

## Amerikan 10 Yıllık Tahvil Faiz Oranlarına Dayanılarak BİST 100 Endeks Tahmininde Ağaç Tabanlı Regresyon Modelleri Uygulaması

*Application of Tree-Based Regression Models or BİST 100 Index Estimate Based on American 10-Year Bond Interest Rates*

Salim Sercan SARI<sup>1</sup>

### öz

Bu çalışmada Borsa İstanbul'da işlem gören BİST 100 endeksinin Amerikan hazine 10 yıllık gösterge tahvil faiz oranları aracılığıyla tahmin edilmesi amaçlanmıştır. Elde edilen 258 adet veri literatürde son yıllarda kullanılan iki adet matematiksel yöntem ile analiz edilmiştir. Zaman serisi alanında kullanılan Rastgele Orman (RF) Modeli ve Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS) Modeli bu çalışmada kullanılan ağaç tabanlı regresyon modelleridir. Kullanılan modellerde BİST 100 endeksi kapanış fiyatları bağımlı değişken; Amerikan hazine 10 yıllık gösterge tahvil faiz oranları bağımsız değişken olarak belirlenmiştir. Analiz aşamasında 206 adet veri modellerin eğitilmesinde, 52 adet veri ise modellerin test edilmesinde kullanılmıştır. Modellerin istatistiksel olarak başarılı olup olmadıkları, hata kareleri ortalaması (HKO) ve Nash-Sutcliffe model verimlilik katsayısı (NSE) başarı kriterleri ile test edilmiştir. Sonuçlar incelendiğinde, MARS modelinin en yüksek NSE değerine sahip olduğu ve Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarının BİST 100 endeksini tahmin edebildiği görülmüştür. Ülkemizde finans alanında yapılan tahminlerde yeni olarak kullanılan bu yöntemler sayesinde daha başarılı yatırım kararlarının alınabileceği düşünülmektedir. Ayrıca çalışma ile oluşturulan modellerin daha sonra geliştirilerek diğer araştırmacılara ışık tutacağı düşünülmektedir.

**Anahtar Kelimeler:** BİST 100 Endeksi, Tahvil Faizi, Rastgele Orman (RF) Modeli, Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS) Modeli.

### ABSTRACT

In this study, it is aimed to estimate the BIST 100 index traded in Borsa Istanbul using the US Treasury 10-year benchmark bond interest rates. The 258 data obtained were analyzed with two mathematical methods used in the literature in recent years. Random Forest (RF) Model and Multivariate Adaptive Regression Spline (MARS) Model used in the time series field are tree-based regression models used in this study. In the models used, the BIST 100 index closing prices are the dependent variable; The US Treasury 10-year benchmark bond interest rates were determined as the independent variable. During the analysis phase, 206 data were used in training the models and 52 data were used in testing the models. Whether the models were statistically successful or not was tested with the success criteria of mean squares of error (MSE) and Nash-Sutcliffe model efficiency coefficient (NSE). When the results are analyzed, it is seen that the MARS model has the highest NSE value and the US Treasury 10-year benchmark bond interest rates can predict the BIST 100 index. It is thought that more successful investment decisions will be made thanks to these new methods used in the estimations made in the field of finance in our country. In addition, it is believed that the models created by the study will be developed later and shed light on other researchers.

**Keywords:** BIST 100 Index, Bond Rate, Random Forest (RF) Model, Multivariate Adaptive Regression Spline (MARS) Model.

Tür: Araştırma makalesi

Gönderim tarihi: 25.09.2021

Kabul tarihi: 28.12.2021

<https://doi.org/10.51945/cuiibfd.1000827>

<sup>1</sup>Arş. Gör. Dr., Erzincan Binali Yıldırım Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, [salim.sari@erzincan.edu.tr](mailto:salim.sari@erzincan.edu.tr) (ORCID: 0000-0003-2607-5249)

## 1.GİRİŞ

Reel ekonomik aktiviteleri çeşitli yollardan etkileyen hisse senedi piyasasına, sermaye piyasalarının geliştiği ülkelerde daha yoğun ilgi gösterilmektedir. Sermaye piyasalarında bireysel ve kurumsal yatırımcılar yaşanan fiyat dalgalanmaları ve temettü getirilerini dikkate alarak yatırım tercihinde bulunmaktadır. Diğer yandan finansal piyasalarda yaşanan olumlu veya olumsuz gelişmelerde, hisse senedi fiyatlarında değişime neden olarak yatırımcıların kararlarını etkileyebilmektedir. Yatırımcılar hisse senedi yatırımlarında karşılaşılabilecek risk ve belirsizlikler karşısında değer kaybının önüne geçebilmek için enflasyonun üzerinde getiri elde etmeyi hedeflemektedirler. Yatırımcılar söz konusu hedefe ulaşabilmek ve portföylerini çeşitlendirebilmek için hisse senedinin yanı sıra borçlanma araçları içerisinde yer alan tahvil de satın almaktadırlar. Tahvil sabit getirili menkul kıymet olarak kabul edildiği için hisse senedine göre daha az riskli olarak değerlendirilebilir.

Finansal piyasalardaki ve yatırım araçlarındaki gelişmelere göre hisse senedi ve tahvile yapılan yatırım tercihi değişebilmektedir. Para politikasında alınan önlemler sonrasında faizlerin artmasıyla yatırımcıların risk ve belirsizlikten korunmak için tahvile olan ilgileri artmaktadır. Ekonomik istikrarda yaşanan büyüme sonrasında ise faizlerin düşmesiyle hisse senedine olan ilgi artış göstermektedir.

1980'den sonra hisse senedi ve tahvil piyasaları arasındaki hareketleri inceleyen birçok araştırma yapılmıştır. Fama ve French (1993), Elton vd. (2001)'in literatüre kazandırdıkları çalışmalardan yola çıkılarak yapılan sonraki çalışmalarda hisse senedi ve tahvillerin fiyat dinamiklerinin bütünleştirilmesi amaçlanmıştır.

