

Finansal Zaman Serisi Tahminlerinde Makine Öğrenmesi Uygulamaları

Özge Dinç Cavlak¹

Özet

Mevcut çalışma, finansal zaman serisi tahminlerinde kullanılan çeşitli makine öğrenmesi algoritmalarını inceleyerek, genel bir çerçeve çizmeyi amaçlamaktadır. Çalışmada, finansal zaman serisi tahminlerinde yaygın olarak kullanılan ağaç tabanlı Rastgele Ormanlar (Random Forests, RF), Aşırı Gradyan Artırma (Extreme Gradient Boosting, XGBoost), CatBoost (Categorical Boosting), Hafif Gradyan Artırma Makinesi (Light Gradient Boosting Machine, LightGBM) ve Doğal Gradyan Artırma (Natural Gradient Boosting, NGBoost) modelleri hakkında bilgi verilmektedir. Ardından, girdi değişkenlerin katkısını ölçerek, makine öğrenmesi modellerinin yorumlanabilirliğini arttıran SHAP yaklaşımı açıklanmaktadır. Son olarak, makine öğrenmesi algoritmaları kullanılarak gerçekleştirilen finansal zaman serisi tahminleri, hisse senedi, tahvil ve emtia piyasaları üzerinden tartışılarak, bu modellerin tahmin performansı ve etkililiği gösterilmektedir. Makine öğrenmesi modellerinin hisse senedi ve tahvil piyasalarının yanı sıra, ham petrol, doğal gaz, altın, tarımsal ürünler gibi emtia fiyatlarını tahminde üstün performans gösterdikleri belirtilmektedir. Ayrıca, bu modellerin yeşil enerji hisse senedi ve yeşil tahvil fiyatlarını tahmin etmede etkili oldukları öne sürülerek; yatırım kararları, politika oluşturma ve sürdürülebilir enerji dönüşümü için önemi vurgulanmaktadır. Sonuç olarak yenilikçi makine öğrenmesi modelleri, finansal zaman serisi tahminlerinde oldukça üstün performans göstermekte; tahmin doğruluğu, yorumlanabilirlik ve doğrusal olmayan örüntüleri yansıtabilme gibi avantajlar sunması bakımından önem arz etmektedir.

1 Dr. Öğr. Üyesi, Ankara Hacı Bayram Veli Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, ozge.dinc@hbv.edu.tr, ORCID: <https://orcid.org/0000-0002-7728-983X>

1. Giriş

Tahmin, bireylerin ve örgütlerin proaktif karar alma ve politika geliştirme süreçlerinde önemli rol oynamaktadır (Kumar vd., 2023). Tahmin doğruluğu kadar, tahmin modeli çıktılarını doğru bir biçimde yorumlayabilmenin de kritik bir öneme sahip olduğu ifade edilmektedir (Lundberg ve Lee, 2017). Özellikle finansal zaman serisi modellerinde, yüksek düzeyde rastgelelik ve artıkların bulunmaması gibi durumlar nedeniyle tahmin zorlukları söz konusu olabilmektedir (Kumar vd., 2023). Bu doğrultuda, finansal zaman serisi tahminlerinde doğruluğu artıran ve karmaşık yapıları öngörebilen makine öğrenmesi modelleri tercih edilmektedir. Murphy (2012) makine öğrenmesini, verideki örüntüleri otomatik bir biçimde tespit edebilen bir dizi yöntemi içeren ve bu örüntüleri, gelecek verileri tahmin etmede veya diğer belirsizlik altında karar verme durumlarında kullanan bir yöntem olarak tanımlamaktadır. Bertolini vd. (2021) bu tanımlamayı genişleterek makine öğrenmesini, veriden bilgi üreten bir dizi yöntem ve algoritma olarak tanımlarken; deneyimden öğrenme yeteneğini sürekli olarak geliştirdiğini ifade etmektedir. Geleneksel istatistiksel yöntemler, örnekleme ilişkin dağılımlar ve doğrusallık gibi birtakım varsayımlar gerektirdiğinden, bu yöntemlerin durağan olmayan ve doğrusal olmayan zaman serisi verileri üzerinde etkisi sınırlı olmakta; makine öğrenmesi modellerinin, örneklem dağılımına ilişkin az sayıda varsayım ile doğrusal olmayan örüntüleri yansıtılması bakımından, zaman serileri analizlerinde güvenilir yöntemler olduğu belirtilmektedir (Lv vd., 2022). Bunun yanı sıra, ağaç tabanlı modeller, yaygın kullanılan denetimli makine öğrenmesi yöntemleri arasında yer almakta; tahmin modellerinin doğruluğunu ve yorumlanabilirliğini iyileştirme, doğrusal olmayan ilişkileri yansıtılma, sınıflandırma ve regresyon problemlerine uyarlanabilme yetenekleriyle ön plana çıkmaktadır (Rady vd., 2021).

