

HİSSE SENEDİ GETİRİLERİNDEKİ VOLATİLİTENİN TAHMİNLENMESİNDE DESTEK VEKTÖR MAKİNELERİNE DAYALI GARCH MODELLERİNİN KULLANIMI



Kafkas Üniversitesi İktisadi ve
İdari Bilimler Fakültesi Dergisi
Cilt 5, Sayı 8, 2014
ISSN : 1309 - 4289

Mutlu GÜRSOY
Yrd. Doç. Dr.,
İstanbul Arel Üniversitesi,
İİBF İşletme Bölümü
mutlugursoy@arel.edu.tr

Mehmet Erdal BALABAN
Prof. Dr.,
İstanbul Üniversitesi,
İşletme Fakültesi
balaban@istanbul.edu.tr

ÖZ | Belirsiz bir değişkenin alabileceği olası tüm değerlerin dağılımının ifadesi olarak volatilité, finansal piyasalardaki varlıkların getirilerinin sergilediği değişkenliği dikkate alması gereken bir yatırımcı için hayati bir olgudur. Sonuç olarak volatilitenin modellenmesi ve tahminlenmesi finansal risk yönetiminde önemli rol oynar. Bu çalışmada GARCH tipi modellerden GARCH(1,1), EGARCH(1,1) ve GJR-GARCH(1,1) modellerine, son yıllarda gittikçe popüleritesi artan güçlü bir makine öğrenmesi metodu olan Destek Vektör Makineleri (DVM) ile yaklaşılmıştır. Bu amaçla 04.01.2007 – 31.12.2012 dönemine ait günlük İMKB ulusal 100 endeksi-kapanış fiyatları kullanılmış ve modellerin klasik çözümü ile DVM çözümlerinin tahminleme performansları kıyaslanmıştır. Elde edilen sonuçlara göre, DVM'ye dayalı karma GARCH modellerinin daha iyi performans gösterdiği gözlenmiştir.

Anahtar Kelimeler: PDestek Vektör Makineleri, GARCH, EGARCH, GJR-GARCH, Volatilité

Jel Kodu: C45, C53



Makaleyi çevrimiçi görüntülemek için QR
Kodu okutunuz.

VOLATILITY FORECASTING IN STOCK RETURNS USING SUPPORT VECTOR MACHINES BASED GARCH MODELS



Kafkas University Journal of
Economics and Administrative
Sciences Faculty
Vol. 5, Issue 8, 2014
ISSN : 1309 - 4289

Mutlu GÜRSOY

Assoc. Prof. Dr.

Istanbul Arel University,
Business Administration Department
mutlugursoy@arel.edu.tr

Mehmet Erdal BALABAN

Prof. Dr.,

Istanbul University
School of Business
Department of Quantitative Methods
balaban@istanbul.edu.tr

ABSTRACT | Volatility, as a spread of all likely outcomes of an uncertain variable, is crucial phenomenon for the investors who must consider the spread of asset returns in financial markets. Therefore, volatility modelling and forecasting plays an important role in financial risk management. Support Vector Machine (SVM) is an efficient learning technique for classification and regression problems, including financial time series forecasting. In this study, we aimed to compare the forecasting performance of SVM based GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) models with their corresponding classical models using daily returns in Istanbul Stock Exchange for the period 04.01.2007 – 30.12.2012. The results confirmed the remarkable generalization performance of SVM, as shown in the SVM literature.

Keywords: Support Vector Machines, GARCH, EGARCH, GJR-GARCH, Volatility.

Jel Code: C45, C53



Scan QR Code to see this article online

1. GİRİŞ

Genel olarak zaman serileri analizi ve tahminleme, özel olarak volatilitenin modellenmesi ve tahminlenmesi, finans alanında hayati bir rol oynamaktadır. Finansal zaman serilerini üreten süreçler üzerinde deterministik unsurlardan ziyade stokastik unsurların daha baskın olması nedeni ile, zaman serileri ile tahminleme geleneksel olarak oldukça zor bir problem olarak görülür. Finansal varlıkların getiri serilerinin dağılımı ve davranışı birçok araştırmaya konu olmuş ve iyi bilinen bazı örüntülere (pattern) sahip oldukları gözlenmiştir. Volatilitenin kümelenmesi (volatility clustering), kaldıraç etkisi (leverage effects), aşırı basıklık (leptokurtosis), asimetric fiyat hareketleri (gain/loss asymmetry) vb. karakteristikleri dikkate alan ekonometrik araçlar 1980'li yıllarda geliştirilmeye başlamıştır. Volatilitenin modellenmesi ve kestiriminde finansal ekonometri alanında geliştirilmiş birçok model mevcut olmakla birlikte, baskın ekonometrik modeller GARCH ailesi olarak adlandırılan modellerdir. Çalışmanın kapsamı Engle (1982) tarafından geliştirilen ARCH modelinin bir genelleştirimi olan ve Bollerslev (1986) tarafından geliştirilen GARCH (Generalized ARCH) modeli ve onun bazı eksiklerini tamamlayan EGARCH (Nelson, 1991) ve GJR-GARCH (Glosten vd., 1993) modelleri ile sınırlandırılmıştır.

Finansal varlıkların getirilerinin dağılımı esas olarak sıfır ortalama, yüksek basıklık ve düşük otokorelasyon ile karakterize edilir. Otokorelasyonun olmaması getirilerdeki artış ya da azalışların (birbirinden) bağımsız olmasını gerektirmez. Bağımsızlık getirilerin herhangi bir doğrusal olmayan fonksiyonunun da otokorelasyon içermemesini gerektirir (Cont, 2001: 230). Pratikte getirilerin, mutlak değerleri ve kareleri gibi basit doğrusal olmayan fonksiyonlarının anlamlı otokorelasyon ya da ısrarcılık (persistence) sergiledikleri görülmektedir. Bu durum volatilitenin kümelenmesi olarak adlandırılan olgudur ve Akgiray (1989), West ve Cho (1995), Andersen ve Bollerslev (1998) tarafından yapılan çalışmalarda GARCH modelinin bu özelliği iyi yakaladığı belirtilmiştir. Ancak Black (1976) tarafından menkul getirilerinin volatilitenin olumsuz haberler karşısında olumlu haberlere nazaran daha fazla artış gösterdiği tespit edilmiştir. Kaldıraç etkisi olarak adlandırılan bu olgu Nelson (1991), Engle ve Ng (1993) ve Campbell ve Kyle (1993) tarafından ampirik olarak doğrulanmıştır. Negatif ve pozitif volatilitenin şoklarına simetrik tepki veren GARCH modeli kaldıraç etkisinin modellenmesinde yetersiz kalmaktadır. EGARCH ve GJR-GARCH modelleri asimetric modeller olup iyi ve kötü haberlerin volatilitenin üzerindeki farklı etkilerini dikkate almaktadır.

İstatistiksel Öğrenme Teorisi-İÖT (Statistical Learning Theory)'ye dayanan ve ileri beslemeli bir ağ çeşidi olan DVM, yüksek genelleme (generalization) kapasitesi, aşırı uyum (Overfitting) ve yerel minimum problemlerinin aşılması gibi özellikleri nedeniyle umut verici bir makine öğrenmesi tekniği haline gelmektedir.

Bu çalışmada Pérez-Cruz vd. (2003) ve Ou ve Wang (2010) yaklaşımları izlenerek; DVR, DVM Regresyon problemini göstermek üzere, karma DVR-GARCH, DVR-EGARCH ve DVR-GJR modellerinin İMKB ulusal 100 fiyat endeksindeki volatilitenin tahminleme performansları,

klasik maksimum olabilirlik-ML (Maximum Likelihood) metoduna dayanan kestirimlerinin tahminleme performansları ile kıyaslanmıştır.

2. LİTERATÜR TARAMASI

DVM oldukça geniş bir alanda sınıflandırma ve regresyon problemlerine uygulanmasının yanısıra, bir çok araştırmacı tarafından finansal zaman serilerinin tahminlenmesi problemlerine de başarı ile uygulanmıştır (Trafalis ve İnce, 2000; Tay ve Cao, 2001; Kim, 2003). Ayrıca İMKB üzerine yapılan çalışmalardan sınıflandırma (Öğüt vd., 2009; Özdemir vd., 2011; Kara vd., 2011 ; Timor vd., 2012) ve regresyon (İnce, 2005) problemi olarak oluşturulan DVM ve DVR modelleri oldukça başarılı sonuçlar vermiştir.