Hisse senedi ve tahviller, yatırımcıların portföylerinde en fazla ağırlığı taşıyan ve yatırımcıların aldıkları pozisyonlarda belirleyici olan yatırım araçlarıdır. İki varlığın arasındaki ilişki, bir portföydeki ağırlıklarını ve dolayısıyla risk çeşitlendirmesinin performansını büyük ölçüde etkilemektedir (Fang vd., 2018: 127). Hisse senedi ve tahvil piyasaları bilgi açısından verimli olduğunda, tahvil ve hisse senedi getirileri arasında eş zamanlı bir ilişki gözlemlenebilmektedir. Tahvil piyasası daha az verimliyse, hisse senetleri, dayanak varlıkların değeri hakkındaki bilgileri daha hızlı yansıtmaktadır ve hisse senedi getirilerinin gelecekteki tahvil getirileri için tahmin gücüne sahip olduğu gözlemlenmektedir.(Hotchkiss ve Ronen, 2002: 1326).

Finansal piyasalar üzerinde makroekonomik değişkenler arasında kabul edilen faiz oranlarının da etkide bulunduğu kabul edilmektedir. Fon arz eden veya fon talep eden finansal aktörlerin yatırım kararlarında daha rasyonel hareket edebilmeleri için faiz oranındaki değişimleri dikkate alarak süreci değerlendirmeleri yüksek öneme sahiptir. Amerikan tahvil piyasasında işlem gören Amerikan hazine 10 yıllık gösterge tahvil faiz oranları da küresel olarak risk ve belirsizliğin ölçülmesinde takip edilen en temel değişkenlerdendir. Dünyadaki en büyük tahvil piyasası olarak kabul edilen Amerikan tahvil piyasası finansal piyasalarda yaşanan olumsuzluklarda yatırımcılar tarafından güvenli liman olarak görülmektedir. Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarındaki değişiklikler gelişmekte olan ülkeler tarafından para giriş veya çıkışı şeklinde algılanmakta ve sermaye hareketlerinin yönünü belirlemektedir.

Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarının etkilediği değişkenlerden biri de gelişmekte olan ülke borsalarıdır. Literatürde Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarının ülke borsa endekslerine etkisi veya aralarındaki ilişkinin incelendiği birçok çalışma bulunmaktadır. Diğer taraftan makroekonomik değişkenler arasında yer alan faiz oranları kullanılarak borsa getirilerinin tahmin edildiği çalışmalara da rastlanmaktadır. Bu çalışmada ise ağaç tabanlı regresyon yöntemlerinden olan Rastgele Orman (RF) Modeli ve Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS) Modeli kullanılarak BİST 100 endeksinin Amerikan hazine 10 yıllık gösterge tahvil faiz oranları aracılığıyla tahmin edilmesi amaçlanmıştır. Hisse senedi piyasalarını tahmin etmek için Amerikan tahvil piyasasının kullanıldığı daha önceki çalışmalarda ağaç tabanlı regresyon yöntemlerinden yararlanan herhangi bir çalışmaya ulaşılamamıştır. Bu açıdan çalışmanın literatüre önemli katkı sağlayacağı düşünülmektedir.

Çalışmanın sonraki bölümünde literatür taraması yapılarak benzerlik taşıyan çalışmalar hakkında bilgilere yer verilmiştir. Literatür taramasından sonra çalışmanın veri seti, değişkenler ve yöntemi açıklanmıştır. Analizler sonucu elde edilen bulgulara ise çalışmanın son bölümü olan sonuç ve önerilerden önce değinilmiştir. Bulguların değerlendirilmesi, yorumlar ve gelecekte yapılabilecek çalışmalar için sunulan öneriler ile çalışmaya son verilmiştir.

## 2. LİTERATÜR TARAMASI

Amerikan tahvil piyasası aracılığıyla hisse senedi piyasası ilişkisinin ve tahmininin çeşitli matematiksel modeller yardımıyla finansal piyasalar üzerindeki etkisini inceleyen birçok çalışma gerçekleştirilmiştir. Literatür taraması kapsamında ilgili araştırmalar dikkate alınarak çalışmanın temel yapısı oluşturulmuştur. Söz konusu literatür aşağıda özetlenmiştir.

Keim ve Stambaugh (1986) çalışmalarında tahvil ve hisse senedi fiyatlarının seviyelerini yansıtan önceden belirlenmiş birkaç değişkenin, firmaların hisse senetlerinin getirilerini, temerrüt risklerine sahip uzun vadeli tahvilleri ve temerrütsüz tahvilleri tahmin ettiğini belirtmişlerdir. Campbell (1987) çalışmasında ele aldığı dönemde, faiz oranlarının dönem yapısı dikkate alınarak, hisse senedi getirilerinin yanı sıra bono ve bonoların fazla getirilerinin öngörüldüğünü belgelemiştir. Green (1991) çalışmasında devlet tahvillerinin fiyatını tahmin etmek için metodoloji geliştirmiştir. Fama ve French (1993) çalışmalarında hisse senedi getirilerinin, tahvil piyasası faktörlerindeki paylaşılan varyasyon yoluyla tahvil getirileriyle bağlantılı olduğunu göstermişlerdir. Kirby (1997) çalışmasında hisse senedi ve tahvil getirilerini tahmin etmek için uzun vadeli getirilerin tahmin edilebilir olduğunun aksini öne sürmektedir. Desai ve Bharati (1998) çalışmalarında GARCH modelleri yerine yapay sinir ağları modelleri kullanarak büyük hisse senedi ve şirket tahvilleri arasındaki ilişkilerde daha iyi performans elde etmişlerdir.

Duran vd. (2010) çalışmalarında para politikasının Türkiye'deki hisse senedi fiyatları üzerindeki etkisini, eşzamanlılık sorununa etkili bir çözüm olan değişen oynaklığa

dayalı belirleme yöntemiyle tahmin etmişlerdir. Politika faizlerindeki artışların başta mali sektör endeksi olmak üzere hisse senedi fiyatlarını düşürdüğü sonucuna ulaşmışlardır.