Makine öğrenmesi modellerinin zaman serisi tahminlerinde sunduğu bu avantajlar değerlendirildiğinde, bu modellerin, doğruluk, yorumlanabilirlik ve etkili tahminler sunarak finansal zaman serisi tahmini literatürüne derin bir anlayış kazandırdığı görülmektedir. Bu kapsamda, mevcut çalışma, ağaç tabanlı makine öğrenmesi algoritmalarını özetleyerek; bu algoritmaların, hisse senedi, tahvil ve emtia piyasalarını tahmin performanslarını değerlendirmektedir. Ayrıca, yeşil enerji hisse senedi, yeşil tahvil ve temiz enerji gibi araçların tahmin performanslarına odaklanan çalışmalara yer vererek, yatırımcılar ve politika yapımcılar için önemli çıkarımlar sunmaktadır.

2. Ağaç Tabanlı Makine Öğrenmesi Modelleri

Mevcut çalışma, finansal zaman serisi tahminlerinde yaygın olarak kullanılan ağaç tabanlı makine öğrenmesi modellerini ele alarak, tahmin bağlamında özetlemeyi amaçlamaktadır.

2.1. Rastgele Ormanlar

Breiman (2001) tarafından geliştirilen Rastgele Ormanlar (Random Forests, RF), her bir karar ağacı için bağımsız ve aynı dağılıma sahip rastgele vektörler üretmekten bir çok karar ağacı oluşturan bir topluluk öğrenimi yöntemidir. Buna göre, bu vektör ve eğitim seti kullanılarak ağaç büyütülür ve bir sınıflandırıcı üretilir. Ormandaki ağaç sayısı arttıkça, genelleme hatası belirli bir değere yakınsamaktadır. Bir ormandaki ağaç sınıflandırıcılarının genelleme hatası, ormandaki bireysel ağaçların gücüne ve aralarındaki korelasyona bağlı olmakta; daha düşük genelleme hatası, rastgele ormanların sınıflandırıcılar arasında daha düşük korelasyona ve daha yüksek güce sahip olduğunu göstermektedir. Büyük sayılar kanunu sayesinde daha fazla ağaç eklendikçe aşırı öğrenme söz konusu olmamakta; bu da RF algoritmasının etkili bir tahmin aracı olduğunu göstermektedir. Rastgele bir vektöre bağlı birden çok karar ağacından oluşan ve bu ağaçların ortalaması alınarak oluşturulan RF tahmin modeli, Eşitlik 1'de verilmektedir;

$$\hat{Y} = \frac{1}{q} \sum_{i=1}^q g_i(x) \quad (1)$$

Burada $g_i(x)$, i . rastgele ağaçlar kümesini ve x , girdi özellikleri vektörünü ifade etmektedir.

2.2. Gradyan Artırma Modelleri

Friedman (2002) tarafından geliştirilen Gradyan Artırma (Gradient Boosting), fonksiyon tahminini parametre uzayından ziyade fonksiyon uzayında sayısal optimizasyon perspektifinden ele alan; sınıflandırma ve regresyon temelli bir topluluk öğrenimi yöntemidir. Gradyan Artırma, en küçük kareler yöntemi ile bir zayıf öğreneni ardışık olarak mevcut sözde artıklar üzerine modelleyerek, eklemeli regresyon modelleri oluşturmaktadır. Sözde artıklar, her bir eğitim veri noktasındaki model değerlerine göre minimize edilmekte olan kayıp fonksiyonunun gradyanı olarak ifade edilmektedir. Sürece rastgeleliğin dahil edilmesi, yakınsama doğruluğunu ve çalışma hızını arttırmaktadır. Her bir adımda eğitim verisi setinin tümünden rastgele olarak bir alt küme çekilerek, bu rastgele seçilen altkümenin zayıf öğreneni eğitmek ve model güncellemesini hesaplamak amacıyla kullanıldığı

ifade edilmektedir. Gradyan Artırma algoritmasının doğruluğunun bu rastgeleştirme süreci sayesinde arttığı; bu artışın derecesinin eğitim seti büyüklüğü, doğru hedef fonksiyon, koşullu dağılım ve zayıf öğrenenin kapasitesine bağlı olduğu belirtilmektedir.

Gradyan Artırma, farklı durumlarda son teknoloji sonuçlar elde edebilen güçlü bir makine öğrenmesi tekniği olarak heterojen özellikler, gürültülü veriler ve karmaşık bağımlılıkların bulunduğu öğrenme problemleri için temel bir yöntem olarak uygulanmaktadır (Prokhorenkova vd., 2018). Bu yöntem, işlevsel bir uzayda zayıf modellerin yinelemeli olarak birleştirilmesiyle güçlü tahminleyicilerin nasıl oluşturulabileceğini teorik olarak açıklamaktadır (Dorogush vd., 2018).

Mevcut çalışmada, *Aşırı Gradyan Artırma* (Extreme Gradient Boosting, XGBoost), *CatBoost* (Categorical Boosting), *Hafif Gradyan Artırma Makinesi* (Light Gradient Boosting Machine, LightGBM) ve *Doğal Gradyan Artırma* (Natural Gradient Boosting, NGBost) modellerine yer verilmektedir.

