

BORSA İSTANBUL-100 ENDEKSİ OYNAKLIK TAHMİNİ: MAKİNE ÖĞRENMESİ YAKLAŞIMLARININ KARŞILAŞTIRMALI ANALİZİ

Hatice BAŞKAYA¹

Gönderim tarihi: 02.06.2026

Kabul tarihi: 24.09.2026

Öz

Türkiye sermaye piyasalarının ana göstergesi olmasının yanı sıra, BIST 100 endeksi, yapısal kırılmalara ve ani fiyat sıçramalarına olan eğilimi nedeniyle volatilité modellemesi için özgün bir ampirik çerçeve teşkil etmektedir. Bu çalışmanın amacı, yüksek asimetrik bilgi duyarlılığı sergileyen gelişmekte olan piyasaların karakteristik bir örneği olan BIST 100 endeksinin günlük volatilité dinamiklerini, parametrik GARCH modelleri ve parametrik olmayan makine öğrenme algoritmaları aracılığıyla karşılaştırmalı olarak analiz etmektedir. Araştırmada, koşullu varyans sürecini tanımlamak için beş farklı GARCH spesifikasyonu (sGARCH, EGARCH, GJR-GARCH, IGARCH, TGARCH) ve dokuz makine öğrenme algoritması (MLP, SVR, RF, KNN, DT, GBDT, XGBoost, SGD, ExtraTrees) kullanılmıştır. Elde edilen ampirik bulgular, TGARCH-Ged modelinin, BIST 100'ün kaldırıcı etkisini en iyi temsil eden GARCH ailesi içindeki model olduğunu göstermektedir. Ayrıca makine öğrenmesi modellerinin hata ölçütleri (MAE ve RMSE) bakımından GARCH modellerini geride bıraktığı; ancak en başarılı makine öğrenmesi algoritmalarının kendi aralarında istatistiksel olarak anlamlı bir performans ayrışması sergilemediği tespit edilmiştir.

Anahtar Kelimeler: Oynaklık Tahmini, Makine Öğrenmesi, Medyan Gerçekleşen Varyans (MedRV), GARCH Modelleri, BİST-100.

Jel Sınıflaması: G17, C58, C45

VOLATILITY FORECAST FOR THE ISTANBUL STOCK EXCHANGE- 100 INDEX: A COMPARATIVE ANALYSIS OF MACHINE LEARNING APPROACHES

Abstract

In Turkey, the BIST 100 index serves not only as the primary indicator of the capital markets but also provides a unique empirical framework for volatility modeling due to its tendency toward structural breaks and sudden price spikes. The aim of this study is to conduct a comparative analysis of the daily volatility dynamics of the BIST 100 index—a characteristic example of emerging markets exhibiting high asymmetric information sensitivity—using parametric GARCH models and non-parametric machine learning algorithms. The study utilized five different GARCH specifications (sGARCH, EGARCH, GJR-GARCH, IGARCH, TGARCH) and nine machine learning algorithms (MLP, SVR, RF, KNN, DT, GBDT, XGBoost, SGD, ExtraTrees) to model the conditional variance process. The empirical findings reveal that the TGARCH-GED specification outperforms other variants within the GARCH family in capturing the leverage effect of the BIST 100 index. Furthermore, although machine learning models systematically outperform their GARCH counterparts across conventional error metrics (MAE and RMSE), the top-performing algorithms exhibit no statistically significant divergence in predictive accuracy among themselves.

Keywords: Volatility Forecasting, Machine Learning, Median Realized Variance (MedRV), GARCH Models, BIST-100.

JEL Classification: G17, C58, C45

¹ Dr. Öğr. Üyesi Aydın Adnan Menderes Üniversitesi, Nazilli MYO, Yönetim ve Organizasyon Bölümü, Aydın, hbaskaya@adu.edu.tr, ORCID: 0000-0002-6098-3999.

1. Giriş

Finans piyasalarında volatilité, finansal varlık fiyatlamalarında zaman içindeki dalgalanmalara karşı risk yönetimi, opsiyon fiyatlandırması, portföy optimizasyonu, ve piyasanın sermaye ihtiyacının belirlenmesinde önemli bir yol oynamaktadır. Finansal piyasada yer alan yatırımcıların yatırım kararları ya da finansal piyasadaki kurumların risk ölçüm modelleri belirlemesi, piyasa istikrarlı olması ile ilgilidir. Borsa İstanbul (BİST), hem gelişmekte olan bir piyasa ekonomisinin yapısal özelliklerini hem de küresel finansal şoklara duyarlılığını yansıtan volatilitesi yüksek bir yapıya sahiptir. Örneğin, 2008 küresel finans krizi, 2018 döviz krizi ve 2020 COVID-19 pandemisi gibi gelişmeler, BIST 100 endeksinde yapısal kırılmalar yaratarak sert dalgalanmalara yol açmıştır. Bu bağlamda, BIST 100 oynaklığının güvenilir bir şekilde modellenmesi, piyasa aktörleri için öncelikli bir konudur. Volatilité modellerinde geleneksel yaklaşım, Engle (1982) tarafından önerilen ARCH ve Bollerslev (1986) tarafından genişletilen GARCH modelleri ailesidir. Bu modeller, getiri serilerindeki volatilité kümelenmesi, basık dağılım veya negatif şok veya haberlerin yarattığı volatilité artışının pozitif şoklara kıyasla daha şiddetli olduğunu gösteren kaldıraç etkisi gibi olguları yakalamada başarılı olmalarına karşın doğrusal parametre kısıtlamaları nedeniyle karmaşık olmakta ve modellemede sınırlı kalabilmektedir. Söz konusu kısıtlamalara karşın; yapay sinir ağları (ANN), destek vektör makineleri (SVM), rastgele orman (RF) ve k-en yakın komşu (KNN) gibi makine öğrenmesi algoritmaları ya doğrudan oynaklık kestiriminde kullanılmış ya da GARCH modelleri ile entegre edilerek hibrit yapılar aracılığıyla tahmin performansında kayda değer iyileşmeler sağlamıştır (Aras, 2021: 1). Geleneksel varyans modellerinin aksine, oynaklığın sürekli bileşenini süresiz şoklardan (ani yükselişlerden) ayıran Gerçekleşmiş Oynaklık (RV) ve özellikle aşırı değerlere karşı dirençli bir yapı sunan Medyan Gerçekleşmiş Oynaklık (MedRV) gibi sağlam tahmincilerin kullanımı, yüksek frekanslı verilerde oynaklık analizine yeni bir boyut kazandırmıştır. Yüksek frekanslı veri setiyle hesaplanan RV, gün boyunca toplanan kare getirilerden elde edilmekte; MedRV gibi atlama-dayanıklı ölçütler ise ani fiyat sıçramalarının neden olduğu ölçüm hatalarını azaltmaktadır (Barndorff-Nielsen ve Shephard, 2004: 5; Andersen vd., 2012: 76). Bu ölçütlerin makine öğrenmesi modelleriyle birlikte kullanılması, hem değişken kalitesini hem de model girdi bilgisini artırmaktadır.

Bu çalışma; BIST 100 endeksi bağlamında beş farklı GARCH modelini, dört ayrı hata dağılımı varsayımı altında sistematik olarak karşılaştırmaktadır. Ayrıca, dokuz makine öğrenmesi algoritmasının oynaklık tahmin başarımı, koşullu varyansın gözlemlenemeyen gerçek değerine yakınsayan iki temsil değişkeni olan standart gerçekleşen varyans (RV) ile fiyat sıçramalarına (atlama bileşenlerine) karşı dirençli medyan gerçekleşen varyans (MedRV) üzerinden değerlendirilmektedir. Bu kapsamıyla çalışma, geleneksel parametrik yaklaşımlar

ile veri güdümlü yöntemleri tek bir metodolojik çerçeve altında bütünleştirerek oynaklık tahmin literatürüne karşılaştırmalı ve çok boyutlu bir katkı sunmayı amaçlamaktadır. Çalışmanın ikinci bölümünde finansal volatilitenin ekonomik anlamı, GARCH modellerinin teorik gerekçesi ve makine öğrenmesinin konumlanması ele alınmakta; üçüncü bölümde ilgili literatür özetlenmekte; dördüncü bölümde veri seti, tahmin modelleri ve hata ölçütleri tanıtılmakta; beşinci bölümde GARCH ve makine öğrenmesi modellerine ait ampirik bulgular sunulmakta ve elde edilen sonuçlar literatürle karşılaştırmalı olarak değerlendirilmekte; altıncı bölümde ise temel bulgular özetlenerek politika önerileri ve gelecek araştırma yönelimleri tartışılmaktadır.

2. Teorik Çerçeve

2.1. Finansal Volatilitenin Ekonomik Anlamı

Finansal volatilité, sadece istatistiksel bir dağılım ölçüsü olmanın ötesinde, piyasa katılımcılarının belirsizlikle olan ilişkisini tanımlayan çok boyutlu bir ekonomik kavramdır. Fama (1970)'nın etkin piyasa hipotezi çerçevesinden bakıldığında, fiyat hareketleri yeni bilginin piyasaya yansımalarının doğrudan bir sonucudur; bu durumda volatilité, bilgi akışının yoğunluğunu ve piyasanın bu bilgiyi ne kadar hızlı ve eksiksiz sindirebildiğini göstermektedir. Ekonomik belirsizliğin artması, yatırımcıların talep ettiği risk primini yükseltmekte, bu da varlık fiyatlarında daha geniş dalgalanmalara yol açmaktadır. Dolayısıyla volatilité; yalnızca geçmiş fiyat hareketlerinin bir özeti değil, piyasanın gelecekteki belirsizliğe atfettiği değerinde cisimleşmiş bir yansımasıdır. Bu çerçevede opsiyon piyasalarında fiyatlara yansıyan zımnî volatilité, piyasa katılımcılarının geleceğe ilişkin risk beklentilerinin bir göstergesi iken; tarihsel volatilité, geçmiş dönemlerde yaşanan belirsizliğin geriye dönük bir kaydı niteliği taşımaktadır.