Ohmi ve Okimoto (2016) çalışmalarında hisse senedi-tahvil getiri korelasyonlarındaki olası eğilimleri incelemişlerdir. Sonuçta, birçok gelişmiş, daha güvenli ülke için hisse senedi-tahvil korelasyonlarında önemli düşüş eğilimlerinin varlığını göstermişlerdir.

Cihangir ve Tanrıöven (2016) çalışmalarında ABD 10 yıllık devlet tahvili faiz oranlarındaki değişimin, incelenen ülkelerin para birimlerinin değeri üzerindeki etkisi ve derecesini Alman Filtre tahmin yöntemi ile analiz etmişlerdir. Tahvil faiz oranlarında meydana gelen bir birimlik değişimin bir ülke dışında incelenen diğer ülkelerin kur değişimleri üzerinde negatif bir etki yaptığı görülmüştür.

Scholz vd. (2016) çalışmalarında parametrik olmayan modeller kullanarak yıllık tahvil getirisini, fazla hisse senedi getirilerini tahmin etmek için kullanmışlardır. Mevcut tahvil getirisi bilinmediğinden, onu yine parametrik olmayan teknikler kullanarak önceki bir adımda oluşturmuşlardır. Sonuç olarak tahmin edilen tahvil getirilerinin hisse senedi tahminini önemli ölçüde iyileştirdiği görülmüştür.

Cao vd. (2017) çalışmalarında Kanadalı firmalar tarafından otuz yılı aşkın süredir ihraç edilen tahvilleri içeren yeni bir veri tabanı kullanarak hisse senedi ve tahvil piyasaları arasında firmaya özgü bilgi akışının nasıl olduğunu incelemişlerdir. Sonuç olarak, hisse senedinden tahvil piyasasına akan bilgi akışının güçlü kanıtlarına ulaşmışlardır.

Eyüboğlu ve Eyüboğlu (2018) çalışmalarında 2006 Ocak-2016 Şubat dönemi aylık verilerini kullanarak Görünürde İlişkisiz Regresyon yöntemi ile birlikte Amerikan 10 yıllık tahvil faizleri ile gelişmekte olan ülke borsaları arasında herhangi bir ilişki olup olmadığını test etmeyi amaçlamışlardır. Sonuç olarak ABD 10 yıllıklarının 5 gelişmekte olan ülkenin borsa endeksi üzerinde etkili olduğunu göstermişlerdir.

Fang vd. (2018) çalışmalarında DCC-MIDAS modeline dayalı olarak ABD hisse senedi ve tahvil piyasalarının zamanla değişen uzun vadeli korelasyonu üzerindeki yatırımcı duyarlılığı bileşik endeksinin etkisini araştırmışlardır. Analizde 1997 Asya mali krizi ve 2008 küresel mali krizin yapısal kırılma noktalarını dikkate almışlardır. Sonuçta yatırımcı duyarlılığı bileşik endeksinin uzun vadeli hisse senedi-tahvil korelasyonu üzerinde önemli ölçüde olumlu bir etkiye sahip olduğunu ve kriz şokunun ortalama korelasyonu önemli ölçüde azalttığını, ancak duyarlılığın etkisinin önemli ölçüde değişmediğini göstermişlerdir.

Devpura vd. (2021) 12 tahmin değişkenini kullanarak 25 Ekonomik İşbirliği ve Kalkınma Örgütü ülkesinde tahvil fazlası getiri öngörülebilirliğini test etmişlerdir. Analizde zaman serisi ve panel veri modelleri kullanılmıştır. Elde edilen sonuçlarda petrol spot ve vadeli işlem fiyatları gibi emtialar ile dünya emtia fiyat endeksinin tahvil fazlası getiri sağladığı tahmin edilmiştir. Ulaşılan ikinci bulgu ise, makroekonomik değişkenlerin tahvil fazlası getirileri başarılı bir şekilde tahmin ettiğiidir.

Lee (2021) çalışmasında Linear OLS ve lineer olmayan kantil regresyonları kullanarak Avrupa'daki hisse senetleri ve hazine bonoları arasındaki dinamik birlikteliğin doğasını,

finansal belirsizliğin ana itici güçlerini ve yatırımcıların gelecekteki ekonomik duruma ilişkin beklentilerini incelemiştir. Bulgularda, yatırımcıların ekonominin gelecekteki durumu ve finansal (hisse senedi) piyasa belirsizliği hakkındaki algılarının, Avrupa hisse senetleri ve hazine bonolarının ortak fiyatlandırılmasında kilit faktörler olduğu görülmüştür.

Nazemi vd. (2021) çalışmalarında ABD şirket tahvillerinin geri kazanım oranlarının zamanlar arası analizi için makine öğrenimi yaklaşımlarını uygulamışlardır. Makine öğrenimi tekniklerinin çeşitli zaman dışı tahmin kurulumlarında geleneksel yaklaşımlardan daha iyi performans sergilediğini göstermişlerdir.

Konuya ilişkin literatür taramasından sonra, çalışmada BİST 100 endeksinin Amerikan hazine 10 yıllık gösterge tahvil faiz oranları aracılığıyla tahmin edilmesi amaçlanmıştır.

### 3.MATERYAL ve YÖNTEM

Bu çalışmada investing.com web adresinden temin edilen, Amerikan hazine 10 yıllık gösterge tahvil faiz oranları ve BİST 100 endeksi kapanış fiyat değerleri kullanılmıştır. Zaman serisi değerleri 2000-2021 yıllarına ait 258 adet veriden oluşmaktadır. Ayrıştırma oranı olarak ise, literatürde yapılmış çalışmalar ışığında Bauder ve Khoshgoftaar (2017), Demirkol vd. (2018) zaman serisi değerleri yaklaşık olarak %80'i eğitim, %20'si test olmak üzere iki gruba ayrılmıştır.