2.2.1. Aşırı Gradyan Artırma

Chen ve Guestrin (2016) tarafından geliştirilen Aşırı Gradyan Artırma (Extreme Gradient Boosting, XGBoost), ölçeklenebilir ağaç artırımına dayalı bir topluluk öğrenimi yöntemidir. Böylelikle, seyrek verilerin işlenebilmesi, öğrenme sürecinin hızlanması ve daha hızlı model keşfine olanak vermektedir. XGBoost, seyrek veriye duyarlı algoritma, ağırlıklı çeyreklik özetleme, önbellek erişim örüntüleri, veri sıkıştırma ve parçalama gibi önemli sistemler ve algoritmik optimizasyonlar kullanarak en az küme kaynağı ile daha büyük çapta örnekleri modelleyebilmektedir. XGBoost algoritmasına ilişkin tahmin modeli, Eşitlik 2'de belirtilmektedir;

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in F \quad (2)$$

Burada, \hat{y}_i model tahminini, K toplam ağaç sayısını; f_k , k . ağaç tahminini, x_i , girdi özellik vektörünü ve F olası regresyon ağaçları kümesini ifade etmektedir.

2.2.2. CatBoost

Yandex araştırmacıları tarafından geliştirilen gradyan artırma tabanlı CatBoost (Categorical Boosting) algoritması, klasik algoritmaya alternatif olarak permütasyona dayalı sıralı artırma tekniğinin uygulanması ve kategorik özelliklerin işlenmesine yönelik yenilikçi bir alternatif sunmaktadır

(Prokhorenkova vd., 2018). Bu yöntem, kategorik özellikleri başarılı bir şekilde işleyebilen ve bu özellikleri ön işleme aşamasında ele almak yerine eğitim süreci sırasında ele alan bir gradyan artırma algoritması sunmaktadır. CatBoost algoritması, ağaç yapısı seçiminde yeni bir şema kullanarak aşırı öğrenmeyi azaltmakta; hızlı bir eğitim süreci sağlayarak benzer büyüklükteki topluluk modelleri uygulamalarından daha iyi performans gösteren puanlama uygulamaları içermektedir (Dorogush vd., 2018). Sıralı artırma prosedürü uygulanarak geliştirilen tahmin modeli, Eşitlik 3'te verilmektedir;

$$h^t = \arg \min_{h \in H} \frac{1}{n} \sum_{k=1}^n \left(-g^t(x_k, y_k) - h(x_k) \right)^2 \quad (3)$$

Burada $g^t(x_k, y_k)$, koşullu gradyan dağılımını ve $h(x_k)$, temel tahminleyiciyi göstermektedir.

2.2.3. Hafif Gradyan Artırma Makinesi

Hafif Gradyan Artırma Makinesi (Light Gradient Boosting Machine, LightGBM) algoritması, geleneksel gradyan artırma karar ağaçlarının büyük veriyi işlerken karşılaştığı hesaplama karmaşıklığı ve işlem süresi problemlerini gidermek amacıyla Ke vd. (2017) tarafından geliştirilmiştir. Bu algoritma, gradyan tabanlı tek taraflı örnekleme ve ayrıcalıklı özellik desteleme olmak üzere iki yenilikçi teknik kullanarak, hesaplama hızı, doğruluk ve verimli bellek kullanımı açısından üstünlük sağlamaktadır. LightGBM, karar ağaçlarının eğitilmesi sürecinde diğer gradyan artırma algoritmalarında kullanılan seviye bazlı büyüme stratejisi yerine, yaprak bazlı büyüme stratejisini kullanmaktadır (Ke vd., 2017; Üstüner vd., 2020). LightGBM modeli, Eşitlik 4'te verilmektedir.

$$\hat{Y} = \sum_{i=1}^T g_i(x) \quad (4)$$

Burada $g_i(x)$, regresyon ağaçlarını; T ağaç sayısını göstermektedir (Kocaarslan ve Mushtaq, 2024).

2.2.4. Doğal Gradyan Artırma

Duan vd. (2020) tarafından geliştirilen Doğal Gradyan Artırma (Natural Gradient Boosting, NGBoost) algoritması, gradyan artırma temelli olasılıksal regresyon tahmin modelleri ortaya koymaktadır. Buna

göre NGBoost, çok parametrelili bir artırma algoritmasını doğal gradyan ile birleştirerek, varsayılan çıktı dağılımı parametrelerinin gözlemlenen özelliklere göre nasıl değiştiğini verimli bir biçimde ortaya koyan, olasılıksal bir tahmin yöntemidir. NGBoost, gradyan artırma yöntemini olasılıksal regresyona genelleyerek, koşullu dağılımın parametrelerini çok parametrelili bir artırma algoritmasının hedefleri olarak ele almaktadır. Bu bağlamda algoritma, temel öğrenen, sürekli parametrelili herhangi bir dağılım ailesi ve herhangi bir puanlama kuralı ile kullanılabilir; esneklik, ölçeklenebilirlik ve genel uygulanabilirlik gibi avantajlar sağlamaktadır. NGBoost, gerçek değerli parametreler kümesiyle tanımlanan bir dağılım ailesinin seçimine ve parametrelerin değişkenler arasında değişimine olanak vermektedir. Ayrıca NGBoost, geleneksel artırma algoritmalarının aynı verimli karmaşıklığı ile çok sayıda özellik veya gözleme ölçeklenmekte; sınıflandırma ve regresyon problemlerini etkili bir biçimde ele alabilmektedir. Duan vd. (2020) gerçekleştirdikleri deneylerinde çok parametrelili artırma ve doğal gradyanın birlikte çalışarak performansı artırdığını belirtmektedir. Bu algoritma, olasılıksal tahmin için denetimli bir öğrenme yöntemi olarak tanımlanmakta; koşullu olasılık dağılımı $P_\theta(y|x)$ parametrelerini, artırma kullanarak x 'in bir fonksiyonu olarak tahmin etmektedir.