Volatilite modellemesi kapsamında; Donaldson ve Kamstra (1997) tarafından Yapay Sinir Ağı (YSA) 'na dayalı GJR-GARCH modeli önerilmiş ve S&P500, NIKKEI, TSEC ve FTSE endeksleri için klasik parametrik modele göre daha iyi performans elde edilmiştir. Yümlü vd. (2005), İMKB endeks serisindeki volatilitenin tahminlenmesi problemini üç farklı YSA mimarisi ile ele almış ve başarılı sonuçlar elde etmişlerdir. Bildirici ve Ersin (2009), dokuz farklı GARCH ailesi model için YSA yaklaşımı uygulamış ve çoğu karma modelin İMKB endeks serisindeki volatiliteni daha iyi tahminlediği belirtilmiştir. Pérez-Cruz vd. (2003) dört uluslararası hisse senedi piyasası ve iki hisse senedi için DVR'ye dayalı karma GARCH(1,1) modeli ile volatiliteni tahminlemesi yapmış ve klasik modele göre IBEX serisi haricinde tümü için R-kare istatistiğine göre daha iyi performans elde etmiştir. Van Gestel vd. (2001) DAX30 serisi için bir DVM çeşidi olan En Küçük Kareler DVR metodunu Bayescil çerçevede ele alarak volatiliteni tahminlemesi yapmış ve klasik modele göre daha iyi performans elde etmiştir. Ou ve Wang (2010) üç Asya hisse senedi piyasası için En Küçük Kareler DVR'yi kullanarak karma GARCH(1,1), EGARCH(1,1) ve GJR-GARCH(1,1) modelleri ile volatiliteni tahminlemesi yapmış ve klasik modellere göre daha iyi performans elde etmiştir. Gavrishchaka ve Banerjee (2006) S&P500 serisi için DVR(10) modelini 4 farklı klasik GARCH ailesi model ile kıyaslamış ve oldukça başarılı sonuç elde etmiştir. Bildirici ve Ersin (2012) GARCH, APGARCH, FIGARCH ve FIAPGARCH modellerini (Logistic) Smooth Transition Autoregressive ve YSA yaklaşımı ile güçlendirmiş ve YSA'ya dayalı modellerin, volatiliteni kümelenebilmesi, asimetri ve doğrusal olmayan özellikleri klasik modellere göre daha etkin olarak ortaya çıkardığını göstermiştir. Ayrıca dalgacık (wavelet) ve spline dalgacık DVR modellerinin GARCH kestirimlerinde kullanımı ile başarılı sonuçlar elde edilmiştir (Tang, Tang ve Sheng, 2009; Tang, Sheng ve Tang, 2009).

3. DESTEK VEKTÖR MAKİNELERİ

Destek Vektör Makineleri ilk olarak 1960'lı yıllarda Vapnik ve arkadaşları tarafından başlatılan çalışmalarla ortaya çıkmış ve İÖT'ye dayanan bir makine öğrenmesi algoritmasıdır. İÖT'nin ve buna dayandırılan DVM'nin temel kaynağı olarak Vapnik (1995, 1998), Burges

(1998), Smola ve Schölkopf (2004) incelenebilir.

3.1. Öğrenme Problemi ve İstatistiksel Öğrenme Teorisi

$x \in X \subseteq \mathbb{R}^d$ ve $y \in Y \subseteq \mathbb{R}$ ve aralarında olasılıklı bir ilişki olan iki rassal değişken kümesi ve $D_n := \{(x_i, y_i) \in X \times Y\}_{i=1}^n$ girdi-çıkı çiftlerinden oluşan eğitim seti olsun. Denetimli öğrenme (supervised learning) ya da örneklerden öğrenme bağlamında öğrenen makine, programlanmaktan ziyade, eğitim seti üzerinde verilen bir işlemi (örneğin sınıflandırma ya da regresyon kestirimi) yapması için eğitilir. Bu yaklaşımda eğitim, girdi ve çıktılar arasındaki ilişkiyi en iyi tanımlayan $f(x)$ fonksiyonun seçilmesi işlemidir. İÖT'nin temel problemi seçilen fonksiyonun ne kadar iyi genelleme yapacağıdır. Diğer bir ifade ile seçilen fonksiyonun yeni girdi değerleri için çıktıları ne kadar iyi tahminleyebileceğidir. Burada girdi-çıkı çiftleri arasındaki ilişki olasılıksaldır çünkü genellikle X 'in bir elemanı Y 'nin bir elemanını tek değerli (unique) olarak belirleyememektedir. X ve Y uzayları üzerinde herhangi bir belirgin dağılım varsayımı olmayıp, gözlem noktaları zamanla değişmeyen, bilinmeyen ve $X \times Y$ uzayında tanımlı bir ortak olasılık (joint probability) dağılımından $\{P(x, y) = P(x)P(y|x)\}$ örneklenmiştir.

Eğer y değerini $f(x)$ olarak tahminlemekten dolayı oluşan hatanın ölçüsünü $L(y, f(x))$ kayıp (loss) fonksiyonu ile tanımlarsak, kayıp fonksiyonun riski; $P(x, y)$ dağılımından gelen her bir nokta için oluşan kaybın ortalama değeri diğer bir ifade ile beklenen değeri olacaktır ve $R(f) := E[L(y, f(x))]$ şeklinde gösterilir. Riskin beklenen değeri oldukça geniş bir fonksiyonlar kümesi $H := \{f: X \rightarrow Y\}$ (hipotez uzayı) üzerinde tanımlanır. Buradaki sorun gözlem çiftlerini üreten dağılımının bilinmemesidir ve dolayısıyla beklenen riski minimize eden ve hedef fonksiyon olarak da adlandırılan fonksiyonunun belirlenememesidir. Bu durumda elimizdeki sınırlı sayıdaki eğitim verisini kullanarak öğrenme algoritmasını uygulayabileceğimiz bir tümevarım prensibine gerek duyulur ve bu İÖT içerisinde Ampirik Risk Minimizasyonu (ARM) (Empirical Risk Minimization) metodu ile sağlanır.

ARM metodu riskin beklenen değerininin, D_n veri setini kullanarak, stokastik bir yaklaşımının (approximation) elde edilmesini sağlar ve ampirik risk $R_{amp}(f) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i))$ olarak adlandırılır. Ampirik riskin H hipotez uzayı üzerinde doğrudan minimizeasyonu olan bu problem, ampirik riski minimize edebilecek birden çok fonksiyon olabilmesi nedeniyle kötü konumlanmış bir problemdir (Vapnik, 1998). Ayrıca ampirik riskin minimum değeri sıfır ya da sıfıra çok yakın olsa dahi, gerçek riskin (beklenen risk) değeri çok daha büyük olabilir ve bu durum aşırı uyuma yol açabilir.

İÖT ise ampirik riskin beklenen riskten sapmasına olasılıksal bir sınır koyar. Böylece ampirik risk minimize edilirken aşırı uyum problemi kontrol edilir. Bu sınır, gözlem sayısı n ve

hipotez uzayının karmaşıklığını gösteren kapasite değeri olmak üzere en azından $1-\eta$ olasılık ile

$$R(f) < R_{amp}(f) + \Phi\left(\sqrt{\frac{h}{n}}, \eta\right)$$

şeklinde olacaktır. Burada Φ fonksiyonu η ve $\frac{h}{n}$ parametrelerine bağlı artan bir fonksiyondur ve risk fonksiyonelinin düzenleme (regularization) terimidir. Beklenen riskin küçük olması için eşitsizliğin sağ tarafındaki iki terimde aynı anda minimize edilmesi gerekmektedir. İÖT’de hipotez uzayının kapasite ölçüsü olarak bazı uygun nicelikler verilse de en popüler olanı VC-Boyutu’dur (Vapnik-Chervonenkis Dimension).

