

Volatility Prediction Using Hybrid Deep Learning and GARCH Family Models: A Case Study on Bitcoin

Saeed Sadrzadeh Moghadam^{*}, Kambiz Hojabr Kiani^{}**

Kambiz Peykarjou^{*}, Alireza Moradi^{****}**

Abstract

Predicting cryptocurrency price volatility is a crucial yet challenging task. Due to the nonlinear characteristics and time-varying nature of factors influencing cryptocurrency prices, this study introduces a novel approach for volatility forecasting. The proposed method integrates two key techniques. The first is the classic GARCH models, which provide concise statistical insights into price volatility through GARCH-based forecasts. The second technique involves machine learning models. The superior performance of combining GARCH models with machine learning in forecasting volatility across various markets, such as energy, base metals, and particularly stock markets, has been demonstrated compared to using either approach independently. To test this hypothesis in cryptocurrency markets, this study designs and evaluates the performance of various models based on the GARCH family and LSTM networks in predicting the volatility of a selected cryptocurrency. Subsequently, multiple hybrid models are constructed, where the

^{*} Ph.D. Candidate of Department of Economic, Faculty of Management and Economics, Islamic Azad University, Science and Research Branch, Tehran, Iran, saeed.sadrzadeh@srbiau.ac.ir, ORCID: 0000-0001-9930-0385.

^{**} Professor, Department of Economics, Faculty of Economics, Shahid Beheshti University, Tehran, Iran (Corresponding Author), kianikh@yahoo.com, ORCID: 0000-0002-0191-2138.

^{***} Assistant Professor, Department of Economics, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran, k.peykarjou@srbiau.ac.ir, ORCID :0000-0002-7468-6549.

^{****} Assistant Professor, Department of Economics, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran, Alirezaradin@yahoo.com, ORCID :0000-0001-9655-8612.

Date received: 14/11/2023, Date of acceptance: 16/02/2025



outputs of four GARCH-family models GARCH, EGARCH, GJR-GARCH, and TARCH under three different residual distribution assumptions are fed into the LSTM network. In other words, the GARCH models serve as feature extractors, while the machine learning models leverage these extracted features as sequential inputs to predict future volatility. The results indicate that standalone machine learning models not only outperform GARCH models under any residual distribution assumption but also demonstrate that the GARCH forecasts, as significant informational features, markedly enhance the hybrid machine learning models' ability to predict future volatility levels.

Keywords: GARCH models, Machine Learning, Long Short-Term Memory, Bitcoin, Volatility.

JEL classification: C22, C89, G17.

پیش‌بینی نوسان با استفاده از ترکیب مدل‌های یادگیری عمیق و مدل‌های خانواده GARCH: مطالعه موردی بیت‌کوین

سعید صدرزاده مقدم*

کامبیز هژبر کیانی**، کامبیز پیکارجو*** علیرضا مراد****

چکیده

پیش‌بینی نوسانات قیمت رمزارزها یک موضوع مهم و در عین حال چالش برانگیز است. با توجه به خصوصیات غیرخطی و ویژگی تغییرات زمانی عوامل مختلفی که بر قیمت رمزارزها تأثیر می‌گذارند، این مطالعه یک روش جدید برای پیش‌بینی نوسانات قیمت ارائه می‌نماید. در این روش، دو تکنیک مهم ترکیب شده‌اند. یکی از آن‌ها مدل‌های خودرگرسیون واریانس ناهمسان‌تعمیم‌یافته (GARCH) کلاسیک است که اطلاعات آماری مفیدی را درباره نوسانات قیمت، به صورت فشرده و از طریق پیش‌بینی‌های GARCH ارائه می‌کند. تکنیک دوم مدل‌های یادگیری ماشین است. عملکرد بهتر ترکیب مدل‌های GARCH و یادگیری ماشین در پیش‌بینی نوسانات در بازارهای مختلف مانند انرژی، فلزات اصلی و به خصوص بازارهای سهام، نسبت به هر یک از مدل‌ها بصورت جداگانه ثابت شده است. برای تأیید این فرضیه در بازار رمزارزها، در این مطالعه مدل‌های مختلفی بر اساس خانواده

* دانشجوی دکتری گروه اقتصاد، دانشکده مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی،

تهران، ایران، saeed.sadrzadeh@srbiau.ac.ir، (شناسهٔ ارکید: ۰۳۸۵-۹۹۳۰-۰۰۰۱-۰۰۰۰)

** استاد گروه اقتصاد، دانشکده اقتصاد دانشگاه شهید بهشتی، تهران، ایران (نویسندهٔ مسئول)،

kianikh@yahoo.com، (شناسهٔ ارکید: ۲۱۳۸-۰۱۹۱-۰۰۰۲-۰۰۰۰)

*** استادیار گروه اقتصاد، دانشکده مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران،

ایران، k.peykarjou@srbiau.ac.ir، (شناسهٔ ارکید: ۶۵۴۹-۷۴۶۸-۰۰۰۲-۰۰۰۰)

**** استادیار گروه اقتصاد، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران، Alirezaradin@yahoo.com،

(شناسهٔ ارکید: ۸۶۱۲-۹۶۵۵-۰۰۰۱-۰۰۰۰)

تاریخ دریافت: ۱۴۰۲/۰۸/۲۳، تاریخ پذیرش: ۱۴۰۳/۱۱/۲۸



GARCH و شبکه حافظه کوتاه مدت طولانی (LSTM) طراحی و عملکرد آن‌ها در پیش‌بینی نوسانات یک رمزارز منتخب ارزیابی شده است. سپس، مدل‌های ترکیبی مختلف ساخته شده که خروجی‌های چهار مدل GARCH، EGARCH، GJR-GARCH و TARCH، با سه فرض مختلف برای توزیع باقیمانده‌ها به شبکه LSTM تغذیه شده است. به عبارت دیگر، مدل‌های GARCH به عنوان استخراج کننده ویژگی استفاده شده اند و مدل‌های یادگیری ماشین یک دنباله از ویژگی‌های استخراج شده را به عنوان ورودی خود برای تولید نوسانات آتی به کار گرفته‌اند. نتایج نشان می‌دهد که پیش‌بینی‌های مدل‌های یادگیری ماشین بصورت جداگانه، نه تنها پیش‌بینی‌های مدل‌های GARCH را با هر فرض توزیع بهبود می‌بخشند، بلکه پیش‌بینی‌های مدل‌های GARCH به عنوان ویژگی‌های اطلاعاتی قابل توجه، قابلیت بهبود محسوسی را در پیش‌بینی مقادیر نوسان آتی توسط مدل‌های ترکیبی یادگیری ماشین دارند.

کلیدواژه‌ها: مدل‌های گارچ، یادگیری ماشین، حافظه کوتاه مدت طولانی، بیت‌کوین، نوسان

طبقه‌بندی JEL : C22, C89, G17

۱. مقدمه

در علوم مالی، نوسان به انحراف استاندارد مشروط (یا واریانس مشروط) بازده دارایی اطلاق می‌شود. در میان بازارهای مختلف مالی، رشد سریع بازار رمزارزها، نوسانات بالا و کاربردهای آن در معاملات تجاری مختلف، توجه دانشگاهیان و سرمایه‌گذاران را به خود جلب کرده است. مانند بسیاری از سری‌های زمانی مالی، رمزارزها نیز خصوصیت واریانس ناهمسان مشروط را دارند و پیش‌بینی نوسان آن‌ها به عنوان یک متغیر اساسی در مدیریت ریسک، وظیفه‌ای چالش برانگیز است. از دیدگاه اقتصاد مالی، ادبیات پیش‌بینی نوسان عمدتاً متشکل از خودرگرسیون واریانس ناهمسان (ARCH) و حالت تعمیم یافته آن (GARCH) می‌باشد. چنین مدل‌هایی به طور خاص، برای مدل‌سازی داده‌های بازار مالی که سری‌های زمانی در آن‌ها بسیار پرنوسان هستند، کاربرد دارند (حاجی زاده و همکاران، ۲۰۱۲). مدل GARCH توسط بولرسلف (Bollerslev, 1986) به عنوان تعمیم مدل ARCH معرفی شده توسط انگل (Engle, 1982) ارائه شد و یکی از محبوب‌ترین مدل‌ها برای پیش‌بینی نوسانات سری‌های زمانی است. مدل GARCH یک مدل متقارن است که در آن واریانس شرطی بر اساس مقادیر مربع باقیمانده‌ها و واریانس شرطی دوره‌های قبل تعیین می‌شود. براساس اثر

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۴۱

اهرم (Leverage Effect) در سری‌های زمانی مالی، نوسانات تمایل دارند پس از یک شوک منفی بیشتر از شوک مثبت به همان بزرگی، افزایش یابند (Yu, 2019). مدل GARCH قادر به مدل‌سازی اثر اهرم نیست، زیرا ویژگی‌های این مدل فرض می‌کند که واریانس به بزرگی شوک بستگی دارد و مستقل از علامت شوک است. (Khaledi et al 2019). بعدها چند مدل با انعطاف بیشتر مانند مدل‌های گارچ نمایی (EGARCH) توسط نلسون (Nelson, 1991) و مدل ناهمسانی توان‌دار نامتقارن (APARCH) توسط دینگ و همکاران (Ding et al, 1993) معرفی شدند که با گنجانیدن پارامترهای جدید در مدل‌های GARCH قادر به در نظر گرفتن رفتار نامتقارن در سری‌های زمانی بودند. به تازگی و با استفاده فرآیندهای هوش مصنوعی (AI) در بسیاری از زمینه‌ها، کاربرد چنین مدل‌هایی در حال توسعه در بازارهای مالی به ویژه بازارهای سهام و رمزارزها می‌باشد. مطالعات بسیاری نشان داده است که مدل‌های شبکه عصبی مصنوعی (ANN) در پیش‌بینی‌های نوسانات از مدل‌های سستی GARCH پیشی گرفته‌اند. زیرا می‌توانند غیر خطی بودن سری زمانی را مدل‌سازی کرده و نیازی به ثابت بودن سری زمانی جهت مدل‌سازی نیست.

