücresinin bir benzetimidir. Girdi ve çıktı birimleri arasındaki ilişkilerin tespiti ve tahmini için bazı aktivasyon fonksiyonlarını kullanır. YSA, finansal modellemede sıklıkla kullanılan başarılı bir tahmin modelidir. Borsa endeksi hareket yönlerinin tahmin edilmesi için bu çalışmada üç katmanlı ileri beslemeli model kullanılmıştır. Girdi katmanında dokuz sinir hücresi, gizli katmanda yedi sinir hücresi ve çıktı katmanında bir sinir hücresi kullanılmıştır. Girdi katmanındaki her sinir hücresi bir teknik göstergiyi temsil etmektedir. Ara katmanda yer alan her bir sinir hücresinde transfer fonksiyonu olarak tanjant sigmoid, çıktı katmanında ise logaritmik sigmoid fonksiyon kullanılmıştır. Çıktı değerleri için artış ve azalış yönünün belirlenmesinde eşik değeri olarak 0,5 kullanılmıştır. YSA parametreleri ve mimarisi için çok sayıda deneme yapılmıştır. YSA modeli parametreleri, gizli katmandaki sinir hücresi sayısı ( $n$ ), öğrenme oranı ( $lr$ ), momentum sabiti ( $mc$ ) ve iterasyon sayısı ( $ep$ ) dir. Tablo 3’de YSA için kullanılan parametreler ve kullanım düzeyleri sunulmaktadır. Çok sayıda parametre kombinasyonu denenmiş ve en iyi parametre kümesi belirlenmeye çalışılmıştır. YSA ile ilgili finansal tahminler konusunda ayrıntılı bilgi için aşağıdaki kaynaktan<sup>1</sup> faydalanılabilir.

**Tablo 3: Eğitim sürecinde kullanılan parametre ayarları**

Parametre	Kullanılan düzey
Gizli katman nöron sayısı ( $n$ )	5,...,50
İterasyon ( $ep$ )	250, 500,...,2000
Momentum sabiti ( $mc$ )	0.1, 0.2,...,0.9
Öğrenme oranı ( $lr$ )	0.1, 0.2, 0.3

### 1.2. Destek Vektör Makineleri

Yapısal risk minimizasyonuna bağlı olarak Vapnik (1995) tarafından geliştirilen sınıflandırma, karakter tanıma ve regresyon gibi kullanım alanları olan çok güçlü bir algoritmadır. Veri kümeleri arasında maksimum sınıf aralıklarına sahip hiper düzlemler tanımlayarak, pozitif ve negatif örnekler arasındaki ayırım marjını maksimize etmek fikrine dayanır. DVM, noktaları desen uzayında ya da daha yüksek boyutlu bir uzayda iki ayrık uzaya atayarak sınıflandırır (Khemchandani ve diğ., 2009). DVM’nin ana fikri, pozitif ve negatif örnekler

<sup>1</sup> Patel, J., Shah, S., Thakkar, P., ve Kotecha, K. (2015a). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259–268. <http://doi.org/10.1016/j.eswa.2014.07.040>.

arasındaki ayırım marjının maksimize edileceği şekilde karar yüzeyi olarak bir hiper düzlem oluşturmaktır (Xu ve diğ., 2009).

Örneklerin bir eğitim kümesi için,  $x_i \in R^d$  girdi vektörleri ve karşılığı olan  $y_i \in \{1, -1\}$  görüntü kümesi ile nesnelere iki sınıfa dâhil ederek nasıl sınıflandıracağını öğrenir. İki sınıflı bir sınıflandırma problemi için;  $x_i \in R^d (i = 1, 2, \dots, N)$  girdi vektörleri ve  $y_i \in \{1, -1\} (i = 1, 2, \dots, N)$  karşılık kümesi olduğu varsayalım. Burada (1) ve (-1) iki sınıfı ifade etsin. Burada amaç ikili bir sınıflandırıcı oluşturmak veya mevcut örneklerden, daha önce görülmemiş bir örnekteki birimleri yanlış sınıflandırma olasılığı düşük olan bir karar fonksiyonu türetmektir. DVM, giriş vektörlerini  $x_i \in R^d$  yüksek boyutlu bir özellik alanına  $\Phi(x_i) \in H$  haritalar ve en uygun ayırma hiper düzlemini oluşturur. Hiper düzlem ile  $H$  uzayındaki her bir sınıfın en yakın veri noktaları arasındaki mesafe maksimize edilir.  $\Phi$  Haritalaması  $H$  iç çarpım uzayında tanımlanan  $K(x_i, x_j)$  çekirdek fonksiyonları tarafından gerçekleştirilir. Elde edilen sınıflandırıcı  $\alpha_i$  katsayılarını belirlemek için eşitlik 1 de verilen karar fonksiyonu ve Eşitlik (2)'de verilen kuadratik programlama problemine dayalıdır (Hua ve Sun, 2001).