Asimetrik enformasyon açısından konu değerlendirildiğinde, piyasada aynı varlığa ilişkin farklı bilgi setleriyle hareket eden aktörlerin varlığı, fiyat keşif sürecini karmaşık ve süresiz bir niteliğe büründürmektedir. Bilgilenmiş yatırımcılar pozisyon aldıkça, piyasa yapıcılar alım-satım marjını genişleterek asimetrik enformasyondan kaynaklanan riski fiyatlamakta; bu dinamik doğrudan volatilitéye katkıda bulunmaktadır. Piyasa mikro yapısı teorisi, volatilitenin kısmen sipariş akışı dinamiklerinden, kısmen de işlem maliyetleri ve likidite koşullarındaki değişimlerden beslendiğini ortaya koymaktadır. Öte yandan davranışsal finans literatürü, aşırı güven, sürü davranışı ve bilişsel önyargılar ve sapmaların dönemsel fiyat fazla tepkimeleri (overreaction) ve eksik tepkimelere (underreaction) yol açarak volatilitéyi sistematik biçimde artırabildiğini göstermektedir (Shiller, 2003: 101). Bu teorik çerçeveden

hareketle, BIST 100 gibi gelişmekte olan piyasaların volatilité örüntüleri; kurumsal yatırımcı tabanının sığılığı, kur kanalıyla gelen dışsal şoklar ve makroekonomik politika belirsizlikle-riyle ek bir karmaşıklık kazanmaktadır. Tüm bu nedenlerle, volatilitéyi doğru modelleyebil-mek yalnızca istatistiksel bir sorun değil, temelde ekonomik bir sorudur ve kullanılacak mo-delin bu teorik derinliğı yansıtacak esneklikte olması gerekmektedir.

2.2. GARCH Modelleri: Teorik Problemler ve Model Seçiminin Gerekeşesi

Zaman serisi finansal verisinin modellenmesinde karşılaşılan ilk temel sorun, getirilerin var-yansının zaman içinde sabit olmadığıdır; bu olgu heteroskedastisite (değişen varyans) olarak adlandırılmakta ve Klasik En Küçük Kareler tahmin edicisinin etkinliğini zayıflatmaktadır. Engle (1982) tarafından geliştirilen ARCH modeli, bu teorik soruna yanıt olarak koşullu var-yansın geçmiş hata karelerine bağılı olduğu bir yapı kurmuş ve volatilitenin öngörülebilir ol-duğunu istatistiksel açıdan ortaya koymuştur. Bollerslev (1986) ise ARCH'ın yüksek para-metresayısına neden olan yapısını çözmek için gecikmeli koşullu varyans terimlerini modele dahil ederek GARCH'ı geliştirmiştir. GARCH'ın asıl katkısı, gözlemlenen volatilité kümelenmesi olgusunu ve yüksek volatilité kalıcılığını tutarlı biçimde sağlayabilmesidir. Buna karşılık standart GARCH modeli, negatif ve pozitif şokların volatilité üzerinde simetrik etki yarattığını varsaymaktadır; bu varsayım, kötü haberlerin iyi haberlerden daha güçlü bir vo-latilité tepkisine yol açtığını belgeleyen kaldıraç etkisiyle çelişmektedir. Bu teorik eksikliği gidermek amacıyla Nelson (1991) EGARCH'ı ve Glosten vd. (1993) GJR-GARCH'ı geliştirmiş-tir; her iki model de yönlü asimetriyi yakalayan parametreler içermektedir. Zakoian (1994)'ın önerdiği TGARCH modeli de eşik değerlerine dayalı yapısıyla şok yönünün vola-tilité dinamiğindeki asimetrik rolünü modellemektedir. Bu çalışmada kullanılan beş GARCH spesifikasyonu, birbirinin eksikliklerini gidermek amacıyla geliştirilmiş modellerin bir araya getirilmesiyle oluşmuştur: değişen varyans sorunu ARCH ile ele alınmış; volatilité kümelen-mesi ve kalıcılık problemi GARCH ile çözülmüş; asimetri ve kaldıraç etkisi ise EGARCH, GJR-GARCH ve TGARCH ile modellenmiştir. Model seçimi bu teorik hiyerarşinin gerek-tirdiğı kapsayıcılığı sağlamak amacıyla yapılmıştır.

2.3. Makine Öğrenmesi Modellerinin Teorik Konumlanması ve Sınırlılıkları

GARCH ailesinin temel özelliğı parametrik yapısıdır: model, önceden belirlenmiş bir var-yans denkleminde dayanmakta ve tahmin süreci bu denklemin parametrelerini veri içinden öğrenmeye indirgemektedir. Bu yapı teorik açıdan şeffaf ve yorumlanabilir olmakla birlikte, doğrusal olmayan, yüksek boyutlu ya da önceden bilinmeyen fonksiyonel bağımlılıkları mo-dellemekte yetersiz kalabilmektedir (Drost ve Klaassen, 1997: 194). Buna karşılık makine

öğrenmesi (ML) modelleri, veri güdümlü yapıları sayesinde herhangi bir fonksiyonel form varsayımına ihtiyaç duymaksızın doğrusal olmayan ilişkileri öğrenebilmektedir. Random Forest ve Gradient Boosting gibi topluluk öğrenme yöntemleri, değişkenler arası yüksek düzeyli etkileşimleri ve doğrusal olmayan örüntüleri yakalamada üstün performans sergilemektedir.

Bununla birlikte ML modellerinin bu esnekliği beraberinde ciddi teorik sınırlılıklar da getirmektedir. Birincisi, çoğu ML modeli kara kutu² (black-box) niteliği taşımakta; tahmin sürecinin ekonomik yorumlaması kısıtlı kalmaktadır, karar ağaçları ve sinir ağlarında hangi değişkenin volatilitiyi hangi mekanizma aracılığıyla etkilediğini analitik bir denklemle ifade etmek mümkün değildir (Stiglic vd., 2016: 3). İkincisi ve en kritik olanı, aşırı öğrenme riskidir: model eğitim verisinin kendine özgü gürültüsünü genelleştirilebilir bir örüntü olarak öğrenebilmekte, bu da örneklem dışı tahmin performansının yanıltıcı biçimde düşmesine yol açabilmektedir. Bu riski azaltmak için bu çalışmada zaman serilerine uygun çapraz doğrulama tekniklerine ve düzenleme parametrelerine başvurulmuştur (Prado, 2018: 11). Üçüncüsü, hiperparametre optimizasyonu hem hesaplama maliyeti hem de seçim belirsizliği açısından ek zorluklar doğurmaktadır. Bu sınırlılıklar göz önüne alındığında ML ve GARCH yaklaşımları birbirinin rakibi değil, tamamlayıcısı olarak değerlendirilmeli; her iki paradigmanın güçlü yönleri karşılaştırmalı bir çerçevede analiz edilmelidir.

2.4. Gerçekleşmiş Varyansın Ölçüm Teorisi: RV ve MedRV

Finansal volatilitenin modellenebilmesi için önce ölçülmesi gerekmektedir; ancak volatilité, doğrudan gözlemlenebilir bir değişken değildir. Günlük kapanış fiyatları arasındaki kare getiri gibi geleneksel oynaklık vekilleri, yüksek ölçüm hatası içermekte ve gün içi bilgiyi büyük ölçüde yok saymaktadır. Andersen ve Bollerslev (1998: 888) tarafından başlatılan gerçekleşmiş varyans (Realized Variance, RV) literatürü bu ölçüm teorisi sorununa köklü bir çözüm sunmaktadır: yüksek frekanslı gün içi getirilerin karelerinin toplanmasıyla elde edilen RV, örneklem aralığı sonsuz küçüldükçe teorik anlık varyansa yakınsayan tutarlı bir tahmincidir. Bu sonuç, gerçek varyansın in-fill asimptotik çerçevesinde gerçekleşen varyans aracılığıyla tutarlı biçimde tahmin edilebileceğine dair önemli bir teorik güvencedir. Bununla birlikte, gerçek piyasa verisi ani fiyat sıçramaları ve piyasa mikro yapısı gürültüsü sebebiyle iki önemli tehdit içermektedir. Sıçrama bileşeni, varlığın sürekli difüzyon sürecinin dışındaki kesintili fiyat hareketlerini yansıtırken; mikro yapı gürültüsü bid-ask bounce, ayrık fiyat tık-

² Algoritmalar sıklıkla birer ‘kara kutu’ olarak nitelendirilir; zira sahip oldukları karmaşıklık ve teknik şeffaflıktan yoksun oluşları, iç işleyişlerini anlaşılabilir ve belirsiz hale getirir” (Diakopoulos, 2015: 404).

ları ve asenkron işlem gibi mekanizmalardan kaynaklanmaktadır. Her iki unsur da RV tahminini yukarı yönlü önyargıya sürüklemektedir. Bu teorik güçlüğü aşmak amacıyla Andersen vd. (2012: 78) tarafından önerilen Medyan Gerçekleşmiş Varyans (MedRV), ardışık üç gün içi getirinin mutlak değerleri arasındaki medyanı temel alarak hem sıçrama bileşenini hem de mikro yapı gürültüsünü filtreleyen sağlam (robust) bir tahmincidir. MedRV'nin teorik üstünlüğü iki kanaldan kaynaklanmaktadır: medyan istatistiğinin aşırı değerlere karşı doğal direnci ve hesaplama yapısının sıçrama bileşenini sistematik olarak dışarıda bırakması. Bu çerçevede MedRV, volatilitenin sürekli bileşenine daha saf bir yakınsama sağlamakta; bu da model tahmin performansı açısından daha düşük gürültü içeren bir hedef değişken anlamına gelmektedir.

Yüksek frekanslı ekonometri literatüründe sıçramaya dirençli ilk tahminci olan Çift Güç Varyasyonu (BV) (Barndorff-Nielsen ve Shephard, 2004), sonlu örneklemelerde sıçramaların ardışık gözlemlere sızması nedeniyle kısıtlı bir dayanıklılık sergilemektedir. Bu kısıtı aşmak üzere Andersen vd. (2012) tarafından önerilen MinRV, ardışık üç mutlak getirinin minimumunu esas alarak BV'ye kıyasla daha yüksek sonlu örneklem etkinliği sunmaktadır. Gürültü kaynaklı sapmayı düzeltmeyi amaçlayan İki Ölçekli Gerçekleşmiş Varyans (TSRV) (Zhang vd., 2005) ise sıçrama bileşenini doğrudan hedef almamaktadır. Bu çalışmada MedRV'nin tercih edilme gerekçesi; hem fiyat sıçramalarına hem de piyasa mikro yapı gürültüsüne karşı eşzamanlı direnç sağlaması ve BV'ye kıyasla sonlu örneklemelerde daha yüksek istatistiksel etkinlik sunmasıdır (Andersen vd., 2012).