3. SHAP

Oyun teorisine dayalı (Shapley, 1953) SHAP (SHapley Additive exPlanations) değerleri, Lundberg ve Lee (2017) tarafından geliştirilerek, topluluk ya da derin öğrenme gibi karmaşık modellerin yorumlanabilmesine olanak sağlayan bütünlük bir SHAP yaklaşımı sunmaktadır. Yazarlar, basit modellerin daha rahat yorumlanabildiğini ancak bu tür modellerde tahmin doğruluğu sorunları yaşanabildiğini belirterek; karmaşık modellerin daha doğru sonuçlar ürettiğini ifade etmektedirler. Bu kapsamda, karmaşık modellerin kullanılmasıyla birlikte model çıktılarının doğruluğu ve yorumlanabilirliği arasındaki dengenin ön plana çıktığı; bütünlük yenilikçi bir yaklaşım sunularak, model tahminlerinin yorumlanabilmesine olanak verildiği belirtilmektedir. SHAP yaklaşımının, eklemeli özellik katkı yöntemlerini birleştirerek, yerel doğruluk, eksiklik ve tutarlılık özelliklerini barındıran bir yapı ortaya koyduğu ifade edilmektedir.

Lundberg vd. (2018), SHAP değerlerini etkileşim öğelerini kapsayacak şekilde geliştirerek, ağaç topluluk yöntemlerinden elde edilen tahminlerin yorumlanmasında kullanmakta; üstel hız kazanımları, denetimli kümeleme, SHAP bağımlılık grafikleri ve SHAP özet grafikleri ile bu modellere daha derin bir anlayış kazandırmaktadır.

SHAP yaklaşımı, her bir açıklayıcı değişkenin katkısını ölçerek makine öğrenmesi modellerini açıklamakta; doğruluk ve tutarlılık gibi avantajlar sağlayarak modelin yorumlanabilmesine olanak sunmaktadır (Choi vd., 2025).

4. Finansal Piyasaların Tahminlerinde Makine Öğrenmesi Modelleri

Finansal piyasalar, makroekonomik risk faktörleri, jeopolitik riskler, piyasa oynaklıkları ve petrol fiyatı belirsizlikleri gibi pek çok faktörden etkilenmekte; bu nedenle bu piyasaların doğru ve etkin bir biçimde tahmini zorlu bir süreç halini almaktadır. Finansal piyasaların geleneksel zaman serisi modelleri kullanılarak tahmin edilmesinin yanı sıra, yenilikçi makine öğrenmesi yöntemlerinin özellikle hisse senedi, tahvil, enerji, emtia piyasaları ve oynaklık tahminlerinde kullanıldığı örnekler, ilgili yazında yer almaktadır (bkz. Ghallabi vd., 2025; Jabeur vd., 2021; Kocaarslan ve Soytaş, 2023; Lahmiri, 2024; Prasad ve Bakhshi, 2022; Soltani vd., 2024). Makine öğrenmesi algoritmaları, ekonomik ve finansal faktörlerin karmaşık ve doğrusal olmayan özelliklerini dikkate alarak önemli avantajlar sağlamaktadır (Jabeur vd., 2021). Bu nedenle mevcut çalışma, ilgili literatürde son yıllarda yapılan çalışmaları inceleyerek, yenilikçi makine öğrenmesi algoritmalarının finansal zaman serisi tahminlerindeki performanslarına odaklanmaktadır.

4.1. Hisse Senedi Piyasaları

Hisse senedi piyasaları, küresel finansal sistemin temel unsurlarından biri olarak oldukça karmaşık ve kaotik bir yapıya sahip olup; hisse senedi tahminleri, yatırımcıların, analistlerin ve araştırmacıların odak noktası haline gelmiştir (Gupta vd., 2025). Ancak bu karmaşık yapı, hisse senedi fiyat hareketlerinin etkili bir şekilde tahminini güçleştirmektedir (Kara vd., 2011). Bu bağlamda, yenilikçi makine öğrenmesi modellerinin hisse senedi tahminindeki performanslarını değerlendirmek, etkili portföy ve riskten kaçınma stratejileri geliştirebilmek için kritik bir öneme sahiptir. Ghallabi vd. (2025) temiz enerji hisse senedi fiyatlarını ESG (Environmental, Social, Governance) hisse senedi piyasalarını analiz ederek makine öğrenmesi modelleriyle tahmin etmişlerdir. Bu bağlamda, ESG hisse senedi fiyatlarının temiz enerji fiyatları üzerindeki tahmin edici gücünü karşılaştırmak amacıyla RF, XGBoost, CatBoost, LightGBM, SVM (Destek Vektör Makineleri) ve NGBoost algoritmaları kullanılmaktadır. Model performansları, Doğruluk Skoru, F1 Skoru, ortalama karesel hata (Mean Squared Error, MSE) ve Eğri Altında Kalan Alan metrikleri ile karşılaştırılarak, en iyi performans gösteren modelin NGBoost olduğu gösterilmektedir. Araştırmanın sonuçları,