$$f(x) = \text{sgn}(\sum_{i=1}^N y_i \alpha_i \cdot K(x, x_i) + b) \quad (1)$$

Kuadratik problem ise aşağıdaki gibi tanımlanır:

$$\text{Maks. } \sum_{i=1}^N \alpha_i - 1/2 \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \cdot y_i y_j \cdot K(x, x_i), 0 \leq \alpha_i \leq c$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad i = 1, 2, \dots, N \quad (2)$$

Burada  $c$ ; marjin ile sınıflandırma hataları arasındaki dengeyi kontrol eden düzenleme parametresidir. DVM'de yaygın kullanılan polinomial ve radyal tabanlı olmak üzere iki adet çekirdek fonksiyon vardır.

$$\text{Polinomial fonksiyon: } K(x_i, x_j) = (x_i \cdot x_j + 1)^d \quad (3)$$

$$\text{Radyal tabanlı fonksiyon: } K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (4)$$

Modellerin başarısı çekirdek fonksiyonların ve parametrelerin seçimi ile doğrudan ilişkilidir. DVM parametreleri, çekirdek fonksiyonu (polinomial veya radyal tabanlı), çekirdek fonksiyonu derecesi ( $d$ ) veya gamma katsayısı ( $\gamma$ ) ve düzenleme parametresi ( $c$ )'dir. Bu çalışmada kullanılan DVM için kullanılacak olan parametre kümeleri Tablo 4'te sunulmaktadır.

**Tablo 4: Eğitim sürecinde kullanılan parametre ayarları**

Parametre	Düzyey (polinomial)	Düzyey (radial tabanlı)
Çekirdek fonksiyonu derecesi ( $d$ )	1, 2, 3, 4	
Çekirdek fonksiyonu Gamma Katsayısı ( $\gamma$ )		0, 0.1, 0.2, ..., 5.0
Düzenleme parametresi ( $c$ )	1, 10, 100	1, 10, 100

### 1.3. Naive Bayes

Naive-Bayes sınıflandırma algoritması şartlı olasılık kuralları üzerine yapılandırılmış makine öğrenme algoritmalarından bir tanesidir. Parametre seçimi, kullanım kolaylığı ve hızlı çözümler sunması bakımından diğer bazı yöntemlere göre üstünlükleri vardır. Algoritma diğer makine öğrenme algoritmaları gibi eğitim için ayrılan tarihsel zaman serisi veya yatay kesit veri setleri üzerinden gözlemlerin hangi sınıflara dâhil olduğunu öğrenir ve daha önce görmediği gözlemleri genelleştirme yeteneği ile ilgili sınıflara atar. Bu süreçte olasılıkların tahmini için Bayes teoremi kullanılır. Bayes teoremi sonsal (posterior)  $P(C|X)$  olasılıkların,  $P(C)$ ,  $P(X|C)$  ve  $P(X)$  olasılıkları kullanılarak hesaplanmasını sağlar. Bayes teoremi şu şekilde ifade edilebilir (Eşitlik 5).