3. Literatür Taraması

Volatilité tahmininde GARCH modellerinin kullanımına ilişkin literatür son yıllarda hızla geleneksel GARCH, hibrit yaklaşımlar (GARCH+ML), gerçekleşen varyans+ML olarak değişmektedir. Geleneksel GARCH yaklaşımları içerisinde Bollerslev (1986) tarafından önerilen standart GARCH modeli oynaklık tahmininin temel yapı taşıdır. Nelson (1991) asimetrik oynaklığı yakalamak için EGARCH'ı, Glosten vd. (1993) kaldıraç etkisini modelleyen GJR-GARCH'ı geliştirmiştir. İlgili literatürde öne çıkan çalışmalar ve bu çalışmaların temel bulguları, Tablo 1'de karşılaştırmalı olarak özetlenmiştir.

Tablo 1. Literatür Taraması

Yazar	Kullanılan Değişkenler	Yöntem	Bulgular
Çelik ve Ergin (2014)	BİST 30 endeks vadeli işlem sözleşmeleri gün içi yüksek frekanslı verileri	HAR-RV, HAR-RV-CJ, MIDAS, GARCH, Realized GARCH	HAR-RV modeli, GARCH'a kıyasla daha yüksek örneklem dışı oynaklık tahmini başarımı sunduğu saptanmıştır.
Kristjanpoller ve Minutolo (2015)	Altın fiyatı, döviz kurları, faiz oranı	Hibrit ANN-GARCH	Hibrit model, GARCH'a kıyasla MAE'yi %25 azalttığı, ANN-GARCH üstün tahmin kapasitesi sergilediği tespit edilmiştir.
Kim ve Won (2018)	Hisse senedi endeksi günlük getirileri (KOSPI, KOSDAQ)	LSTM + GARCH-tipi çoklu hibrit model	LSTM-GARCH hibrit modeli, tek GARCH ve standart hibrit modellerden daha iyi örneklem dışı tahmin sunduğu belirtilmiştir.
Tektaş ve Karataş (2021)	BİST 30 finansal oranları, hisse kapanış fiyatları (2010–2019)	YSA, Rastgele Orman, XGBoost	XGBoost en iyi tahmin başarımını sergilediği, RF de YSA'ya kıyasla üstün performans gösterdiği belirtilmiştir.
Aras (2021)	Bitcoin günlük getirisi; RV, BV, MinRV, MedRV	ANN, SVM, RF, KNN + sGARCH(p,q), istifleme topluluğu (LASSO)	Yüksek model mertebelerinin hibrit performansı artırdığı; LASSO tabanlı istifleme en iyi sonuçları verdiği bulunmuştur.
Christensen vd. (2022)	Dow Jones bileşen hisseleri yüksek frekanslı getirileri (RV)	Düzenleştirme, regresyon ağaçları, yapay sinir ağları, HAR	ML modelleri HAR ailesini geride bıraktı; özellikle uzun ufuklarda öngörü kazanımları tespit edilmiştir.
Zhang vd. (2022)	NASDAQ hisse senetleri yüksek frekanslı verisi (RV)	LSTM, CNN, HAR; gün içi ortak oynaklık	Sinir ağları ağaç tabanlı ve doğrusal modelleri geride bıraktığı saptanmıştır.
Çelik ve Ergin (2022)	BİST endeks günlük kapanış fiyatları	Destek Vektör Regresyonu, Rastgele Orman, XGBoost	Topluluk yöntemleri (RF, XGB), SVR ile rekabetçi sonuçlar üretti; ensemble yaklaşımların potansiyeli gösterildiği saptanmıştır.
Henriques ve Sardorsky (2023)	Küresel enerji emtia fiyatları (ham petrol, doğalgaz, vb.)	GARCH ailesi + ML (RF, GBDT, XGB), SHAP açıklanabilirlik	ML modelleri örneklem dışında GARCH'ı geride bıraktı; GARCH aşırı, ML modelleri eksik tahmin eğilimi sergilediği belirtilmiştir.
Xu vd. (2024)	Hisse senedi piyasası günlük getirileri	GARCH bilgili sinir ağı (GINN): GARCH + LSTM hibrit	GINN, GARCH ve standart LSTM'yi R^2 , MSE, MAE açısından geride bıraktığı saptanmıştır.

Tablo 1 genel olarak değerlendirildiğinde, finansal piyasalarda oynaklık ve getiri tahmininde makine öğrenmesi ve hibrit modellerin geleneksel ekonometrik yöntemlere kıyasla daha başarılı sonuçlar verdiği görülmektedir. Özellikle XGBoost, Rastgele Orman, LSTM ve GARCH tabanlı hibrit yaklaşımların örneklem dışı tahmin performansını artırdığı; topluluk öğrenmesi ve derin öğrenme yöntemlerinin finansal tahmin çalışmalarında öne çıktığı anlaşılmaktadır.

3.1. GARCH Temelli Çalışmalar

Volatilite tahminine ilişkin erken dönem literatürün büyük bölümü, doğrudan GARCH ailesi içindeki spesifikasyonların karşılaştırılmasına odaklanmaktadır. Bu çalışmalar arasında, Çelik ve Ergin (2014), Türkiye’de hisse senedi piyasasında HAR-RV modelinin standart GARCH'a kıyasla belirgin biçimde daha yüksek örneklem dışı tahmin başarımı sunduğunu ortaya koymuştur. Bu bulgu, gün içi yüksek frekanslı bilgiyi kullanan gerçekleşen varyans tabanlı yaklaşımların geleneksel günlük GARCH modellerine olan üstünlüğüne işaret etmektedir. Nelson (1991) ve Glosten vd. (1993) tarafından geliştirilen EGARCH ve GJR-GARCH gibi asimetrik modellerin, simetrik standart GARCH'a kıyasla kaldıraç etkisini daha iyi yakaladığına dair teorik öngörüler, izleyen ampirik literatürde geniş ölçüde doğrulanmıştır. Bu çalışmalar göstermektedir ki parametrik modeller volatilite kümelenmesini ve asimetriyi içselleştirdiğinde açıklama güçleri artmakla birlikte, doğrusal parametrik yapı doğrusal olmayan ve yüksek boyutlu piyasa dinamiklerini tam olarak yakalayamamaktadır. Modelin teorik varsayımları piyasa gerçekliğinden uzaklaştıkça tahmin hataları sistematik biçimde büyümekte; bu da daha esnek yaklaşımlara duyulan ihtiyacı doğrulamaktadır.

3.2. Hibrit Model Çalışmaları (GARCH + Makine Öğrenmesi)

Parametrik modellerin sınırlılıklarını aşmak amacıyla GARCH'ı makine öğrenmesi algoritmalarıyla birleştiren hibrit yapılar giderek yaygınlaşmıştır.(bkz. Tablo 1.) Farklı varlık sınıfları (altın, hisse senedi, kripto para, enerji emtiaları) ve farklı mimariler (ANN-GARCH, LSTM-GARCH, GARCH-bilgili sinir ağları) üzerinden yürütülen bu çalışmalar tutarlı bir bulguya işaret etmektedir: hibrit yapılar, GARCH'ın teorik altyapısı ile ML'nin esnekliğini bir araya getirerek saf GARCH modellerine kıyasla örneklem dışı tahmin hatasını azaltmaktadır. Kristjanpoller ve Minutolo (2015), altın fiyatı oynaklığının tahmininde ANN-GARCH hibrit modelinin MAE'yi saf GARCH'a göre kayda değer ölçüde azalttığını raporlamıştır. Kim ve Won (2018), LSTM'yi birden fazla GARCH-tipi modelle birleştirmiş ve hibrit yapının örneklem dışı başarımı önemli ölçüde artırdığını göstermiştir. Aras (2021), Bitcoin oynaklığında ANN, SVM, RF ve KNN'i farklı GARCH(p,q) mertebeleri ile birleştirerek

LASSO tabanlı topluluk öğrenmesinin neredeyse tüm baz modelleri geride bıraktığını ortaya koymuştur. Henriques ve Sadorsky (2023), enerji piyasası oynaklığı üzerinde GARCH ailesi ve ML modellerini SHAP açıklanabilirlik analiziyle birlikte değerlendirmiş; ML modellerinin örneklem dışında GARCH'ı geride bıraktığını göstermiştir. Xu vd. (2024) tarafından önerilen GARCH-bilgili sinir ağı (GINN) ise yapısal GARCH dinamiklerini sinir ağı çerçevesine entegre ederek hem yorumlanabilirliği hem de tahmin performansını iyileştirmeyi hedeflemiştir. Bu çalışmalar göstermektedir ki hibrit modeller, GARCH'ın teorik altyapısını ve ML'nin esnekliğini bir arada sunarak tek başına kullanılan her iki yaklaşımı da genellikle geride bırakmaktadır. Bununla birlikte hibrit yapıların model karmaşıklığını ve hiperparametre optimizasyon yükünü artırdığı; aşırı öğrenme riskinin yönetilmesinin daha da güçleştiği teorik olarak not edilmelidir.

3.3. Saf Makine Öğrenmesi Yaklaşımları

Son dönemde pek çok çalışma, GARCH altyapısına başvurmaksızın yalnızca makine öğrenmesi algoritmalarını kullanarak volatilité tahmini gerçekleştirmiş ve bu yaklaşımın rekabetçi sonuçlar ürettiğini göstermiştir (bkz.Tablo1).Gelişmiş piyasalarda Christensen vd. (2022), Dow Jones bileşen hisselerinin gerçekleşen varyansını tahmin etmek amacıyla düzenlileştirme, regresyon ağaçları ve yapay sinir ağlarını HAR ailesiyle karşılaştırmış; minimum hiperparametre ayarına rağmen ML modellerinin HAR'ı geride bıraktığını, özellikle uzun öngörü ufuklarında kazanımların belirginleştiğini göstermiştir. Zhang vd. (2022), gün içi RV tahmininde sinir ağlarının ağaç tabanlı ve doğrusal modelleri performans açısından aştığını ve oynaklıkta hisseler arası evrensel mekanizmanın varlığına dair kanıt sunduğunu bulgulamıştır. Türk piyasasına yönelik çalışmalar arasında Tektaş ve Karataş (2021) BİST 30 hisse senetleri için YSA, RF ve XGBoost'u karşılaştırmış; XGBoost en iyi sonuçları üreten RF'nin de YSA'ya kıyasla üstün performans gösterdiğini bulmuştur. Çelik ve Ergin (2022) ise zaman serisi bağlamında topluluk yöntemlerinin (RF, XGBoost) SVR ile rekabetçi sonuçlar ürettiğini ortaya koymuştur. Bu çalışmalar, model varsayımlarından bağımsız olarak volatilité dinamiklerini öğrenme kapasitesine sahiptir; ancak bu başarı büyük ölçüde gürültüsüz ve bilgi-zengin hedef değişkene (RV, MedRV gibi gerçekleşen varyans ölçütlerine) erişime bağlıdır. Yüksek frekanslı veriye erişimin kısıtlı olduğu ortamlarda saf ML yaklaşımlarının üstünlüğü azalmakta, hibrit yapılar ise daha güvenli bir seçenek olmaktadır.