NGBoost modelinin diğer modellere üstünlük sağladığını; temiz enerji hisse senedi fiyatları ile ESG piyasaları değişkenleri arasında güçlü bir ilişki olduğunu ileri sürmektedir. Ayrıca araştırmacılar, SHAP yöntemi kullanılarak şeffaflığın ve yorumlanabilirliğin arttırıldığını belirtmektedirler. Son olarak, makine öğrenmesi modellerinin yeşil enerji hisse senedi fiyatlarını tahmin etmede başarılı olduğu, bu bulgunun yatırım kararları, politika oluşturma ve sürdürülebilir enerji dönüşümü için önemi vurgulanmaktadır.

Bunun yanı sıra, çeşitli ülkelerin hisse senedi piyasaları, makine öğrenmesi yöntemleriyle tahmin edilmekte ve bu modellerin üstün performans gösterdikleri ortaya konmaktadır. Zhao (2021), Nepal örneği üzerinden hisse senedi piyasalarının tahmininde SVM ve XGBoost'un en iyi performans gösteren modeller olduğunu belirtmektedir. Choi vd. (2024), Kore hisse senedi getirilerini çeşitli makine öğrenmesi modelleri kullanarak tahmin etmekte; LightGBM modelinin üstün performans sergilediğini belirtmektedir. Sharma ve Jain (2023), hisse senedi tahminlerinde çeşitli makine öğrenmesi modelleri kullanarak, XGBoost'un hisse senedi tahmininde etkin bir model olduğunu ortaya koymaktadırlar.

4.2. Tahvil Piyasaları

Kocaarslan ve Soytaş (2023) geleneksel tahvil, hisse senedi ve enerji emtia piyasaları ile ekonomik risk faktörlerinin ABD belediye yeşil tahvilleri üzerindeki etkilerini RF, XGBoost, CatBoost ve LightGBM algoritmaları kullanarak COVID-19 öncesi ve sonrası dönem için tahmin etmektedirler. Model performanslarının değerlendirilmesi amacıyla, ortalama mutlak hata (Mean Absolute Error, MAE), belirtme katsayısı (R^2), ortalama karesel hata (Mean Squared Error, MSE) ve kök ortalama karesel hata (Root Mean Squared Error, RMSE) metrikleri kullanılmış ve modellerin tahmin edici güçleri ortaya konmuştur. Veri, eğitim ve test setleri olarak ayrılarak en uygun parametre kombinasyonunu belirlemek amacıyla çapraz doğrulama tekniği uygulanmıştır. Araştırmanın sonuçları, COVID-19 öncesinde en iyi performans gösteren modelin XGBoost; COVID-19 dönemi için ise en iyi performans gösteren modelin CatBoost olduğunu göstermektedir. Ardından SHAP yaklaşımı kullanılarak girdi değişkenlerin yanıt değişkeni üzerindeki göreceli etki büyüklükleri ve önem düzeyleri ortaya konmakta; kısmi etki grafikleri ile piyasalar arasındaki etkileşimler gösterilmektedir. COVID-19 pandemisi öncesinde hisse senedi ve yeşil tahvil piyasaları arasında zayıf bir ilişki bulunurken, geleneksel tahvil piyasaları ile yeşil tahvil piyasaları arasında güçlü bir ilişki olduğu belirtilmektedir. COVID-19 döneminde ise hisse senedi ve enerji emtia piyasalarının yeşil tahvil piyasaları üzerinde

oldukça önemli etkileri olduğu; ayrıca piyasaların makroekonomik risk ve belirsizlik faktörleri ile etkileşim içinde olduğu ifade edilmektedir.

Kocaarslan ve Mushtaq (2024), ABD belediye yeşil tahvil piyasaları ile enerji emtia ve hisse senedi piyasaları arasındaki dinamik koşullu korelasyonları ortaya koyarak, CatBoost, XGBoost ve LightGBM algoritmaları ile dinamik korelasyonları tahmin etmeyi amaçlamaktadırlar. Bunun yanı sıra, döviz kuru, çeşitli ekonomik ve belirsizlik faktörleri ile SHAP yaklaşımı kullanarak, değişkenlerin göreceli önem düzeylerini ortaya koymaktadırlar. Araştırmacılar SHAP analizinin, girdi değişkenlerine sayısal değerler atayarak makine öğrenmesi modellerinin ürettiği tahminlere ilişkin daha derin bir anlayış kazandırdığını; her bir girdi değişkenin tahmine ne ölçüde ve hangi yönde katkı sunduğunun ortaya konduğunu ifade etmektedirler. Araştırmanın sonuçları, XGBoost ve CatBoost modellerinin tahmin doğruluğunun LightGBM modelinden daha üstün olduğunu ortaya koymaktadır.