$$P(C|X) = \frac{P(C)P(X|C)}{P(X)} \quad (5)$$

Burada;  $P(C|X)$ ,  $X$  olayı gerçekleştiğinde  $C$  olayının meydana gelmesi için hesaplanan sonsal olasılıktır. Bu çalışmada  $P(C)$ , borsa hareketleri (yükseliş/düşüş) sınıfına dâhil olma olasılığı ve  $X$  ise test verisini ifade

eder.  $P(X|C)$  C olayının gerçekleşmesi durumunda X olayının gerçekleşmesinin koşullu olasılığını gösterir. Şartlı olasılık eğitim verisi kullanılarak tahmin edilir. Naive Bayes sınıflandırıcı algoritmasının işleyişi şu şekilde özetlenebilir (Patel ve diğ., 2015a):

$C_1, C_2, \dots, C_m$  olmak üzere  $m$  sınıf ve test verisinin gerçekleşme olayı  $X$  verilsin. Bayes sınıflandırıcı, Bayes teoremini (Eşitlik 6) kullanarak test verilerini en yüksek olasılıklı bir sınıfa sınıflandırır.

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (6)$$

$A_1, A_2, \dots, A_m$  gibi çok özellikli bir veri seti için  $P(X|C_i)$  olasılıklarını hesaplamak oldukça güç ve zahmetli olacaktır. Hesaplamaları kolaylaştırmak için sınıf koşullu bağımsızlık varsayımı (Naive assumption) yapılır. Bu varsayım, özelliklerin değerlerinin diğerlerinden koşullu olarak bağımsız olduklarını ifade eder. Eşitlik (7) şartlı olasılığın genelleştirilmiş halidir.

$$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i) = P(x_1|C_i) * P(x_2|C_i) * \dots * P(x_n|C_i) \quad (7)$$

Burada,  $x_k$  X olayı için  $A_k$  özelliğinin aldığı değerdir.  $P(x_k|C_i)$ 'nin hesaplanması kategorik veya sürekli olmasına göre değişiklik gösterir. Eğer  $A_k$  kategorik ise,  $P(x_k|C_i)$ ,  $A_k$  özelliği için eğitim veri setinde  $x_k$  değerini alan  $C_i$  sınıfının gözlem sayısının, eğitim veri setindeki  $C_i$  sınıfının tüm gözlem sayısına bölümüdür. Eğer  $A_k$  sürekli ise, Gauss dağılım fonksiyonu veri setine ayarlanır ve  $P(x_k|C_i)$  Eşitlik (8) ile hesaplanır. Bu çalışmada tüm değişkenler kategorik olduğu için Gauss fonksiyonu kullanılmamıştır.

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (8)$$

$$P(x_k|C_i) = f(x_k, \mu_{C_i}, \sigma_{C_i}) \quad (9)$$

Burada,  $\mu_{C_i}, \sigma_{C_i}$  sırasıyla ortalama ve standart sapmayı ifade etmektedir. X Gözleminin sınıfı  $C_i$  olarak tahmin edilir.

## 2.AMPİRİK BULGULAR

Modellerin başarılarını test etmek ve birbirleriyle karşılaştırabilmek için sınıflandırma başarıları yanında, F istatistiği de hesaplanmıştır. F istatistiği, doğru pozitif (DP), yanlış pozitif (YP), doğru negatif (DN) ve yanlış negatif (YN) istatistikleri kullanılarak Eşitlik (10)-(13) ile hesaplanmaktadır. Yüksek F istatistiği tahmin modelinin performansının diğer modellere göre daha iyi olduğu anlamına gelmektedir.

$$Kesinlik_{pozitif} = \frac{DP}{DP+YP} \quad (10)$$

$$Kesinlik_{negatif} = \frac{DN}{DN+YN} \quad (11)$$

$$Hassasiyet_{pozitif} = \frac{DP}{DP+YN} \quad (12)$$

$$Hassasiyet_{negatif} = \frac{DN}{DN+YP} \quad (13)$$

Kesinlik, negatif ve pozitif kesinliğin ağırlıklı ortalaması iken hassasiyet pozitif ve negatif hassasiyetin ağırlıklı ortalamasıdır. F istatistiği Eşitlik (14) ile hesaplanmıştır.

$$F = \frac{2 * Kesinlik * Hassasiyet}{Kesinlik + Hassasiyet} \quad (14)$$

Tüm veri seti  $\{-1,1\}$  ile kodlandığından veri seti normalize edilmemiştir. Daha önce de ifade edildiği gibi her parametre için belirlenen aralıklarda denemeler yapılmış ve en iyi sınıflandırıcı araştırılmıştır. En düşük hata ve en yüksek sınıflandırma performansına sahip üç YSA modeli parametreleri Tablo 5'te sunulmaktadır.