3.4. Araştırma Boşluğu ve Çalışmanın Katkısı

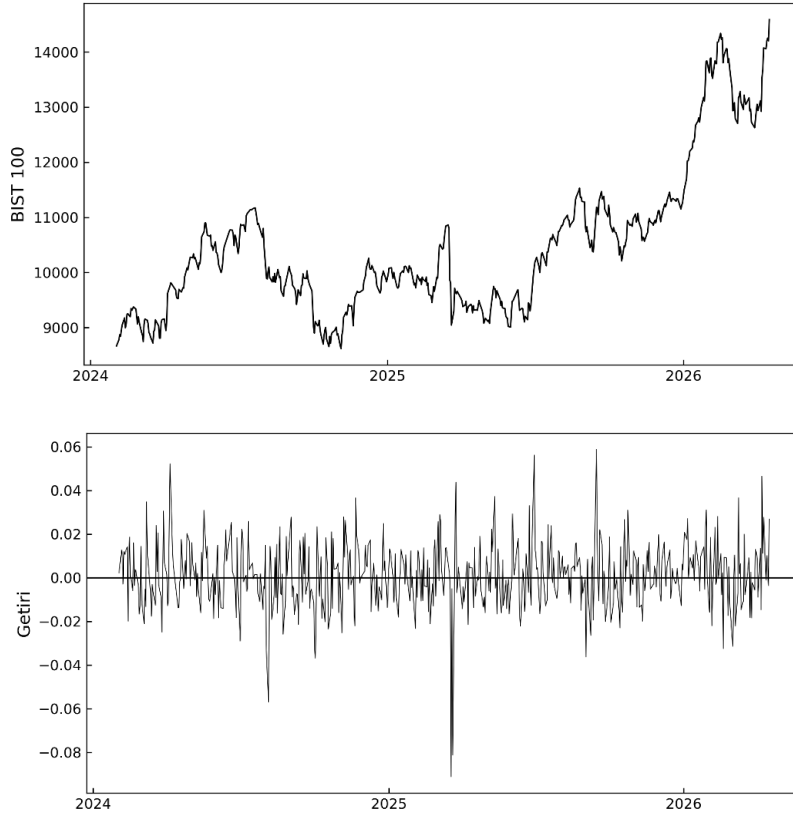
Yukarıda özetlenen literatür bütünü değerlendirildiğinde üç kritik boşluk dikkat çekmektedir. Birincisi, BIST 100 özelinde hem GARCH hem de ML algoritmalarını sistematik biçimde

karşılaştıran çalışmalar son derece sınırlıdır; mevcut Türk piyasası çalışmalarının büyük çoğunluğu fiyat ya da yön tahminine odaklanmakta, oynaklık tahminini ise ikincil bir konu olarak ele almaktadır. İkincisi, MedRV gibi atlama-dayanıklı varyans ölçütlerinin BIST 100 bağlamında hedef değişken olarak denenmediği görülmektedir; oysa teorik açıdan MedRV, mikro yapı gürültüsünü filtreleyen daha saf bir varyans tahmincisi sunmaktadır. Üçüncüsü, mevcut hibrit çalışmaların çoğu ya belirli bir ML algoritmasını GARCH ile eşleştirmekte ya da sistematik bir hiperparametre optimizasyon çerçevesi sunmamaktadır. Bu çalışma tam olarak bu boşlukları doldurmayı hedeflemektedir: beş GARCH spesifikasyonunu dört hata dağılımı varsayımı altında ve dokuz ML algoritmasını zaman serilerine uygun çapraz doğrulama ve genişleyen pencere yaklaşımıyla karşılaştıran bu çalışma, hem RV hem de MedRV hedef değişkeni üzerinden BIST 100 oynaklık literatürüne yöntemsel açıdan kapsamlı ve çok boyutlu bir katkı sunmaktadır.

4. Materyal ve Yöntem

4.1. Veri Seti

Analiz için kullanılan veriler, 1 Şubat 2024 ile 17 Nisan 2026 tarihleri arasındaki günlük kapanış fiyatlarından oluşmaktadır. Toplam yaklaşık 554 günlük gözlem ele alınmıştır. Getiriler, $rt = \log(P_t/P_{t-1})$ şeklinde hesaplanmıştır; burada P_t , t günündeki kapanış fiyatını ifade etmektedir. Çalışma kapsamında gözlenemeyen bir bileşen olan gerçek oynaklığa yakınsamak için gün içi 30 dakikalık yüksek frekanslı getirilerden faydalanılmıştır. Bu amaç için, literatürde sıklıkla kullanılan iki vekil değişken olan gerçekleşen varyans (RV) ve medyan gerçekleşen varyans (MedRV) ölçüleri sonuçların tutarlılığını incelemek için bağımlı değişkenler olarak alınmıştır. Bu çalışmanın temel odağı, günlük getiri serilerine dayalı geleneksel parametrik GARCH ailesi ile parametrik olmayan makine öğrenmesi algoritmalarını karşılaştırmaktır. Bu nedenle, çoklu zaman ölçekli dinamikleri doğrudan modelleyen Heterojen Otoregresif (HAR-RV) ve sıçrama ayrıştırmalı türevleri (HAR-J, HAR-CJ) analiz kapsamı dışında tutulmuş; model başarımları doğrudan RV ve MedRV temsil değişkenleri üzerinden değerlendirilmiştir.

Şekil 1: BIST 100 indeksi ve getirilerinin zaman yolu grafikleri

Şekil 1, günlük fiyatları ve getirileri göstermektedir. Görsel olarak da anlaşılacağı üzere, endeks verileri durağan olmayan özellikler ve oldukça dalgalı bir seyir sergilediğinden getiri serileri üzerinden analiz gerçekleştirilmiştir. Serinin tanımlayıcı istatistikleri Tablo 2’de sunulmaktadır.

Tablo 2. Tanımlayıcı istatistikler

	Ort	Std. Sap.	Min	Q1	Med	Q3	Mak	Çarpıklık	Basıklık	Normallik
Getiri	0.109	1.505	-8.720	-0.810	0.050	1.042	6.060	-0.256	6.950	366.3
RV	0.201	0.435	0.011	0.073	0.116	0.182	5.822	9.112	105.1	248300
MedRV	0.130	0.175	0.002	0.056	0.089	0.142	2.237	6.559	61.33	82510

Not: RV ve MedRV için gerçekleşen varyans değerleri 1000 ile çarpılmıştır.

Tablo 2, serinin tanımlayıcı istatistiklerini sunmaktadır. Beklendiği gibi, Jarque-Bera (JB) testine göre seri normal dağılım göstermemektedir ve standart sapması oldukça yüksektir. Ayrıca, Tablo 2’de, serinin hafifçe negatif çarpık olduğu ve büyük uç değerleri gösteren leptokurtik davranış sergilediği anlaşılmaktadır. Dikkat edileceği üzere, MedRV değeri uç değerlere bir başka deyişle gün içindeki aşırı iniş çıkışlara karşı daha dirençli olduğundan daha küçük standart sapma, çarpıklık ve basıklık değerlerine sahiptir. Öngörü performansını değerlendireceğimiz test seti son altı aylık dönemden olmak üzere toplamda 120 gözlemde oluşmaktadır. Oynaklık tahminleri bir adım ileriye doğru yapısında oluşturulmuş ve tüm modeller bir haftaya denk gelen test setindeki her 5 gözlemde yeniden eğitilerek güncellenmiştir. Bu esnada, veri setinin boyutu göz önünde bulundurularak genişleyen pencere yaklaşımından faydalanılmıştır. Bir başka deyişle, tüm modeller test seti boyunca 24 ‘er kez yeni verilerdeki değişimi model parametrelerine yansıtmak amacıyla tekrar oluşturulmuştur. Bu her bir adımda, makine öğrenme modelleri için en uygun hiperparametre aramaları tekrar yapılmıştır.

4.2. Tahmin Modelleri

4.2.1. GARCH Modelleri

Bu çalışmada beş farklı GARCH spesifikasyonu değerlendirilmiştir: sGARCH, EGARCH, GJR-GARCH, IGARCH ve TGARCH. Her bir model; normal (Norm), çarpık normal (Snorm), genelleştirilmiş hata (Ged) ve çarpık genelleştirilmiş hata dağılımı (Sged) olmak üzere dört farklı hata dağılımı varsayımı altında tahmin edilmiştir. Söz konusu modellerin oynaklık denklemleri Tablo 3’te sunulmaktadır.

Tablo 3. GARCH modellerinin oynaklık denklemleri

Modeller	Oynaklık Denklemleri
sGARCH	$\sigma^2_{t+1} = \omega + \alpha r^2_t + \beta \sigma^2_t$
EGARCH	$\log(\sigma^2_{t+1}) = \omega + \alpha e_t + \gamma e_t + \beta \log(\sigma^2_t)$
GJR-GARCH	$\sigma^2_{t+1} = \omega + \alpha r^2_t + \gamma I(r_t < 0) r^2_t + \beta \sigma^2_t$
IGARCH	$\sigma^2_{t+1} = \omega + \alpha r^2_t + (1 - \alpha) \sigma^2_t$
TGARCH	$\sigma_{t+1} = \omega + \alpha \sigma_t (e_t - \eta_1 e_t) + \beta \sigma_t$

4.2.2. Makine Öğrenmesi Modelleri

Çalışma kapsamında hibrit GARCH modellerini oluşturmak amacıyla 9 farklı makine öğrenme tekniği kullanılmıştır. Bu modellere ait kısa açıklamalar aşağıda olup oluşturulan deney kapsamında her bir makine öğrenme modelinin en uygun hiperparametrelerini belirlemek üzere faydalanılan değer uzayı Tablo 4’te verilmiştir. Her modele ait hiperparametre

kombinasyonun farklı olması ve aynı zamanda bazı modellerde tüm kombinasyonları denemenin aşırı zaman alabileceği göz önünde bulundurularak mümkün hiperparametre kombinasyonları arasından 30 'ar örnek çekilerek modellemeler gerçekleştirilmiştir. Ayrıca, veri setine en uygun hiperparametre kombinasyonunu bulmak amacıyla zaman serilerinde çapraz doğrulama yapmayı olanaklı hale getiren Python 'nin Scikit-learn kütüphanesindeki TimeSeriesSplit fonksiyonundan faydalanılmıştır. Bu esnada 5 katlı çapraz doğrulama yapılmıştır. Bunun sonucu olarak, genişleyen pencere yaklaşımı da göz önünde bulundurulduğunda, her bir makine öğrenme tekniğinin toplamda 120 kez oluşturulduğudur. Makine öğrenmesi modelinde kullanılan algoritmalar Tablo 4'te sunulmuştur:

MLP: Çok katmanlı algılayıcı mimarisinde gizli katman boyutu, aktivasyon fonksiyonu, optimizasyon algoritması (*solver*), düzenleme parametresi (*alpha*), başlangıç öğrenme oranı (*learning_rate_init*), yığın boyutu (*batch_size*) ve maksimum yineleme sayısı (*max_iter*) optimize edilmiştir.