Aşağıda çalışmada kullanılan ağaç tabanlı regresyon modellerine bilgiler verilmektedir.

#### 2.1 Rastgele Orman (RF)Yöntemi

Son yirmi yılda, daha doğru ve daha güvenilir sınıflandırmalar elde etmek için çeşitli parametrik olmayan sınıflandırma algoritmaları geliştirilmiştir. Bu algoritmaların en yoğun olarak araştırılan ve yaygın olarak kullanılanları, Sinir Ağları, Destek Vektör Makineleri ve Rastgele Orman'dır. Genel olarak, bu gelişmiş sınıflandırma yaklaşımlarının performansına ilişkin önceki karşılaştırmalı çalışmalarda, Sinir Ağlarının, Destek Vektör Makinelerinin ve RF'nin geleneksel sınıflandırma yöntemlerinden daha iyi performans sergilediği gösterilmiştir. Bu tür gelişmiş sınıflandırıcılar, geleneksel sınıflandırıcılar tarafından sağlanmayan avantajlar sunmaktadır. Örneğin, Sinir Ağları, Destek Vektör Makineleri ve RF'de eğitim verilerinin dağılımı açısından hiçbir istatistiksel varsayımda bulunulmamaktadır. Söz konusu yöntemlerde daha az sayıda eğitim verisi veya spektral olarak karışık eğitim verileri ile bile tatmin edici sınıflandırma sonuçları üretebilmektedir. RF'nin Sinir Ağları ve Destek Vektör Makineleri'nden farklı ise, karmaşık bir sınıflandırma sürecini bir dizi karar sürecine veya ağaçlara bölerek grup şeklinde sınıflayabilmesidir (He vd., 2015: 2253-2254). RF yöntemi regresyon problemlerinin yanı sıra iki sınıflı ve çok sınıflı sınıflandırma problemlerini tek ve çok yönlü bir çerçevede ele almaya izin veren güçlü bir parametrik olmayan istatistiksel yöntemdir (Breiman, 2001). Rastgele Orman (RF) yönteminin tutarlılığı yakın zamanda (Scornet vd., 2015) tarafından kanıtlanmıştır. Pratik bir bakış açısıyla, RF yaygın olarak kullanılmaktadır ve ayarlanacak yalnızca birkaç parametre ile son derece yüksek performans sergilemektedir. RF, birkaç

bağımsız ağacın tanımına dayandığından, birçok ağacın farklı çekirdekler üzerine paralel olarak inşa edildiği bu yöntemin paralel ve daha hızlı bir uygulamasını elde etmek kolaydır (Genuer ve ark., 2017).

RF yönteminde kullanıcılardan ağaç sayısı ve oluşturulması planlanan ağaç yapısı için kullanılacak değişken sayısı istenmektedir. RF yöntemi isteğe bağlı olarak yordayıcı değişkenlerin önemi ve verilerin iç yapısının bir ölçüsü şeklinde iki ek bilgi üretmektedir. Değişkenlerin tanımlanması zor bir kavramdır. Çünkü bir değişkenin önemi diğer değişkenlerle etkileşiminden kaynaklanıyor olabilir. RF algoritması, bir değişkenin önemini, diğerlerinin tümü değişmeden bırakırken o değişken için verilere izin verildiğinde tahmin hatasının ne kadar arttığına bakarak tahmin yapmaktadır (Liaw ve Wiener, 2002). RF tarafından üretilen yakınlık matrisinin  $(i, k)$  ögesi, ağaçların  $i$  ve  $k$  ögelerinin aynı terminal düğümüne düştüğü ağaçların fraksiyonudur. Buradaki mantık, "benzer" gözlemlerin, benzer olmayanlardan daha sık olarak aynı uç düğümlerde olması gerektiğidir (Liaw ve Wiener, 2002). Tüm bu prosedürlerdeki ortak unsur,  $k$ 'inci ağaç için, geçmiş rastgele vektörler  $Q_1, \dots, Q_{k-1}$ 'den bağımsız olarak, ancak aynı dağılımla rastgele bir  $Q_k$  vektörünün üretilmesidir. Bir ağaç, eğitim seti ve  $Q_k$  kullanılarak büyütülmekte ve  $x$ 'in bir girdi vektörü olduğu bir sınıflandırıcı  $h(x, Q_k)$  ile sonuçlandırılmaktadır. Örneğin, bagging rastgele vektöründe  $Q$ , kutulara rastgele atılan  $N$  dardan kaynaklanan  $N$  kutu içindeki sayılar olarak üretilmekte, burada  $N$ , eğitim setindeki örneklerin sayısını ifade etmektedir. Rastgele bölünmüş seçimde  $Q$ , 1 ile  $k$  arasında bir dizi bağımsız rastgele tamsayıdan oluşmaktadır.  $Q$ 'nun doğası ve boyutluluğu, ağaç yapımındaki kullanımına bağlı olarak değişmektedir. Çok sayıda ağaç oluşturulduktan sonra, en popüler sınıfa oy verilmektedir. Bu prosedürlere de rastgele ormanlar denilmektedir (Breiman, 2001).