**Tablo 5: En iyi üç YSA modeline ilişkin parametreler**

Model	lr (öğr. oranı)	ep (iterasyon)	mc (momentum sbt)	n(gizli katman sinir hücre sayısı)	Sınıflandırma performansı.
1	0.1	500	0.1	7	0.9834
2	0.1	500	0.6	7	0.9816
3	0.1	500	0.2	6	0.9704

En iyi YSA modeli olarak öğrenme oranı 0,1; momentum sabiti 0,1 ve tek gizli katmanda yedi sinir hücresi olan ağ belirlenmiştir. Ayrıca, ağ 500 iterasyonda eğitilmiş ve test verisi üzerinde sınıflandırma performansı 0,9834 olarak hesaplanmıştır. Eğitim esnasında 10 kat çapraz doğrulama yöntemi kullanıldığı için, sınıflandırma performansı olarak hesaplanan istatistik, her katta test verisine modelin verdiği cevapların ortalamasıdır. Tüm tahmin modellerinde eğitim seti için ortalama mutlak hata (MAE) istatistiği hesaplanmış ve seçilen modellerin tamamında eğitim hatası 0,01 düzeyinin altında gerçekleşmiştir. Sunulan sınıflandırma performansları ve hata istatistikleri test verisine ait istatistiklerdir. Modellerin seçimi yapılırken de test verisi üzerindeki başarıları karar kriteri olmuştur. İlk üç model birbirine yakın sınıflandırma performansları göstermiştir.

YSA modeline benzer şekilde Tablo 3'te belirtilen parametreler ve kullanılacak düzeylere göre çok sayıda parametre kombinasyonu oluşturulmuş ve denemeler yapılmıştır. Kullanılan çekirdek fonksiyonuna göre polinomial ve radyal tabanlı olmak üzere en iyi performans gösteren üçer model belirlenmiştir. Farklı parametrelerle aynı performansı gösteren bazı modeller bulunmaktadır. Polinomial çekirdek fonksiyonlu DVM modeli radyal tabanlı çekirdek fonksiyonlu modele göre daha iyi bir sınıflandırma performansı göstermiştir. Tablo 6'da sunulan altı model arasında birinci derece polinomial çekirdek fonksiyonlu DVM modeli için düzenleme parametresi  $c=100$  ile en iyi sınıflandırıcı olarak tespit edilmiştir.

**Tablo 6: En iyi üç DVM modeline ilişkin parametreler**

	Çekirdek fonk.	$d$	$\gamma$	$c$	Perf.
1	Polinomial	1	-	100	0.9845
2	Polinomial	1	-	10	0.9841
3	Polinomial	2	-	10	0.9824
4	Radyal tabanlı	-	0.01	100	0.9841
5	Radyal tabanlı	-	1.0	10	0.9823
6	Radyal tabanlı	-	0.01	1	0.9690

NB algoritması ile yapılan sınıflandırmada kategorik verilerin kullanılmış olmasından dolayı Gauss dağılım fonksiyonu kullanılmamıştır. NB algoritmasının sınıflandırma başarısı 0,9506 olarak hesaplanmıştır. Eğitim sürecinde 10-kat çapraz doğrulama yöntemi kullanılmıştır. NB algoritması için kesinlik ve hassasiyet istatistikleri kullanılarak F istatistiği hesaplanmış ve Tablo 7'de görüldüğü gibi modeller karşılaştırılmıştır.

**Tablo 7: Modellerin karşılaştırılması**

	DP	YP	F-İst..	Perf.
<b>YSA</b>	0,984	<b>0,015</b>	<b>0,984</b>	<b>0,9830</b>
<b>DVM</b>	0,985	<b>0,014</b>	<b>0,985</b>	<b>0,9845</b>
<b>NB</b>	0,951	<b>0,048</b>	<b>0,951</b>	<b>0,9506</b>



Tablo 7’de en iyi YSA, DVM ve NB modelleri karşılaştırılmaktadır. F istatistiği YSA için 0,984 DVM için 0,985 ve NB için 0,951 olarak hesaplanmıştır. Sınıflandırma performansı açısından da DVM modeli 0,9845 ile en iyi sınıflandırma algoritması olarak tespit edilmiştir.

