

Makine Öğrenmesi Teknikleri ile Banka Hisse Senetlerinin Fiyat Tahmini

Prediction of Bank Stocks Price with Machine Learning Techniques

Zinnet Duygu AKŞEHİR
Ondokuz Mayıs Üniversitesi
Mühendislik Fakültesi
Bilgisayar Mühendisliği Bölümü
Samsun, TÜRKİYE
duygu.aksehir@bil.omu.edu.tr
ORCID: 0000-0002-6834-6847

Erdal KILIÇ
Ondokuz Mayıs Üniversitesi
Mühendislik Fakültesi
Bilgisayar Mühendisliği Bölümü
Samsun, TÜRKİYE
erdal.kilic@bil.omu.edu.tr
ORCID: 0000-0003-1585-0991

Öz

Banka hisse senetlerinin bir gün sonraki kapanış fiyatını tahmin etmek zor bir problemdir. Çünkü hisse senetleri fiyatlarını etkileyen çok sayıda etken vardır. Bu çalışmada 1 Ocak 2016 - 9 Mayıs 2019 tarihleri arasındaki veriler ele alınarak bazı banka hisse senetlerinin bir gün sonraki kapanış fiyatları tahmin edilmeye çalışılmıştır. Hisse senedi hareketlerinde doğrusal örüntülere rastlanması sebebiyle tahmin modeli geliştirilirken karar ağacı, çoklu regresyon ve rassal orman yöntemlerinden yararlanılmıştır. Bu yöntemlerle oluşturulan modeller için iki farklı girdi değişken kümesi kullanılmıştır. Birinci veri kümesinde, 46 teknik ve 4 temel göstergeden oluşan toplam 50 gösterge vardır. İkincisi ise teknik göstergelerde yapılan indirgemeler sonucunda 29 teknik gösterge ve 4 temel göstergeden oluşan toplam 33 göstergeden meydana gelmektedir. Bu iki farklı girdi değişken kümesi için her iki modelin tahmin performansı R^2 ölçütü ile değerlendirilmiştir. R^2 sonuçları incelendiğinde, teknik göstergede yapılan indirgemenin modellerin tahmin performansını olumlu yönde etkilediği görülmüştür.

Anahtar Sözcükler: tahmin, hisse senedi fiyatı, karar ağacı, çoklu regresyon, rassal orman, teknik ve temel gösterge.

Abstract

It is a difficult problem to predict the one-day next closing price of bank stocks. Because there are many factors affecting stock prices. In this study, using data from 1 January 2016 to 9 May 2019 date and some bank stocks have been tried to predict the closing prices of the next day. The decision tree, multiple regression, and random forest methods were used in developing the estimation model due to finding linear patterns in stock movements. Two different sets of input variables were used for the models created with these methods. There are 50 indicators consisting of 46 techniques and 4 fundamental indicators in the first input variable set. The second input variable set, as a result of the reduction in technical indicators there is a total of 33 indicators consisting of 29 technical indicators and 4 basic indicators. For these two different input variables sets, the estimation performance of both models was evaluated by the R^2 criteria. When the R^2 results were analyzed, it was seen that the reduction in the technical indicator had a positive effect on the predictive performance of the models.

Keywords: prediction, stock price, decision tree, multiple regression, random forest, technical and basic indicators.

1. Giriş

Hisse senedi getirilerinin tahmin edilebilirliği, yatırımcılar açısından büyük önem teşkil etmekle birlikte araştırmacılar tarafından da en çok araştırılan, merak edilen bir konu haline gelmiştir. Hisse senedi

Gönderme ve kabul tarihi: 22.10.2019 - 11.12.2019

Makale türü: Araştırma

fiyatlarının kesin bir şekilde belirlenmesi oldukça güç bir problemdir. Bunun sebebi olarak, yüksek belirsizlik ve fiyatlardaki hareketlilik (oyunluk-volatilite) gösterilebilir. Ayrıca siyasi olaylar, genel ekonomik durum, diğer borsalardaki hareketler, yatırımcıların beklentileri gibi birçok değişken de fiyat hareketlerini etkilemektedir.

Borsa ve hisse senedi endeks tahmini kapsamında var olan çalışmalar incelendiğinde, yapay zeka ve veri madenciliği tekniklerinin ön planda olduğu görülmektedir. Bu tekniklerden olan yapay sinir ağları (ANN) yöntemlerinin diğer yöntemlere oranla daha sık kullanıldığı göze çarpmaktadır. Yapılan çalışmalar genellikle borsa endeks değerinin tahmin edilmesi yönündedir [1-9].

Örneğin, Akel ve Bayramoğlu, ANN ile IMKB 100 endeksinin tahminini gerçekleştirmek için 4 Ocak 1999-28 Şubat 2001 arasındaki verileri kullanmışlardır. Ağa girdi değişkenleri olarak USD/TL kuru, IMKB işlem hacmi, merkez bankası döviz rezervleri, merkez bankası bir aylık mevduat faiz oranı ve altın borsası değişkenlerini vermişlerdir [1].

Karaatlı ve arkadaşları ise, yapmış oldukları çalışmada IMKB 100 endeks değerini tahmin etmek için regresyon ve yapay sinir ağı modelini kullanmıştır. Tahmin tarihi olarak Ocak 1960-Aralık 2002 belirlenmiş olup veriler aylık olarak ele alınmıştır. Çalışmada tahmin modellerine girdi değişkenleri olarak hazine bonusu faiz oranı, cumhuriyet altın fiyatı, enflasyon oranı, sanayi üretim endeksi, tasarruf mevduatı faiz oranı, döviz kuru ve zaman değişkenleri verilmiştir. Modeller karşılaştırıldığında regresyon modelinin daha başarılı olduğu görülmüştür [2].