## 2.2 Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS) Yöntemi

MARS yöntemi çok değişkenli parametrik olmayan regresyona yönelik çok değişkenli uyarlamalı regresyon eğrileri yaklaşımı olarak bilinmektedir. Friedman (1991), MARS yöntemini bir dizi bağımsız değişken ile bağımlı değişken arasındaki ilişkiyi gösteren istatistiksel bir yöntem olarak tanıtmaktadır. MARS yöntemi, doğrusal olmayan ve parametrik olmayan bir regresyon yöntemidir. Eğitim veri setlerinin ayrı bölgelere bölündüğü bu yöntemde, bölünen her bir veri setinin kendi regresyon doğrusuna sahip olduğu bir böl ve yönet stratejisi bulunmaktadır. Bağımsız ve bağımlı değişkenler arasındaki temel işlevsel ilişki hakkında özel bir varsayım gerekli görülmemektedir. Bölümlerin uç noktalarına düğüm denilmektedir. Her bağımsız değişken için düğüm, bu parametrenin aralığındaki herhangi bir veri noktası olabilmekte ve düğümlerin seçimi, uyarlamalı bir regresyon algoritmasına dayalı olarak tüm olası tek değişkenli düğüm konumları üzerindeki aramaları içermektedir. Bir düğüm, bir veri bölgesinin sonunu ve diğerinin başlangıcını işaretlemektedir. Temel işlevler olarak bilinen sonuçtaki parçalı eğriler, modele daha fazla esneklik vermektedir (Goh vd., 2017: 150). MARS yöntemi ile yapılan tahminlerde bir dizi bağımsız değişken ile bağımlı değişken arasındaki ilişki en kolay şekilde, modelin bir varyans analizi sunumundan yorumlanabilmektedir. Burada uydurulan fonksiyon, tek değişkenlerdeki toplamsal temel fonksiyonların ve değişkenler arasındaki etkileşimlerin doğrusal bir kombinasyonu olarak ifade edilmektedir. MARS'ta, düşük ve yüksek dereceli modellerin karşılaştırılması,

değişkenlerin tek tek mi yoksa kombinasyon halinde mi girmesine izin verileceğinin belirlenmesine yardımcı olmaktadır (Sephton, 2001). Amaç, bir yanıt değişkeni  $y$ 'nin bir veya daha fazla bağımsız değişkene  $x_1, \dots, x_n$  bağımlılığını yalnızca gözlemlenen gerçekleştirmeler  $\{y_i, x_{1i}, \dots, x_{ni}\}_{i=1}^N$  verildiğinde modellemektir. Verileri oluşturan sistemin veriyi içeren bir etki alanı  $(x_1, \dots, x_n) \subset D \in R^n$  üzerinden aşağıdaki şekilde tanımlandığı varsayılmaktadır:

$$y = f(x_1, \dots, x_n) + \epsilon \quad (1)$$

$N$  boyutlu argümanın tek değerli deterministik fonksiyonu  $f$ ,  $x_1, \dots, x_n$  üzerindeki  $y$ 'nin birleşik tahmin ilişkisini yakalamaktadır. Beklenen değeri sıfır olarak tanımlanan ek stokastik bileşen  $\epsilon$ , genellikle  $y$ 'nin  $x_1, \dots, x_n$  dışında ne kontrol edilen ne de gözlemlenen miktarlara bağımlılığını yansıtmaktadır. Regresyon analizinin amacı, verileri ilgilenilen  $D$  alanı üzerinde  $f(x_1, \dots, x_n)$ 'ye uygun bir yaklaşım olarak hizmet edebilecek bir  $\hat{f}(x_1, \dots, x_n)$  fonksiyonu oluşturmak için kullanmaktır (Friedman ve Roosen, 1995).

MARS tekniği, yinelemeli bölümlenme regresyon stratejisinin bir genellemesi olarak veya eklemeli modellemenin bir genellemesi olarak görülebilmektedir. Yinelemeli bölümlenmenin aksine, bu yöntemde sürekli türevlerle sürekli modeller üretilmektedir. Katkı modelleri, MARS tarafından uygulanabilen modellerin bir alt kümesidir. Ancak MARS aynı zamanda araştırmacı tarafından belirlenen bir sıraya kadar etkileşimlere izin vermekte ve tek bir genel kriter kullanarak katkı işlevlerinin ve etkileşimlerinin etkileşim sırasını ve karmaşıklığını değiştirmektedir. MARS ayrıca kategorik değişkenleri, iç içe geçmiş değişkenleri ve eksik değerleri modellemek için doğal bir yaklaşım sağlamaktadır (Friedman ve Roosen, 1995).

### 2.3. Model Başarı Ölçütleri

Bu çalışmada, tahmin modellerinin performansının değerlendirilmesinde, hata kareleri ortalaması (HKO), Nash–Sutcliffe model verimlilik katsayısı (NSE) ve belirleme katsayısı ( $R^2$ ) kullanılmıştır. Değerlendirme için kullanılan ölçütler Eşitlik 1, 2 ve 3 (a,b,c)'te verilmiştir.

$$HKO = \frac{1}{n} \sum_{i=1}^n (B_{gi} - B_{ti})^2 \quad (2)$$

$$NSE = 1 - \left[ \frac{\frac{1}{n} \sum_{i=1}^n (B_{ti} - B_{gi})^2}{\frac{1}{n} \sum_{i=1}^n (B_0 - B_{gi})^2} \right] \quad (3)$$

$$R^2 = 1 - \frac{\text{Regresyon Kareler Toplamı (RKT)}}{\text{Genel Kareler Toplamı (GKT)}} \quad (4)$$

$$RKT = \sum (B_{ti} - B_0)^2 \quad (5)$$

$$GKT = \sum (B_{gi} - B_0)^2 \quad (6)$$

Eşitlik 1, 2 ve 3 (a,b,c)'te NSE, Verimlilik Katsayısını;  $n$ , toplam gözlem ve tahmin değeri sayısını;  $B_{ti}$ ,  $i$ . verim tahmin değerini;  $B_{gi}$ ,  $i$ . gerçek verim değerini;  $B_0$ , gerçek verim değerleri ortalamasını ifade etmektedir. NSE, hata karelerinin ortalaması ve gerçekleşen değerlerin varyansına bağlı bir parametre olarak adlandırılmaktadır.



Verimlilik katsayısının 1 olması tahmin başarısının en yüksek değerde olduğunu gösterir (Nash ve Sutcliffe, 1970). Bu nedenle 1'e çok yakın değerler çok başarılı model olarak kabul edilmektedir (Başakın vd., 2019: 758-759).

## BULGULAR

Bu bölümde ağaç tabanlı regresyon modelleriyle yapılan analizlerden önce değişkenlere ilişkin elde edilen tanımlayıcı istatistiklerden Tablo 1'de bahsedilmiştir.