Sınıflandırılan örneklere ilişkin hesaplanan değerler ise Tablo 8’de sunulmaktadır. YSA modeli toplamda 38 gözlemi yanlış sınıflandırırken, DVM 36, NB ise 1115 gözlemi yanlış sınıflandırmıştır.

**Tablo 8: Gözlemlerin sınıflandırılması**

	YSA		DVM		NB	
Sınıflar	0	1	0	1	0	1
0	1078	7	1080	5	1041	44
1	31	1212	31	1212	71	1172

## SONUÇ

Borsa veya hisse senetlerinin hareket yönünün tahmini, etkin stratejiler geliştirmek ve yatırımcılara yol göstermek için önemli olmakla birlikte barındırdıkları belirsizlik ve kaotik yapı gereği oldukça güçtür. Borsa endeksi hareketlerinin tahmininde kullanılan modellerin tahmin başarısı kullanılan değişkenlerin temsil kabiliyetine ve modellerin doğru ve iyi tanımlanmış parametre kümelerine sahip olmasına bağlıdır.

Bu araştırmada borsa endeksinin geleceğe dönük hareketlerini tahmin edebilmek için yöntem denemeleri yapılmıştır. YSA, DVM ve NB sınıflandırma algoritmalarının sınıflandırma performansları BIST 100 endeksi 2009-2018 tarihsel verisi kullanılarak araştırılmış ve karşılaştırılmıştır. Borsa endeksinin hareketine ilişkin belirleyici göstergeler olarak dokuz teknik gösterge hesaplanmış ve devamında içerdikleri trende göre yeniden kategorize edilerek yeni bir veri seti oluşturulmuştur. Sürekli veriler yerine yeni oluşturulan trend belirleyici veri seti modellere girdi olarak sunulmuştur. YSA, DVM ve NB sınıflandırıcı algoritmaları ile BIST 100 endeksi yönü için iyi bir sınıflandırıcı araştırılmıştır.

Yapılan analizler sonucunda, üç makine öğrenme algoritmasının da BIST 100 endeks hareketlerinin yönünü tahmin edebileceği tespit edilmiştir. Tahmin performanslarının da birbirine yakın olduğu görülmektedir. Her yanlış sınıflandırılan birimin bile önemli olduğu düşünüldüğünde, trend belirleyici veri seti ile edinilen sonuçlara göre DVM modelinin en iyi sınıflandırıcı olduğu, devamında ise sırasıyla YSA ve NB sınıflandırıcılarının geldiğini ifade etmek mümkündür. Bu çalışma için kaynak teşkil eden çalışmalarda ulaşılan sonuçlara göre çok daha iyi bir başarı performansı yakalanmıştır. Örneğin Kara ve diğ. (2011); Patel ve diğ. (2015a)’de en yüksek test tahmin performansı 0,62 ile 0,83 değerleri arasında gerçekleşmiştir. Parametre seçimi, kullanım kolaylığı gibi durumlardan bakıldığında ise NB algoritmasının diğer yöntemlere göre avantajlı olduğu söylenebilir. Bu noktada özellikle YSA’nın daha dezavantajlı bir model olduğu pek çok çalışmada vurgulanmaktadır. DVM için ise deneysel hataların değil yapısal hataların minimize edilmesi, yerel minimumlara takılmadan küresel minimumların yani en uygun çözümlerin elde edilebilmesi modelin avantajlarındandır.

Son olarak, borsa tahminlerinde teknik belirleyicilerin kullanıldığı pek çok çalışma literatürde mevcuttur. Ancak, teknik belirleyiciler kullanılarak yeni bir veri setinin türetilmesi ile ilgili sınırlı sayıda çalışma olduğunu söylemek gerekmektedir. Bu çalışma ile YSA, DVM ve NB algoritmaları yazılımlar ile uzman sistemlere dönüştürüp yatırımcılara al-sat pozisyonu almada yardımcı bir sistem geliştirilebilir.