Hipotez uzayı oldukça büyüktür ve DVM’nin de dayandırıldığı Yapısal Risk Minimizasyonu (YRM) (Structured Risk Minimization) yaklaşımı ile uzayın, kapasite ölçülerine göre, m adet denklik sınıfı oluşturulur. Böylece daha küçük H hipotez uzayı elde edilir. Her bir H_i denklik sınıfı için VC-Boyutu h_i ve $h_i \leq h_{i+1}$ olmak üzere, hipotez uzayı iç içe geçmiş bir yapıya kavuşur ve problem $H_1 \subset H_2 \subset \dots \subset H_m$

$R_{amp}(f) + \Phi(\sqrt{h_i/n}, \eta)$ ifadesinin H uzayında minimizasyonu haline dönüşür.

3.2. DVM Regresyon Modelinin Formülasyonu

ε -DVR’de amaç, verilen D_n eğitim seti için, y_i değerlerinden en fazla ε kadar sapacak doğrusal $f(x)$ fonksiyonunu bulmaktır. Diğer bir ifade ile girdi uzayında gözlem noktalarına olabildiğince yakın uzanan bir hiperdüzlemin kurulmasıdır. Kayıp fonksiyonu olarak Vapnik tarafından önerilen ε -duyarsız (ε -insensitive) kayıp fonksiyonu kullanılır. Tanımlanan kayıp fonksiyonu değerlerinden sapmaları bir $[-\varepsilon, \varepsilon]$ aralığında tolere edip, aralığın dışında doğrusal olarak cezalandırır. $\langle \cdot, \cdot \rangle$ nokta çarpım operatörünü göstermek üzere, hedef fonksiyon

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b, \mathbf{w} \in X, b \in \mathbb{R} \quad (1)$$

olarak verildiğinde fonksiyonun düzlüğünü (flatness) kontrol etmek için, diğer bir ifade ile aşırı uyumdan kaçınmak için minimize edilmelidir. Bu durumda risk fonksiyoneli DVR

problemi için $R_{amp}(f) + \frac{\lambda}{2} \|\mathbf{w}\|^2$ halini alır ve regresyon problemi doğrusal durum için,

$$\min. \frac{1}{2} \|\mathbf{w}\|^2$$

$$\text{öyleki} \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \quad (2)$$

konveks optimizasyon problemi olarak ifade edilir. Model yapısal olarak eğitim setindeki her noktanın, ε toleransı içerisinde, yaklaştırımını sağlayan bir $f(x)$ fonksiyonunun varlığına işaret eder. Daha matematiksel olarak, optimizasyon probleminin uygun bir çözümünü varsayar. Ancak bu her zaman mümkün olmayabilir ya da bazı hatalara izin verilmek istenebilir. Bu noktada “yumuşak marj” (soft margin) yaklaşımı kayıp fonksiyonu kullanılır. (2)’deki her bir kısıt için ξ ve ξ^* aylak değişkenleri kullanarak model,

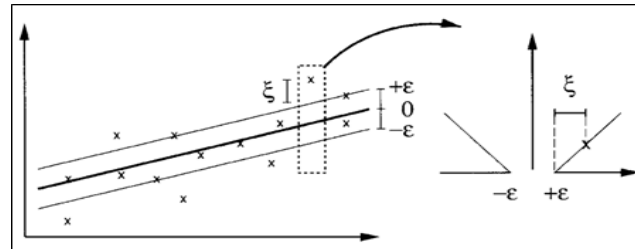
$$\min. \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3)$$

$$\text{Öyleki} \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \end{cases} \quad \xi_i, \xi_i^* \geq 0$$

şeklinde yeniden kurulur. Bu durumda ε -duyarsız kayıp fonksiyonu (Şekil 1),

$$|\xi|_\varepsilon := \begin{cases} 0 & \text{eğer } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & \text{aksi taktirde} \end{cases}$$

şeklinde tanımlanır. Modeldeki C sabiti eğitim hatası üzerine konulan bir ceza parametresidir ve $f(x)$ fonksiyonun düzlüğü ile ε toleransından sapmalar arasındaki ödünleşimi (tradeoff) belirler.



Şekil 1: Yumuşak Marjlı Kayıp Fonksiyonu Kaynak: Schölkopf vd. (2000)

Amaç fonksiyonu ve kısıtlardan Langrange fonksiyonu oluşturulup KKT (Karush-Kuhn-Tucker) koşulları uygulandığında, modelin dual çözümünden

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i \quad \text{ve} \quad f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \quad (4)$$

elde edilir. α_i ve α_i^* Langrange çarpanları olup, $\alpha_i, \alpha_i^* \in (0, C)$ dir. (4) ifadesi Destek Vektör Açılımı (expansion) olarak adlandırılır. Diğer bir ifade ile w , girdi değerlerinin doğrusal kombinasyonu olarak ifade edilir. KKT koşulları gereği; $|f(x_i) - y_i| \geq \varepsilon$ için Langrange çarpanları sıfır olmayabilir ancak, $|f(x_i) - y_i| < \varepsilon$ koşulunu sağlayan durumlarda Langrange çarpanları sıfırdır. Langrange çarpanları sıfır olmayan noktalar “Destek Vektörler” olarak adlandırılır. Böylece (4)’de w , girdi değerlerinin “seyrek” (sparse) açılımı ile ifade edilir. Bunun

anlamı \mathbf{w} vektörünü tanımlamak için tüm \mathbf{x}_i değerlerine gereksinim olmadığı, sadece langrange çarpanları sıfır olmayan noktaların yani destek vektörlerin yeterli olduğudur. Öğrenme Teorisi açısından bakıldığında destek vektörler, eğitim setinin taşıdığı bilginin yoğunlaştırıldığı noktalar yani bilginin temsil (knowledge representation) edildiği yapıdır ve doğal olarak destek vektör olmayan noktalar çıkarılıp tekrar eğitim işlemi gerçekleştirildiğinde sonuç değişmeyecektir.

Doğrusal olmayan durum için girdi uzayından doğrusal olmayan bir dönüşümle (5) yüksek boyutlu özellik uzayına (F) geçilir ve problem özellik uzayında bir hiperdüzlem araştırılmasına dönüşür.

$$\phi(\mathbf{x}): X \in \mathbb{R}^d \rightarrow \mathcal{F} \in \mathbb{R}^h, d < h \quad (5)$$

Bu durumda (1) ifadesi,

$$f(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle + b \quad (6)$$

ve (4) ifadesi

$$\mathbf{w} = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(\mathbf{x}_i, \cdot) \quad \text{ve} \quad f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle k(\mathbf{x}_i, \mathbf{x}) \rangle + b \quad (7)$$

halini alır. Burada $k(\mathbf{x}_i, \mathbf{x}) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle$ Hilbert Uzayında bir kernel fonksiyonudur. Mercer koşullarını sağlayan her hangi bir simetrik pozitif yarı-tanımlı kernel fonksiyonu kullanılabilir (Schölkopf ve Smola, 2002) ve en yaygın kullanılan kernel fonksiyonları

$$\begin{aligned} k(\mathbf{x}_i, \mathbf{x}_j) &= \mathbf{x}_i^T \mathbf{x}_j && \text{Doğrusal Kernel} \\ k(\mathbf{x}_i, \mathbf{x}_j) &= (\mathbf{x}_i^T \mathbf{x}_j + 1)^k && \text{Polinom Kernel} \\ k(\mathbf{x}_i, \mathbf{x}_j) &= \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) && \text{RBF (Radial Basis Function) Kernel} \end{aligned}$$

olarak tanımlanır. DVR modelinde C, ε ve kernel parametreleri (örn. RBF kernel için γ parametresi) eğitim setine (uygulamaya) bağlıdır ve eğitim öncesinde kullanıcı tarafından belirlenmelidir. Alternatif bir formülasyon Schölkopf vd. (2000) tarafından geliştirilen ν -DVR (nu-DVR) dir. Bu yaklaşımda ε aralığının dışında kalan gözlem noktalarının kesrine (fraction) bir üst sınır getirilerek (8), ilgili ε parametresinin otomatik olarak hesaplanması sağlanmaktadır.

$$\text{Hataların Sayısı} / n \leq \nu \leq \text{Destek Vektörlerin Sayısı} / n \quad (8)$$

(8) ifadesinde hataların sayısı ε aralığının dışında kalan gözlem noktaları sayısı olup, $\nu \in [0,1]$ parametresi hataların ve destek vektörlerin kesrine sırasıyla üst ve alt sınır getiren bir sabittir. Çalışmamızda ν -DVR kullanılmıştır.