Modellerinde farklı giriş değişkenleri kullanan Kutlu ve Badur, yapay sinir ağları yaklaşımı ile IMKB 100 endeksinin tahmin etmek için 2 Temmuz 2001-13 Temmuz 2006 tarihleri arasındaki verileri kullanılmışlardır. Çalışmada girdi değişkenleri olarak önceki günün endeks, Amerikan doları, gecelik faiz oran değerleri ile birlikte önceki günün Fransa, Almanya, İngiltere, S&P500, Brezilya ve Japonya borsa endeks değerlerini kullanılmışlardır. Bu değişkenler kullanılarak 3 farklı model oluşturulmuştur. Önceki günün endeks değeri, Amerikan doları ve gecelik faiz oranı değişkenlerinden oluşturulan birinci modelin daha başarılı sonuçlar ürettiği görülmüştür [3].

Diler, çalışmasında IMKB 100 endeksinin bir gün sonraki yönünü ANN yöntemiyle tahmin etmeye çalışırken modelin girdi değişkenleri olarak 10 günlük basit hareketli ortalama, 5 ve 10 günlük ağırlıklandırılmış hareketli ortalama, 10 günlük momentum, stokastik gösterge (%K), göreceli güç endeksi (RSI), MACD (12 ve 26 günlük üssel ortalamalar) kullanmış ve IMKB 100 endeksinin günlük getiri değerleri belirlenmiştir. Oluşturulan model için %60,81 başarı oranı elde edilmiştir [4].

Altay ve Satman, IMKB 30 endeksini ANN ve regresyon yöntemleri ile tahmin etmeye çalışırken verileri hem günlük hem de aylık olarak ele almışlardır. Modeller karşılaştırıldığında, her iki durum için (verilerin günlük ve aylık olarak ele alınması) regresyon modelinin daha başarılı olduğu görülmekle birlikte ANN modelinin de genel olarak endeksin yönünü doğru tahmin ettiği belirtilmiştir [5].

Yakut ve arkadaşları, ANN ve destek vektör makineleri (SVM) yöntemleri ile BIST 100 endeks değerini tahmin etmeye çalışmışlardır. Tahmin modeli oluşturulurken analiz dönemi olarak 3 Ocak 2005 - 31 Aralık 2012 belirlenmiş olup girdi değişkenleri; bir gün, iki gün ve üç gün önceki endeks değerleri, önceki günlere ait faiz oranları ve Amerikan doları değeri, haftanın günlerini temsil eden 5 kukla değişkeni ve önceki günlere ait Japonya, Brezilya, İngiltere, Fransa, Almanya borsa endeks değerleri seçilmiştir. Bu girdi değişkenleri ile 9 farklı model oluşturulmuştur. SVM ile yapılan tahminde, önceki günlere ait endeks, Amerikan doları, faiz değeri, diğer yabancı ülkelerin endeks değerlerini ve 5 kukla değişkenini girdi değişkeni olarak alan modelin daha başarılı bir tahmin yaptığı görülmüştür. ANN ile yapılan tahminde ise, önceki günlere ait endeks değerleri, Amerikan doları ve faiz değerlerini girdi değişkeni olarak alan model ile bu modeldeki girdi değişkenlerine 5 kukla değişkeninin de eklendiği modelin diğer modellere göre daha başarılı bir tahmin yaptığı görülmüştür [6].

Benzer şekilde, yurtdışı borsa endeks tahmin çalışmaları da literatürde yer almaktadır. Örneğin Sui ve arkadaşları, yapmış oldukları çalışmada Shangay borsası endeksini destek vektör makineleri ile bir önceki günkü fiyatı baz alarak o günkü yönünü tahmin etmeye çalışmışlardır. Çalışmada, Alexander filtresi, RSI (Relative Strength Index), MFI (Money Flow Index), BB (Bollinger Bands), Chaikin osilatörü (CO), MACD (Moving Average Coverage/Divergence), Stokastik %K, A/D (Accumulation/Distribution) osilatörü ve Williams'

%R teknik göstergeleri kullanılmıştır. Bu teknik göstergelerle yapılan tahmin çalışmasında %54,25 başarı oranı elde edilmiştir [7].

Kim, çalışmasında SVM yöntemini kullanarak KOSPI endeksinin günlük yönünü tahmin etmeye çalışmıştır. Kullandığı tahmin modeline girdi değişkenleri olarak 12 adet teknik gösterge vermiş olup analiz dönemini Ocak 1989 - Aralık 1998 olarak belirlenmiştir. Ayrıca SVM yöntemini geri yayılım sinir ağları (BPN) ve durum tabanlı çıkarsama (CBR) yöntemleri ile karşılaştırmıştır. SVM'nin, diğer iki yöntemle oranla daha iyi performans gösterdiği belirtilmiştir [8].

Inthachot ve arkadaşları ise, 10 adet teknik gösterge kullanarak ANN ve SVM yöntemleri ile Tayland Borsası endeksini tahmin etmeye çalışmışlardır. Bu iki modelin performansı değerlendirildiğinde ANN yöntemiyle oluşturulan modelin daha başarılı olduğu görülmüştür [9].

Yukarıda incelenen yurt içi ve yurt dışı borsa endeks tahmin çalışmaları göz önüne alındığında veri kümesi olarak kullanılan hisse senedi piyasaları, doğrusal olmayan veri olarak kabul edilmiştir. Bu sebeple, tahminlemede doğrusal olmayan veriler için geliştirilmiş olan destek vektör makinesi, yapay sinir ağları, bulanık mantık gibi yöntemlerin daha çok tercih edilmiştir.