SVR: Regularizasyon parametresi C, epsilon (ϵ) ve çekirdek genişliği (*gamma*) temel hiperparametrelerini oluşturmaktadır.

RF: Ağaç sayısı (*n_estimators*), maksimum ağaç derinliği (*max_depth*), düğüm bölünmesi için gereken minimum örnek sayısı (*min_samples_split*) ve yaprak düğümündeki minimum örnek sayısı (*min_samples_leaf*) optimize edilmiştir (Breiman, 2001).

KNN: Komşu sayısı (*k*) ve mesafe ağırlıklandırma şeması (*weights*) temel hiperparametrelerdir.

DT: Maksimum derinlik (*max_depth*), minimum bölünme örnekleri (*min_samples_split*) ve yaprak düğüm örnekleri (*min_samples_leaf*) optimize edilmiştir.

GBDT: Ağaç sayısı (*n_estimators*), öğrenme oranı (*learning_rate*), maksimum derinlik (*max_depth*), alt örnekleme oranı (*subsample*), bölünme örnekleri (*min_samples_split*) ve kayıp fonksiyonu (*loss*) optimize edilmiştir.

XGBoost: Model karmaşıklığını ve genelleştirme gücünü dengelemek amacıyla; ağaç sayısı (*n_estimators*), öğrenme oranı, derinlik, alt örnekleme oranları (*subsample*, *colsample_bytree*), budama eşiği (*gamma*), *min_child_weight* ile L1/L2 ceza terimleri (*reg_alpha*, *reg_lambda*) optimize edilmiştir.

SGD: Düzenleme katsayısı (*alpha*), ceza türü (*penalty*), elastik ağ karışım oranı (*l1_ratio*) ve öğrenme oranı politikası optimize edilmiştir.

ExtraTrees: RF ile benzer hiperparametre yapısına (n_estimators, max_depth, min_samples_split, min_samples_leaf) sahip olup, bölünme eşiklerinin tamamen rassal seçilmesiyle model çeşitliliği artırılmaktadır.

Tablo 4. Makine öğrenmesi modelleri için oluşturulan hiperparametre uzayı

Model	Hiperparametre	Deney uzayı
SVR	C	{0.001, 0.01, 0.1, 1, 10, 30, 50, 100, 500, 1000}
	gamma	{scale, auto, 0.001, 0.01, 0.1, 1}
	epsilon	{0.0001, 0.001, 0.01, 0.05, 0.1, 0.5, 1, 5, 10}
RF	n_estimators	{50, 100, 200}
	max_depth	{4, 6, 8, 10}
	min_samples_split	{2, 5, 7, 10, 20, 30, 50, 100}
	min_samples_leaf	{1, 2, 5, 7, 10, 20, 30, 50}
KNN	n_neighbors	{3, 5, 7, 10, 20, 30, 50, 75, 100, 150, 200}
	weights	{uniform, distance}
DT	max_depth	{4, 6, 8, 10}
	min_samples_split	{2, 5, 7, 10, 20, 30, 50, 100}
	min_samples_leaf	{1, 2, 5, 7, 10, 20, 30, 50}
GBDT	n_estimators	{50, 100, 200, 300}
	learning_rate	{0.001, 0.01, 0.05, 0.1}
	max_depth	{3, 4, 5, 6, 7}
	min_samples_split	{2, 5, 7, 10, 20, 30, 50, 100}
	subsample	{0.8, 0.9, 1.0}
	loss	{squared_error, absolute_error, huber}
XGBoost	n_estimators	{100, 200, 300, 500}
	learning_rate	{0.01, 0.05, 0.1}
	max_depth	{3, 5, 7}
	subsample	{0.8, 0.9, 1.0}
	colsample_bytree	{0.8, 0.9, 1.0}
	gamma	{0, 0.1, 0.5}
	min_child_weight	{1, 3, 5}
	reg_alpha	{0, 0.01, 0.1}
	reg_lambda	{1, 1.5, 3}
SGD	alpha	{0.0001, 0.001, 0.01, 0.1, 0.5, 1, 5, 10, 30, 50, 100, 1000}
	l1_ratio	{0.1, 0.3, 0.5, 0.7, 0.9}
	penalty	{l2, l1, elasticnet, None}
	learning_rate	{invscaling, optimal, adaptive}
ExtraTrees	n_estimators	{50, 100, 200}
	max_depth	{4, 6, 8, 10}
	min_samples_split	{2, 5, 7, 10, 20, 30, 50, 100}
	min_samples_leaf	{1, 2, 5, 7, 10, 20, 30, 50}
MLP	hidden_layer_sizes	{{(1),(5),(10),(50),(100),(5,5),(10,10),(50,50),(100,100)}
	activation	{relu, tanh}
	solver	{adam, sgd}
	alpha	{1e-5, 1e-4, 1e-3, 0.01, 0.1}
	learning_rate_init	{0.0001, 0.001, 0.005, 0.01}
	batch_size	{32, 64, auto}
	max_iter	{500, 1000, 2000}

4.3. Hata Ölçütleri

Oluşturulan oynaklık modellerinin performansı en çok kullanılan MAE ve RMSE iki standart hata ölçüyle değerlendirilmiştir. Bu ölçülere ait formülasyonlar Denklem (1) ve (2) 'de sunulmuştur.

$$MAE = \frac{1}{q} \sum_{t=m+1}^{m+q} |h_t - \hat{\sigma}_t^2| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{q} \sum_{t=m+1}^{m+q} (h_t - \hat{\sigma}_t^2)^2} \quad (2)$$

Burada $\hat{h}_t$ model tahmini, σ_t^2 ise ilgili gerçek varyans için kullanılan vekil değişkeni (RV veya MedRV) göstermektedir.

MAE (Mean Absolute Error - Ortalama Mutlak Hata): Tahmin edilen değerler ile gerçekleşen değerler arasındaki mutlak farkların ortalamasıdır. Hataların büyüklüğünü doğrusal bir ölçekte ölçtüğü için modelin ortalama hata performansını yansıtmakta ve uç değerlere karşı RMSE'ye kıyasla daha dirençli (robust) bir gösterge sunmaktadır.

RMSE (Root Mean Square Error - Hata Kareler Ortalamasının Karekökü): Hata terimlerinin karelerinin ortalamasının karekökü alınarak hesaplanır. Hataların karesini baz alması nedeniyle, büyük ölçekli hataları daha ağır cezalandırmakta; bu yönüyle modelin yüksek sapmalı tahminlerden kaçınma başarısını ve varyansını ölçen duyarlı bir metrik olarak kabul edilmektedir.

Düşük MAE ve RMSE değerleri, modelin öngörü kabiliyetinin yüksekliğine işaret eder; özellikle bu iki ölçüt arasındaki farkın az olması, model çıktılarında uç değerlerden kaynaklı varyansın düşük olduğunu teyit etmektedir (Hyndman & Koehler, 2006).

$$RV_t = \sum_{i=1}^N r_t^2(i) \quad (3)$$

(3) nolu denklemde yer alan gerçekleşen varyans ilgili güne ait yüksek frekanslı getirilerin karelerinin toplamı üzerinden hesaplanmaktadır. Formülde yer alan $r_t(i)$ t günündeki i. yüksek frekanslı alt döneme ait getiriyi, N ise günlük toplam gözlem sayısını temsil etmektedir; bu çalışmada veriler 30 dakikalık periyotlarla örneklediği için N=17 olarak belirlenmiştir.

$$MedRV_t = \frac{\pi}{6-4\sqrt{3}+\pi} \left(\frac{N}{N-2} \right) \sum_{i=1}^{N-1} \text{med}(|r_t(i-1)|, |r_t(i)|, |r_t(i+1)|)^2 \quad (4)$$

(4) no'lu denklem MedRV, Andersen vd. (2012) tarafından ani fiyat sıçramalarına ve mikro-yapı gürültüsüne karşı dayanıklı bir gerçekleşen varyans ölçütü olarak önerilmiştir.

Formülde yer alan $\frac{\pi}{6-4\sqrt{3}+\pi}$ terimi, yanlışlığı gidermek amacıyla türetilmiş teorik bir ölçeklendirme sabitidir; $\left(\frac{N}{N-2}\right)$ ifadesi ise sonlu örneklem için uygulanan düzeltme çarpanıdır. Her üç ardışık yüksek frekanslı getirinin mutlak değerleri arasındaki medyanı alarak bunu kareleyerek toplamaktır. Medyanın kullanılması, olağandışı büyük tek bir gözlemin yani sıçramanın ölçütü aşırı biçimde yükseltmesini engellemektedir. Bu nedenle MedRV, gün içi oynaklığın daha sağlam ve güvenilir bir tahmincisi olarak kabul edilmektedir.

$$\sigma_t^2 = f_{ML}(r_{t-1}^2, \sigma_{t-1}^2) \quad (5)$$

5 no'lu denklem ise, hibrit GARCH modelinin girdi ve çıktılarını fonksiyonel yapıda temsil etmektedir. Bu denklem, geleneksel GARCH modellerinin tahmin ettiği koşullu varyansın parametrik olmayan bir işlev aracılığıyla elde edildiği hibrit yaklaşımı ifade etmektedir. Standart GARCH modelinde σ_t^2 , gecikmeli kare getiri r_{t-1}^2 ve gecikmeli koşullu varyans (σ_{t-1}^2) üzerine kurulan doğrusal bir denklemle modellenmektedir. Hibrit yaklaşımda ise bu doğrusal ilişki, f_{ML} ile gösterilen bir makine öğrenmesi algoritmasının esnekliğiyle ikame edilmektedir. Böylece GARCH modelinin yapısal varsayımlarına bağlı kalınsızlık, gecikmeli bilgi kümesinden koşullu oynaklık doğrusal olmayan biçimde öğrenilebilmektedir. Bu çerçevede f_{ML} ; RF, GBDT, KNN, SVR ve MLP gibi farklı algoritmalar aracılığıyla tahmin edilmiş ve her birinin oynaklığı açıklama kapasitesi karşılaştırmalı olarak değerlendirilmiştir.