4. DVR VOLATİLİTE MODELLERİNİN KURULUMU

Ampirik bulgular GARCH modelinin, kestirimi gereken parametre adedi açısından, ARCH modeline göre daha tutumlu (parsimonious) olduğu ve GARCH(1,1) modelinin ise uzun bir gözlem periyodu içerisinde değişen varyansın modellenmesinde yeterli olduğu yönünde bir destek sağlamaktadır. Bu nedenle finansal volatilitenin modellenmesinde popüler hale gelmiştir (Akgiray, 1989; Gökcan, 2000; Poon ve Granger, 2003). Çalışmamız GARCH(1,1) süreci ile

kısıtlandırılmıştır.

P_t ve P_{t-1} sırasıyla t ve $t-1$ zamanındaki kapanış endeks değerleri olmak üzere, endeksin yüzde olarak getirisi (R_t) (9) ifadesi ile verilmiştir. Çalışmamızda t indisi günlük gözlemleri göstermektedir.

$$R_t = 100 \times \ln \left(\frac{P_t}{P_{t-1}} \right) \quad (9)$$

R_t serisi için baskın olan ampirik model AR(p)-GARCH(1,1) dir ve (10) ifadesi ile tanımlanır.

$$\Phi_p(B)R_t = \mu + \varepsilon_t, \quad \Phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B \quad (10.1)$$

$$\varepsilon_t = \sigma_t z_t, \quad \varepsilon_t \sim D(0, \sigma_t^2) \text{ ve } z_t \sim D(0,1) \text{ (beyaz gürültü)} \quad (10.2)$$

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2, \quad \omega > 0 \quad (10.3)$$

$B, B^k R_t = R_{t-k}$, şeklinde tanımlanan kaydırma operatörüdür. μ parametresi bir sabittir ve pratikte sıfıra yakın ya da sıfır olarak kestirilir. R_t 'nin kendi geçmiş değerleri ile tahminlenmesinin genellikle mümkün olmadığı düşüncesinin sonucu olarak (Etkin Piyasa Hipotezi varsayımları çerçevesinde yürütülen tartışmalar için Granger, 1992 ve burada referans gösterilen çalışmalar incelenebilir), p için sıfır ya da küçük bir değer alınır (Franses ve Dijk, 1996). Çalışmamızda $\mu=p=0$ olarak alınmıştır. Bu durumda (10.1) ifadesindeki koşullu ortalama denklemi

$$R_t = \varepsilon_t = \sigma_t z_t \quad (11)$$

şeklini alır.

Koşullu varyansın (σ^2) kesin-pozitif olabilmesi için, $\alpha \geq 0$ ve $\beta \geq 0$ koşullarının sağlanması gerekmektedir. $(\alpha + \beta)$ toplamı ise şokların volatilité üzerindeki etkisinin hangi hızda azalacağını gösterir (Chou, 1988). Diğer bir ifade ile şokların volatilité üzerindeki ısrarcılığının bir ölçüsüdür. Şokların volatilité üzerinde ısrarcılığı arttıkça toplam 1'e yaklaşacaktır. Durağanlık koşulu için 1'den küçük olmalıdır. Toplamın 1'den büyük olması, bir dönemde ortaya çıkan şokun bir sonraki dönemde daha büyük bir volatilitéye yol açacağını gösterir. Toplamın 1'e eşit olması ise şokların volatilité üzerinde ısrarcılığının kalıcı (permanent) olduğunu gösterir ve bu model ile koşulsuz varyans belirlenemez. Bu tür süreçleri Engle ve Bollerslev (1986) "Bütünleşik(Integrated)-GARCH" olarak adlandırmıştır.

D dağılımı için sıklıkla normal dağılım kullanılsa da, bir çok çalışma göstermiştir ki finansal varlıkların getirisi nadiren normal dağılmaktadır ve student's t, kararlı pareto dağılımı, vb. gibi diğer dağılımlar kullanılabilir (Nelson, 1991).

DVR-GARCH(1,1) modelinin girdileri olan $\sigma_{(t-1)}^2$ ve $\varepsilon_{(t-1)}^2$ değerlerinin öğrenme makinesinin eğitimine başlanmadan önce verilmiş olması gerekmektedir. Sözkonusu değerlerden ilki getirilerin karesinin hareketli ortalamasından ($\hat{\sigma}_t^2 = \frac{1}{5} \sum_{k=0}^4 R_{t-k}^2$) elde edilmiştir (Pérez-Cruz vd., 2003). Diğerleri için ise (11) ifadesinden şokların karesi, getirilerin karesi ($R_t^2 = \varepsilon_t^2$) alınarak

elde edilmiştir. Buradan DVR-GARCH(1,1) modeli

$$\hat{\sigma}_t^2 = h(\hat{\sigma}_{t-1}^2, R_{t-1}^2) \quad (12)$$

olarak belirlenir.

EGARCH(1,1) modeli (13) logaritmik formda ifade edilir.

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha \left[\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - E\left(\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}}\right) \right] + \gamma \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) \quad (13)$$

Bu model iyi ve kötü haberlerin volatilité üzerinde farklı etkilere sahip olmasına γ parametresini kullanarak izin verir. Ayrıca koşullu varyansın kesin-pozitif olması için parametreler üzerinde herhangi bir kısıtlama getirmez.

Son olarak GJR(1,1) aşağıdaki gibi ifade edilir.

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2 + \gamma S_{t-1}^- \varepsilon_{t-1}^2 \quad (14)$$

Burada $\varepsilon_{(t-1)} < 0$ (kötü haber) durumu için $S_{t-1}^- = 1$ aksi takdirde 0 dır. Durağanlık varsayımı için bu model $\alpha + \beta + \frac{1}{2}\gamma < 1$, $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \gamma \geq 0$ olmasını gerektirmektedir. Bu modelde kötü haberler iyi haberlere göre daha yüksek bir etkiye sahiptir. İyi haberler volatilité üzerinde α kadar etki ederken, kötü haberler $\alpha + \gamma$ kadar bir etkiye sahiptir. Son olarak şokların ısrarcılığı $\alpha + \beta + \frac{1}{2}\gamma$ ifadesi ile ölçülür.

DVR-GARCH(1,1) kuruluşuna benzer şekilde, DVR-EGARCH(1,1)

$$\ln(\hat{\sigma}_t^2) = h(\ln \hat{\sigma}_{t-1}^2, \left[\frac{|R_{t-1}|}{\hat{\sigma}_{t-1}} - E\left(\frac{|R_{t-1}|}{\hat{\sigma}_{t-1}}\right) \right], \frac{R_{t-1}}{\hat{\sigma}_{t-1}}) \quad (15)$$

ve DVR-GJR(1,1),

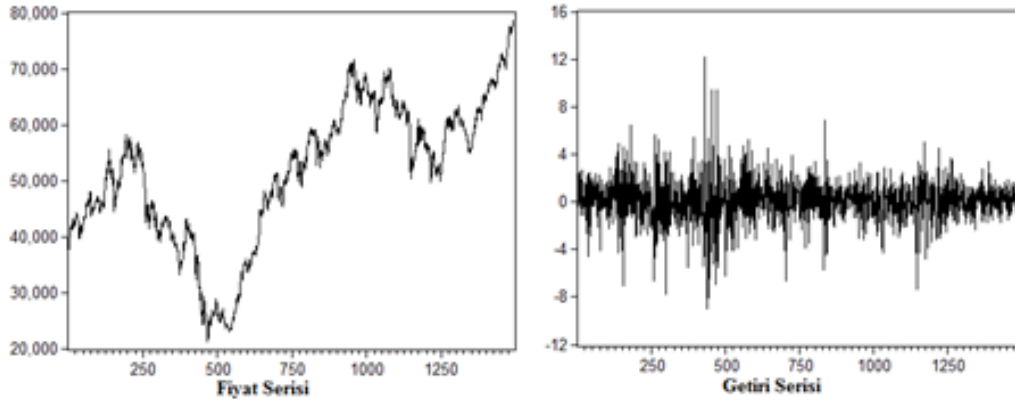
$$\hat{\sigma}_t^2 = h(\hat{\sigma}_{t-1}^2, R_{t-1}^2, S_{t-1}^- R_{t-1}^2) \quad (16)$$

olarak ifade edilir. Burada $E\left(\frac{|R_{t-1}|}{\hat{\sigma}_{t-1}}\right)$ beklenen değeri için ortalama değeri alınır (Ou ve Wang, 2010). Böylece getiri serisinin dağılımının bilinmesine gerek kalmamaktadır.