İşgüzar vd. (2024), Japonya yeşil tahvil endeksini, çeşitli makine öğrenmesi algoritmaları kullanarak tahmin etmekte; RF ve XGBoost modellerinin test veri setinde daha başarılı tahminler ürettiklerini ortaya koymaktadır.

4.3. Emtia Piyasaları

Jabeur vd. (2021), COVID-19 pandemisi sürecinde petrol fiyatlarını tahmin etmek amacıyla LightGBM, CatBoost, XGBoost, RF, Yapay Sınır Ağları (Artificial Neural Networks, ANN) modelleri kullanarak model performanslarını karşılaştırmışlardır. Bu bağlamda, yeşil enerji kaynakları, ESG endeksleri, hisse senedi piyasaları, oynaklık endeksi, çeşitli emtialar, döviz kuru ve Bitcoin gibi faktörlerin petrol fiyatlarını tahmin etmedeki önem düzeyleri, SHAP analizi kullanılarak ortaya konmuştur. Çalışmanın sonuçları, XGBoost modelinin petrol fiyatlarını tahmin etmede diğer makine öğrenmesi modellerinden daha üstün olduğu bulgusuna ulaşırken; makine öğrenmesi modellerinin genel performanslarının, diskriminant analizi ve lojistik regresyona kıyasla daha iyi olduğunu göstermektedir. Ayrıca, petrol, yeşil enerji kaynakları ve altın birbirleriyle oldukça yüksek korelasyon ortaya koymakta; araştırmacılar piyasalar arasındaki bu etkileşimlerin risk yayılımı ve portföy yönetimi bağlamında önemine dikkat çekmektedir. Jabeur vd. (2024), altın fiyatlarını tahmin etmek amacıyla çeşitli makine öğrenmesi modelleri kullanarak, XGBoost algoritmasının diğer modellere üstün geldiğini göstermektedir. Ardından, SHAP yaklaşımı kullanılarak, ham petrol, gümüş, demir fiyatları, döviz kurları, S&P 500 endeksi ve ABD enflasyon oranı açıklayıcı değişkenlerinin, altın fiyatlarını açıklama

kapasitelerini ve XGBoost model çıktılarını yorumlamadaki etkinliği ortaya konmaktadır.

Lahmire (2024), ham petrol, doğal gaz, benzin gibi fosil enerji piyasası emtialarının tahminlerinde çeşitli makine öğrenmesi sistemlerini, Bayes optimizasyon yöntemi kullanarak optimize etmekte; Gauss regresyon sürecinin fosil enerji piyasalarını tahmin etmede en üstün yöntem olduğunu ifade etmektedir. Jin ve Xu (2025), Çin emtia fiyat endeksinin tahmininde, eğitim süreci için çapraz doğrulama ve Bayes optimizasyonu, analiz süreci için ise Gauss süreç regresyonları kullanarak; belirtilen zaman periyodu için modellerin doğru bir şekilde tahmin edildiğini belirtmektedir. Ludovico vd. (2022), şeker, canlı sığır, kahve, etanol, mısır ve soya fasulyesi gibi ürünleri içeren tarımsal emtia piyasalarının tahminlerinde, RF, ANN, SVM, XGBoost gibi çeşitli makine öğrenmesi modelleri kullanmakta; SVM algoritmasının en doğru tahmin veren model olduğunu ifade etmektedirler. Rady vd. (2021) altın fiyatlarını tahmin etmede, karar ağaçları, rastgele ormanlar ve gradyan artırma ağaçlarını kullanmakta; rastgele ormanların, tahmin doğruluğu bakımından en iyi model olduğunu belirtmektedir.

5. Sonuç

Makine öğrenmesi algoritmalarının kullanıldığı finansal zaman serileri çalışmaları incelendiğinde, yatırımcılara, yöneticilere ve politika yapıcılara yönelik birtakım çıkarımlar ve öneriler sunulmakta; çalışma sonuçları bu bağlamlarda tartışılmaktadır. Ghallabi vd. (2025), araştırmalarından elde edilen bulguların, makine öğrenmesi algoritmalarının yeşil enerji hisse senedi fiyatlarını tahmin etmedeki doğruluğunu vurguladığını; yatırımcılara, portföy yöneticilerine ve politika yapıcılara portföy çeşitlendirmesi bağlamında çıkarımlar sunduğunu belirtmektedir. Jabeur vd. (2024), doğru bir tahmin yöntemi kullanmanın yatırımcılara ve politika yapıcılara uygulamada çıkarımlar sunduğunu; merkez bankalarının işlem güvenliği ve rezerv stratejileri bağlamında altın fiyatlarında meydana gelen dalgalanmalar hakkında bilgi sahibi olmaları gerekliliğini vurgulamaktadır. Tang vd. (2022), finansal verinin doğrusal olmayan, karmaşık ve düzensiz yapısı nedeniyle finansal zaman serisi tahminlerinin önemine dikkat çekmekte; yenilikçi modeller geliştirilerek çeşitli araştırmalar yapıldığını belirtmektedir. Ayrıca araştırmacılar, doğru tasarlanmış tahmin modellerinin daha iyi stratejiler geliştirmede ve kararlar almada hükümetlere yardımcı olabileceğini; karlılığı artırmada ise bireylere ve kurumlara yol gösterebileceğini belirtmektedir. Kocaarslan ve Mushtaq (2024), yeşil tahvil piyasalarının tahmininde makine öğrenmesi modellerinin geleneksel tahmin modellerinden daha doğru çıkarımlar sunduğunu belirterek, COVID-19 pandemi dönemi için yeşil