۲. مبانی نظری

۱.۲ مدل‌های GARCH

انگل، نشان داد که می‌توان به صورت هم‌زمان میانگین و واریانس شرطی یک سری زمانی را مدل‌سازی کرد. هدف اصلی مدل‌سازی نوسانات، استفاده از یک فرآیند خودرگرسیون ARMA به منظور شناسایی هرگونه روند پایدار، در تغییرات جزء خطا است. اگر یک سری زمانی خودرگرسیون مرتبه اول ($p=1$) را در نظر بگیریم، خواهیم داشت:

$$var(y_t) = E[(y_{t-1} - \alpha_0 - \alpha_1 y_t)^2] = E_t(\varepsilon_{t+1})^2 \quad (1)$$

$E_t(\varepsilon_{t+1})^2$ یک فرآیند خودرگرسیون از مرتبه p در نظر گرفته می‌شود و بر اساس مدل زیر برآورد خواهد شد.

$$\hat{\varepsilon}_t^2 = \beta_0 + \beta_1 \hat{\varepsilon}_{t-1}^2 + \dots + \beta_p \hat{\varepsilon}_{t-p}^2 + u_t \quad (2)$$

اگر $\beta_1 = \beta_2 = \dots = \beta_p = 0$ باشد، واریانس جمله اختلال ثابت و برابر با β_0 خواهد بود. در غیر این صورت واریانس شرطی y_t مطابق با فرآیند خودهمبسته معرفی شده در رابطه (۲) تغییر می‌کند. به این الگو ARCH گفته می‌شود. جملات اختلال می‌توانند از

الگوهای مختلفی پیروی نمایند. لذا شکل‌های متفاوتی از الگوی ARCH وجود دارد. الگوی معرفی شده توسط انگل برای واریانس شرطی توسط بولرسف توسعه داده شد. در این مدل، الگوی واریانس شرطی از اجزای $ARCH(\varepsilon_{t-i}^2)$ و نیز $GARCH(\sigma_{t-i}^2)$ به صورت ترکیبی و با مرتبه‌های متفاوت تعیین می‌شود.

$$\sigma_t^2 = \alpha_0 + \sum_1^q \alpha_i \varepsilon_{t-i}^2 + \sum_1^p \beta_i \sigma_{t-i}^2 \quad (۳)$$

که در آن σ_t واریانس شرطی ε_t است. یک مدل $GARCH(1,1)$ که بعنوان مدل پایه یا استاندارد شناخته می‌شود، به صورت زیر تعریف می‌شود.

$$y_t = \mu_t + u_t \quad (۴)$$

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (۵)$$

$$u_t = \sigma_t \varepsilon_t, \varepsilon_t \sim N(0,1) \quad (۶)$$

رابطه شماره (۴) معادله میانگین شرطی و رابطه (۵) واریانس شرطی را نشان می‌دهد. اولین محدودیت در مدل $GARCH(1,1)$ استاندارد، شرط غیرمنفی بودن ضرایب ثابت α ، β و ω است و دومین محدودیت این مدل $\alpha + \beta < 1$ است (Bollerslev, 1986). مدل $GARCH$ خطی پایه نمی‌تواند خواص نامتقارن سری‌های زمانی را نمایش دهد. لذا مدل‌هایی مانند $EGARCH$ ، $TGARCH$ ، $GJR-GARCH$ متعلق به خانواده $GARCH$ معرفی می‌شوند که توانایی بررسی و نمایش خواص نامتقارن سری‌های زمانی را داشته باشند. مدل $GJR-GARCH$ از نظر معادله میانگین با مدل $GARCH$ پایه یکسان و تفاوت آنها در معادله واریانس شرطی است. گلاستن، جگنزان و رانکل (Glosten, Jaganathan, Runkle, 1994) فرض عدم تقارن اثر شوک‌های مثبت و منفی بر واریانس‌های شرطی را با اضافه کردن یک تابع شاخص به معادله واریانس مشروط مطرح کردند. در این الگو $\varepsilon_{t-1} = 0$ یک آستانه (Threshold) در نظر گرفته می‌شود که شوک‌های بزرگتر از شوک آستانه، تأثیرات متفاوتی نسبت به شوک‌های کوچکتر از شوک آستانه دارند. مدل $GJR-GARCH$ به صورت زیر خواهد بود.

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I_{[\varepsilon_{t-1}] \quad (۷)$$

$I_{[\varepsilon_{t-1}]}$ یک تابع شاخص است که با ازای $\varepsilon_{t-1} < 0$ برابر با یک و به ازای $\varepsilon_{t-1} \geq 0$ برابر با صفر می‌باشد. (Enders, 1948) مدل $TGARCH$ شبیه مدل $GJR-GARCH$ است با این

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۴۳

تفاوت که از قدر مطلق باقیمانده‌ها به جای توان دوم آن‌ها استفاده می‌کند. معادله واریانس شرطی مدل TGARCH به صورت زیر است.

$$\sigma_t^2 = \omega + \alpha|\varepsilon_{t-1}| + \beta\sigma_{t-1} + \gamma|\varepsilon_{t-1}|I_{[\varepsilon_{t-1}]}$$
 (۸)

حالت عمومی‌تر این مدل که می‌تواند هر دو وضعیت را توصیف کند به شکل زیر است.

$$\sigma_t^k = \omega + (\alpha + \gamma I_{[\varepsilon_{t-1}]})|\varepsilon_{t-1}|^k + \beta\sigma_{t-1}^k$$
 (۹)

واریانس شرطی در این مدل برابر با $(\sigma_t^k)^{2/k}$ است. این حالت عمومی به مدل A-PARCH معرفی شده توسط دینگ، گرنجر و انگل (Ding, Granger & Engle, 1993) شبیه است. اگر k مساوی با ۲ باشد مدل به حالت GJR-GARCH خواهد بود و برای $k=1$ مدل به حالت TGARCH تعریف می‌شود. مدل دیگری که برای پیش‌بینی‌های غیرخطی استفاده می‌شود مدل EGARCH است. نلسون (Nelson, 1991) به منظور بررسی اثر عدم تقارن در اطلاعات یا چولگی‌هایی که اغلب در بازده‌های مالی مشاهده می‌شوند، مدل EGARCH را معرفی کرد که شرط نامنفی بودن پارامترها حذف می‌کند. در اینجا بازهم معادله میانگین شرطی مانند حالت‌های قبلی است و معادله واریانس شرطی به شکل زیر معرفی می‌شود.

$$\ln(\sigma_t) = \omega + \alpha \frac{u_{t-1}}{\sqrt{\sigma_{t-1}}} + \beta \left(\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}}} - \sqrt{\frac{2}{\pi}} \right)$$
 (۱۰)

لگاریتم‌گیری در سمت چپ معادله تضمین واریانس غیر منفی است. α ، β و ω پارامترهای ثابت مدل هستند. برای $\alpha = 0$ اثر شوک‌های منفی و مثبت یکسان است. در صورتیکه $\alpha > 0$ باشد، شوک‌های مثبت واریانس شرطی را افزایش داده و اگر $\alpha < 0$ باشد، شوک‌های مثبت باعث کاهش واریانس شرطی خواهند شد. لذا پارامتر α عدم تقارن‌های مدل را مشخص می‌کند. از دیگر خصوصیات مدل‌های GARCH خصوصیت لپتوکورتوسیس (Leptokurtosis) توزیع تجربی بازده‌های مالی است. محققان اغلب به منظور ایجاد توزیع‌های دنباله‌دار، از توزیع خطای تعمیم‌یافته (Generalized Error Distributions) یا توزیع تی‌استیودنت (t-Student) استفاده می‌کنند، که علاوه بر فرض گوسی سستی، خطاهای ε_t براساس توزیع GED یا توزیع تی‌استیودنت توزیع می‌شود. تابع چگالی احتمال توزیع تی‌استیودنت با در نظر گرفتن درجه آزادی ν و با این فرض که برای وجود گشتاور مرتبه دوم، باید درجه آزادی بزرگتر از دو باشد، به صورت زیر است.

$$f(\varepsilon_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})} (\nu-2)^{-\frac{1}{2}} (h_t)^{-\frac{1}{2}} \left[1 + \frac{\varepsilon_t^2}{h_t(\nu-2)} \right]^{-\frac{(\nu+1)}{2}} \quad (11)$$

تابع توزیع GED نیز به صورت زیر تعریف می شود.

$$f(\varepsilon_t) = \frac{v \exp \left[-\left(\frac{1}{2} \right) \left| \frac{\varepsilon_t}{\lambda h_t^{\frac{1}{\nu}}} \right|^{\frac{1}{\nu}} \right]}{h_t^{1/2} \lambda 2^{(1+\frac{1}{\nu})} \Gamma(\frac{1}{\nu})} \quad (12)$$

در اینجا ν پارامتری است که ضخامت دنباله توزیع را در مقایسه با توزیع نرمال نشان می دهد. زمانی که $\nu = 2$ باشد توزیع GED تبدیل به توزیع نرمال استاندارد خواهد شد. برای $\nu < 2$ و $\nu > 2$ توزیع GED به ترتیب دنباله ضخیم تر و نازک تر از توزیع نرمال خواهد داشت.