5. UYGULAMA

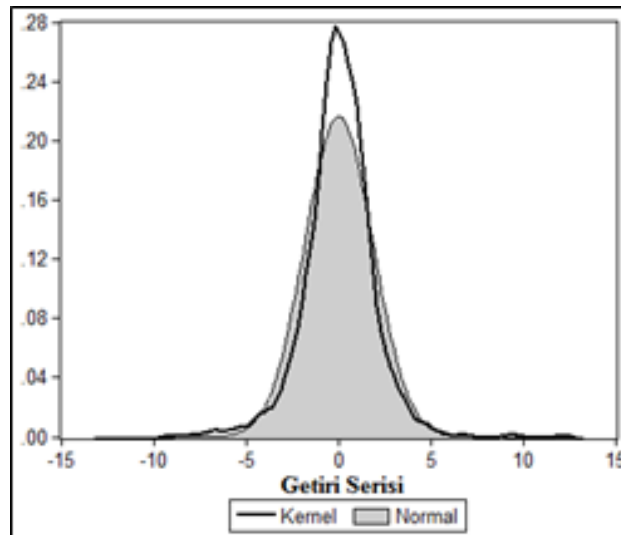
5.1. Veri Seti ve İstatistiksel Analiz

Çalışmada, 04.01.2007 – 31.12.2012 dönemi için 1497 veriden oluşan günlük İMKB ulusal 100 endeksi kapanış fiyatları kullanılmıştır. Veriler Merkez Bankası web sitesinden alınmıştır. Kaydırma ve hareketli ortalama işlemleri sonucunda veri sayısı 1491'e düşmüş ve 12.01.2007 – 20.10.2011 dönemini kapsayan 1193 adet veri eğitim seti, 21.10.2011 – 31.12.2012 dönemini kapsayan 298 adet veri test seti olarak ayrılmıştır. Bu analizlerde ve klasik modellerin çözümünde Eviews 6.0 yazılımı, DVM modellerin çözümünde RapidMiner 5.2 yazılımı kullanılmıştır.



Şekil 2: İMKB 100 endeksi fiyat verisi ve logaritmik getiri (tüm veri)

Endeks serisinin grafiği (Şekil 2) incelendiğinde ortalama durağan olmadığı ve düzensiz iniş çıkışlardan değişen varyans olabileceği görülmektedir. (9) ifadesi kullanılarak oluşturulan getiri serisi incelendiğinde, serinin ortalama durağan olduğu söylenebilir. Yine grafikten büyük değişimleri büyük, küçük değişimleri küçük hareketlerin izlediği, yani bir volatilitite kümelenmesi olduğu gözlenmektedir. Getiri serisinin parametrik olmayan (kernel) olasılık yoğunluk fonksiyonu kestirimi ile aynı ortalama ve varyansa sahip normal dağılımın grafiği (Şekil 3) incelendiğinde, getiri serisinin basıklık katsayısının 3'den büyük (kalın kuyruk) ve çarpıklık katsayısının küçük de olsa negatif (sola çarpık) olduğu gözlenmektedir. Görsel olarak elde edilen bu bulgular istatistiksel analiz sonuçları (Tablo 1) ile uyumludur.



Şekil 3: Getiri Serisinin Olasılık Yoğunluk Fonksiyonu Parametrik Olmayan Kestirimi

Tablo 1: Fiyat ve getiri serilerinin tanımsal istatistikleri

	Fiyat Serisi	Getiri Serisi(Tüm Veri)
Gözlem sayısı	1492	1491
Ortalama	51260.70	0.049048
Medyan	53234.60	0.076681
Maksimum	78579.08	12.12721
Minimum	21228.27	-9.013694
Standard Sapma	12732.90	1.846541
r(1)	0.996	0.047
r(10)	0.961	0.057
r(20)	0.919	-0.031
Ljung-Box Q(20)	27621(0.0000)	46.838(0.001)
ADF	1.051723(0.9238)	-36.78073(0.0000)
Phillips-Peron	1.011734(0.9184)	-36.83590(0.0000)
Çarpıklık	-0.403255	-0.104916
Basıklık	2.493089	6.743774
Jarque-Bera	56.4111(0.0000)	873.4698(0.0000)

Açıklama: parantez () içerisindeki değerler olasılıklardır. $r(k)$, k gecikmesi için otokorelasyon katsayısıdır. ADF ve PP testleri için %1 ve %5 anlam düzeyi için kritik değerler sırasıyla -2.56651 ve 1.94104 tür. Jarque-Bera testi için %5 anlam düzeyinde kritik Chi-kare(2) değeri 5.99146'dır. Q(20) istatistiği için %1 ve %5 anlam düzeyleri için kritik Chi-kare(20) değerleri sırasıyla 37.5662 ve 31.4104 tür.

Tablo 1'den fiyat serisinin tüm gecikmeler için yüksek korelasyon katsayısına sahip olduğu 20 gecikme için Ljung-Box istatistiği ile görülmektedir. Geliştirilmiş Dickey-Fuller (ADF) ve Phillips-Peron (PP) istatistikleri incelendiğinde fiyat serisi düzeyinde durağan değildir. Ljung-Box Q(20) istatistiği sonucu pekiştirmektedir.

Getiri serisinin hemen hemen sıfır ortalamaya, ortalama etrafında aşırı basıklığa ve hafif sola çarpıklığa sahip bir seri olduğu görülmektedir. Yüksek Jarque-Bera istatistik değeri serinin normal dağılmadığını göstermektedir. ADF ve PP istatistiklerine göre getiri serisi düzeyinde durağandır. Ancak Ljung-Box istatistiği düşük bir otokorelasyona işaret etmektedir. Getiri serisinin düşük ya da sıfır otokorelasyona sahip olduğunu göstermemizin yanısıra, serinin karelerinden elde edilen "getiri-kare" serisinin ise yüksek otokorelasyon katsayılarına sahip olduğunu göstermemiz gerekmektedir.

Tablo 2'de Ljung-Box Q istatistiği getiri-kare serisinin getiri serisinden daha yüksek ve daha ısrarlı (persistent) otokorelasyon katsayılarına sahip olduğunu göstermektedir. Getiri-kare serisinin her gecikme değeri için anlamlı otokorelasyona sahip olması, sürecin volatilitate kümelenmesine sahip olduğunu yansıtmaktadır. Serinin düzeyinde durağan olması ve getiri-kare serisinin tüm gecikme değerleri için yüksek bir otokorelasyon göstermesi GARCH modelinin

uygulanabilirliğini göstermektedir.

Tablo 2: Getiri-kare serisinin tanımsal istatistikleri

R_t^2	Tüm Veri
Ortalama	3.409833
Standard Sapma	8.163116
r(1)	0.125
r(10)	0.120
r(20)	0.116
Ljung-Box Q(20)	635.86(0.0000)

Açıklama: Q istatistiği için %1 ve %5 anlam düzeyleri için kritik Chi-kare(20) değerleri sırasıyla 37.5662 ve 31.4104 tür.