tahviller ile riskli varlıklar arasındaki bağlantılılığa dikkat çekmektedirler. Jabeur vd. (2021), finansal piyasaların ve emtia piyasalarının, COVID-19 pandemisinin başlamasıyla birlikte önemli kayıplar verdiğini; petrol fiyatlarındaki dalgalanmaların tahmin doğruluğunun, geleneksel modellerden ziyade gelişmiş makine öğrenmesi modelleri kullanılarak artırılmasının önemine vurgu yapmaktadır. Soltani vd. (2024), temiz enerji, emtia, yeşil tahvil ve ESG endeksi fiyatlarını, makine öğrenmesi modelleri kullanarak COVID-19 salgını ve Rusya-Ukrayna çatışması dönemleri kapsamında tahmin ederek, risk yönetimi uygulamaları, yeşil varlıkların denetlenmesi ve politika yapıcılar bağlamında değerlendirmektedir. Ayrıca, etkili tahminlerin, optimal portföy dağılımı ve riskten korunma stratejileri bağlamında önemini vurgulamaktadırlar.

Mevcut çalışma, yenilikçi makine öğrenmesi algoritmalarının finansal zaman serisi uygulamaları üzerindeki performanslarını değerlendirerek, yatırımcılara, uzmanlara ve politika yapıcılara yönelik birtakım çıkarımlar sunmayı amaçlamaktadır. Buna göre, mevcut araştırmalar incelendiğinde, yenilikçi makine öğrenmesi algoritmalarının, finansal zaman serisi uygulamalarında tahmin doğruluğu ve model çıktılarının yorumlanabilirliği bağlamında oldukça üstün performans gösterdikleri ortaya konmaktadır. Bu modellerin, geleneksel yatırım araçlarının yanı sıra, yeşil tahvil, temiz enerji hisse senetleri, ESG hisse senetleri gibi sürdürülebilir yatırım araçlarının tahminlerinde de iyi performans sergilediği anlaşılmaktadır. Özellikle iklim riskleri, finansal krizler ve jeopolitik riskler söz konusu olduğunda; bu araçlar yatırımcılara ve uzmanlara portföy çeşitlendirmesi, riskten korunma ve güvenli liman stratejileri sunabilmektedir. Bu bağlamda ESG varlıkları, iklim riski ve jeopolitik riske karşı çeşitlendirme aracı olarak kullanılmakta (Cepni vd., 2023; Yang vd., 2024); ESG hisse senetlerinin Rusya-Ukrayna çatışması sürecinde güvenli liman olduğu belirtilmekte (Katsampoxakis vd., 2024) ve ESG hisse senetlerinin yüksek jeopolitik risk dönemlerinde riskten korunma ve güvenli liman olma özellikleri taşıdığı ifade edilmektedir (Nittayakamolphon vd., 2024). İklim kaynaklı ve jeopolitik küresel risklerin yüksek olduğu dönemlerde finansal piyasaların oynaklığı artmakta; bu durum yatırımcıları önemli ölçüde etkilemektedir. Piyasadaki belirsizliklerin arttığı bu dönemlerde politika yapıcıların sürdürülebilir yatırımları teşvik etmeleri önem arz edecektir.

Kaynakça

- Bertolini, M., Mezzogori, D., Neroni, M., & Zammori, F. (2021). Machine Learning for industrial applications: A comprehensive literature review. *Expert Systems with Applications*, 175, 114820.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Cepni, O., Demirer, R., Pham, L., & Rognone, L. (2023). Climate uncertainty and information transmissions across the conventional and ESG assets. *Journal of International Financial Markets, Institutions and Money*, 83, 101730.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Choi, J. E., Shin, J. W., & Shin, D. W. (2025). Vector SHAP Values for Machine Learning Time Series Forecasting. *Journal of Forecasting*, 44(2), 635-645.
- Choi, W., Jang, S., Kim, S., Park, C., Park, S., & Song, S. (2024). Return prediction by machine learning for the Korean stock market. *Journal of the Korean Statistical Society*, 53(1), 248-280.
- Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*.
- Duan, T., Anand, A., Ding, D. Y., Thai, K. K., Basu, S., Ng, A., & Schuler, A. (2020, November). Ngboost: Natural gradient boosting for probabilistic prediction. In *International conference on machine learning* (pp. 2690-2700). PMLR.
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4), 367-378.
- Ghallabi, F., Souissi, B., Du, A. M., & Ali, S. (2025). ESG stock markets and clean energy prices prediction: Insights from advanced machine learning. *International Review of Financial Analysis*, 97, 103889.
- Gupta, S., Nachappa, S., & Paramanandham, N. (2025). Stock market time series forecasting using comparative machine learning algorithms. *Procedia Computer Science*, 252, 893-904.
- İşgüzar, S., Fendoğlu, E., Şimşek, A. İ., & Türkoğlu, M. (2024). Green Bond Index Price Forecasting: Comparative Analysis of Machine Learning Models. *Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 46(3), 568-589.
- Jabeur, S. B., Khalfaoui, R., & Arfi, W. B. (2021). The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: Evidence from explainable machine learning. *Journal of Environmental Management*, 298, 113511.