۲.۲ یادگیری ماشین (Machine Learning)

یادگیری ماشین زمینه ای از مطالعه در هوش مصنوعی است که الگوریتم ها در آن توسعه یافته اند تا رایانه ها بدون برنامه ریزی صریح و روشن الگوهای موجود در داده ها را یاد بگیرند. روش های یادگیری ماشین در سال های اخیر اهمیت بسزایی پیدا کرده و تعداد افراد یا سازمان هایی که آنها را توسعه می دهند افزایش یافته است. علاوه بر این، منابع محاسباتی برای انجام این نوع الگوها در دسترس تر شده اند. روش های آماری و اقتصادسنجی اغلب در زمان مدل سازی متغیرهای نامانا و یا متغیرهایی که وابستگی های پیچیده دارند، مشکلاتی را به وجود می آورند. (Dixon et al, 2015) خوشبختانه تکنیک های یادگیری ماشین می توانند این نوع ساختارهای پیچیده را شناسایی و با آن مقابله کنند. (Fischer & Krauss, 2018) شبکه های عصبی مصنوعی (ANN) بخشی از مدل های یادگیری ماشین هستند. بصورت کلی این شبکه ها بر اساس مجموعه ای از واحدهای موسوم به نورون ها (Neurons) که با یکدیگر تعامل دارند و اطلاعات بین آنها به اشتراک گذاشته می شوند، تشکیل شده است. نورون یک بردار اطلاعاتی مانند x را دریافت می کند که می تواند داده ورودی و یا خروجی یک نورون دیگر باشد. اطلاعات دریافت شده، در یک بردار وزنی مانند w ، که در طول فرآیند یادگیری توسط الگوریتم تخمین زده می شود، ضرب و بدین ترتیب یک مقدار خروجی مانند $f(x)$ ، از طریق یک تابع تبدیل (Transformation Function) مانند g که به تابع فعال سازی

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۴۵

(Activation Function) معروف است، ایجاد می‌شود. این فرآیند را می‌توان به صورت ساده تر از طریق عبارت زیر نشان داد:

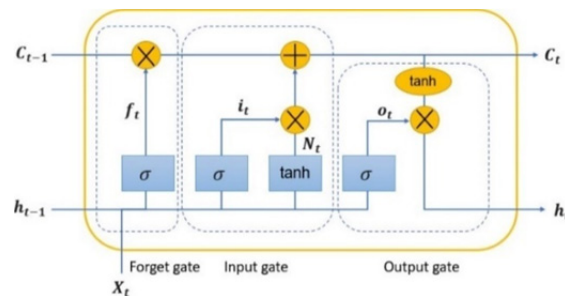
$$f(x) = g(x \cdot w) + b \quad (۱۳)$$

پرسپترون چندلایه (MultiLayer perceptron /MLP) ساده‌ترین شکل شبکه‌های عصبی است که از یک لایه ورودی، یک لایه پنهان و یک لایه خروجی تشکیل شده‌است. شبکه‌های عصبی بازگشتی (Recurrent Neural Network/RNN) یک روش کامل‌تر برای مدل‌سازی داده‌های متوالی است. این شبکه از لایه‌های پنهان برای جمع‌بندی اطلاعات تاریخی داده‌ها استفاده و چرخه‌ای برای نگهداری اطلاعات در طول زمان ایجاد می‌کند. یک واحد در شبکه عصبی بازگشتی را می‌توان با x_t نشان داد که یک مقدار h_t تولید می‌کند و در ساختار آن، خروجی زمان $t-1$ بعنوان ورودی در زمان t استفاده می‌شود. زمانی که پیش‌بینی از خروجی مورد نظر در زمان t ، به ورودی ارائه شده در زمان‌های طولانی‌تر از $t-1$ وابسته باشد، الگوریتم‌های بازگشتی با مشکلاتی روبرو می‌شوند، چراکه پارامترهای آن‌ها، در راه حل‌های بهینه برای اثرات کوتاه‌مدت بدست می‌آیند. یکی از اصلی‌ترین مشکلات در شبکه‌های عصبی بازگشتی، گرادیان‌های محو شونده (Vanishing Gradient) است، که باعث توقف یا کندی یادگیری شبکه می‌گردد. این مشکل زمانی اتفاق می‌افتد که گرادیان‌ها یا مشتقات تابع خطا نسبت به وزن‌های شبکه، در طول فرآیند یادگیری به سرعت کوچک می‌شوند و در نهایت به صفر می‌رسند. این مشکل باعث می‌شود تغییرات کمی در وزن‌های شبکه رخ داده و بهینه‌سازی متوقف شود. در واقع مشتق‌هایی که از لایه‌های انتهایی شبکه می‌آیند باید ضرب‌های متعددی را پشت سر بگذارند تا به لایه‌های ابتدایی برسند، لذا اگر این ضرب‌ها مقادیر کوچک (کمتر از یک) داشته باشند گرادیان‌ها رفته‌رفته کوچک و محو می‌شوند.

۱.۲.۲ شبکه‌های حافظه کوتاه مدت طولانی (Long Short-Term Memory /LSTM)

شبکه‌های LSTM، یکی از انواع نوع شبکه‌های عصبی بازگشتی است، که توسط هوکرایتر و اش‌مید‌هوبر (Hochreiter & Schmidhuber, 1997) و برای حل مشکل گرادیان‌های محو شونده، معرفی شد. شبکه‌های LSTM اجازه می‌دهند اطلاعات در شبکه باقی بمانند و وابستگی‌های کوتاه‌مدت و طولانی‌مدت را در دنباله ورودی یاد بگیرند. در واقع مشکل محو شوندگی تدریجی اطلاعات در شبکه‌های عصبی بازگشتی در اینجا برطرف شده است.

شکل شماره ۱ بخش‌های مختلف یک شبکه LSTM را نشان می‌دهد. در یک شبکه LSTM، سه دروازه (Gate) ورودی، فراموشی و خروجی وجود دارد.



شکل ۱. شمای یک واحد از شبکه عصبی حافظه کوتاه مدت طولانی
منبع: یافته‌های تحقیق

دروازه فراموشی: شبکه LSTM ابتدا در مورد اینکه چه بخشی از اطلاعات دوره قبل باید نادیده گرفته شوند، تصمیم می‌گیرد. این موضوع را با رابطه زیر می‌توان نمایش داد:

$$f_t = \sigma(X_t U_f + H_{t-1} W_f) \quad (14)$$

X_t ورودی در زمان جاری (t)، U_f وزن تخصیص داده شده به ورودی، H_{t-1} بلوک اطلاعاتی دوره قبلی و W_f یک ماتریس وزن‌دهی برای اطلاعات دوره قبلی است. یک تابع زیگموند (Sigmoid) روی f_t اعمال و مقدار صفر یا یک را برمی‌گرداند. مقدار صفر نشان می‌دهد همه اطلاعات فراموش شوند و مقدار یک باعث نگهداری اطلاعات دوره قبلی خواهد شد.

دروازه ورودی: این دروازه اهمیت اطلاعات جدید را بصورت کمی تعیین می‌کند.

$$i_t = \sigma(X_t U_i + H_{t-1} W_i) \quad (15)$$

U_i و W_i به ترتیب وزن‌های مربوط به اطلاعات ورودی و اطلاعات موجود در شبکه است. یک تابع زیگموند دیگر مقادیر را در بازه صفر و یک قرار داده و هر چه مقادیر مذکور به یک نزدیک‌تر باشند، نشان دهنده اهمیت بیشتر اطلاعات ورودی است. اطلاعات جدید به اطلاعات قبلی افزوده شده و یا از آن کسر می‌گردد. در اینجا از یک تابع تانژانت هایپربولیک استفاده می‌شود که اطلاعات را بین -1 برای کسر کردن از اطلاعات موجود در شبکه و $+1$ برای اضافه کردن به اطلاعات موجود در شبکه قرار می‌دهد. N_t اطلاعات

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۴۷

جدید است که با استفاده از دروازه فراموشی و دروازه ورودی، قبل از اضافه شدن به اطلاعات موجود بروزرسانی می‌شود.

$$N_t = \tanh(X_t U_c + H_{t-1} W_c) \quad (16)$$

$$C_t = f_t C_{t-1} + i_t N_t \quad (17)$$

C_t سلول اطلاعات یا حافظه بلند مدت مدل در زمان $t-1$ است.

دروازه خروجی: این دروازه تصمیم می‌گیرد که چه بخشی از اطلاعات بعنوان خروجی وارد مرحله بعدی شود.

$$O_t = \sigma(X_t U_o + H_{t-1} W_o) \quad (18)$$

$$h_t = Q_t \tanh(C_{t-1}) \quad (19)$$

۲.۲.۲ شبکه حافظه کوتاه مدت طولانی دو طرفه (Bidirectional LSTM)

شبکه عصبی LSTM دوطرفه یکی از روش‌های یادگیری ماشین است که اغلب برای یادگیری داده‌های ترتیبی مورد استفاده قرار می‌گیرد. همانگونه که در بخش قبل گفته شد روش LSTM با استفاده از دروازه‌هایی برای یادگیری و فراموش کردن داده‌ها، به نوعی به صورت حافظه‌دار عمل می‌کند. به این ترتیب با دور شدن از داده‌های اولیه، ممکن است تاثیر آن داده‌ها کم نشده و به این ترتیب عملکرد بهتری مشاهده می‌شود. در مدل LSTM دوطرفه، در زمان آموزش، یک بار از ابتدا به انتها، داده‌ها مورد بررسی قرار گرفته و وزن‌ها و مقادیر خطا اصلاح می‌شوند، که به آن حالت پسین (Forward states) گفته می‌شود. سپس یک بار از انتها به ابتدا فرآیند آموزش انجام و ماتریس‌های مربوطه تنظیم می‌شوند، که به آن حالت پیشین (Backward States) گفته می‌شود. در نهایت این دو بخش از محاسبات، در کنار یکدیگر بعنوان خروجی قرار می‌گیرند. خروجی این دو حرکت به ورودی در حرکت‌ها در جهت مخالف متصل نیستند.