5.2. Modellerin Kestirimi

Getiri serisinin aşırı basıklığı ve Jarque-Bera istatistiği, serinin normal dağılmadığını göstermiştir. Bu nedenle normal dağılım varsayımı altında yapılan model kestirimleri için, ML metodu yerine Bollerslev ve Wooldridge (1992) tarafından geliştirilen “Quasi-Maximum Likelihood” (QML) metodu kullanılmıştır. Nelson (1991) EGARCH modelinin GED özelliği gösterdiğini varsaymaktadır. Her bir modelin normallik varsayımı altında QML kestirimi yanısıra, GED varsayımı altında ML kestirimleri de yapılmıştır.

Tablo 3 GARCH modellerinin kestirim sonuçlarını göstermektedir. LL, AIC ve SIC bilgi kriterleri incelendiğinde normal varsayım altında EGARCH-Normal ve GED varsayımı altında EGARCH-GED modellerinin en iyi kestirim performansını gösterdikleri görülmektedir. En düşük performans, her iki dağılım varsayımı altında, GARCH modellerindedir.

GARCH modellerinin $(\alpha+\beta)$ değerlerine bakıldığında GARCH süreçlerinin yüksek ısrarcılığa sahip ve kovaryans durağan oldukları görülmektedir. Standartlaştırılmış hata terimlerinin karelerine ait Q2(20) istatistiğine bakıldığında, diğer modellere göre getiri-kare serisinin dinamiklerini daha iyi temsil ettiği görülmektedir. Ancak standartlaştırılmış hata terimlerinin basıklık katsayısına baktığımızda, getiri serisinin aşırı basıklığını diğer modeller kadar iyi temsil edemediği görülmektedir.

EGARCH modellerinin (β) değerleri bize EGARCH süreçlerinin de yüksek ısrarcılığa sahip ve kovaryans durağan olduğunu göstermektedir. Basıklık katsayısı getiri serisinin aşırı basıklığını en iyi temsil eden model olduğunu göstermektedir. Ancak getiri-kare serisinin dinamiklerini (görece) en kötü temsil eden model olarak görülmektedir.

GJR modellerinin $(\alpha + \beta + \frac{1}{2}\gamma)$ değerleri sürecin yaklaşık EGARCH süreçleri kadar yüksek ısrarcılığa sahip ve kovaryans durağan olduğunu göstermektedir. Getiri serisinin aşırı basıklığını temsil etmede EGARCH modellerine göre daha kötü, getiri-kare serisinin dinamiklerini temsil

etmede ise EGARCH modelinden daha iyi olduğu görülmektedir.

Tablo 3: Parametrik Modellerin Kestirimi

	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
	Normal Dağılım			Genelleştirilmiş Hata Dağılımı		
μ	0.136445	0.122984	0.096918	0.122989	0.114718	0.094689
ω	0.165586	-0.110015	0.237477	0.165061	-0.108048	0.240444
α	0.125713	0.217455	0.062360	0.121516	0.218382	0.063014
β	0.833363	0.948490	0.803123	0.836258	0.946061	0.800006
γ	-	-0.089866	0.138522	-	-0.094098	0.141196
$\alpha + \beta$	0.959076	-	-	0.957774	-	-
LL	-2391.766	-2379.903	-2382.996	-2373.251	-2364.898	-2366.966
AIC	4.019741	4.001515	4.006704	3.990354	3.978016	3.981487
SIC	4.036799	4.022838	4.028027	4.011677	4.003604	4.007075
	Standartlaştırılmış hata terimi $\hat{\varepsilon}_t = R_t/\hat{\sigma}_t$					
Ort.	-0.049410	-0.036615	-0.022488	-0.041988	-0.031722	-0.021121
S.S	0.999178	0.999919	1.000416	1.001791	1.001730	1.002279
Çarp.	-0.122634	-0.013904	-0.021933	-0.127973	-0.009536	-0.018783
Bas.	4.284594	4.039904	4.150661	4.292278	4.045416	4.153353
r(1)	0.058	0.052	0.057	0.058	0.052	0.057
Q(20)	30.232 (0.066)	28.887 (0.090)	29.466 (0.079)	30.453 (0.063)	28.987 (0.088)	29.451 (0.079)
r2(1)	0.002	-0.022	-0.030	0.004	-0.024	-0.032
r2(10)	-0.016	-0.000	0.000	-0.016	0.001	0.001
r2(20)	0.027	0.029	0.029	0.028	0.030	0.030
Q2(20)	11.189 (0.941)	16.502 (0.685)	15.878 (0.724)	11.096 (0.944)	17.097 (0.647)	16.189 (0.705)

Açıklama: LL, log likelihood değeri. AIC, Akaike Bilgi Kriteri. SIC, Schwarz Bilgi Kriteri. Modeller normal dağılım varsayımında Quasi-Maximum Likelihood yöntemi, Genelleştirilmiş Hata Dağılımı varsayımında Maximum Likelihood yöntemi ile çözülmüştür. Parentez () içerisindeki değerler olasılıklardır. r(1), birinci gecikmede otokorelasyon katsayısı. Q(20), 20. gecikmede Ljung-Box istatistiği. r2(k), standartlaştırılmış hata terimlerinin karesinin k. gecikme için otokorelasyon katsayısı. Q2(20), standartlaştırılmış hata terimlerinin karesinin 20. gecikmede Ljung-Box istatistiği.

DVR modelleri RBF kernel kullanılarak eğitime sokulmuştur. Tüm modellerde C parametresi C=10 olarak alınmıştır. Nu(v) ve Gamma(γ) parametrelerinin optimum değerleri uygun bir aralıkta “grid search” tekniği kullanılarak belirlenmiştir. Aşırı uyumdan kaçınmak için çapraz doğrulama yöntemi uygulanmıştır. Bunun için eğitim seti 10 eşit alt sete bölünmüş ve setlerden 9 adedi eğitim 1 adedi test seti olarak kullanılmıştır. Böylece 10 farklı eğitim ve test seti (10-fold cross validation) için “grid search” tekniği kullanılmıştır.

Tablo 4 eğitim sonucu DVR modellerinin optimal parametrelerini içermektedir. Burada

RMSE değerleri her bir model için yapılan 10 adet çapraz doğrulamaya ait değerlerin ortalamasıdır. Eğitim seti için DVR-EGARCH modelinin iyi bir performans gösterdiği söylenebilir, ancak bu durum genelleme başarısı hakkında bir fikir vermez. Öngörülse (predictive) modellerin başarısı, genelleme performansı (test seti) ile ölçülür.

Tablo 4: DVR Modellerin Optimum Parametre Değerleri

	DVR- GARCH	DVR- EGARCH	DVR-GJR
Gamma(γ)	0.29545455	0.04932614	0.41875
Nu (ν)	0.402	0.3886016	0.366367
C	10	10	10
Tolerans(ϵ)	1.0E-4	1.0E-4	1.0E-4
RMSE	8.137	2.240	8.293

Açıklama: RMSE, ortalama hata karesinin kökü (Root Mean Square Error)

5.3. Modellerin Tahmin Performansları

Volatilitenin tahminlerinin etkinliğinin saptanması basit bir işlem değildir. Bunun nedeni koşullu varyansın gizli (latent) bir değişken olmasıdır. Son yıllarda gerçekleşen volatilitenin, gün içinde oluşan fiyat oluşumlarının yüksek frekanslı örnekleme (örneğin 15 dk. aralıklarla) yoluyla tahmin edilmesi daha güvenilir bir ölçü olduğundan tercih edilmektedir (McMillan ve Garcia, 2009). Çalışmamızda gün içi verileri temin etmemiz mümkün olmadığı için volatilitenin tahminleme çalışmalarında standard olarak kullanılan (17) ifadesi kullanılmıştır (Franses ve Dijk, 1996). Burada μ terimi, ele alınan dönem için getirilerin ortalamasını göstermektedir.

$$\hat{\sigma}_t^2 = (R_t - \mu)^2 \quad (17)$$

Tahmin performansının değerlendirilmesinde simetrik ve simetrik olmayan tahmin hatası istatistiklerinden faydalanılmıştır. Simetrik ölçü kriterleri olarak ortalama hata karesinin karekökü (root mean square error:RMSE), ortalama mutlak yüzde hata (mean absolute percentage error:MAPE), ortalama mutlak hata (mean absolute error:MAE) ve 0-1 aralığında değer alabilen Theil eşitsizlik katsayısı (Theil inequality coefficient:TIC) alınmıştır.