**Tablo 1. Değişkenler Tablosu**

Değişkenler	Minimum	Maksimum	Ortalama	Standart Sapma	Gözlem Sayısı
US10YILLIK	0,533	6,662	3,234	1,340	258
BIST100	76,260	1476,720	579,350	352,231	258

Tablo 1'de görüldüğü üzere bağımlı değişken olan BİST 100 Endeksi kapanış değerlerinin minimum değeri 0,533, ortalaması 3,234 ve maksimum değeri ise 6,662'dir. BİST 100 Endeksi kapanış değerlerinin standart sapması ise 1,340'dır. Bağımsız değişken olan Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarının minimum değeri 76,260, ortalaması 579,350 ve maksimum değeri ise 1476,720'dir. Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarının standart sapması ise 352,231'dir.

Çalışmada kullanılan 2 yöntemin tahmin performansı modellerin ayırt edici özelliklerine göre aşağıda karşılaştırılmıştır.

### RF Yöntemi Bulguları

RF yöntemi ile yapılan analizlerde verilerin %80'i eğitim, %20'si ise test amaçlı kullanılmıştır. Sonuçların başarısını ölçmek için hata kareleri ortalaması (HKO) ve Nash-Sutcliffe model verimlilik katsayısı (NSE) kullanılmıştır. RF'nin BİST 100'ü tahmin için ürettiği modele ait sonuçlar Tablo 2'de gösterilmiştir.



**Tablo 2. Değişkenler için RF Model Başarı Değerleri**

Model	Minumum Eleman Sayısı	Alt Örneklem Sayısı	Test Değerleri		R <sup>2</sup>
			HKO	NSE	
Model 1	2	0,6	45155,650	0,646	0,655
Model Denklemi	BIST100=0,741*US10YILLIK+131,11				

Tablo 2 incelendiğinde kurulan modellerin NSE değerlerinin test kümesinde 0,646 olduğu ve yapılan tahminlerin başarılı sonuçlar verdiği görülmektedir. Ayrıca ağaç mimarisinde artık alt dala ayırma işleminin yapılmayacağı son eleman sayısı ve çapraz doğrulama kullanılacak ayırma yüzdesini belirten alt örnek yüzdesi değerinin değişiminin tahmin sonuçlarına etkisi incelenmiştir.

#### **MARS Yöntemi Bulguları**

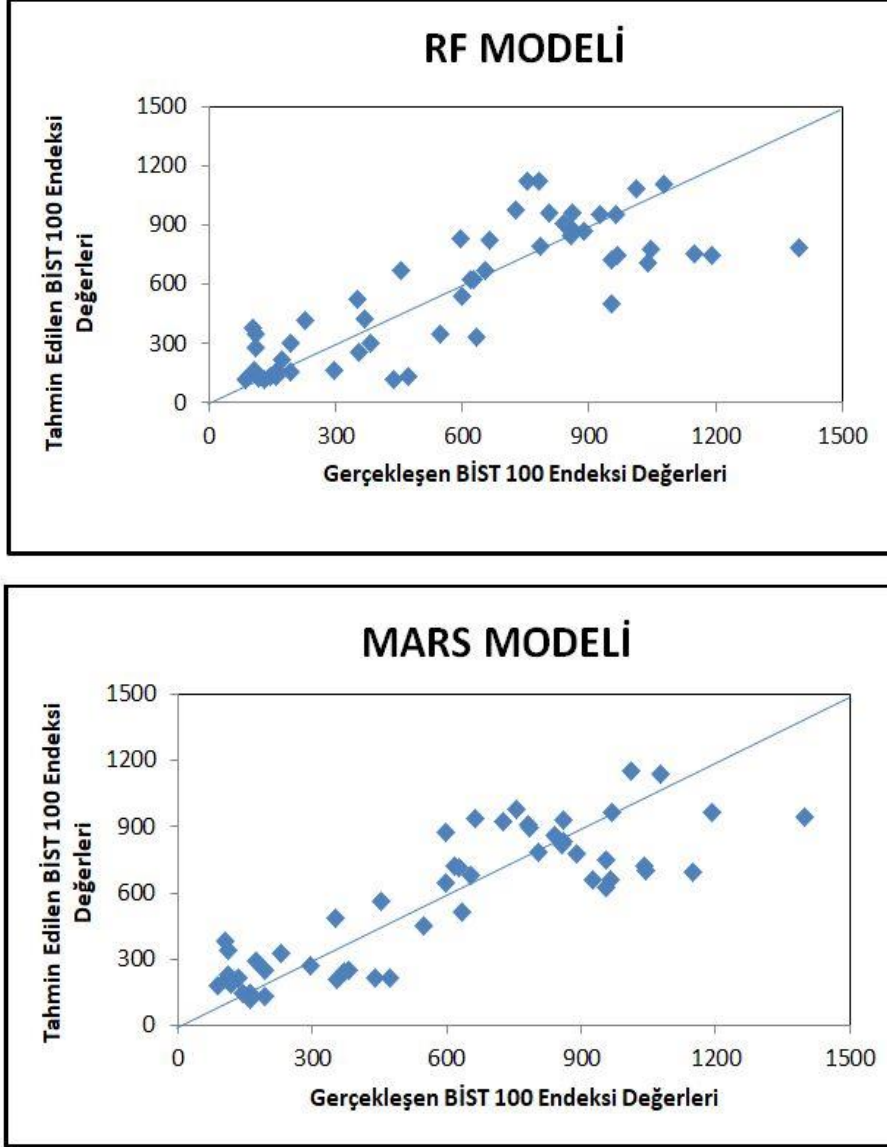
MARS yöntemi ile yapılan analizlerde verilerin %80'i eğitim, %20'si ise test amaçlı kullanılmıştır. Sonuçların başarısını ölçmek için hata kareleri ortalaması (HKO) ve Nash-Sutcliffe model verimlilik katsayısı (NSE) kullanılmıştır. MARS'ın BIST 100'ü tahmin için ürettiği modele ait sonuçlar Tablo 3'de gösterilmiştir.