۳.۲.۲ شبکه عصبی پیچشی (Convolutional Neural Network)

شبکه عصبی پیچشی، یک الگوریتم یادگیری ماشین است که می‌تواند یک مجموعه از داده‌ها را بعنوان ورودی دریافت نموده، به جنبه‌ها یا ویژگی‌های مختلف موجود در داده‌ها، اهمیت (وزن‌ها و خطاهای قابل یادگیری) اختصاص داده و در نهایت یکی را از دیگری

متمایز می‌کند. پیش‌پردازش مورد نیاز در یک شبکه کانولوشنی، به مراتب کمتر از سایر الگوریتم‌های یادگیری ماشین است. همچنین این شبکه‌ها، توانایی یادگیری فیلترها یا ویژگی‌ها را دارند، در حالیکه در روش‌های دیگر، این فیلترها به صورت دستی طراحی و مهندسی می‌شدند. یک شبکه عصبی پیچشی از لایه‌های پیچشی (Convolutional Layer)، لایه ادغام (Pooling Layer) و لایه تماماً متصل (Fully Connected Layer) تشکیل شده است. در فرآیند مدل‌سازی یادگیری ماشین، تعریف فیلترهای مشخص توسط کاربر، به منظور کشف یک ساختار در داده‌ها و آموزش شبکه براساس ساختار مذکور، انجام می‌شود. جستجو در داده‌ها برای یافتن تنها یک الگو، منجر به نتایج خوبی نشده و باعث می‌شود شبکه از لحاظ کارایی محدود باشد. هر لایه کانولوشن در شبکه عصبی کانولوشنی شامل مجموعه‌ای از فیلترها یا کرنل (kernel) است و خروجی توسط کانولوشن بین فیلترها و لایه ورودی ساخته می‌شود. تعداد فیلترها، تعداد ویژگی‌های تشخیص داده شده را نشان می‌دهد. این پارامتر معمولاً بصورت نمایی از ۲ (معمولاً بین ۳۲ تا ۴۰۹۶) انتخاب می‌شود. اغلب استفاده از فیلترهای بیشتر، باعث ایجاد یک شبکه عصبی قدرتمندتر می‌شود. اما این افزایش پارامترها ممکن است باعث بیش برآورد (Overfitting) گردد. در لایه پیچشی یک گام (Stride) تعریف می‌شود که نشان می‌دهد فیلتر بعد از محاسبه در یک پنجره از ورودی، چند درایه یا خانه باید جلو برود تا دوباره محاسبات را انجام دهد. در این لایه تمام نورون‌های لایه قبلی به تمام نورون‌های لایه فعلی متصل نیستند. لایه ادغام یکی دیگر از لایه‌های مهم در شبکه عصبی پیچشی است. هدف لایه ادغام کاهش اندازه مکانی خروجی لایه پیچشی است. این لایه پارامتر قابل آموزش نداشته و صرفاً یک نمونه برداری ساده و موثر انجام می‌دهد. رایج‌ترین نمونه لایه ادغام، لایه ادغام مقدار حداکثر و لایه ادغام مقدار متوسط است. لایه تماماً متصل در یک شبکه عصبی پیچشی، بعنوان لایه انتهایی شناخته می‌شود که تمام اطلاعات موجود در لایه قبلی را دریافت و براساس ویژگی‌های موجود دسته بندی می‌کند. هر نورون در این لایه، به تمام نورون‌های لایه قبلی متصل است. در صورت استفاده از تنها یک لایه کاملاً متصل، براساس نوع تابع فعال ساز انتخاب شده، این لایه می‌تواند شبیه یک رگرسیون خطی و یا غیرخطی عمل نماید.

۳.۲ معیارهای ارزیابی مدل

هدف اصلی طراحی مدل‌های پیش‌بینی‌کننده در سری‌های زمانی، پیش‌بینی مقادیر آتی براساس داده‌های گذشته است. این امکان وجود دارد که بهترین مدل برآورد شده با داده‌های آموزش، بهترین مدل در پیش‌بینی‌های برون نمونه‌ای نباشد. لذا از معیارهایی برای ارزیابی دقت پیش‌بینی مدل استفاده می‌شود. معیارهای مورد استفاده جهت ارزیابی مدل برآورد شده و دقت پیش‌بینی، در این پژوهش مجذور میانگین مربعات خطا (RMSE)، مجذور درصد میانگین مربعات خطا (RMSPE) و قدرمطلق میانگین خطا (MAE) می‌باشند.

$$RMSE = \sqrt{\frac{\sum_{t=T+1}^{T+m} (y_t^f - y_t)^2}{m}} \quad (20)$$

$$MAE = \frac{\sum_{t=T+1}^{T+m} |y_t^f - y_t|}{m} \quad (21)$$

$$MAPE = \sqrt{\frac{100 \sum_{t=T+1}^{T+m} \left(\frac{y_t^f - y_t}{y_t} \right)^2}{m}} \quad (22)$$

m طول دوره پیش‌بینی، y_t مقادیر واقعی و y_t^f مقادیر پیش‌بینی شده است. در این پژوهش پس از برآورد مدل، با استفاده از داده‌های بخش آموزش، مقادیر برآورد شده با داده‌های بخش ارزیابی مورد مقایسه قرار گرفته و خطاهای پیش‌بینی، براساس روابط (۲۰) الی (۲۲) محاسبه می‌شود.

۳. پیشینه تحقیق

مدل‌های GARCH در ابتدا توسط بولرسف (Bollerslev, 1986) معرفی شدند. بر همین اساس انگل (Engel, 2002) با استفاده از یک مدل GARCH(1,1) و داده‌های نزدیک و داو جونز، یک مدل مالی برای تشکیل پرتفولیو طراحی و از متغیر ارزش در معرض خطر، جهت ارزیابی مدل استفاده نمود. به همین ترتیب، مدل‌های GARCH چند متغیره با خاصیت همبستگی شرطی پویا (Dynamic Conditional Correlation / DCC) توسط انگل پیشنهاد گردید. این مدل‌ها دارای کمترین پارامتر نسبت به سایر مدل‌های چند متغیره خانواده GARCH هستند. لاموروکس و لاستراپ (Lamoureux & Lastrapes, 1990) مدل‌های GARCH تشکیل پرتفولیو

را با در نظر گرفتن حجم معاملات به عنوان متغیر برونزا، گسترش دادند. نتایج نشان داد که اضافه کردن این متغیر، قدرت توضیح‌دهندگی مدل را بهبود می‌بخشد. در موضوع رمزارزها، مطالعه چن و همکاران (Chu et al, 2017) یکی از اولین مطالعاتی است که از حالت‌های مختلف مدل GARCH برای پیش‌بینی بازده روزانه ۷ رمزارز با بالاترین ارزش بازار استفاده کرده‌است. آنها مطالعات قبلی انجام شده روی بیت‌کوین را مدل کرده و نشان دادند در هنگام استفاده از مدل‌های GARCH نامتقارن، ویژگی‌های یکسانی بین رمزارزهای منتخب و طلا و دلار وجود دارد. با این وجود بائور و همکاران (Baur et al, 2018)، مطالعه قبلی را توسعه دادند و به نتایج متضادی دست یافتند. به عبارت دیگر، نتایج آنها نشان داد دینامیک بیت‌کوین بسیار متفاوت از طلا و دلار است. ویجای لاکشمی (Vijayalakshmi, 2024) در یک مطالعه جالب پس از جمع‌آوری داده‌های سهام شرکت اپل و استفاده از شبکه اجتماعی توییتر و اجرای تحلیل احساسات (Sentiment Analysis) روی آن، با استفاده از مدل‌های خانواده GARCH برای ارائه یک پیش‌بینی یک ساله از نوسان قیمت سهام مذکور اقدام کرد. نتایج مطالعه نشان داد مدل GJR-GARCH عملکرد بهتری در مقایسه با دیگر مدل‌های خانواده GARCH دارد. با معرفی و توسعه مدل‌های یادگیری ماشین، محققین مطالعات بسیاری را برای مدل‌سازی سری‌های زمانی انجام دادند. فیشر و کوراس (Fischer & Krauss, 2018) با استفاده از مدل LSTM پیش‌بینی‌هایی برای حرکت‌های جهت‌دار در سهام را ارائه و با این اطلاعات، یک استراتژی سرمایه‌گذاری پیشنهاد نمودند. نتایج آنها نشان داد استفاده از یادگیری ماشین در پیش‌بینی روند، حتی در بازارهای باتلاطم، امکان به دست آوردن اطلاعات مفید را بوجود می‌آورد. در این راستا، رقابت‌هایی در استفاده از روش‌های آماری و اقتصادسنجی و یادگیری ماشین برای ارزیابی پیش‌بینی متغیرهای کلان اقتصادی بوجود آمده و در برخی از موارد، روش‌های آماری کلاسیک، روش‌های یادگیری ماشین را پشت سر می‌گذارند (Makridakis, 2018). در حوزه رمزارزها، مدینا و هوئن (Medina & Huynh, 2021) از مدل‌های یادگیری ماشین برای پیش‌بینی روند حرکتی بیت‌کوین، با در نظر گرفتن متغیرهای غیرسستی و استفاده از روش غیرپارامتریک انتخاب متغیر، استفاده نمودند. بانیک و همکاران (Banik et al, 2022) سامانه پشتیبان تصمیم‌گیری (DSS) را با استفاده از رویکرد LSTM برای پیش‌بینی بازار سهام هند ایجاد نمودند. با این حال، در هیچ یک از این مطالعات به صورت صریح به نوسان (Volatility) پرداخته نشده‌است. آصاف و همکاران (Assaf et al, 2022) روش جدیدی را برای بهبود دقت