Kullanılan modellerin performansları arasında istatistiksel olarak anlamlı bir fark bulunup bulunmadığının belirlenmesi, elde edilen bulguların doğru yorumlanabilmesi için önemlidir. Bu amaçla Hansen, Lunde ve Nason (2011) tarafından geliştirilen Model Güven Kümesi (Model Confidence Set, MCS) testi uygulanmıştır. İki rakip model arasında istatistiksel üstünlüğü belirlemek için geleneksel olarak Diebold ve Mariano (2002) testi gibi ikili karşılaştırma yöntemleri tercih edilirken; MCS testi ise çoklu karşılaştırmaya imkân tanınması ve daha az, daha zayıf varsayımlar gerektirmesi bakımından bu yöntemlerden ayrılmaktadır. MCS testi, belirlenen bir güven düzeyinde (α) anlamlılık eşiği belirleyerek H_0 (sıfır) hipotezi reddedilmeyene kadar bir dizi istatistiksel test uygulamakta ve kümedeki en kötü performanslı modeller elenmektedir. Bu işlem, kalan modeller arasında anlamlı bir fark kalmayana dek sürmektedir. Test istatistiklerinin asimptotik dağılımları standart olmadığından, bu dağılımlar bootstrap yöntemleriyle yaklaşık olarak elde edilmektedir. Volatilite tahmini literatüründe MCS testi hem modeller arası istatistiksel karşılaştırma hem de üstün tahmin gücüne sahip model alt kümelerinin belirlenmesi amacıyla başarıyla kullanılmıştır (Aras, 2021).

5. Ampirik Bulgular

5.1. GARCH Modellerinin Tahmin Performansı

Araştırmada beş GARCH modelinin dört farklı hata dağılımı varsayımı altındaki örneklem dışı tahmin performansını sırasıyla MAE ve RMSE ölçütleriyle Tablo 5 ve Tablo 6’da sunulmaktadır.

Tablo 5. GARCH modellerinin MAE'ye göre tahmin performansları

Modeller	RV				MedRV			
	Norm	Snorm	Ged	Sged	Norm	Snorm	Ged	Sged
sGARCH	0.1505	0.1508	0.1358	0.1336	0.1500	0.1504	0.1346	0.1326
EGARCH	0.1291	0.1291	0.1237	0.1240	0.1277	0.1277	0.1216	0.1219
GJR-GARCH	0.1331	0.1389	0.1339	0.1331	0.1319	0.1375	0.1325	0.1317
IGARCH	0.1419	0.1441	0.1498	0.1498	0.1411	0.1435	0.1492	0.1492
TGARCH	0.1277	0.1276	0.1221	0.1222	0.1264	0.1263	0.1202	0.1203

Not: Hata ölçüleri 1000 ile çarpılmıştır. Koyu fon en düşük değeri göstermektedir.

Tablo 6. GARCH modellerinin RMSE'ye göre tahmin performansları

Modeller	RV				MedRV			
	Norm	Snorm	Ged	Sged	Norm	Snorm	Ged	Sged
sGARCH	0.1978	0.2006	0.1896	0.1885	0.1722	0.1750	0.1600	0.1587
EGARCH	0.1860	0.1860	0.1825	0.1828	0.1591	0.1591	0.1540	0.1543
GJR-GARCH	0.1880	0.1920	0.1893	0.1888	0.1593	0.1640	0.1603	0.1597
IGARCH	0.1940	0.1949	0.1980	0.1980	0.1669	0.1684	0.1723	0.1723
TGARCH	0.1842	0.1840	0.1808	0.1809	0.1565	0.1563	0.1519	0.1520

Not: Hata ölçüleri 1000 ile çarpılmıştır. Koyu fon en düşük değeri göstermektedir.

Tablo 5 ve 6’da, her bir hata ölçütünde (sütunda), %95 anlamlılık düzeyinde istatistiksel olarak üstün performans gösteren modellerin oluşturduğu küme gölgelendirilerek gösterilmiştir; bu küme içerisinde en düşük hata değerine sahip model(ler) ise koyu fon ile işaretlenmiştir. Tablo 5 ve Tablo 6 birlikte incelendiğinde, TGARCH modeli, genelleştirilmiş hata dağılımı (Ged) varsayımı altında her iki hata ölçütü ve her iki temsil değişkeni için en yüksek tahmin doğruluğunu sergilemektedir (RV-MAE: 0.1221; RV-RMSE: 0.1808; MedRV-MAE: 0.1202; MedRV-RMSE: 0.1519). Her bir gruptaki gölgelendirilmiş modeller, MCS testi sonucunda istatistiksel olarak benzer öngörü performansı sağlamaktadır. Burada dikkat çekici olan, Ged ve Sged dağılımları altında EGARCH ve TGARCH modellerinin diğer modellere üstünlük kurduğudur. Her iki performans ölçüsü ve her iki oynaklık ölçüsünde de bu durum tutarlı bir şekilde gerçekleşmiştir. IGARCH, genelleştirilmiş hata dağılımları (Ged ve Sged)

altında en düşük başarıyı sergilerken; normal ve çarpık normal (Snorm) dağılımları altında bu konumu sGARCH almaktadır. sGARCH ve GJR-GARCH modelleri, özellikle Ged ve Sged dağılımları altında birbirine yakın değerler üretmektedir. EGARCH, TGARCH ve GJR-GARCH gibi asimetriyi yakalayan modellerin, sGARCH ve IGARCH gibi simetrik modellere göre daha iyi performans göstermesi asimetrisinin BIST 100 oynaklık modellemesinde önemli rol oynadığına işaret eder. Bu durum, veri setinde şokların yönüne bağlı farklı etkilerin (kaldıraç etkisi) önemli olduğu anlamına gelmektedir. Bununla birlikte, kalın kuyruk özelliklerini dikkate alan dağılımlar, özellikle GED ve SGED, normal ve skew-normal dağılımlara göre daha başarılı sonuçlar üretmekte, böylece finansal getirilerin aşırı değerler (uç gözlemler) içeren yapısını daha iyi yakalayarak model tahmin doğruluğunu artırmaktadır.

5.2. Makine Öğrenmesi Modellerinin Tahmin Performansı

Araştırmada, makine öğrenmesi modelinin performans istatistikleri ile en iyi GARCH modeli (TGARCH-Ged) karşılaştırmalı biçimde Tablo 7’de sunulmaktadır.

Tablo 7. Makine öğrenmesi modellerinin genel performans istatistikleri

Modeller	RV		MedRV	
	RMSE	MAE	RMSE	MAE
Best GARCH	0.1808	0.1221	0.1519	0.1202
KNN	0.1697	0.0797	0.1205	0.0581
RF	0.1697	0.0777	0.1241	0.0575
XGB	0.1704	0.0865	0.1270	0.0660
DT	0.1723	0.0799	0.1287	0.0772
GBDT	0.1725	0.0689	0.1290	0.0628
ExtraTrees	0.1739	0.0836	0.1301	0.0640
MLP	0.1790	0.1023	0.1328	0.0643
SGD	0.1807	0.0846	0.1363	0.0640
SVR	0.2691	0.1887	0.1922	0.1410

Not: Hata ölçüleri 1000 ile çarpılmıştır. Koyu fon en düşük değeri göstermektedir.

Tablo 7’de, her bir hata ölçütünde (sütunda), %95 anlamlılık düzeyinde istatistiksel olarak üstün performans gösteren modellerin oluşturduğu küme gölgelendirilerek gösterilmiştir; bu küme içerisinde en düşük hata değerine sahip model(ler) ise koyu fon ile işaretlenmiştir. Tablo incelendiğinde, Destek Vektör Regresyonu (SVR) dışındaki tüm makine öğrenmesi algoritmalarının, Ortalama Mutlak Hata (MAE) kriteri temelinde en uygun Genelleştirilmiş Otoregresif Koşullu Değişken Varyans (GARCH) spesifikasyonundan belirgin şekilde daha yüksek öngörü performansı sergilediğini ortaya koymaktadır. Gerçekleşen Oynaklık (RV) serisi analiz edildiğinde; Gradyan Artırımlı Karar Ağaçları (GBDT), MAE = 0,0689 değeri

ile en düşük hata payına ulaşarak üstünlük sağlamıştır. Hata karelerinin ortalamasının karekökü (RMSE) perspektifinden ise Rastgele Orman (RF) ve K-En Yakın Komşu (KNN) algoritmaları, $RMSE = 0,1697$ değeri ile en yüksek tahmin tutarlılığına ulaşan modeller olmuştur. Medyan Gerçekleşen Oynaklık (MedRV) temsil değişkeni kullanıldığında; RF ($MAE = 0,0575$) ve KNN ($MAE = 0,0581$) en başarılı modeller olarak ön plana çıkmaktadır. Buna karşın SVR modeli; RV serisi için elde ettiği $MAE = 0,1887$ ve $RMSE = 0,2691$ değerleri ile tüm modelleme setleri içerisinde en düşük öngörü performansını sergilemiştir. Makine öğrenmesi algoritmalarının sağladığı kestirim üstünlüğünün MAE ve RMSE kayıp fonksiyonları bağlamında belirgin bir asimetri sergilediği görülmektedir. Referans GARCH spesifikasyonuna kıyasla MAE düzeyinde yaklaşık %44'lük (0.1221'den 0.0689) kayda değer bir hata azalışı gerçekleşmiştir. Buna karşılık, aynı karşılaştırma RMSE ölçütü temelinde yapıldığında iyileşme oranı yaklaşık %6 ile sınırlı kalmaktadır (0,1808'den 0,1697'ye). MAE ve RMSE arasındaki bu belirgin makas, makine öğrenmesi modellerinin tipik/sıradan günlerde GARCH'a kıyasla belirgin bir üstünlük sergilediğine, ancak büyük tahmin hatalarının ortaya çıktığı uç günlerde bu üstünlüğün önemli ölçüde daralarak GARCH modeliyle yakınsadığına işaret etmektedir. RMSE, büyük hataları kareye alarak ağırlıklandıran bir ölçüt olduğundan, uç gözlemlerdeki performans farkına MAE'den daha duyarlıdır. Risk yönetimi açısından kritik olan dönemlerin tam da bu tür uç/stres günleri olduğu göz önüne alındığında, bu bulgu Bölüm 6'da sunulan politika çıkarımlarının ampirik temelini oluşturmaktadır.