Ancak bu çalışma kapsamında kullanılacak olan 5 büyük bankanın hisse senedi piyasası verileri incelendiğinde artış ve azalış trendlerinin doğrusal örüntülere sahip olduğu görülmüştür. Bu nedenle, değer tahmininde ANN ve SVM gibi yöntemlerin yanı sıra doğrusal verileri tahmin etmeye yönelik yöntemlerin de ele alınması gerektiği kanaatine varılmıştır.

Bu çalışmada, BIST 100 içinde yer alan 5 büyük banka hisse senedinin (AKBNK, GARAN, HALKB, ISCTR ve YKBNK) bir gün sonraki kapanış fiyatı, geliştirilen modele girdi teşkil edecek 46 teknik gösterge [10], ardından teknik göstergeler üzerinde bir öznitelik azaltma çalışması yapılarak belirlenen 29 teknik gösterge ve bunlara eklenen 4 temel gösterge birlikte karar ağaçları, çoklu regresyon ve rassal orman yöntemleri ile fiyatları tahmin edilmeye çalışılmış ve elde edilen tahmin sonuçları karşılaştırılmıştır.

Çalışma şu şekilde organize edilmiştir: Bölüm 2'de verilerin nasıl elde edildiği ve uygulamada kullanılacak veri kümesinin özellikleri hakkında bilgi

verilmiştir. Bölüm 3'de uygulamada kullanılan karar ağaçları, çoklu regresyon ve rassal orman yöntemleri detaylandırılmıştır. Bölüm 4'te tahmin yöntemlerine ait bulgular tablolar ve grafikler halinde verilerek yorumlanmış, son bölümde ise elde edilen tahmin sonuçları verilerek yapılacak olan gelecek çalışmalardan söz edilmiştir.

2. Veri Kümesinin Hazırlanması

Bu bölümde çalışmada tahminleme modellerinde kullanılacak özniteliklerin seçildiği veri kümesinin nasıl oluşturulduğu hakkında bilgiler verilecektir.

3. Öznitelik Seçimi

Tahmin modelinde kullanılacak özniteliklerin belirlenmesi çalışmanın en can alıcı noktalarından bir tanesidir. Öznitelikler, literatürde var olan çalışmalardan esinlenilerek [4-7, 10] belirlenmiştir. Bu incelemeler sonucunda, 5 büyük banka hisse senedinin bir gün sonraki kapanış fiyatının belirlenmesinde bu bankalara ait hisse senedi verilerinden (açılış, kapanış, gün içi en yüksek, gün içi en düşük ve hacim değerleri) elde edilen 46 adet teknik analiz, S&P 500, USD/TL paritesi, petrol ve MSCI Turkey ETF kapanış değerleri ise 4 temel gösterge olarak belirlenmiştir.

4. Teknik Analiz ve Göstergeler

Piyasadaki fiyat hareketlerini ve işlem hacmini inceleyerek piyasanın bir sonraki eğilimini ve yasanın iyimserlik-kötümserlik seviyesini tahmin etmeye çalışan analiz yöntemidir. Teknik analiz, her şeyi fiyat belirler ve geçmişte meydana gelen bir davranışın gelecekte de meydana geleceği yaklaşımlarına dayanmaktadır. Teknik analiz, çeşitli yöntemlerle tahmin edilmeye çalışılan varlığın fiyatlardaki gidişatını ve bu varsayımları temel alarak, bize bu varlık ile ilgili fikirler sunmaktadır.

Teknik analiz yapılırken, trendleri belirlemek ve oluşumları çözümlenmek zordur. Bu zorluğun üstesinden gelmek için fiyat ve işlem verilerini kullanıp bazı matematiksel hesaplamalar yapılarak fiyat hareketlerinin sayısal değerlerle ifade edilmeye çalışıldığı teknik analiz aracına gösterge adı verilmektedir [10]. Çok fazla teknik analiz göstergeleri mevcut olup, bu göstergeler verdikleri bilgiler dahilinde momentum, güç, oynaklık ve trend olmak üzere 4 ana başlık altında toplanmaktadır.

Momentum Göstergeleri

Ele alınan varlığın seyretmiş olduğu hareket yönünden ziyade belirlenmiş zaman dilimi içerisinde bu varlığa ait fiyatların ne kadar hızlı değiştiğini ölçmeyi amaçlamaktadır. Mass Index, Relative Strength Index, Stochastic Oscillator, Trix göstergeleri momentum göstergelerine örnek olarak verilebilir.

Trend Göstergeleri

Varlığın, belirlenen zaman dilimi içerisinde göstermiş olduğu trendin devam edip etmeyeceğini belirlemeye çalışmaktadır. Kısaca, o varlığa ait fiyatların istikrarlı bir şekilde aşağı, yukarı veya yatay yönde gidip gitmediklerini anlamaya çalışan göstergelerdir. Trend göstergelerine Moving Average Covergence-Divergence, Parabolic SAR, Commodity Channel Index örnek olarak verilebilir.

Oynaklık Göstergeleri

Fiyatlardaki oynaklığı ölçen göstergelerdir. Oynaklığın değişmesi fiyatların da yön değiştireceği anlamına gelmektedir. Bir başka deyişle, fiyatlardaki iniş çıkışların ne derece agresif olduklarını ölçmeye yarayan göstergeler olarak tanımlayabiliriz. Bu göstergelerden bazıları şunlardır: Average True Range, Bollinger Bands, Relative Volatility Index, Chaikin's Volatility, Moving Average.

Güç Göstergeleri

Piyasanın gücünü ölçmeye yarayan göstergelerdir. Güç göstergelerine örnek olarak Accumulation/Distribution Index, Chaikin Money Flow, Demand Index, Volume Rate of Change, Price Volume Trend göstergeleri örnek olarak gösterilebilir [11].