پیش‌بینی نوسانات بازار از طریق معرفی یک تنظیم جدید الگوریتم شبکه عصبی بازگشتی پیشنهاد نمودند. مدل پیشنهادی یک LSTM کامل شده با ورودی چندمتغیره که از قیمت‌های روزانه چندین دارایی با وقفه‌های زمانی مختلف، تشکیل شده است. مدل مذکور برای پیش‌بینی نوسانات استفاده شده و نتایج آن با پیش‌بینی‌های حاصل از GARCH و همچنین نوسانات واقعی دوره پیش‌بینی مقایسه شده است. نتایج نشان می‌دهد که پیش‌بینی نوسانات آینده با استفاده از مدل LSTM دارای دقت قابل قبول نسبت به رویکرد مدل‌سازی GARCH است. همچنین نتایج این مطالعه نشان می‌دهد که مدل LSTM زمانی که با چند دارایی و وقفه‌های زمانی بیشتر از یک روز طراحی شود، می‌تواند دقت پیش‌بینی نوسانات را به طور قابل توجهی بهبود بخشد. یکی از حوزه‌های مطالعاتی در این خصوص، ترکیب مطالعه نوسانات با استفاده از مدل‌های GARCH و یادگیری ماشین است، که منجر به پیشنهاد مدل‌های ترکیبی (Hybrid Model) شده، تا دقت پیش‌بینی نوسانات را بهبود بخشد. براساس مطالعات انجام شده معمولاً ترکیبی از مدل‌های یادگیری عمیق و مدل‌های نوع GARCH در مقایسه با تک مدل یادگیری عمیق یا مدل سازی سری‌های زمانی به تنهایی، عملکرد بهتری پیدا می‌کنند. کریست‌جانپلر و همکاران (Kristjanpoller et al, 2014) مدل هیبریدی و GARCH را در پیش‌بینی سه شاخص بورس آمریکای لاتین از برزیل، شیلی و مکزیک آزمایش کردند. آنها از آزمون میانگین استفاده و بهبود چشمگیری در دقت پیش‌بینی در مقایسه با مدل GARCH را نشان دادند. کریست‌جانپلر و مینوتولو (Kristjanpoller & Minutolo, 2018) با استفاده از ترکیب‌های مختلف پیش‌بینی‌های خانواده GARCH، قیمت بیت‌کوین، شاخص‌های فنی و اجزای اصلی حاصل از شاخص‌های فنی را به عنوان ورودی مدل ترکیبی انتخاب نمودند. در تمام موارد مدل ترکیبی نتایج بهترین مدل GARCH را بهبود بخشید. کاربرد مدل‌های ترکیبی در پیش‌بینی نوسان نفت خام را می‌توان در کار کریست‌جانپلر و مینوتولو (Kristjanpoller & Minutolo, 2016) یافت که پیش‌بینی‌های مدل GARCH به همراه داده‌های سری زمانی مالی (نرخ ارز و شاخص بازار بورس) به عنوان ورودی در مدل مورد استفاده قرار گرفتند. آنها به این نتیجه رسیدند که پیش‌بینی‌های نوسانات به دست آمده توسط مدل ترکیبی، نسبت به مدل‌های پیش‌بینی سنتی بهبود یافته است. بعنوان یک نمونه در زمینه بازار کالاهای فلزی، کریست‌جانپلر و همکاران (Kristjanpoller et al, 2017) پیش‌بینی بهترین مدل‌های GARCH را به همراه برخی دیگر از متغیرهای مالی به عنوان ورودی دیگر مدل‌ها بکار بردند، تا نوسانات طلا، نقره و مس را

پیش‌بینی کنند. نتایج آنها نشان داد که مدل GARCH که از متغیرهای برون‌زا استفاده می‌کند، پیش‌بینی بهتری نسبت به مدل GARCH می‌دهد که تنها با بازدهی متغیر مورد مطالعه برآورد شده است. علاوه بر این نشان دادند که مدل ترکیبی در مقایسه با بهترین پیش‌بینی مدل‌های GARCH پیش‌بینی‌ها را به صورت قابل توجهی بهبود می‌بخشد. در این راستا، کیم و وون (Kim & Won, 2018) مدل‌های LSTM را با مدل‌های GARCH مختلف ترکیب کردند تا نوسانات شاخص قیمت سهام کره جنوبی را پیش‌بینی نمایند. تازگی کار آنها این است که به جای استفاده از پیش‌بینی‌های نوع GARCH از پارامترهای تخمینی دو یا چند مدل GARCH-TYPE به عنوان ورودی‌های مدل LSTM استفاده می‌کنند. نتایج آن‌ها نشان داد که مدل GEW-LSTM آنها که از مدل GARCH، EGARCH با LSTM ترکیب شده، کمترین خطاهای پیش‌بینی را دارد. یک مدل ترکیبی جالب توسط پنگ و همکاران (Peng et al, 2018) طراحی شده است. در این مطالعه از یک مدل SVR برای تخمین معادلات میانگین و واریانس مدل سستی GARCH استفاده شده است. مطالعه آنها نشان داد که برای تمام نرخ ارزها و رمزارزهای مورد مطالعه و در داده‌های با فرکانس بالا و پایین (High and Low Frequency Data) مدل SVR-GARCH در مقایسه با سایر مدل‌های نوع گارچ با پیش‌فرض‌های مختلف برای توزیع‌های باقیمانده، کمترین میزان خطاهای پیش‌بینی را دارد. کوئینترو (Quintero, 2019) از یک روش مشابه تلفیق مدل‌های LSTM و GARCH برای پیش‌بینی نوسانات نرخ تبادل دلار به پزوی کلمبیا استفاده کرد، که در آن ترکیبی از نرخ بهره و قیمت کالا به عنوان متغیرهای برون‌زا اضافه شده است. خالدی و همکاران (Khalidi et al, 2019) در مطالعه‌ای به بررسی تطبیقی مدل‌های GARCH به عنوان مدل پارامتریک و مدل‌های یادگیری عمیق به عنوان مدل‌های غیرپارامتریک، برای پیش‌بینی نوسان پرداختند. در مطالعه آنها هفت مدل از نوع GARCH، LGARCH، EGARCH، APARCH، TGARCH، GJR-GARCH، CGARCH و IGARCH از نظر توانایی در پیش‌بینی نوسان بیت‌کوین تحلیل شد. با فرض چهار توزیع مختلف برای باقیمانده‌های استاندارد شده، مدل TGARCH بهترین رویکرد پارامتریک، در میان تمام مدل‌های هفت‌گانه GARCH انتخاب شد. نویسندگان دو مدل یادگیری ماشین شامل MLP و RNN را نیز ارزیابی کردند. آنها برتری MLP را اثبات نمودند. مقایسه آن با مدل TGARCH نشان داد که از TGARCH نیز بهتر عمل می‌کند. در زمینه بازار سهام راموس‌پرز و همکاران (Pérez et al, 2019) برای پیش‌بینی نوسان شاخص S&P500، مدل انباشته دو سطحی را بر اساس انواع مدل‌های یادگیری ماشین پیشنهاد کردند.

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۵۳

مدل انباشته شده آنها از مدل‌های ترکیبی که بر اساس مدل GARCH، EGARCH و یادگیری ماشین ساخته شده، برتر بود. در همین زمینه، لیو (Liu, 2019) ثابت کرد که مدل LSTM و مدل رگرسیون پشتیبان برداری (SVR) می‌تواند برای یک بازه زمانی بزرگ، پیش‌بینی نوسانات S&P500 را نسبت به مدل GARCH بهبود دهند. ذوالفقاری و همکاران (۲۰۲۰) یک مدل ترکیبی را با استفاده از مدل‌های GARCH و EGARCH و مدل‌های یادگیری ماشین طراحی و نسبت به پیش‌بینی کوتاه‌مدت بازدهی شاخص کل بودرس اوراق بهادار تهران اقدام نمودند. همچنین دو متغیر قیمت نفت و نرخ ارز را نیز در مدل وارد کردند. نتایج تحقیق نشان داد که مدل‌های ترکیبی دقت پیش‌بینی بالاتری نسبت به مدل‌های تکی داشته و همچنین براساس معیارهای ارزیابی خطای پیش‌بینی RMSE و MAPE، مدل RNN-LSTM-EGARCH برپایه توزیع GED دارای خطای پیش‌بینی کمتری نسبت به کلیه مدل‌های مورد استفاده در این تحقیق دارد. دو مدل یادگیری ماشین را می‌توان در کار ویدال و کریست‌جانپلر (Vidal and Kristjanpoller, 2020) مشاهده کرد که در آن از یک مدل ویژه شبکه عصبی پیچشی، برای داده‌های بازدهی طلا استفاده و نتایج را با یک شبکه عصبی حافظه کوتاه مدت طولانی ترکیب کردند. به غیر از بازار طلا، اثربخشی مدل‌های ترکیبی در پیش‌بینی نوسانات مس در مطالعه هو و همکاران (Hu et al, 2020) ثابت شد. نتایج این مطالعه نشان داد که نه تنها مدل‌های شبکه عصبی می‌توانند از مدل GARCH بهتر عمل کنند، بلکه استفاده از پیش‌بینی‌های GARCH نیز به عنوان ورودی، می‌تواند قدرت پیش‌بینی مدل‌ها را تقویت کند. علاوه بر این نتایج نشان داد که گنجاندن مدل‌های LSTM و LSTM دوطرفه به مدل‌های ترکیبی، پیش‌بینی‌های نوسانات قیمت مس را بهبود می‌بخشد. سئو و کیم (Seo and Kim, 2020) مدل عصبی مرتبه‌های بالا (HONN) را برای ساخت مدل‌های ترکیبی به منظور پیش‌بینی نوسانات بیت‌کوین به کار گرفتند. آنها خروجی مدل‌های GARCH به همراه مقادیر نوسان تحقق‌یافته و برخی دیگر از متغیرهای مربوطه را به عنوان داده ورودی انتخاب و در مدل HONN استفاده کردند. آنها نشان دادند که مدل‌های ترکیبی بر اساس HONN پیش‌بینی‌های دقیق‌تری نسبت به مدل‌های دیگر ارائه می‌دهند. ژنگ و همکاران (Zeng et al, 2023) به منظور پیش‌بینی فشار گاز طبیعی در خطوط لوله در راستای ایمن‌سازی خطوط لوله‌های خانواده GARCH و ترکیب آنها با مدل‌های یادگیری ماشین استفاده کردند. همچنین آنها داده‌های هاشناسی را با استفاده از الگوریتم CatBoost در مدل ترکیبی خود وارد نمودند. نتایج آنها نشان داد مدل ترکیبی (GARCH-CatBoost)