$\hat{\sigma}_t$: Tahminlenen volatilitenin ve σ_t : Gerçekleşen volatilitenin göstergesidir. t: tahmin edilen zaman ve N: tahmin sayısı (test setindeki gözlem sayısı) olmak üzere,

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{\sigma}_t - \sigma_t)^2} \quad MAE = \frac{1}{N} \sum_{t=1}^N |\hat{\sigma}_t - \sigma_t|$$

$$MAPE = \frac{1}{N} \left| \frac{\hat{\sigma}_t - \sigma_t}{\sigma_t} \right| \quad TIC = \frac{RMSE}{\sqrt{\frac{1}{N} \sum_{t=1}^N \hat{\sigma}_t^2} + \sqrt{\frac{1}{N} \sum_{t=1}^N \sigma_t^2}}$$

olarak tanımlanır.

Bu dört kriter en küçük değere sahip olan modeli en doğru tahminleyen olarak görür. Ancak bir yatırımcı için karlılık, en küçük istatistiksel tahmin hatasının yanı sıra, doğru pozisyon almayada bağlıdır. Dolayısıyla değişimlerin yönünü doğru olarak tahminlemek önemli hale gelmektedir. Bu amaca uygun bir kriterimizde doğru yönde değişim (Correct Directional Change, CDC) yüzdesi olacaktır. Burada amaç tahmin edilen volatilitelerdeki günlük yön değişimlerinin, gerçekleşen volatilitelerdeki günlük yön değişimlerini yakalama yüzdesini görebilmektir.

$$CDC = \frac{100}{N} \sum_{t=1}^N D_t$$

$$D_t = \begin{cases} 1 & \text{eğer } (\hat{\sigma}_t - \hat{\sigma}_{t-1})(\sigma_t - \sigma_{t-1}) > 0 \\ 0 & \text{aksi takdirde} \end{cases}$$

Simetrik olmayan ölçü kriteri için ortalama bileşik hata (mean mixed error:MME) kullanılmıştır. Ortalama bileşik hata, tahminlenen volatilitenin gerçekleşen volatilitenin altında (U) ya da üstünde (O) kalmasının farklı şekilde ağırlıklandırılması prensibine dayanır.

$$U = \{t : \hat{\sigma}_t - \sigma_t < 0\} \text{ ve } O = \{t : \hat{\sigma}_t - \sigma_t > 0\} \text{ olmak üzere,}$$

Gerçekleşen volatilitenin altında kalan tahminlerin cezalandırılması istendiğinde

$$MME(U) = \frac{1}{N} \left[\sum_{t \in O} |\hat{\sigma}_t - \sigma_t| + \left[\sum_{t \in U} |\hat{\sigma}_t - \sigma_t| \right]^{2A} \right]$$

kriteri kullanılacak, gerçekleşen volatilitenin üstüne çıkan tahminlerin cezalandırılması istendiğinde ise

$$MME(O) = \frac{1}{N} \left[\sum_{t \in U} |\hat{\sigma}_t - \sigma_t| + \left[\sum_{t \in O} |\hat{\sigma}_t - \sigma_t| \right]^{2A} \right]$$

kriteri anlamlı olacaktır. Burada A, $|\hat{\sigma}_t - \sigma_t| \leq 1$ ise -1, aksi takdirde 1 değerini alır.

Tablo 5: Modellerin Performans Değerlemesi

	RMSE	S	MAE	S	MAPE	S	TIC	S	MME(O)	S	MME(U)	S	CDC	S	TS
GARCH-N	0.87991	5	0.75651	6	6.67767	6	0.33233	6	0.92731	5	0.83409	6	27.03	6	40
GARCH-GED	0.87960	4	0.75649	5	6.68901	7	0.33238	7	0.92632	4	0.83459	7	27.03	6	40
E GARCH-N	0.88577	6	0.75553	4	6.51646	4	0.33155	3	0.95179	8	0.82446	4	35.81	1	30
E GARCH-GED	0.88681	7	0.75672	7	6.53127	5	0.33174	5	0.95317	9	0.82577	5	35.81	1	39
GJR-N	0.88985	9	0.76707	9	6.88473	9	0.33301	9	0.94717	7	0.84006	9	32.43	4	56
GJR-GED	0.88859	8	0.76576	8	6.87160	8	0.33280	8	0.94523	6	0.83902	8	32.43	4	50
DVR-GARCH	0.81101	2	0.66984	2	5.20252	3	0.33038	2	0.82789	2	0.78424	3	29.97	5	19
DVR-E GARCH	0.79813	1	0.64799	1	4.76915	1	0.33158	4	0.79356	1	0.77701	1	34.68	2	11
DVR-GJR	0.81284	3	0.67035	3	5.14412	2	0.32956	1	0.83132	3	0.78334	2	33.33	3	17

Açıklama: S: ilgili kritere göre sıralama (skor) (1:en iyi skoru gösterir), TS: Toplam skor

Tablo 5 modellerin test seti üzerindeki tahmin hata istatistiklerini içermektedir. Tablo şu şekilde düzenlenmiştir: Her bir kritere ait sütunun sağında ilgili kritere göre sıralama (skor) değerleri verilmiştir. Tablo'nun son sütununda ise tüm kriterler için modellerin skor toplamı (TS) verilmiştir. Buna göre en iyi toplam skor 11 değeri ile DVR-EGARCH, en kötü toplam skor 56 değeri ile GJR-Normal modeline aittir. RMSE, MEA, MAPE ve MME(O/U) kriterlerine göre ilk üç sırada DVR modelleri bulunmaktadır. TIC ve CDC kriterlerine göre ise ilk üç başarılı model arasında iki DVR modeli bulunmaktadır. Klasik modeller kendi içerisinde değerlendirildiğinde EGARCH modelleri en iyi modeller olarak görülmektedir. DVR modelleri ise kendi içerisinde, son iki model oldukça yakın TS puanı almakla birlikte, DVR-EGARCH, DVR-GJR ve DVR-GARCH olarak sıralanmaktadır. Doğru yönde değişimi (CDC) tüm EGARCH modellerinin iyi yakaladığı görülmektedir. DVR-EGARCH modelinin ise, EGARCH modellerine oldukça yakın durarak, ikinci en iyi doğru yönde değişimi yakalayan model olmuştur.

6. SONUÇ

Bu çalışmada standart GARCH, EGARCH ve GJR-GARCH modelleri için klasik ML metodunun performansı ile DVR algoritmasının performansı kıyaslanmıştır. Literatürde yaygın olarak kullanılan ve tahminleme hatasının büyüklüğünü ölçen kriterlerin yanısıra, doğru yönde tahmin açısından da kıyasladığımızda nu-DVR algoritmasına dayalı karma modellerin genel olarak klasik ML çözümüne göre daha iyi bir performans sergilediği görülmüştür. Sonuçlar ikinci bölümde verilen diğer çalışmalar ile uyumludur. Belirgin bir dağılım varsayımına dayanmayan ve yüksek genelleme kapasitesine sahip olan DVR metodunun zaman serileri analizinde de dikkate alınması gereken bir yöntem olduğu uygulamamızda görülmektedir.

Tablo 5, DVR metodunun klasik modellerin belirlediği model kalıplarının içerisinde öğrenme işlemini gerçekleştirdiği dikkate alınarak okunmalıdır. Dolayısıyla sonuçların, klasik modellerin volatilitiyi iyi modelleyip modellemediği tartışması dışında, klasik ML metodu ile DVR metodolojisinin kıyaslanması şeklinde ele alınması daha gerçekçidir. Girdi-çıkı değerlerinin eğitim öncesinde sağlanmış olması gerektiğinden DVR modellerinin başarısı için, koşullu varyans denklemindeki değişkenlerin uygun değerlerinin elde edilmesinin öneminin de altı çizilmelidir. Bunun anlamı, sürecin örüntülerini daha iyi temsil eden eğitim setinin daha iyi sonuç doğurabileceğidir.