Bu çalışmada ilk önce, yakın zamanda gerçekleştirilmiş olan [10] numaralı tez çalışması esas alınarak 8 tane güç, 8 tane momentum, 21 tane trend ve 9 tane oynaklık teknik göstergesi oluşturulmuştur. Ardından, oluşturulan bu göstergeler hisse senetlerinin fiyatlarını tahmin etmede gerçekten anlamlı mı, daha az teknik göstergeler ile bu tahminleme işlemi gerçekleştirilebilir mi sorularına yanıt aramak için detaylı bir araştırma yapılmıştır. Yapılan araştırmalarla, hisse senetlerindeki hareketliliği tespit etmek için oynaklık ve trend göstergelerine daha fazla ağırlık verilmesi kanaatine varılmıştır. Bu doğrultuda bazı göstergelerin aslında fiyat tahminine çok fazla bir etkisinin olmadığı görülmüştür. Önemli olan teknik göstergelerin azaltılması ile kalan 4 tane güç (Chaikin Money Flow, On Balance Volume, vb.), 7 tane oynaklık (Average

True Range, Bollinger Band, Kelter Channel Central, vb.), 6 tane momentum (Relative Strength Index, Stochastic Oscillator, Money Flow Index, vb.) ve 12 tane trend (Mass Index, Aroon Indicator, Commodity Channel Index, Exponential Moving Average, vb.) göstergesi modellerde kullanılacak teknik göstergeler olarak belirlenmiştir.

4.1. Verilerin Elde Edilmesi

Çalışmada kullanılan banka verileri (açılış, kapanış, hacim, gün içindeki en yüksek ve en düşük değer) Borsa İstanbul'dan alınırken S&P 500, USD/TL paritesi, petrol ve MSCI Turkey ETF kapanış değerleri ise Polonya merkezli yayın yapan stooq.com web sitesinden temin edilmiştir.

5. Yöntem

Çalışmada, 5 büyük banka hisse senedinin bir gün sonraki kapanış fiyatı çoklu regresyon, karar ağaçları ve rassal orman yöntemleri kullanılarak tahmin edilmeye çalışılmıştır.

5.1. Çoklu Regresyon (Multiple Linear Regression - MLR)

Bağımlı değişken ve iki veya daha fazla bağımsız değişken arasındaki ilişkiyi belirlememizi sağlayan çoklu regresyon yöntemi, başta istatistik alanı olmak üzere diğer pek çok alanda sıklıkla kullanılan yöntemlerden birisidir [12].

Öncelikle, bağımlı değişken ve bu bağımlı değişkeni etkileyen iki veya daha fazla bağımsız değişkenin bulunduğu veri kümesi için kurulan modeldeki parametreler tahmin edilmektedir. Bu parametrelerin tahmin edilmesinin ardından bağımsız değişkenin bilinen değerleri ile bağımlı değişkenin alacağı değerler hesaplanır. Bağımlı değişken ve n tane bağımsız değişken arasındaki ilişki (1) numaralı denklem yardımıyla kurulmaktadır.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (1)$$

(1) numaralı denklemde yer alan Y , bağımlı değişkeni ifade ederken X_i ise bağımsız değişkenleri ifade etmektedir. Ayrıca denklemde regresyon katsayıları β_i , hata terimi de ε ile verilmiştir.

Bu çalışmada, banka hisse senetlerinin geçmiş değerleri (açılış, kapanış, hacim, gün içi en yüksek ve en düşük değer), 4 temel gösterge ve teknik göstergeler (önce 46 ardından 29 teknik gösterge) bağımsız değişkenleri oluştururken hisse senedinin bir gün sonraki kapanış fiyatı ise bağımlı değişkeni oluşturmaktadır.

5.2. Karar Ağacı (Decision Tree - DT)

Karar ağaçları kavramı ile sınıflandırma problemlerinde sıklıkla karşılaşırsak da aslında bu yöntem regresyon modelleri üzerine de inşa edilmektedir. Regresyon modeli sayısal bir sonuca ulaşmaya çalışmak için kullanılırken sınıflandırma modeli ise kategorik veriler üzerinde kullanılmaktadır.

Karar ağacı regresyon modeli öncelikle veri kümesindeki bağımsız değişkenleri bilgi kazancına göre aralıklara bölmektedir. Tahmin işlemi gerçekleştirilirken, sorgulanan değer hangi aralıkta ise o aralığın ortalama değeri geri döndürülmektedir. Belli aralıklarda istenen tahminlerin aynı sonucu üretmesi, karar ağacı regresyon modelinin, regresyon modelindeki gibi sürekli olmadığını göstermektedir.

Regresyon modelinde sınıflandırma söz konusu olmadığı için entropi, Gini index gibi ayırım kuralları kullanarak ağaç oluşturulmamaktadır. Regresyon-karar ağacındaki ayırımlar, (2) numaralı denklemden yararlanılarak artıkların karelerini azaltma algoritmasına göre gerçekleştirilmektedir [13].

$$\operatorname{argmin}[P_l \operatorname{Var}(Y_l) + P_r \operatorname{Var}(Y_r)]x_j \leq x_j^R \quad (2)$$

$$j = 1, \dots, M$$

(2) numaralı denklemde yer alan P_l ve P_r değişkenleri sırasıyla sol ve sağ düğümlerin olasılıklarını ifade etmektedir. Değişkenler x_j ile gösterilirken x_j^R , x_j değişkeni için en iyi ayırım değerini belirtmektedir. Sağ ve sol alt düğümler için sorumlu vektörler $\operatorname{Var}(Y_l)$ ve $\operatorname{Var}(Y_r)$ 'dir. Ayrıca M , eğitim setindeki değişken sayısını ifade etmektedir.