CNNLSTM) نسبت به تمام مدل‌های دیگر حتی مدل‌های منفرد یادگیری ماشین در معیارهای RMSE، MAE، MAPE برتری دارد. آنها نتیجه‌گیری کردند که بکارگرفتن روش‌های اقتصادسنجی می‌تواند قدرت پیش‌بینی نوسانات در مدل‌های یادگیری ماشین را به شکل محسوسی بهبود بخشد. رایدیگام و مانگالاگر (Rayadurgam & Mangalagiri, 2023) با استفاده از خروجی‌های مدل‌های GARCH بعنوان ورودی در مدل‌های یادگیری عمیق به بررسی میزان بهبود در پیش‌بینی متغیر سرریز تلاطم پرداختند. نتایج نشان داد یک مدل ترکیبی پیش‌بینی‌های بهتری را در مقایسه با مدل‌های یادگیری عمیق ارائه می‌دهد که می‌تواند در تعیین استراتژی‌های سرمایه‌گذاران برای مدیریت پرتفو کاربرد داشته باشد. درحالی‌که مطالعات زیادی در خصوص پیش‌بینی نوسان در بازار داری‌های مختلف وجود دارد، این مطالعه رویکرد جامعی در خصوص طراحی یک مدل ترکیبی برای پیش‌بینی نوسان بیت‌کوین دارد بگونه‌ای که :

- چهار مدل از مدل‌های خانواده GARCH شامل (GARCH، EGARCH، GJR- GARCH، TGARCH) به همراه سه نوع توزیع جملات خطا (نرمال، تی-استیودنت و توزیع خطای تعمیم‌یافته) از میان تمام مرتبه‌های قابل برآورد p و q مورد استفاده قرار گرفته است.
- انواع مختلفی از مدل‌های LSTM تک متغیره و چند متغیره که در هر مرحله طراحی پیچیده‌تری دارند، مورد مطالعه قرار گرفته است.
- مدل سازی نوسان قیمت بیت‌کوین بصورت مدل ترکیبی با استفاده از شاخص‌های مختلف استخراج شده از داده‌های روزانه بیت‌کوین در نظر گرفته شده است.
- مدل‌سازی ترکیبی براساس کلیه نتایج بدست آمده از مدل‌های GARCH بصورت یکجا بعنوان متغیر ورودی و همچنین هر یک از نتایج بصورت جداگانه به همراه نوسان جاری انجام شده است.
- مدل‌های خط مبنا (Baseline Model) و گام تصادفی (Random Walk Model) نیز در کنار دیگر مدل‌ها مورد بررسی قرار گرفته است.

۴. مدل تحقیق و روش برآورد

ارزش دارایی‌های مالی به صورت روزانه تغییر می‌کند و سرمایه‌گذاران برای تعیین مقدار این تغییرات، که اغلب پیش‌بینی آن دشوار است، نیاز به یک شاخص دارند. یکی از شاخص‌های مورد استفاده در این حوزه نوسان است که به سه دسته نوسان ضمنی (Implied Volatility)، نوسان تاریخی (Historical Volatility) و نوسان تحقق یافته (Realized Volatility) تقسیم می‌شود. نوسان ضمنی براساس قیمت‌های مورد انتظار بازار شکل گرفته و نوسانی است که فعالین بازار انتظار دارند در یک دوره مشخص در آینده اتفاق بیفتد. نوسان تحقق یافته براساس تغییرات قیمت در یک دوره زمانی مشخص محاسبه می‌شود و اگر این دوره مشخص مربوط به زمان گذشته باشد، نوسان تاریخی نامیده می‌شود. به منظور محاسبه نوسان از بازدهی روزانه (Daily Return) دارایی‌ها (درصد تغییرات در بازه زمانی روز) استفاده می‌شود که لگاریتم نسبت قیمت در یک دوره زمانی مشخص است. بعنوان مثال اگر P_t و P_{t-1} قیمت یک دارایی در زمان t و $t-1$ (برای دو روز متوالی) باشد، بازدهی روزانه به صورت زیر محاسبه می‌شود.

$$r_t = \ln(P_t/P_{t-1}) \quad (23)$$

نوسان تحقق یافته برای یک پنجره زمانی (Interval Window) مشخص شده، بصورت انحراف معیار بازدهی تعریف می‌شود. برای محاسبه نوسان روزانه براساس یک تواتر زمانی مشخص مانند روزانه، هفتگی، ماهیانه یا سالیانه می‌توان از روابط زیر استفاده نمود.

$$\sigma_{\text{daily}} = \sqrt{\sum_t r_{t-1,t}^2} * \sqrt{\frac{1}{\text{interval}-1}} \quad (24)$$

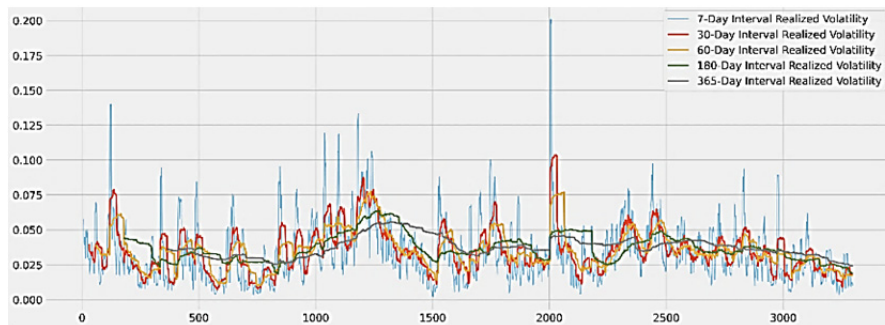
$$\sigma_{\text{annualized}} = \sqrt{\sum_t r_{t-1,t}^2} * \sqrt{\frac{365}{\text{interval}-1}} \quad (25)$$

$$\sigma_{\text{monthly}} = \sqrt{\sum_t r_{t-1,t}^2} * \sqrt{\frac{30}{\text{interval}-1}} \quad (26)$$

$$\sigma_{\text{weekly}} = \sqrt{\sum_t r_{t-1,t}^2} * \sqrt{\frac{7}{\text{interval}-1}} \quad (27)$$

در این مطالعه داده‌های روزانه بیت‌کوین در بازه زمانی ۲۰۱۴/۰۹/۱۷ تا ۲۰۲۳/۰۹/۲۷ شامل قیمت باز شدن، قیمت بسته شدن، بالاترین قیمت، پایین‌ترین قیمت و حجم معاملات از طریق پایگاه اطلاعاتی یاهو فاینانس (Yahoo Finance) جمع‌آوری شده و بازدهی روزانه و

لگاریتم بازدهی روزانه براساس قیمت بسته شدن هر روز مورد محاسبه قرار گرفته است. شکل شماره (۲) نوسان قیمت بیت کوین در بازه‌های زمانی هفتگی، ماهیانه، دوماه، شش ماهه و سالیانه را نشان می‌دهد.



شکل ۲. نوسان تحقق یافته بیت کوین برای پنجره‌های زمانی مختلف

منبع: یافته‌های تحقیق

با توجه به شکل شماره (۲) بازه زمانی ۳۰ روزه برای محاسبه نوسان جاری انتخاب شده است. چراکه به نظر می‌رسد در بازه ۷ روزه مقادیر نوسان اختلال (noise) بیشتری داشته، در حالیکه در بازه زمانی طولانی‌تر هموارتر (Smooth) است و بیشتر به سمت میانگین خود تمایل دارد. لذا با استفاده از بازدهی روزانه و اعمال یک پنجره غلتان ۳۰ روزه، مقدار نوسان جاری محاسبه شده است. هدف اصلی در مدل‌های پیش‌بینی سری زمانی، پیش‌بینی مقادیر آتی براساس مشاهدات گذشته است، لذا مقدار نوسان آتی از طریق یک وقفه زمانی هفتگی و جابجایی زمانی (Shift Forward) برای نوسان جاری در نظر گرفته شده است. پس از انجام محاسبات مذکور، فیله‌های خالی حذف و داده‌ها از نظر زمان مرتب و داده‌های نهایی برای بازه زمانی ۲۰۱۴/۱۰/۱۷ تا ۲۰۲۳/۰۹/۲۰ بکار گرفته می‌شود. خلاصه اطلاعات و ویژگی داده‌های مورد استفاده در جدول شماره (۱) آورده شده است.

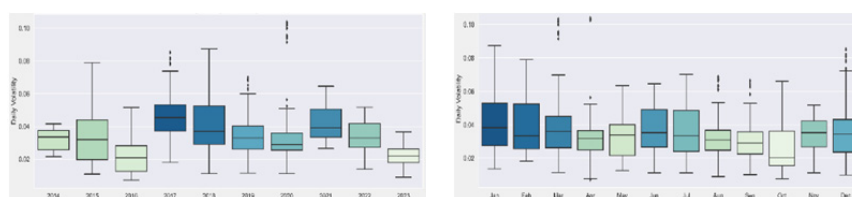
پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۵۷

جدول ۱. خلاصه اطلاعات داده‌های جمع‌آوری شده

نوسان آتی	نوسان جاری	لگاریتم بازدهی	بازدهی	جمع معاملات	قیمت بسته شدن	کمترین قیمت	بیشترین قیمت	قیمت باز شدن	
۳۲۶۱	۳۲۶۱	۳۲۶۱	۳۲۶۱	۳۲۶۱	۳۲۶۱	۳۲۶۱	۳۲۶۱	۳۲۶۱	تعداد
۰/۰۳۴	۰/۰۳۴	۰/۰۰۱	۰/۲۰۱	۱/۶۵۱۰	۱۴۰۶۴	۱۳۶۹۰	۱۴۳۹۲	۱۴۰۵۷	میانگین
۰/۰۱۶	۰/۰۱۶	۰/۰۳۷	۰/۷۳۲	۱/۹۵۱۰	۱۵۹۹۸/۱	۱۵۵۵۲/۵	۱۶۳۹۴/۳	۴۱۶۰۰۱	انحراف معیار
۰/۰۰۷	۰/۰۰۷	-۰/۴۶	-۳۷/۱	۵۹۱۴۵۷۱	۱۷۸/۱۰	۱۷۱/۵۱	۲۱۱/۷۳۱	۱۷۶/۸۹	کمترین
۰/۱۰۳	۰/۱۰۳	۰/۲۲۵	۲/۲۴۷	۳/۵۵۱۱	۶۷۵۶۶/۸	۶۶۳۸۲	۶۸۷۸۹/۶	۶۷۵۴۹	بیشترین

منبع: محاسبات تحقیق

به منظور بررسی ساختار کلی نوسان قیمت بیت‌کوین، نوسان جاری روزانه به صورت گروه‌بندی شده برای بازه زمانی ماهیانه و سالیانه در شکل شماره (۳) نشان داده شده‌است. بیشترین نوسان قیمت بیت‌کوین در ماه‌های ژانویه و دسامبر اتفاق افتاده و در ماه‌های مارچ و آپریل بیشترین میزان داده پرت (Outliers Data) وجود دارد. ماه‌های آپریل، آگوست و سپتامبر کمترین میزان نوسان را داشته‌است. در سال ۲۰۲۰ بیشترین داده پرت اتفاق افتاده و بیشترین نوسانات مربوط به سال‌های ۲۰۱۵ و ۲۰۱۸ بوده‌است. سال‌های ۲۰۱۴ و ۲۰۲۳ کمترین میزان نوسان قیمتی در بیت‌کوین اتفاق افتاده‌است.