- Jabeur, S. B., Mefteh-Wali, S., & Viviani, J. L. (2024). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research*, 334(1), 679-699.
- Jin, B., & Xu, X. (2025). Machine learning predictions of China commodity price indices. *Discrete Mathematics, Algorithms and Applications*, 2550043.
- Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications*, 38(5), 5311-5319.
- Katsamposakis, I., Xanthopoulos, S., Basdekis, C., & Christopoulos, A. G. (2024). Can ESG Stocks Be a Safe Haven during Global Crises? Evidence from the COVID-19 Pandemic and the Russia-Ukraine War with Time-Frequency Wavelet Analysis. *Economies*, 12(4), 89.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30.
- Kocaarslan, B., & Mushtaq, R. (2024). The impact of liquidity conditions on the time-varying link between US municipal green bonds and major risky markets during the COVID-19 crisis: A machine learning approach. *Energy Policy*, 184, 113911.
- Kocaarslan, B., & Soytaş, U. (2023). The role of major markets in predicting the US municipal green bond market performance: New evidence from machine learning models. *Technological Forecasting and Social Change*, 196, 122820.
- Kumar, A., Chauhan, T., Natesan, S., Pham, N. T., Nguyen, N. D., & Lim, C. P. (2023). Towards an efficient machine learning model for financial time series forecasting. *Soft Computing*, 27(16), 11329-11339.
- Lahmiri, S. (2024). Fossil energy market price prediction by using machine learning with optimal hyper-parameters: A comparative study. *Resources Policy*, 92, 105008.
- Ludovico, S. N., Salgado, R. M., Beijo, L. A., Miguel, E. C., & Rezende, M. L. (2022). Commodity price forecasting via machine learning models. *Sigmae*, 11(2), 45-69.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
- Lundberg, S. M., Erion, G. G., & Lee, S. I. (2018). Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*.
- Lv, S. X., Peng, L., Hu, H., & Wang, L. (2022). Effective machine learning model combination based on selective ensemble strategy for time series forecasting. *Information Sciences*, 612, 994-1023.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

- Nittayakamolphon, P., Bejrananda, T., & Pholkerd, P. (2024). Green bonds and ESG stocks as safe haven or hedging asset for other financial assets. *Kasetsart Journal of Social Sciences*, 45(4), 1307-1318.
- Prasad, A., & Bakhshi, P. (2022). Forecasting the direction of daily changes in the India VIX index using machine learning. *Journal of Risk and Financial Management*, 15(12), 552.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31.
- Rady, E. H. A., Fawzy, H., & Fattah, A. M. A. (2021). Time series forecasting using tree based methods. *Journal of Statistics Applications & Probability*, 10(1), 229-244.
- Shapley, L. (1953). A Value for n-Person Games. In: Kuhn, H. and Tucker, A., Eds., *Contributions to the Theory of Games II*, Princeton University Press, Princeton, 307-317.
- Sharma, P., & Jain, M. K. (2023, November). Stock market trends analysis using extreme gradient boosting (XGBoost). In *2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 317-322). IEEE.
- Soltani, H., Taleb, J., Ben Hamadou, F., & Boujelbène-Abbes, M. (2024). Using machine learning to forecast clean energy, commodities, green bonds and ESG index prices: How important is financial stress?. *Euro-Med Journal of Business*.
- Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., ... & Li, J. (2022). A survey on machine learning models for financial time series forecasting. *Neurocomputing*, 512, 363-380.
- Üstüner, M., Abdikan, S., Bilgin, G., & Şanlı, F. B. (2020). Hafif gradyan artırma makineleri ile tarımsal ürünlerin sınıflandırılması. *Türk Uzaktan Algılama ve CBS Dergisi*, 1(2), 97-105.
- Yang, J., Agyci, S. K., Bossman, A., Gubareva, M., & Marfo-Yiadom, E. (2024). Energy, metals, market uncertainties, and ESG stocks: Analysing predictability and safe havens. *The North American Journal of Economics and Finance*, 69, 102030.
- Zhao, S. (2021, May). Nepal stock market movement prediction with machine learning. In *Proceedings of the 2021 5th International Conference on Information System and Data Mining* (pp. 1-7).