5.3. Rassal Orman (Random Forest - RF)

Rassal orman, yaygın olarak kullanılan makine öğrenmesi modellerinden bir tanesidir. Hiperparametre kestirimine gerek kalmadan iyi sonuç veren bu model, sınıflandırma ve regresyon problemlerine de uygulanabilir olması nedeniyle birçok alanda yaygın bir şekilde kullanılmaktadır. Rassal orman yöntemindeki temel mantık, çok sayıda karar ağaçlarının bir araya gelerek adından anlaşılacağı üzere bir orman, ağaç topluluğu oluşturmasıdır. Bu topluluk içerisindeki her bir karar ağacı aslında, kullanılan veri setinden rastgele seçilen alt kümeleri temsil etmektedir. Bu modelle yeterli kadar karar ağacı oluşturulmuş ise modelin performansını olumsuz yönde etkileyen aşırı

ezberleme (overfitting) probleminin de meydana gelme olasılığı azaltılmış olur [14].

6. Bulgular

Çalışma kapsamında oluşturulan modellerin banka hisse senetlerinin bir gün sonraki endeks tahmin sonuçları değerlendirilecektir. Çalışmada analiz dönemi 1 Ocak 2016 - 26 Nisan 2019 olarak belirlenmiş olup veriler günlük olarak ele alınmıştır. Ayrıca 29 Nisan 2019 ile 9 Mayıs 2019 tarihleri arasında 9 günlük gerçek zamanlı tahminleme işlemi gerçekleştirilmiştir.

Modellerin performansı değerlendirilirken aşağıda verilen denklem ile belirtilen R^2 (R Square) ölçütünden yararlanılmıştır.

$$R^2 = 1 - \frac{\text{Hataların kareleri toplamı}}{\text{Her noktanın ortalamaya uzaklığının kareleri toplamı}}$$

Bu ölçüte R^2 değerinin 1'e yakın olması modelin çok başarılı tahminler yaptığını gösterir.

Geliştirilen çoklu regresyon, karar ağacı ve rassal orman modellerinde girdi değişkenleri olarak 46 teknik ve 4 temel gösterge kullanılmıştır. Modelin girdisini oluşturan toplam bu 50 gösterge ile AKBNK, GARAN, HALKB, ISCTR ve YKBNK hisse senetlerinin bir gün sonraki endeks değerleri tahmin edilmeye çalışılmıştır. Ardından teknik göstergelerde detaylı incelemelerle indirgenme yapılmış olup sayıları 46 olan teknik göstergeler 29'a düşürülmüştür. Bu indirgeme işleminden sonra yine aynı hisse senetleri üzerinde 33 girişli model ile endeks tahmini yapılarak ilk durumla olan sonuçlar karşılaştırılmıştır.

Karar ağacı, çoklu regresyon ve rassal orman yöntemleri kullanarak 5 banka hisse senedinin 50 ve 33 gösterge ile geliştirilen modellerin R^2 değerleri Çizelge-1'de verilmiştir.

Çizelge-1 incelendiğinde teknik göstergelerde yapılan indirgeme işleminin başarılı olduğu fakat iki farklı teknik gösterge ile yapılan tahmin sonuçlarının birbirine çok da uzak olmadığı görülmektedir.

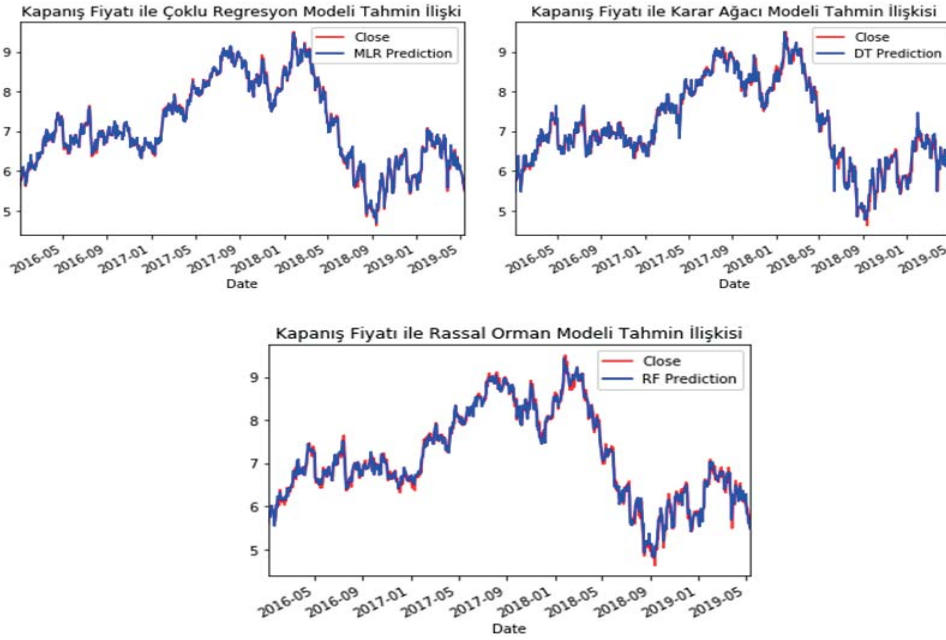
Çizelge-1: Oluşturulan Modellerin R² Değerleri

Hisse Senetleri	R ² Değerleri					
	Karar Ağacı Modeli		Çoklu Regresyon Modeli		Rassal Orman Modeli	
	50 Gösterge	33 Gösterge	50 Gösterge	33 Gösterge	50 Gösterge	33 Gösterge
AKBNK	0,961315	0,960465	0,979386	0,980412	0,977671	0,978688
GARAN	0,960927	0,965797	0,979243	0,980124	0,979468	0,980155
HALKB	0,980088	0,980943	0,989977	0,990256	0,989163	0,989944
ISCTR	0,983219	0,983122	0,987918	0,988527	0,98897	0,988382
YKBNK	0,959043	0,95977	0,984721	0,985127	0,980997	0,980665