7. KAYNAKÇA

- [1] AKGİRAY, V. (1989), "Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts", *The Journal of Business*, 62(1), 55-80.
- [2] ANDERSEN, T.G. ve T. BOLLERSLEV (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do provide Accurate Forecasts", *International Economic Review*, 39(4), 885-905.
- [3] BİLDİRİCİ, M. ve Ö.Ö.ERSİN (2009), "Improving Forecasts of GARCH Family Models with The Artificial Neural Networks: An Application to The Daily Returns in Istanbul Stock Exchange", *Expert Systems with Applications*, 36, 7355-7362.

- [4] BİLDİRİCİ, M. ve Ö.Ö.ERSİN (2012), “Nonlinear Volatility Models in Economics: Smooth Transition and Neural Network Augmented GARCH, APGARCH, FIGARCH and FIAPGARCH Models”, Munich Personal RePEc Archive, MPRA Paper No: 40330.
- [5] BLACK, F. (1976), “Studies in Stock Price Volatility Changes”, in Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association, 177-181.
- [6] BOLLERSLEV, Tim (1986), “Generalized Autoregressive Conditional Heteroskedasticity”, Journal of Econometrics, 31, 307-327.
- [7] BOLLERSLEV, T. ve J.M.WOOLDRIDGE (1992), “Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances”, Econometric Reviews, 11, 143-172.
- [8] BURGESS, Christopher J.C. (1998), “A Tutorial of Support Vector Machines for Pattern Recognition”, Data Mining and Knowledge Discovery, 2, 121-167.
- [9] CHAMPBELL, J.Y. ve A.S.KYLE (1993), “Smart Money, Noise Trading and Stock Price Behaviour”, Review of Economic Studies, 60, 1-34.
- [10] CHOU, R.Y. (1988), “Volatility Persistence and Stock Valuations: Some Empirical Evidence Using GARCH”, Journal of Applied Econometrics, 3(4), 279-294.
- [11] CONT, Rama (2001), “Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues”, Quantitative Finance, 1, 223-236.
- [12] DONALDSON, R.G. ve Mark KAMSTRA (1997), “An Artificial Neural Network –GARCH Model for International Stock Return Volatility”, Journal of Empirical Finance, 4(1), 17-46.
- [13] ENGLE, R.F. (1982), “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, Econometrica, 50(4), 987-1007.
- [14] ENGLE, R.F. ve T.P. BOLLERSLEV (1986), “Modelling the Persistence of Conditional Variances”, Econometric Reviews, 5(1), 1-50.
- [15] ENGLE, R.F. ve V.K. NG (1993), “Measuring and Testing the Impact of News on Volatility”, The Journal of Finance, 48(5), 1749-1778.
- [16] FRANCES, F.H. ve D.V.DIJK (1996), “Forecasting Stock Market Volatility Using (Non-Linear) GARCH Models”, Journal of Forecasting, 15, 229-235.
- [17] GAVRISHCHAKA, V.V. ve S.BANERJEE (2006), “Support Vector Machine as an Efficient Framework for Stock Market Volatility”, Computational Management Science, 3(2), 147-160.
- [18] GLOSTEN, L.R., JAGANNATHAN, R. ve D.E.RUNKLE (1993), “On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks”, The Journal of Finance, 48(5), 1779-1801.
- [19] GÖKCAN, S. (2000), “Forecasting Volatility of Emerging Stock Markets: Linear versus Non-linear GARCH models”, Journal of Forecasting, 19(6), 499-504.
- [20] GRANGER, C.W.J. (1992), “Forecasting Stock Market Prices: Lessons for Forecasters”, International Journal of Forecasting, 8, 3-13.

- [21] İNCE, H. (2005), "Non Parametric Regression Methods and an Application to Istanbul Stock Exchange 100 Index", Yapı Kredi Economic Review, 16(1), 17-28.
- [22] KARA, Yakup, BOYACIOĞLU, Melek A. ve Ö.Kaan BAYKAN (2011), "Prediction 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, 5311-5319.
- [23] KIM, Kyoung-Jae (2003), "Financial Time Series Forecasting Using Support Vector Machines", Neurocomputing, 55, 307-319.
- [24] McMILLAN, D. ve R.Q.GARCIA (2009), "Intra-day Volatility Forecasts", Applied Financial Economics, 19(8), 611-623.
- [25] NELSON, Daniel B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", Econometrica, 59(2), 347-370.
- [26] OU, Phichhang ve H.WANG (2010), "Financial Volatility Forecasting by Least Square Support Vector Machine Based on GARCH, EGARCH and GJR Models: Evidence from ASEAN Stock Markets", International Journal of Economics and Finance, 2(1), 51-64.
- [27] ÖĞÜT, H., DOĞANAY, M.M. ve R.AKTAŞ (2009), "Detecting Stock Price Manipulation in an Emerging Market: The Case of Turkey", Expert Systems with Applications, 36, 11944-11949.
- [28] ÖZDEMİR, A.K., TOLUN, S. ve E.DEMİRCİ (2011), "Endeks Getirisi Yönünün İkili Sınıflandırma Yöntemiyle Tahmin Edilmesi: İMKB-100 Endeksi Örneği", Niğde Üniversitesi İİBF Dergisi, 4(2), 45-59.
- [29] PEREZ-CRUZ, F., AFONSO-RODRIGUEZ, J.A. ve J.GINER (2003), "Estimating GARCH Models Using Support Vector Machines", Quantitative Finance, 3, 1-10.
- [30] POON, Ser-Huang, ve C.W.J. GRANGER (2003), "Forecasting Volatility in Financial Markets: A Review", Journal of Economic Literature, 41, 478-539.
- [31] SCHÖLKOPF, B. ve A.J.SMOLA (2002), Learning With Kernels, MIT Press.
- [32] SCHÖLKOPF, B., SMOLA, A.J., WILLIAMSON, R. ve P.L.BARTLETT (2000), "New Support Vector Algorithms", Neural Computation, 12, 1207-1245.
- [33] SMOLA, A.J. ve B.SCHÖLKOPF (2004), "A Tutorial of Support Vector Regression", Statistics and Computing, 14, 199-222.
- [34] TANG, Ling-Bing, SHENG, Huan-Ye ve Ling-Xiao TANG (2009), "GARCH Prediction Using Spline Wavelet Support Vector Machine", Neural Computing & Applications, 18, 913-917.
- [35] TANG, Ling-Bing, TANG, Ling-Xiao ve Huan-Ye SHENG (2009), "Forecasting Volatility Based On Wavelet Support Vector Machine", Expert Systems with Applications, 36, 2901-2909.
- [36] TAY, Francis E.H. ve Lijuan CAO (2001), "Application of Support Vector Machines in Financial Time Series Forecasting", The International Journal of Management Science, 29, 309-317.
- [37] TİMOR, M., DİNÇER, H. ve Ş.EMİR (2012), "Performance Comparison of Artificial Neural Networks and Support Vector Machines Models for the Stock Selection Problem: An Application on the ISE-30 Index in Turkey", African Journal of Business Management, 6(3), 1191-1198.

- [38] TRAFALIS, T.B. ve Hüseyin İNCE (2000), "Support Vector Machine for Regression and Applications to Financial Forecasting", in Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, 348-353.
- [39] Van GESTEL, T., SUYKENS, J.A.K., BAESTAENS, D.E., LAMBRECHTS, A., LANCKRIET, G., VANDAELE, B., De MOOR, B. ve J.VANDEWALLE (2001), "Financial Time Series Prediction Using Least Square Support Vector Machines within the Evidence Framework", IEEE Transactions on Neural Networks, 12(4), 809-821.
- [40] VAPNIK, V. (1995), The Nature of Statistical Learning Theory, Second Edition, Springer Verlag.
- [41] VAPNIK, V. (1998), Statistical Learning Theory, John Wiley & Sons.
- [42] WEST, K.D. ve D. CHO (1995), "The Predictive Ability of Several Models of Exchange Rate Volatility", Journal of Econometrics, 69, 367-391.
- [43] YÜMLÜ, S., GÜRGEN, F.S. ve N.OKAY (2005), "A Comparison of Global, Recurrent and Smoothed-Piecewise Neural Models for Istanbul Stock Exchange Prediction", Pattern Recognition Letters, 26, 2093-2103.