**Tablo 3. Değişkenler için MARS Model Başarı Değerleri**

Model	Test Değerleri		R <sup>2</sup>
	HKO	NSE	
Model 2	34145,399	0,732	0,731
Model Denklemi	BIST100=0,743*US10YILLIK+132,16		

MARS modeli, R'de "earth" kütüphanesi ile yapılmıştır. MARS modeli parametrelerinde derece sayısı olarak 2,3,4 ve 5 denenmiştir. Çapraz doğrulama işlemi için ise 5 kat ve 10 katlı çapraz doğrulama denenmiştir. Maksimum denklem parametre sayısı ise 50, 100, 150 ve 200 denenmiştir. Budama işleminin yapılması sonrasında Tablo 3'de belirtilen model denklemine ulaşılmıştır. Tablo 3 incelendiğinde kurulan modelin NSE değerinin 0,732 ve olduğu görülmektedir. Çalışmada kullanılan yöntemlere ait bulgular incelendiğinde en başarılı tahmin sonuçlarına MARS yöntemi

ile ulaşıldığı görülmektedir. Her iki yönetime ait modeller aşağıda Şekil 1'de gösterilmiştir.



Şekil 1. Modellerin Gerçekleşen Değer ile Tahmin Değerlerinin Karşılaştırılması

Bu modellerin grafikleri çizilerek görsel olarak karşılaştırma yapılmıştır. Çünkü NSE değerlerinin yanı sıra saçılım grafiklerinin de çizilmesine en uygun modele karar verebilmek için ihtiyaç duyulmaktadır. Tahmin-gözlem saçılım grafiği adıyla tanınan

söz konusu grafiklerde, 45 derecelik bir eğri ile 1 katsayısı temsil edilmektedir. Grafikte gösterilen değerlerin, çizgiye yakın seyretmesi modelin başarılı olduğunu ifade etmektedir. Çizgiden uzaklaşan değerler ise, tahminin başarılı olmadığını ifade etmektedir.

## SONUÇ

Finans dünyasında yaşanan entegrasyonla beraber yatırım tercihi arayışında olan finansal piyasa aktörleri ulusal ve uluslararası kuruluşların paylaştığı verilerden ve göstergelerden yararlanmaktadırlar. Dikkate aldıkları tüm bilgiler ışığında durum analizi yapan finansal aktörler daha sonra yatırım kararı vermektedirler. Analizi veya tahmini yapılan göstergelerden biri de finansal piyasalardaki gelişmelere anında tepki veren hisse senedi piyasalarıdır. Gelişen ülkelerin hisse senedi piyasaları Amerikan hazine 10 yıllık gösterge tahvil faiz oranları ve benzeri makroekonomik göstergelerden etkilenmektedirler. Bu çalışmada Borsa İstanbul'da işlem gören BİST 100 Endeksinin Amerikan hazine 10 yıllık gösterge tahvil faiz oranları aracılığıyla tahmini amaçlanmıştır.

2010-2021 dönemi aylık veri seti ele alınarak tahmin modelleri oluşturulurken verilerin %80'i eğitim, %20'si test için kullanılmıştır. Modellerin oluşturulması sırasında RF ve MARS olmak üzere iki farklı yöntemden faydalanılmıştır. Model başarı sonuçları HKO ve NSE parametreleri yardımıyla yorumlanmıştır. Kullanılan tahmin yöntemlerinin başarı performansları karşılaştırıldığında yöntemler arasında en iyi performansın MARS ile gerçekleştiği görülmüştür.

Finansal piyasalar açısından düşünüldüğünde Covid-19'la birlikte değişen sermaye maliyetlerini dikkate alan yatırımcı aktörlerin ağaç tabanlı yöntemler ile yapacakları tahminlerde etkin sonuçlar elde edebileceği düşünülmektedir. Söz konusu tahmin yöntemlerinin kullanılmasının Türkiye gibi gelişmekte olan ülke yatırımcılarının Amerikan tahvil piyasasında meydana gelen gelişmelere göre risk yönetimi yapabilmesini sağlamaktadır. Risk yönetiminin sermaye giriş ve çıkışlarını kontrol edebilmenin zor olduğu gelişmekte olan ülkelerde, finansal başarının önemli oranda artırılması için fırsat oluşturacağı düşünülmektedir. Amerikan hazine 10 yıllık gösterge tahvil faiz oranları ile ilgili olarak yaptığımız araştırmalarda ülkemizde bu alanda sınırlı çalışma yapıldığı dikkat çekmektedir. Türkiye gibi gelişmekte olan ülkelerin finansal sistemlerinin küresel finansal sisteme ayak uydurabilmesi için Amerikan hazine 10 yıllık gösterge tahvil faiz oranlarından nasıl etkileneceği düşünüldüğünde bu alanda birçok çalışmanın yapılması gerektiği söylenebilir.

Bu çalışmanın ele alınan değişkenler, incelenen konu ve kullanılan yöntemler ile finans literatürüne katkı sağlaması düşünülmektedir. Ulaşılan sonuçlar bundan sonra yapılabilecek çalışmalara yol gösterecektir. Ayrıca her çalışmada olduğu gibi bu çalışmada da sınırlılıklar yer almaktadır. Çalışmada Amerikan hazine 10 yıllık gösterge tahvil faiz oranları aracılığıyla yapılan tahminde Türkiye verileri kullanılmıştır. Bu nedenle farklı yöntemler veya değişkenler ile farklı ülke veya ülke gruplarına ait değişkenler kullanılarak, analizler tekrarlanarak farklı sonuçlar elde edilebilecektir.

Diğer yandan analizin yapıldığı dönem ile analiz yönteminde ki farklı yaklaşımlar ve modeller konunun daha geniş bir çerçeveye de ortaya konulmasına fırsat verebilecektir.