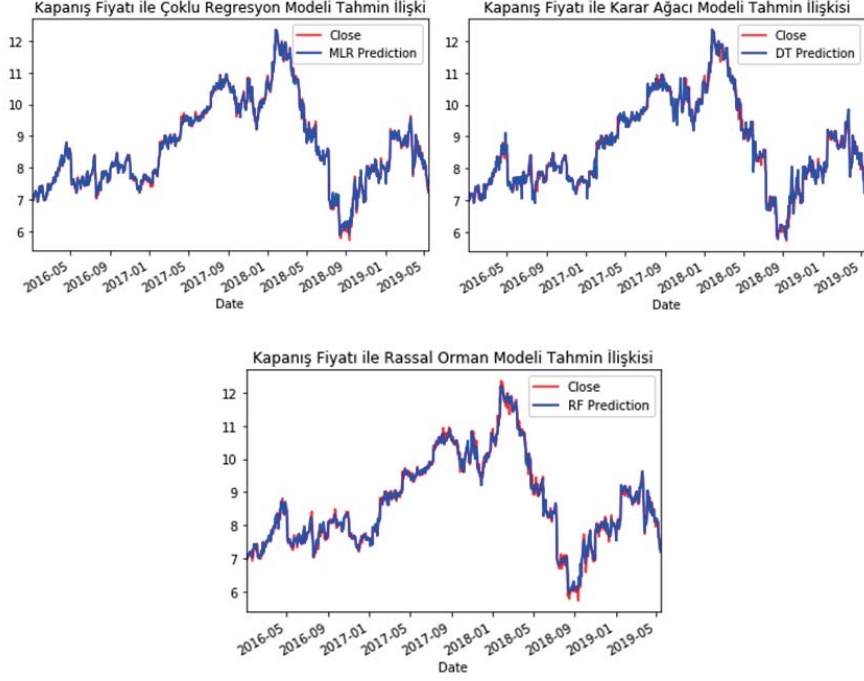
Ayrıca Çizelge-1 incelendiğinde R² değerlerinin 1'e çok yakın olduğu ve dolayısıyla oluşturulan modellerin oldukça başarılı tahminler yaptığı dikkat çekmektedir.

AKBNK, GARAN, HALKB, ISCTR ve YKBNK hisse senetlerinin 29 teknik gösterge ile oluşturulan çoklu regresyon, karar ağacı ve rassal orman modellerindeki başarı grafiği sırasıyla Şekil-1, Şekil-

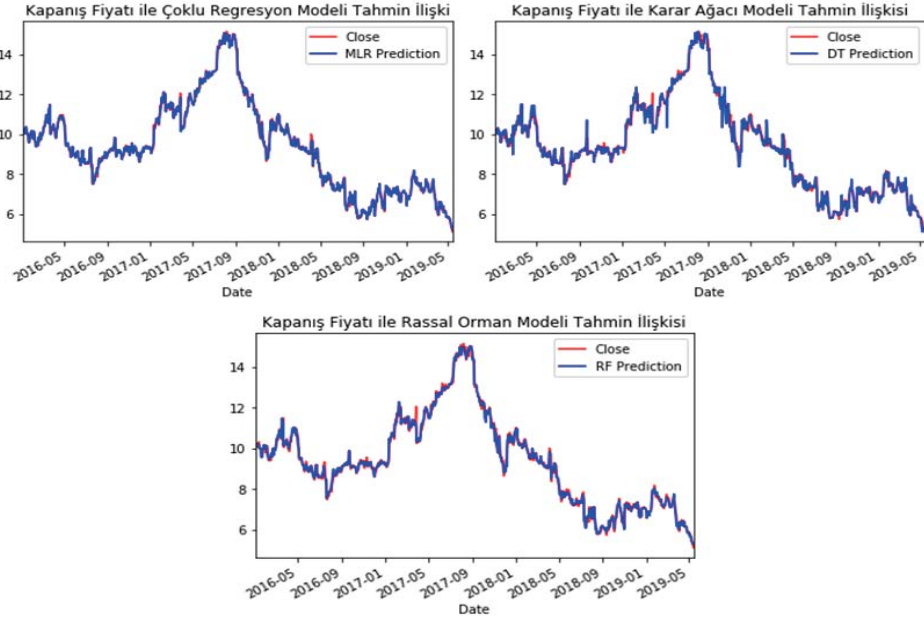
2, Şekil-3, Şekil-4 ve Şekil-5'te verilmektedir. Şekillerde mavi renkle gösterilen eğri modellerin tahmin sonuçlarını ifade ederken kırmızı renkli eğri ise o hisse senedinin kapanış fiyatını ifade etmektedir. Şekiller incelendiğinde iki eğrinin genel olarak üst üste çakıştığı görülmektedir. Bu durum, hisse senetleri üzerindeki tahmin işleminin başarılı olduğunu göstermektedir.