Elde edilen ampirik bulgular arasında öne çıkan kritik bir husus, K-NN algoritmasının sergilediği rekabetçi öngörü gücüdür. Örnek tabanlı (*instance-based*) ve parametrik olmayan yalın mimarisine karşın K-NN; RV serisinde Rastgele Orman (RF) ile özdeş bir RMSE (0.1697) değerine ulaşmış, MedRV serisinde ise en düşük hata düzeyine işaret eden modeller arasında yer almıştır ($MAE=0.0581$). Bu ampirik kanıt, BIST 100 endeksindeki volatilité dinamiklerinin belirgin bir yerel rejim sürekliliği ve durum bağımlılığı sergilediğini doğrulamaktadır. Dolayısıyla, karmaşık parametrik olmayan ve topluluk (*ensemble*) mimarilerinin yanı sıra, faz uzayındaki geçmiş benzerlik örüntülerini esas alan bellek tabanlı yaklaşımların da finansal oynaklığın doğrusal olmayan yapısını yakalamada yüksek modelleme etkinliği sunduğunu göstermektedir.

Makine öğrenmesi modellerinin oynaklık tahmin performansı, referans model olarak alınan TGARCH-GED'e kıyasla genel olarak daha üstün olduğunu göstermektedir. En düşük hata değerleri ölçütü bazında farklı modeller tarafından üretilmiştir: KNN ve RF, RMSE ölçütünde her iki temsil değişkeninde de (RV ve MedRV) en düşük hataya ulaşırken; MAE ölçütünde RV serisinde GBDT, MedRV serisinde ise RF öne çıkmıştır. Bu dağılım, tek bir modelin tüm ölçüt ve temsil değişkeni kombinasyonlarında istikrarlı biçimde üstün

olmadığını, performansın seçilen kayıp fonksiyonuna bağlı olarak değiştiğini göstermektedir. Model Confidence Set (MCS) testi sonuçlarına göre, bu modeller arasında istatistiksel olarak anlamlı bir fark bulunmamakta olup aynı en iyi model kümesi içerisinde yer almaktadır. Bu durum, GARCH'a kıyasla gözlenen öngörü üstünlüğünün belirli bir algoritmaya özgü olmadığını, makine öğrenmesi paradigmasının genel bir özelliği olarak ortaya çıktığını göstermektedir.

Diğer taraftan, MedRV kullanılarak elde edilen sonuçların tüm modeller için daha düşük hata değerleri üretmesi, bu ölçütün mikro yapı gürültüsüne karşı daha dayanıklı olduğunu ve model performansını daha sağlıklı yansıttığını göstermektedir. Hem RV hem de MedRV sonuçları birlikte değerlendirildiğinde, model sıralamalarının büyük ölçüde korunduğu ve özellikle KNN, Random Forest ve Gradient Boosting tabanlı modellerin tutarlı biçimde en iyi performans gösteren modeller arasında yer aldığı görülmektedir. Buna karşın, TGARCH-GED modeli kendi sınıfı içerisinde güçlü bir performans sergilese de makine öğrenmesi modellerinin gerisinde kalmıştır. Bu bulgular, finansal oynaklık tahmininde esnek ve parametrik olmayan yöntemlerin önemini açık biçimde ortaya koymaktadır.

5.3. Bulguların Literatürle Karşılaştırılması

Bu bölümde elde edilen ampirik bulgular, ilgili literatürdeki temel çalışmalarla benzerlik ve farklılıkları açısından değerlendirilmektedir.

5.3.1. Literatürle Örtüşen Bulgular

Makine öğrenmesi modellerinin GARCH'a üstünlüğü bulgusu, literatürdeki pek çok çalışmayla uyum içindedir. Henriques ve Sadorsky (2023), enerji piyasası oynaklığında ML modellerinin örneklem dışında GARCH ailesini geride bıraktığını göstermiş; Christensen vd. (2022) ise minimum hiperparametre ayarına karşın ML yöntemlerinin HAR ailesini aştığını ortaya koymuştur. Bu çalışmada da ağaç tabanlı topluluk yöntemlerinin (GBDT, RF) ile örnek tabanlı bir öğrenici olan KNN'nin en iyi GARCH modelini MAE kriterinde açık farkla geride bırakması, söz konusu bulguları BIST 100 bağlamında da desteklemektedir.

Asimetrik GARCH modellerinin simetrik modellere üstünlüğü de literatürle örtüşmektedir. Nelson (1991) ve Glosten vd. (1993)'ün teorik öngörülerine paralel biçimde, bu çalışmada TGARCH ve EGARCH'ın sGARCH ve IGARCH'a kıyasla belirgin şekilde daha iyi performans göstermesi, kaldıraç etkisinin BIST 100 oynaklık sürecinde belirleyici bir rol oynadığına işaret etmektedir.

Kalın kuyruklu hata dağılımlarının (GED, SGED) normal dağılıma kıyasla daha başarılı sonuçlar üretmesi de finansal getirilerin leptokurtik yapısına ilişkin yaygın literatür bulgusuyla uyumludur. Ağaç tabanlı topluluk yöntemlerinin üstün performansı ise Aras (2021)'in istifleme topluluğunun tekil modellerden daha başarılı olduğu bulgusunu ve Çelik ve Ergin (2022)'in ensemble yaklaşımların potansiyelini öne çıkaran bulgularını desteklemektedir. MedRV'nin hedef değişken olarak kullanılmasının tahmin hatalarını düşürmesi, Andersen vd. (2012)'nin bu ölçütün mikro yapı gürültüsüne karşı sağladığı teorik üstünlüğünü ampirik düzeyde doğrulamaktadır.

5.3.2. Literatürden Ayrışan Bulgular

Bu çalışmanın literatürden ayrışan en dikkat çekici bulgusu SVR modelinin performansına ilişkindir. Aras (2021), Bitcoin oynaklığı üzerine yürüttüğü çalışmada SVM/SVR'ye olumlu sonuçlar raporlarken, bu çalışmada SVR tüm modeller içinde en zayıf tahmin başarımını sergilemiştir. Bu tutarsızlık büyük olasılıkla iki çalışmanın farklı varlık sınıfları (kripto para ve hisse senedi endeksi), farklı piyasa yapısı ve farklı model çerçeveleri (hibrit GARCH ve doğrudan ML) kullanmasından kaynaklanmaktadır; nitekim SVR'nin yüksek boyutlu ve gürültülü zaman serilerindeki genelleme kapasitesi, düzenli fiyat seyreden kripto para verilerine kıyasla daha kırılgan olabilmektedir.

Zhang vd. (2022), gün içi RV tahmininde sinir ağlarının ağaç tabanlı modelleri performans açısından aştığını bulgulamıştır. Ancak bu çalışmada tam tersi bir sıralama gözlemlenmiş; RF, GBDT gibi ağaç tabanlı modeller veya örnek tabanlı bir öğrenici olan KNN modeli, MLP'yi her iki hata ölçütünde de geride bırakmıştır. Bu farklılığın olası açıklaması, Zhang vd. (2022)'nin çok sayıda hisse senedi içeren geniş panelden yararlanarak sinir ağlarının derin temsil kapasitesini daha etkin kullanabilmesi, buna karşın bu çalışmada kullanılan görece kısa tek endeks serisinin MLP için yeterli eğitim sinyali sağlayamamasıdır.

Son olarak, hibrit GARCH+ML modellerini inceleyen Kristjanpoller ve Minutolo (2015), Kim ve Won (2018) ile Xu vd. (2024) hibrit yapıların tek başına GARCH'ı belirgin biçimde geride bıraktığını göstermiştir. Bu çalışmada ise GARCH çıktıları doğrudan ML modellerine girdi olarak aktarılmamış; iki yaklaşım bağımsız olarak kıyaslanmıştır. Üretebildiği ML modellerinin elde ettiği üstün performans, hibrit entegrasyona gerek kalmaksızın parametrik olmayan yaklaşımların tek başına rekabetçi sonuçlar üretebildiğine işaret etmektedir. Bu bulgu, hibrit yapıların sistematik biçimde test edilmesini gelecek araştırmalar için öncelikli bir yön olarak belirleyen metodolojik bir gerekçe de sunmaktadır.

6. Sonuç

Finansal piyasalarda oynaklığın doğru tahmin edilmesi; risk yönetimi, portföy optimizasyonu ve opsiyon fiyatlandırması gibi kritik uygulamalar açısından önemlidir. Bu durum, yapısal kırılmalara ve dışsal şoklara karşı yüksek duyarlılık sergileyen gelişmekte olan piyasalarda daha da belirginleşmektedir. BIST 100 endeksi, küresel finansal krizler, döviz şokları ve pandemi süreçlerinde gözlemlenen keskin fiyat hareketleri bu kırılmanın somut bir örneğini sunmaktadır. Bu bağlamda çalışma, söz konusu piyasaya özgü oynaklık dinamiklerini hem parametrik hem de parametrik olmayan yöntemlerle kapsamlı biçimde modellemeyi amaçlamıştır.

Ampirik bulgular üç temel sonuç etrafında şekillenmektedir. Birincisi, GARCH ailesi içinde TGARCH-GED spesifikasyonu hem RV hem de MedRV temsil değişkenleri için en üstün tahmin başarımını sergilemiştir. Bu sonuç, BIST 100'deki oynaklık sürecinin kaldıraç etkisi ve kalın kuyruklu dağılım yapısını birlikte barındırdığını göstermektedir; negatif şokların pozitif şoklara kıyasla daha şiddetli bir oynaklık tepkisine yol açması, BIST'deki yatırımcı davranışının asimetric bilgi işleme yapısıyla örtüşmektedir. İkincisi, ağaç tabanlı topluluk yöntemleri olan (GBDT, RF) ile örnek tabanlı bir öğrenici olan KNN, MAE ölçütünde en iyi GARCH modelini açık farkla geride bırakmıştır. Özellikle yapısal olarak yalın bir algoritma olan KNN'in karmaşık topluluk modelleriyle eşdeğer RMSE (0.1697) üretmesi, serideki yerel durum bağımlılığının ve hafıza etkisinin gücüne işaret etmektedir. Bu bulgunun ekonomik anlamı şu şekilde ifade edilebilir: BIST 100 oynaklık süreci, sabit parametreliliye doğru bir varyans denklemiyle tam olarak temsil edilemeyen, zamanla değişen ve doğrusal olmayan dinamikler içermektedir. Makine öğrenmesi algoritmalarının bu dinamikleri veri güdümlü biçimde öğrenme kapasitesi, gelişmekte olan piyasalarda risk ölçüm modellerinin daha güvenilir tahminler üretmesine doğrudan katkı sağlayabilir. Üçüncüsü, MedRV'nin hedef değişken olarak benimsenmesi tüm modellerde tahmin hatalarını sistematik olarak azaltmıştır. Bu bulgu, yüksek frekanslı verideki mikro yapı gürültüsünü ve ani fiyat sıçramalarını filtreleyen sağlam varyans ölçütlerinin oynaklık modellemesindeki pratik üstünlüğünü ampirik düzeyde doğrulamaktadır.