## Kaynakça

- Abu-Mostafa, Y. S., & Atiya, A. F. (1996). "Introduction to Financial Forecasting", *Applied Intelligence*, 6(3), s. 205–213. <http://doi.org/10.1007/BF00126626>
- Atsalakis, G. S., & Valavanis, K. P. (2009a). "Forecasting Stock Market Short-Term Trends Using a Neuro-Fuzzy Based Methodology", *Expert Systems with Applications*, 36(7), s. 10696–10707. <http://doi.org/https://doi.org/10.1016/j.eswa.2009.02.043>
- Atsalakis, G. S., & Valavanis, K. P. (2009b). "Surveying Stock Market Forecasting Techniques – Part II: Soft Computing Methods", *Expert Systems with Applications*, 36(3, Part 2), s. 5932–5941. <http://doi.org/https://doi.org/10.1016/j.eswa.2008.07.006>
- Cao, L., & Tay, F. E. H. (2001). "Application of support Vector Machines in Financial Time Series Forecasting", *Omega*, 29(4), s. 309–317. [http://doi.org/10.1016/S0305-0483\(01\)00026-3](http://doi.org/10.1016/S0305-0483(01)00026-3)
- Chen, A.-S., Leung, M. T., & Daouk, H. (2003). "Application of Neural Networks to an Emerging Financial Market: Forecasting and Trading the Taiwan Stock Index", *Computers & Operations Research*, 30(6), s. 901–923. [http://doi.org/https://doi.org/10.1016/S0305-0548\(02\)00037-0](http://doi.org/https://doi.org/10.1016/S0305-0548(02)00037-0)
- Chen, Z., Matousek, R., & Wanke, P. (2018). "Chinese Bank Efficiency During the Global Financial Crisis: A Combined Approach Using Satisficing Dea and Support Vector Machines", *North American Journal of Economics and Finance*, 43(September 2017), s. 71–86. <http://doi.org/10.1016/j.najef.2017.10.003>
- Chun, S.-H., & Kim, S. H. (2004). "Data Mining for Financial Prediction and Trading: Application to Single and Multiple Markets", *Expert Systems with Applications*, 26(2), s. 131–139. [http://doi.org/https://doi.org/10.1016/S0957-4174\(03\)00113-1](http://doi.org/https://doi.org/10.1016/S0957-4174(03)00113-1)
- Enke, D., & Thawornwong, S. (2005). "The Use of Data Mining and Neural Networks for Forecasting Stock Market Returns", *Expert Systems with Applications*, 29(4), s. 927–940. <http://doi.org/https://doi.org/10.1016/j.eswa.2005.06.024>
- Hassan, M. R., Nath, B., & Kirley, M. (2007). "A Fusion Model of HMM, ANN and GA for Stock Market Forecasting", *Expert Systems with Applications*, 33(1), s. 171–180. <http://doi.org/https://doi.org/10.1016/j.eswa.2006.04.007>
- Hsu, S. H., Hsieh, J. P. A., Chih, T. C., & Hsu, K. C. (2009). "A Two-Stage Architecture for Stock Price Forecasting by Integrating Self-Organizing Map and Support Vector Regression", *Expert Systems with Applications*, 36(4), s. 7947–7951. <http://doi.org/10.1016/j.eswa.2008.10.065>
- Hua, S., & Sun, Z. (2001). "Support vector machine approach for protein subcellular localization prediction", *Bioinformatics*, 17(8), s. 721–728.
- Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). "Forecasting Stock Market Movement Direction with Support Vector Machine", *Computers & Operations Research*, 32(10), s. 2513–2522. <http://doi.org/https://doi.org/10.1016/j.cor.2004.03.016>
- Kara, Y., Acar Boyacioglu, M., & 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), s. 5311–5319. <http://doi.org/10.1016/j.eswa.2010.10.027>
- Khemchandani, R., Jayadeva, & Chandra, S. (2009). "Knowledge Based Proximal Support Vector Machines", *European Journal of Operational Research*, 195(3), s. 914–923. <http://doi.org/https://doi.org/10.1016/j.ejor.2007.11.023>
- Kim, K. J. (2003). "Financial Time Series Forecasting Using Support Vector Machines", *Neurocomputing*, 55(1–2), s. 307–319. [http://doi.org/10.1016/S0925-2312\(03\)00372-2](http://doi.org/10.1016/S0925-2312(03)00372-2)
- Kimoto, T., Asakawa, K., Yoda, M., & Takeoka, M. (1990). "Stock Market Prediction System with Modular Neural Networks", In 1990 IJCNN International Joint Conference on Neural Networks (s. 1–6 vol.1). <http://doi.org/10.1109/IJCNN.1990.137535>
- Kumar, M., & Thenmozhi, M. (2005). "Forecasting Stock Index Movement: A Comparison of Support Vector Machines and Random Fores". In Forest, Indian Institute of Capital Markets 9th Capital Markets Conference Paper. (s. 1–16). <http://doi.org/10.2139/ssrn.876544>
- Leung, M. T., Daouk, H., & Chen, A.-S. (2000). "Forecasting Stock Indices: A Comparison of Classification and Level Estimation Models", *International Journal of Forecasting*, 16(2), s. 173–190. [http://doi.org/https://doi.org/10.1016/S0169-2070\(99\)00048-5](http://doi.org/https://doi.org/10.1016/S0169-2070(99)00048-5)
- Malkiel, B. G., & Fama, E. F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work", *The Journal of Finance*, 25(2), s. 383–417. <http://doi.org/10.1111/j.1540-6261.1970.tb00518.x>
- Olson, D., & Mossman, C. (2003). "Neural Network Forecasts of Canadian Stock Returns Using Accounting Ratios", *International Journal of Forecasting*, 19(3), s. 453–465. [http://doi.org/https://doi.org/10.1016/S0169-2070\(02\)00058-4](http://doi.org/https://doi.org/10.1016/S0169-2070(02)00058-4)
- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015a). "Predicting Stock And Stock Price Index Movement Using Trend Deterministic Data Preparation and Machine Learning Techniques", *Expert Systems with Applications*, 42(1), s. 259–268. <http://doi.org/10.1016/j.eswa.2014.07.040>
- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015b). "Predicting Stock Market Index Using Fusion of Machine Learning Techniques. *Expert Systems with Applications*, 42(4), s. 2162–2172. <http://doi.org/10.1016/j.eswa.2014.10.031>