شکل ۳. گروه‌بندی ماهیانه (سمت راست) و سالیانه (سمت چپ) نوسان جاری روزانه

منبع: یافته‌های تحقیق

۱.۴ آزمون مانایی

در اولین گام، مانایی سری زمانی بازدهی و لگاریتم بازدهی قیمت بیت کوین مورد بررسی قرار می‌گیرد. نتایج آزمون دیکی فولر (Dickey-Fuller Test) در جدول شماره (۲) نشان می‌دهد هر دو سری زمانی مذکور مانا هستند.

جدول ۲. آزمون مانایی سری زمانی بازدهی روزانه و لگاریتم بازدهی روزانه

مقدار بحرانی	مقدار آماره آزمون دیکی فولر	
۵/۷۵۴۷ e-۳۰	-۱۷/۲۸۴۴	بازدهی قیمت بیت کوین
۰/۰	-۵۸/۲۵۹۵	لگاریتم بازدهی قیمت بیت کوین

منبع: محاسبات تحقیق

۲.۴ تفکیک داده‌ها

با توجه به انجام معاملات رمزارزها (بیت‌کوین در این پژوهش) بصورت ۲۴ ساعت در ۷ روز هفته، تعداد روزهای معاملاتی برابر با ۳۶۵ روز در سال در نظر گرفته شده است. به منظور برآورد مدل و ارزیابی، داده‌ها به سه بخش آموزش، آزمون و پیش‌بینی تقسیم شده است. (جدول شماره ۳) متغیر هدف نوسان آتی است، که در واقع نوسان تحقق‌یافته روزانه برای n دوره زمانی آینده نسبت به روز جاری تعریف می‌شود. لذا اگر $n=7$ و پنجره زمانی برای محاسبه نوسان جاری ۳۰ روز در نظر گرفته شود، پیش‌بینی برای زمان t بصورت پیش‌بینی نوسان تحقق‌یافته برای بازه زمانی $t-23$ تا $t+7$ خواهد بود.

جدول ۳. تفکیک داده‌ها به سه بخش آموزش، ارزیابی و پیش‌بینی

	از	تا	تعداد	درصد از کل داده‌ها
آموزش	۲۰۱۴/۱۰/۱۷	۲۰۲۱/۰۵/۲۳	۲۴۱۱ روز	٪۷۴
ارزیابی	۲۰۲۱/۰۵/۲۴	۲۰۲۳/۰۵/۲۳	۷۳۰ روز	٪۲۲.۴
پیش‌بینی	۲۰۲۳/۰۵/۲۴	۲۰۲۳/۰۹/۲۰	۱۲۰ روز	٪۳.۶

منبع: محاسبات تحقیق

۳.۴ نرمال سازی داده‌ها

در مدل‌سازی با الگوریتم‌های یادگیری ماشین، نرمال‌سازی داده‌ها اهمیت بسزایی دارد. با توجه به اینکه در این پژوهش مدل‌های مختلف، با الگوریتم‌های متفاوت و داده‌های ورودی مختلف مورد استفاده قرار گرفته است، ابتدا کلیه داده‌های جمع‌آوری شده توسط تابع MinMaxScaler نرمال‌سازی می‌شوند. این تابع مقادیر یک بردار را در بازه صفر و یک قرار می‌دهد. خلاصه اطلاعات مربوط به داده‌های بخش آموزش و ارزیابی در حالت اصلی و نرمال شده در جدول شماره (۴) آورده شده است. داده‌های نرمال شده در تمام مدل‌های این پژوهش استفاده شده و نتایج پس از معکوس سازی مورد بررسی و ارزیابی قرار گرفته است.

جدول ۴. خلاصه اطلاعات بخش آموزش و ارزیابی (اصلی و نرمال شده)

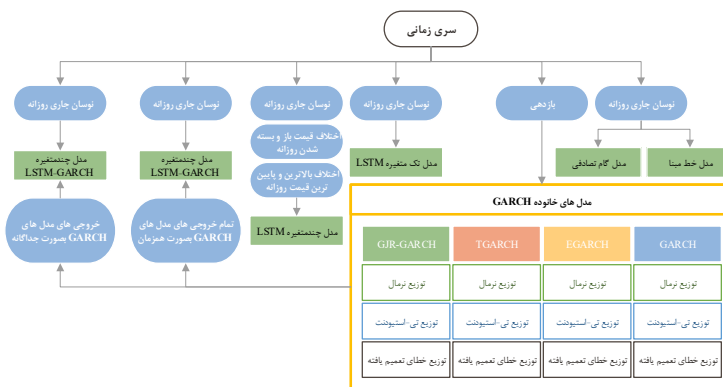
	آموزش		ارزیابی	
	نرمال شده	اصلی	نرمال شده	اصلی
تعداد	۲۸۳۶	۲۸۳۶	۷۳۰	۷۳۰
میانگین	۰/۳۰۰۸۲۹	۰/۰۳۶۳۰۳	۰/۳۰۰۷۱۲	۰/۰۳۶۲۹۲
انحراف معیار	۰/۱۶۹۴۶۱	۰/۰۱۶۳۷۸	۰/۱۶۹۴۸۲	۰/۰۱۶۳۸۰
کمترین	۰/۰	۰/۰۰۷۲۲۸	۰/۰	۰/۰۰۷۲۲۸
بیشترین	۱/۰	۰/۱۰۳۸۷۸	۱/۰	۰/۱۰۳۸۷۸

منبع: محاسبات تحقیق

۵. برآورد مدل

در این بخش با استفاده از داده‌های جمع‌آوری شده ابتدا دو مدل پایه‌ای خط مبنا و گام تصادفی برآورد می‌گردد. سپس از خانواده مدل‌های GARCH، ۱۲ مدل شامل (GARCH استاندارد، EGARCH، TGARCH و GJR-GARCH) با در نظر گرفتن توزیع‌های مختلف برای جمله خطا (توزیع نرمال، توزیع تی‌استیودنت و توزیع خطای تعمیم یافته) برآورد می‌گردد. پس از آن در حوزه شبکه‌های عصبی ۶ مدل شامل (مدل تک‌متغیره LSTM با تنها یک لایه کاملاً متصل، مدل تک‌متغیره LSTM با یک لایه، مدل تک‌متغیره LSTM دوطرفه، مدل تک‌متغیره LSTM دو طرفه با یک لایه کانولوشن و مدل‌های چندمتغیره LSTM) برآورد

می شود. در نهایت خروجی های بدست آمده از کلیه مدل های گارچ (۱۲ مدل برآورد شده) در بهترین مدل LSTM، وارد شده و ۱۳ مدل GARCH-LSTM برآورد می شود. براساس داده های پیش بینی شده برای بازه زمانی بخش ارزیابی و داده های واقعی دوره زمانی مذکور، در هر مرحله معیارهای ارزیابی مدل محاسبه و بهترین مدل از طریق انتخاب کمترین معیار ارزیابی مشخص شده و مدل با کمترین مقدار خطا در معیارهای ارزیابی، بعنوان مدل نهایی در نظر گرفته شده است. در نهایت با استفاده از مدل نهایی، یک پیش بینی برای ۱۲۰ داده انتهایی (داده های بخش پیش بینی) انجام و با مقادیر واقعی مقایسه می گردد. شکل شماره (۴) مدل های طراحی شده در این پژوهش را به همراه متغیرهای مورد استفاده در هر یک نشان می دهد.



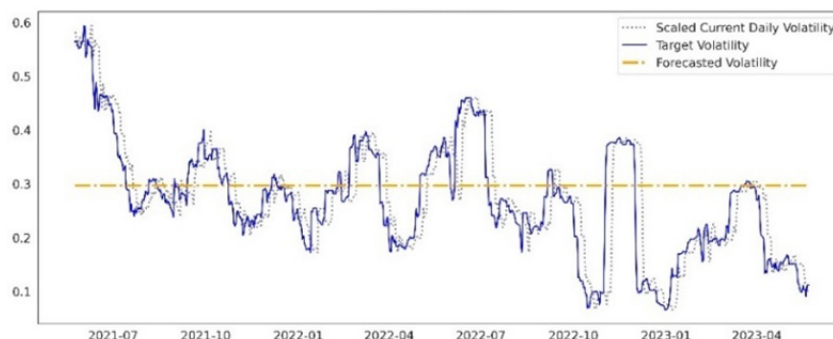
شکل ۴. مدل های مورد بررسی در مطالعه

منبع: یافته های تحقیق

۱.۵ مدل خط مبنا

یکی از ویژگی های اصلی نوسان، بازگشت به مقدار میانگین در بلندمدت است. لذا مدل خط مبنا، براساس مقدار میانگین واریانس تحقق یافته برای دوره آموزش، بعنوان مقدار پیش بینی برای بخش ارزیابی تعریف می شود. شکل شماره (۵) مقدار نوسان جاری پیش بینی شده برای بازه زمانی ارزیابی را نشان می دهد.