Elde edilen bulgular literatürle geniş ölçüde uyum içindedir; ML modellerinin GARCH'a üstünlüğü Henriques ve Sadorsky (2023) ile Christensen vd. (2022)'nin bulgularını, asimetric GARCH modellerinin simetric yapıları üstünlüğü ise Nelson (1991) ve Glosten vd. (1993)'ün teorik öngörülerini desteklemektedir. Bununla birlikte SVR'nin bu çalışmada zayıf performans sergilemesi Aras (2021)'in olumlu bulgularından ayrılmakta; sinir ağlarının ağaç tabanlı modelleri geride bırakmaması ise Zhang vd. (2022) ile ters bir sıralama ortaya

koymaktadır. Bu ayrışmalar, piyasa yapısı ve veri uzunluğunun model seçimi üzerindeki belirleyici rolünü bir kez daha gündeme taşımaktadır.

Bulgular, risk yönetimi ve düzenleyici otoriteler açısından önemli politika çıkarımları sunmaktadır. MAE ve RMSE metrikleri arasındaki performans makası, makine öğrenmesi algoritmalarının doğrudan tüm risk yönetimi süreçlerine tek başına ikame edilmesinden ziyade, hibrit veya çok katmanlı bir çerçevede değerlendirilmesi gerektiğini göstermektedir. Tipik piyasa koşullarında portföy optimizasyonu ve günlük sermaye tahsisi gibi operasyonel kararlarda makine öğrenmesi modellerinin yüksek öngörü gücünden yararlanılması önerilmektedir. Ancak finansal sistemin kırılganlığını belirleyen kriz dönemleri, stres testleri ve aşırı kuyruk riski (Riske Maruz Değer - VaR ve Beklenen Kayıp - ES) hesaplamalarında, kalın kuyruklu hata dağılımlarına sahip asimetrik GARCH spesifikasyonlarının güvenlik marjı sağlayıcı rolü göz ardı edilmemelidir.

Bu çalışmanın bazı kısıtlamalar barındırdığı da göz önünde bulundurulmalıdır. Her şeyden önce, analiz yalnızca BIST 100 endeksini kapsamakta olup bulguların diğer gelişmekte olan piyasalara ya da bireysel hisse senetlerine doğrudan genellenmesi dikkatli bir değerlendirme gerektirmektedir. Çalışmada GARCH çıktıları ML modellerine doğrudan girdi olarak aktarılmamış, iki yaklaşım bağımsız olarak kıyaslanmıştır; dolayısıyla tam hibrit yapıların potansiyel kazanımı bu çerçevede ölçülememektedir. Çalışmada doğrudan bir Riske Maruz Değer (VaR) ve Beklenen Kayıp (ES) geri-test prosedürü (Kupiec, Christoffersen veya dinamik kantil testleri) uygulanmamıştır; dolayısıyla modellerin nihai sermaye yeterliliği performansı ampirik bir kısıt niteliğindedir.

Çalışmada kıyaslama ölçütü olarak parametrik GARCH spesifikasyonları esas alınmış; yüksek frekanslı oynaklık literatüründe sıklıkla tercih edilen HAR-RV, HAR-J ve HAR-CJ tipi modeller dâhil edilmemiştir. Gelecek araştırmalarda, esnek makine öğrenmesi ve derin öğrenme yaklaşımlarının yalnızca GARCH ailesiyle değil, aynı zamanda sürekli ve süreksiz sıçrama bileşenlerini ayırtıran HAR ailesiyle de karşılaştırmalı olarak test edilmesi, volatilité öngörü literatürüne önemli ampirik katkılar sunacaktır.

Gelecek araştırmaların; makine öğrenmesi algoritmalarını sıçrama ayırtırmalı HAR ailesiyle doğrudan kıyaslaması, formal VaR/ES geri-test modellerini ampirik çerçeveye dâhil etmesi, GARCH+ML hibrit mimarilerini test etmesi ve SHAP tabanlı açıklanabilir yapay zekâ (XAI) araçlarıyla ağaç modellerinin ekonomik yorumlanabilirliğini güçlendirmesi önerilmektedir.

Kaynakça

- ANDERSEN, T. G. & BOLLERSLEV, T.; (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts", *International Economic Review*, 39(4), pp. 885–905. <https://doi.org/10.2307/2527343>.
- ANDERSEN, T. G., DOBREV, D. & SCHAUMBURG, E.; (2012), "Jump-Robust Volatility Estimation Using Nearest Neighbor Truncation", *Journal of Econometrics*, 169(1), pp. 75–93. <https://doi.org/10.1016/j.jeconom.2012.01.011>.
- ARAS, S.; (2021), "Stacking Hybrid GARCH Models for Forecasting Bitcoin Volatility", *Expert Systems with Applications*, 174, Article 114747. <https://doi.org/10.1016/j.eswa.2021.114747>.
- BARNDORFF-NIELSEN, O. E. & SHEPHARD, N.; (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps", *Journal of Financial Econometrics*, 2(1), pp. 1–37. <https://doi.org/10.1093/jjfinec/nbh001>.
- BOLLERSLEV, T.; (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31(3), pp. 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1).
- BREIMAN, L.; (2001), "Random Forests", *Machine Learning*, 45(1), pp. 5–32. <https://doi.org/10.1023/A:1010933404324>.
- CHRISTENSEN, K., SİGGAARD, M. & VELİYEV, B.; (2022), "A Machine Learning Approach to Volatility Forecasting", *Journal of Financial Econometrics*, 21(5), pp. 1680–1727. <https://doi.org/10.1093/jjfinec/nbac020>.
- ÇELİK, S. & ERGİN, H.; (2014), "Volatility Forecasting Using High Frequency Data: Evidence from Stock Markets", *Economic Modelling*, 36, pp. 176–190. <https://doi.org/10.1016/j.econmod.2013.09.038>.
- ÇELİK, S. & ERGİN, H.; (2022), "Destek Vektör ve Rassal Orman Regresyonu ile Borsa Endeks Değerlerinin Tahmini", *İktisat Politikası Araştırmaları Dergisi*, 9(2), ss. 153–172.
- DIAKOPOULOS, N.; (2015), "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures", *Digital Journalism*, 3(3), pp. 398–415.
- DROST, F. C. & KLAASSEN, C. A. J.; (1997), "Efficient Estimation in Semiparametric GARCH Models", *Journal of Econometrics*, 81(1), pp. 193–221. [https://doi.org/10.1016/S0304-4076\(97\)00042-0](https://doi.org/10.1016/S0304-4076(97)00042-0).

- ENGLE, R. F.; (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50(4), pp. 987–1007. <https://doi.org/10.2307/1912773>.
- FAMA, E. F.; (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work", *The Journal of Finance*, 25(2), pp. 383–417. <https://doi.org/10.2307/2325486>.
- GLOSTEN, L. R., JAGANNATHAN, R. & RUNKLE, D. E.; (1993), "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks", *The Journal of Finance*, 48(5), pp. 1779–1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>.
- HENRIQUES, B. M. & SADORSKY, P.; (2023), "Modelling and Forecasting Energy Market Volatility Using GARCH and Machine Learning Approach", *arXiv preprint*. arXiv:2405.19849.
- HYNDMAN, R. J. & KOEHLER, A. B.; (2006), "Another Look at Measures of Forecast Accuracy", *International Journal of Forecasting*, 22(4), pp. 679–688.
- KİM, H. Y. & WON, C. H.; (2018), "Forecasting the Volatility of Stock Price Index: A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models", *Expert Systems with Applications*, 103, pp. 25–37. <https://doi.org/10.1016/j.eswa.2018.03.002>.
- KRİSTJANPOLLER, W. & MİNUTOLO, M. C.; (2015), "Gold Price Volatility: A Forecasting Approach Using the Artificial Neural Network–GARCH Model", *Expert Systems with Applications*, 42(20), pp. 7245–7251. <https://doi.org/10.1016/j.eswa.2015.04.058>.
- LÓPEZ DE PRADO, M.; (2018), "Advances in Financial Machine Learning", *Wiley*.
- MADHAVAN, A.; (2000), "Market Microstructure: A Survey", *Journal of Financial Markets*, 3(3), pp. 205–258. [https://doi.org/10.1016/S1386-4181\(00\)00007-0](https://doi.org/10.1016/S1386-4181(00)00007-0).
- NELSON, D. B.; (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, 59(2), pp. 347–370. <https://doi.org/10.2307/2938260>.
- SHİLLER, R. J.; (2003), "From Efficient Markets Theory to Behavioral Finance", *Journal of Economic Perspectives*, 17(1), pp. 83–104.
- STİGLİC, G., KOCBEK, P., FİJACKO, N., ZİTNİK, M., VERBERT, K. & CİLAR, L.; (2020), "Interpretability of Machine Learning-Based Prediction Models in Healthcare", *WIREs Data Mining and Knowledge Discovery*, 10, e1379. <https://doi.org/10.1002/widm.1379>.
- TEKTAŞ, A. & KARATAŞ, A.; (2021), "Makine Öğrenmesi Teknikleri ile Hisse Senedi Fiyat Tahmini", *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 16(1), ss. 1–18. <https://doi.org/10.17153/oguiibf.636017>.

- XU, Z., LIÉCHTY, J., BENTHALL, S., SKAR-GİSLİNGE, N. & McCOMB, C.; (2024), "GARCH-Informed Neural Networks for Volatility Prediction in Financial Markets", *arXiv preprint*. arXiv:2410.00288.
- ZAKOÏAN, J. M.; (1994), "Threshold Heteroskedastic Models", *Journal of Economic Dynamics and Control*, 18(5), pp. 931–955. [https://doi.org/10.1016/0165-1889\(94\)90039-6](https://doi.org/10.1016/0165-1889(94)90039-6).
- ZHANG, C., ZHANG, Y., CUCURİGU, M. & QİAN, Z.; (2022), "Volatility Forecasting with Machine Learning and Intraday Commonality", *Journal of Financial Econometrics*, 22(2), pp. 492–530. <https://doi.org/10.1093/jjfinec/nbad005>.
- ZHANG, L., MYKLAND, P. A., & AİT-SAHALİA, Y. (2005). "A tale of two time scales: Determining integrated volatility with noisy high-frequency data." *Journal of the American Statistical Association*, 100(472), 1394-1411.