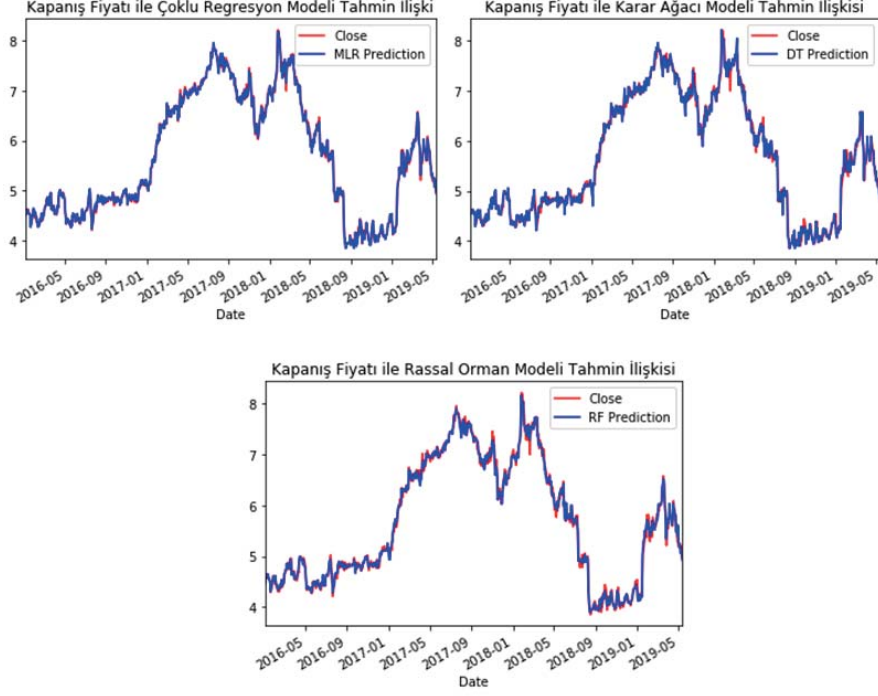
Şekil-1: AKBNK Hisse Senedinin MLR, DT ve RF Modellerindeki Başarı Grafikleri



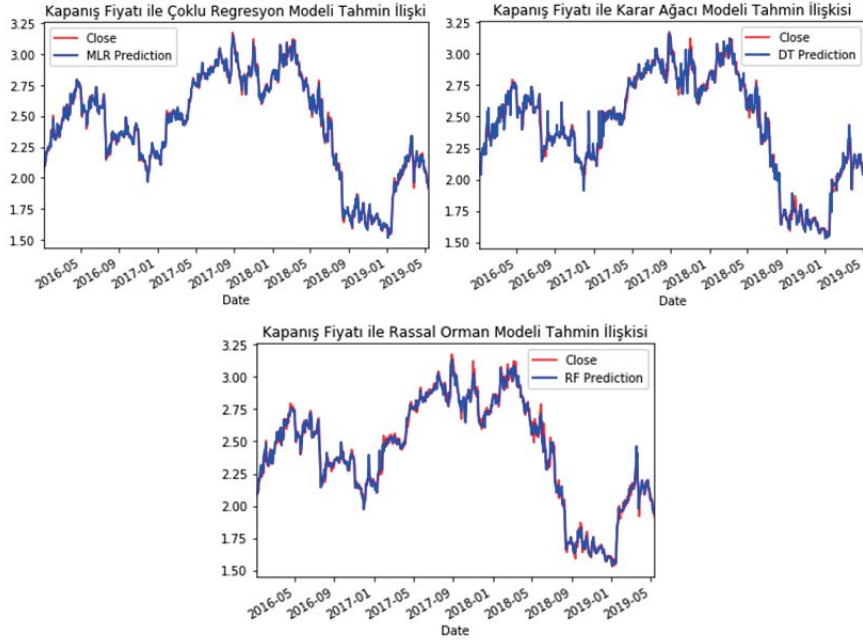
Şekil-2: GARAN Hisse Senedinin MLR, DT ve RF Modellerindeki Başarı Grafikleri



Şekil-3: HALKB Hisse Senedinin MLR, DT ve RF Modellerindeki Başarı Grafikleri



Şekil-4: ISCTR Hisse Senedinin MLR, DT ve RF Modellerindeki Başarı Grafikleri



Şekil-5: YKBNK Hisse Senedinin MLR, DT ve RF Modellerindeki Başarı Grafikleri

Çizelge-2: AKBNK Hisse Senedinin Gerçek Zamanlı Tahminleme Sonuçları

Tarih	Tahmin Sonuçları			Hisse Senedi Kapanış Fiyatı
	Karar Ağacı Modeli	Çoklu Regresyon Modeli	Rassal Orman Modeli	
3 Mayıs 2019	5,31	6,00	5,82	5,99
6 Mayıs 2019	5,31	5,96	5,61	5,93
7 Mayıs 2019	5,31	5,91	5,50	5,84
8 Mayıs 2019	5,31	5,85	5,51	5,81
9 Mayıs 2019	5,31	5,86	5,46	5,67

Çizelge-5: ISCTR Hisse Senedinin Gerçek Zamanlı Tahminleme Sonuçları

Tarih	Tahmin Sonuçları			Hisse Senedi Kapanış Fiyatı
	Karar Ağacı Modeli	Çoklu Regresyon Modeli	Rassal Orman Modeli	
3 Mayıs 2019	5,28	5,24	5,18	5,20
6 Mayıs 2019	5,28	5,25	5,23	5,25
7 Mayıs 2019	5,28	5,28	5,22	5,07
8 Mayıs 2019	5,4	5,08	5,14	5,20
9 Mayıs 2019	5,28	5,24	5,23	5,08

Çizelge-3: GARAN Hisse Senedinin Gerçek Zamanlı Tahminleme Sonuçları

Tarih	Tahmin Sonuçları			Hisse Senedi Kapanış Fiyatı
	Karar Ağacı Modeli	Çoklu Regresyon Modeli	Rassal Orman Modeli	
3 Mayıs 2019	8,14	8,07	8,02	8,14
6 Mayıs 2019	8,61	8,23	8,05	8,02
7 Mayıs 2019	8,14	8,14	8,06	7,77
8 Mayıs 2019	7,99	7,86	7,75	7,75
9 Mayıs 2019	7,51	7,81	7,68	7,60