- Tan, T. Z., Quek, C., & Ng, G. S. (2007). "Biological Brain-Inspired Genetic Complementary Learning for Stock Market and Bank Failure Prediction", *Computational Intelligence*, 23(2), s.236–261. <http://doi.org/10.1111/j.1467-8640.2007.00303.x>
- Tay, F. E. H., & Cao, L. J. (2001). "Improved Financial Time Series Forecasting by Combining Support Vector Machines with Self-Organizing Feature Map", *Intelligent-Data-Analysis*, 5, s. 339–354.
- Thawornwong, S., & Enke, D. (2004). "The Adaptive Selection of Financial And Economic Variables for Use with Artificial Neural Networks", *Neurocomputing*, 56, s. 205–232. <http://doi.org/https://doi.org/10.1016/j.neucom.2003.05.001>
- Vanstone, B., & Finnie, G. (2009). "An Empirical Methodology For Developing Stockmarket Trading Systems Using Artificial Neural Networks", *Expert Systems with Applications*, 36(3, Part 2), s. 6668–6680. <http://doi.org/https://doi.org/10.1016/j.eswa.2008.08.019>
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York, NY: Springer.
- Xu, X., Zhou, C., & Wang, Z. (2009). "Credit scoring Algorithm Based on Link Analysis Ranking with Support Vector Machine", *Expert Systems with Applications*, 36(2, Part 2), s. 2625–2632. <http://doi.org/https://doi.org/10.1016/j.eswa.2008.01.024>
- Zhong, X., & Enke, D. (2017). "Forecasting Daily Stock Market Return Using Dimensionality Reduction", *Expert Systems with Applications*, 67, s. 126–139. <http://doi.org/10.1016/j.eswa.2016.09.027>