#### Kaynakça

- Başakın, E. E., Özger, M., & Ünal, N. E. (2019). Gri tahmin yöntemi ile İstanbul su tüketiminin modellenmesi. *Politeknik Dergisi*, 22(3), 755-761.
- Bauder, R. A., & Khoshgoftaar, T. M. (2017, December). Medicare fraud detection using machine learning methods. In 2017 16th IEEE international conference on machine learning and applications (ICMLA) (pp. 858-865). IEEE.
- Breiman, L. (2001). Random Forests. *Machine learning*, 45(1), 5-32.
- Campbell, J. Y. (1987). Stock Returns and the Term Structure. *Journal of Financial Economics*, 18(2), 373-399.
- Cao, N., Galvani, V., & Gubellini, S. (2017). Firm-Specific Stock and Bond Predictability: New Evidence from Canada. *International Review of Economics & Finance*, 51, 174-192.
- Cihangir, Ç. K., & Tanrıöven, C. (2016). Abd Devlet Tahvili Faiz Oranlarındaki Değişimin Kurlara Etkisi; Kırılgan Paralar, Kırılgan Ekonomiler. *İşletme Araştırmaları Dergisi*, 8(4), 1-14.
- Demirkol, D., Kartal, E., Şeneler, Ç., & Gülseçen, S. (2019). Bir Öğrenci Bilgi Sisteminin Kullanılabilirliğinin Makine Öğrenmesi Teknikleriyle Tahmin Edilmesi. *Veri Bilimi*, 2(1), 10-18.
- Desai, V. S., & Bharati, R. (1998). The efficacy of Neural Networks in Predicting Returns on Stock and Bond Indices. *Decision Sciences*, 29(2), 405-423.
- Devpura, N., Narayan, P. K., & Sharma, S. S. (2021). Bond Return Predictability: Evidence from 25 OECD Countries. *Journal of International Financial Markets, Institutions and Money*, 101301.
- Duran, M., Özlü, P., & Ünalmis, D. (2010). TCMB Faiz Kararlarının Hisse Senedi Piyasaları Üzerine Etkisi. *Central Bank Review*, 10(2), 23.
- Elton, E. J., Gruber, M. J., Agrawal, D., & Mann, C. (2001). Explaining the Rate Spread on Corporate Bonds. *The Journal of Finance*, 56(1), 247-277.
- Eyüboğlu, S., Eyüboğlu, K. (2018). Amerikan 10 Yıllık Tahvil Faizleri İle Gelişmekte Olan Ülke Borsaları Arasındaki İlişkinin Test Edilmesi. *Yönetim Bilimleri Dergisi*, 16(31), 443-459.
- Fama, E. and K. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33, 3-56.

- Fang, L., Yu, H., & Huang, Y. (2018). The role of Investor Sentiment in the Long-Term Correlation Between US Stock and Bond Markets. *International Review of Economics & Finance*, 58, 127-139.
- Friedman, J. H., & Roosen, C. B. (1995). An Introduction to Multivariate Adaptive Regression Splines.
- Genuer, R., Poggi, J. M., Tuleau-Malot, C., & Villa-Vialaneix, N. (2017). Random Forests for Big Data. *Big Data Research*, 9, 28-46.
- Goh, A. T., Zhang, Y., Zhang, R., Zhang, W., & Xiao, Y. (2017). Evaluating Stability of Underground Entry-Type Excavations Using Multivariate Adaptive Regression Splines and Logistic Regression. *Tunnelling and Underground Space Technology*, 70, 148-154.
- Green, C. J. (1991). 'Quick' Methods of Estimating The Price Of Government Bonds. *Oxford Bulletin of Economics and Statistics*, 53(3), 295-311.
- He, J., Harris, J. R., Sawada, M., & Behnia, P. (2015). A comparison of Classification Algorithms Using Landsat-7 and Landsat-8 Data for Mapping Lithology in Canada's Arctic. *International Journal of Remote Sensing*, 36(8), 2252-2276.
- Hotchkiss, E. S., & Ronen, T. (2002). The Informational Efficiency of the Corporate Bond Market: An Intraday Analysis. *The Review of Financial Studies*, 15(5), 1325-1354.
- Investing, <https://m.tr.investing.com> (Erişim Tarihi: 20.08.2021).
- Keim, D. B., & Stambaugh, R. F. (1986). Predicting Returns in The Stock and Bond Markets. *Journal of Financial Economics*, 17(2), 357-390.
- Kirby, C. (1997). Measuring the Predictable Variation in Stock and Bond Returns. *The Review of Financial Studies*, 10(3), 579-630.
- Lee, H. (2021). Time-Varying Comovement of Stock and Treasury Bond Markets in Europe: A Quantile Regression Approach. *International Review of Economics & Finance*, 75, 1-20.
- Liaw, A., & Wiener, M. (2002). Classification and Regression by Random Forest. *R news*, 2(3), 18-22.
- Nash, J.E., Sutcliffe, J. V. (1970), River Flow Forecasting through Conceptual Models. Part 1: A Discussion of Principles, *J. Hydrol.*, 10 (3), 282–290.
- Nazemi, A., Baumann, F., & Fabozzi, F. J. (2021). Intertemporal Defaulted Bond Recoveries Prediction via Machine Learning. *European Journal of Operational Research*.
- Ohmi, H., & Okimoto, T. (2016). Trends in Stock-Bond Correlations. *Applied Economics*, 48(6), 536-552.

- Scholz, M., Sperlich, S., & Nielsen, J. P. (2016). Nonparametric long term Prediction of Stock Returns with Generated Bond Yields. *Insurance: Mathematics and Economics*, 69, 82-96.
- Scornet, E., Biau, G., & Vert, J. P. (2015). Consistency of Random Forests. *The Annals of Statistics*, 43(4), 1716-1741.
- Sephton, P. (2001). Forecasting Recessions: Can We do Better on Mars. *Federal Reserve Bank of St. Louis Review*, 83(March/April 2001).