Çizelge-6: YKBNK Hisse Senedinin Gerçek Zamanlı Tahminleme Sonuçları

Tarih	Tahmin Sonuçları			Hisse Senedi Kapanış Fiyatı
	Karar Ağacı Modeli	Çoklu Regresyon Modeli	Rassal Orman Modeli	
3 Mayıs 2019	2,04	2,04	2,06	2,06
6 Mayıs 2019	2,11	2,05	2,08	2,04
7 Mayıs 2019	2,04	2,03	2,04	1,97
8 Mayıs 2019	2,04	1,97	2,03	1,98
9 Mayıs 2019	2,04	1,98	2,05	1,97

Çizelge-4: HALKB Hisse Senedinin Gerçek Zamanlı Tahminleme Sonuçları

Tarih	Tahmin Sonuçları			Hisse Senedi Kapanış Fiyatı
	Karar Ağacı Modeli	Çoklu Regresyon Modeli	Rassal Orman Modeli	
3 Mayıs 2019	5,98	5,86	6,00	5,81
6 Mayıs 2019	5,98	5,93	5,96	5,68
7 Mayıs 2019	5,98	5,83	5,92	5,61
8 Mayıs 2019	5,81	5,72	5,87	5,55
9 Mayıs 2019	5,81	5,65	5,91	5,41

Ayrıca AKBNK, GARAN, HALKB, ISCTR ve YKBNK hisse senetlerinin, indirgenmiş teknik göstergeler kullanılarak geliştirilen 3 farklı model için 3 Mayıs – 9 Mayıs 2019 tarihleri arasındaki 5 günlük gerçek zamanlı tahminleme sonuçları sırasıyla Çizelge-2, Çizelge-3, Çizelge-4, Çizelge-5 ve Çizelge-6’da verilmiştir.

7. Sonuç ve Gelecek Çalışmalar

Sonuçlar incelendiğinde; borsa endeks tahminlerinden yola çıkılarak, regresyon, karar ağaçları ve rassal orman yöntemleri ile oluşturulan modellerin banka hisse senetlerinin bir gün sonraki kapanış değeri tahmininde oldukça başarılı olduğu görülmüştür. Ayrıca teknik göstergelerde yapılan indirgeme işleminin modelin tahmin başarısını az da olsa arttırdığı gözlemlenmiştir.

Kullanılacak teknik göstergeleri indirgemek için sistematik bir yöntemin geliştirilmesi gelecekte yapılması düşünülen çalışma olarak belirlenmiştir.

Bilgilendirme

Rönesans Holding tarafından desteklenen bu çalışma [15] numaralı çalışmanın genişletilmiş halidir.

8. Kaynakça

- [1] Akel, V., & Bayramoğlu, M. F. *Kriz Dönemlerinde Yapay Sinir Ağları ile Finansal Öngörüle Bulunma: İMKB 100 Endeksi Örneği*. International Symposium on International Capital Flows and Emerging Markets, 24-28 Nisan 2008, Balıkesir, Bandırma.
- [2] Karaatlı, M., Güngör, İ., Demir, Y., & Kalaycı, Ş. *Hisse Senedi Fiyat Hareketlerinin Yapay Sinir Ağları Yöntemi ile Tahmin Edilmesi*. Yönetim ve Ekonomi Araştırmaları Dergisi, 2005, 3(3), 38-48.
- [3] Kutlu, B., & Badur, B. *Yapay Sinir Ağları ile Borsa Endeksi Tahmini*. Yönetim, 2009, 20(63), pp. 25-40.
- [4] Diler, A. İ. *İMKB Ulusal-100 Endeksinin Yönünün Yapay Sinir Ağları Hatayı Geriye Yayıma Yöntemi ile Tahmin Edilmesi*. İMKB Dergisi, 2003, 7(25- 26), pp. 66-82.
- [5] Altay, E., & Satman, M. H. *Stock Market Forecasting: Artificial Neural Network and Linear Regression Comparison in An Emerging Market*. Journal of Financial Management and Analysis, 2005, 18(2), pp. 18-33.
- [6] Yakut, E., Elmas, B., & Yavuz, S. *Yapay Sinir Ağları Ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini*. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 2014, 19(1), pp. 139-157.
- [7] Sui, X., Hu, Q., Yu, D., Xie, Z. & Qi, Z. *A Hybrid Method for Forecasting Stock Market Trend Using Soft-Thresholding De-noise Model and SVM*. Springer Verlag Berlin Heidelberg, 2007, pp. 387-394.
- [8] Kim, K. *Financial Time Series Forecasting Using Support Vector Machines*. Neurocomputing, 2003, 55, pp. 307-319.
- [9] Inthachot, M., Boonjing, V., & Intakosum, S. *Predicting SET50 Index Trend Using Artificial Neural Network and Support Vector Machine*. In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, 2015, pp. 404-414. Springer.
- [10] Aktürk, Z. *BIST100 Endeksinin Günlük Modellenmesi*. Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, 2019.
- [11] Çetinyokuş, T., & Gökçen, H. *Borsada Göstergelerle Teknik Analiz için Bir Karar Destek Sistemi*. Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 2002, 17(1).
- [12] Asiltürk, I., & Çunkaş, M. *Modeling and Prediction of Surface Roughness in Turning Operations Using Artificial Neural Network and Multiple Regression Method*. Expert systems with applications, 2011, 38(5), pp. 5826-5832.
- [13] Breiman, L. *Classification and Regression Trees*. Routledge, 2017.
- [14] Breiman, L. *Random Forests*. Machine learning, 2001, 45(1), pp. 5-32.
- [15] Akşehir, Z. D., & Kılıç, E. *Banka Hisse Senetlerinin İndirgenmiş Teknik Göstergelerle Fiyat Tahmini*. in 2019 4th International Conference on Computer Science and Engineering (UBMK), 2019, IEEE.