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۶۱

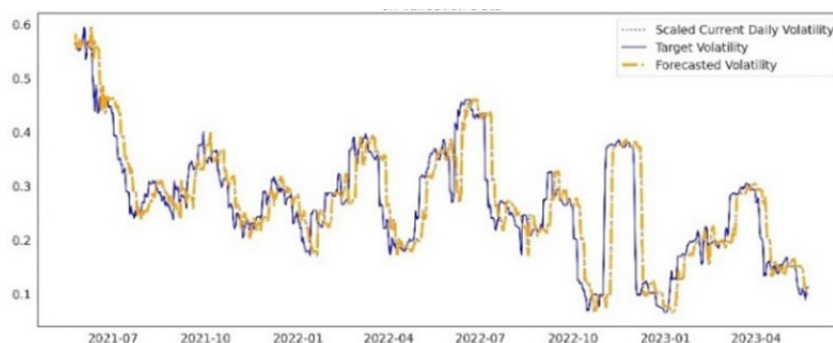


شکل ۵. خروجی مدل خط مبنا

منبع: یافته‌های تحقیق

۲.۵ مدل گام تصادفی

یک واقعیت اصلی در خصوص نوسان، تمایل به خودهمبستگی میان مقادیر این سری زمانی است. با در نظر گرفتن این موضوع، مدل گام تصادفی برای نوسان قیمت بیت‌کوین به صورت پیش‌بینی مقادیر آتی صرفاً براساس مقادیر گذشته برآورد می‌شود. در این حالت با در نظر گرفتن پنجره زمانی ۳۰ روز و پیش‌بینی برای ۷ روز آتی، مقادیر نوسان پیش‌بینی شده برای بازه زمانی داده‌های ارزیابی در شکل شماره (۶) نمایش داده شده است.



شکل ۶. خروجی مدل گام تصادفی

منبع: یافته‌های تحقیق

براساس تمام معیارهای خطای در نظر گرفته شده، مدل گام تصادفی نسبت به مدل خط مبنا پیش‌بینی‌های بهتری را برای بازه زمانی داده‌های ارزیابی ایجاد می‌کند.

جدول ۵. معیارهای ارزیابی مدل خط مبنا و گام تصادفی

نام مدل	RMSPE	RMSE	MAE
مدل خط مبنا	۰/۷۸۶۷۷۸	۰/۱۰۷۶۱۲	۰/۰۸۵۳۰۴
مدل گام تصادفی	۰/۳۲۹۶۳۸*	۰/۰۵۸۰۲۱*	۰/۰۳۷۳۹۴*

منبع: محاسبات تحقیق

۳.۵ مدل‌های خانواده گارچ

در این پژوهش ۴ مدل از خانواده گارچ (GARCH، EGARCH، GJR-GARCH و TARCH) برای پیش‌بینی نوسان رمزارز بیت‌کوین انتخاب شده است. سری‌های زمانی مالی به صورت معمول از حالت نرمال انحراف و چولگی و کشیدگی بیش از حالت عادی را از خود نمایش می‌دهند (Lahmiri & Boukadoum, 2015). همچنین نتایج مطالعه خالدی و همکاران (Khalidi et al, 2019) نشان می‌دهد که مدل‌های EGARCH و TARCH نسبت به نوع توزیع بسیار حساس هستند. بنابراین در این پژوهش سه توزیع شامل، توزیع نرمال برای نشان دادن وضعیت نرمال، توزیع تی استیودنت برای نشان دادن وضعیت چولگی و توزیع خطای تعمیم‌یافته برای نشان دادن کشیدگی سری زمانی بازده بیت‌کوین در نظر گرفته شده است. در اولین گام به منظور انتخاب مرتبه‌های p و q ، با رسم نمودارهای توابع خودهمبستگی و خودهمبستگی جزئی، مقادیر $p=7$ و $q=7$ انتخاب می‌شود. برآورد مدل GARCH(7,7) نشان می‌دهد تنها یکی از ضرایب معنادار است. با توجه به مطالعات بولرسف (Bollerslev, 1987)، هو و همکاران (Hu et al, 2020) که نشان دادند GARCH(1,1) نسبت به مدل‌های با مرتبه‌های بالاتر عملکرد بهتری دارند، در این پژوهش برای مدل‌های GARCH و EGARCH از مرتبه‌های (۱ و ۱) استفاده شده است. به منظور انتخاب بهترین وقفه p و q برای مدل‌های GJR-GARCH و TARCH با استفاده از یک تابع حلقه، مرتبه‌های مختلف در بازه صفر تا سه برای p و q مورد بررسی قرار گرفته و براساس معیارهای بیزین شوارتز (Bayesian Information Criterion/BIC)، حداکثر درست‌نمایی (Maximum Likelihood) و RMSE بهترین مدل انتخاب می‌شود. پیوست شماره (۱) مرتبه‌های p و q را برای مدل‌های GJR-GARCH و TARCH را برای سه نوع توزیع خطای باقیمانده‌ها نشان می‌دهد. نتایج ۱۲ مدل برآورد شده در این مرحله که در جدول شماره (۶) گزارش شده، نشان می‌دهد که براساس

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۶۳

شاخص RMSPE، مدل گارچ نمایی با توزیع خطای تی‌استیودنت و براساس شاخص‌های RMSE و MAE مدل TARCH به ترتیب با توزیع‌های خطای نرمال و تعمیم‌یافته، نسبت به دیگر مدل‌های گارچ با توزیع‌های مختلف برتری دارند.

جدول ۶. معیارهای ارزیابی خطای پیش‌بینی برای مدل‌های GARCH

مدل	RMSPE	RMSE	MAE
GARCH (1,1) توزیع نرمال	۰/۵۲۱۰۸	۰/۱۵۴۰۲۱	۰/۱۳۷۶۹
GARCH (1,1) توزیع تی‌استیودنت	۰/۳۳۳۳۴۱	۰/۰۹۸۷۲۶	۰/۰۸۱۸۵۴
GARCH (1,1) توزیع خطای تعمیم یافته	۰/۳۹۵۰۹۸	۰/۱۱۶۰۹۱	۰/۰۹۹۴
EGARCH (1,1) توزیع نرمال	۰/۳۲۰۵۳۹	۰/۱۰۴۸۸۷	۰/۰۸۳۲۳۵
EGARCH (1,1) توزیع تی‌استیودنت	۰/۲۳۷۵۷۱*	۰/۰۶۹۸۷۹	۰/۰۵۰۹۵۶
EGARCH (1,1) توزیع خطای تعمیم یافته	۰/۲۵۵۸۵۲	۰/۰۷۸۴۸۶	۰/۰۵۸۷۸۳
GJR-GARCH (1,1,0) توزیع نرمال	۰/۵۲۱۱۳۳	۰/۱۵۳۹۱۳	۰/۱۳۷۶۱۷
GJR-GARCH (1,1,2) توزیع تی‌استیودنت	۰/۲۹۳۱۹۶	۰/۰۸۸۹۳۶	۰/۰۶۲۹۴۷
GJR-GARCH (2,1,1) توزیع خطای تعمیم یافته	۰/۳۷۰۰۱۱	۰/۱۰۹۱۱۹	۰/۰۹۱۸۷۱
T-GARCH (1,2,0) توزیع نرمال	۰/۲۷۷۵۷۶	۰/۰۶۳۸۲۳*	۰/۰۴۹۰۹۲
T-GARCH (1,2,1) توزیع تی‌استیودنت	۰/۲۵۰۷۰۵	۰/۰۷۱۷۵۸	۰/۰۵۳۵۸۹
T-GARCH (1,2,1) توزیع خطای تعمیم یافته	۰/۲۴۸۳۰۱	۰/۰۶۴۴۵۷	۰/۰۴۸۲۸۹*

منبع: محاسبات تحقیق

۴.۵ مدل‌های LSTM

طراحی و معماری مدل‌های یادگیری ماشین، نقش مهمی در پیش‌بینی‌های انجام شده توسط این مدل‌ها دارد. به منظور انتخاب بهترین مدل یادگیری ماشین برای پیش‌بینی نوسان رمزارز بیت‌کوین، ابتدا ساده‌ترین مدل طراحی و در هر مرحله پیچیدگی‌ها شامل لایه‌ها و متغیرهای مختلف، به مدل افزوده می‌گردد. جدول شماره (۷) مدل‌های LSTM بکار رفته در این پژوهش را نشان می‌دهد. مدل‌های یادگیری ماشین نتایج متفاوتی را برای یک مجموعه مشخص از آموزش، در هر اجرا نمایش می‌دهند. که این موضوع می‌تواند بعلت مقیاس‌دهی اولیه تصادفی به وزن‌ها یا دیگر عوامل موثر در طراحی مدل باشد. لذا منظور پایداری نتایج،

هر یک مدل‌های LSTM ده مرتبه تکرار و معیارهای خطا بصورت میانگین بدست آمده، گزارش می‌شود. در این پژوهش طراحی تمام مدل‌های LSTM بگونه‌ای است که فرآیند آموزش به محض عدم بهبود شاخص RMSPE قطع می‌گردد.

جدول ۷. مدل‌های LSTM استفاده شده در پژوهش

مدل LSTM (۱)	LSTM - تنها یک لایه کاملاً متصل (متغیر ورودی نوسان جاری)
مدل LSTM (۲)	LSTM تک لایه با ۲۰ واحد (متغیر ورودی نوسان جاری)
مدل LSTM (۳)	مدل تک متغیره LSTM دو طرفه با دو لایه (۶۴ و ۳۲ واحد) (متغیر ورودی نوسان جاری)
مدل LSTM (۴)	مدل تک متغیره LSTM دو طرفه، دو لایه ۶۴ و ۳۲ واحد و یک لایه پیچشی (متغیر ورودی نوسان جاری)
مدل LSTM (۵)	مدل چند متغیره LSTM دو طرفه با سه لایه (متغیرهای ورودی شامل اختلاف قیمت باز و بسته شدن، اختلاف بالاترین و پایین‌ترین قیمت، حجم معاملات به اضافه نوسان جاری)
مدل LSTM (۶)	مدل چند متغیره LSTM دو طرفه با چهار لایه (متغیرهای ورودی شامل اختلاف قیمت باز و بسته شدن، اختلاف بالاترین و پایین‌ترین قیمت، حجم معاملات به اضافه نوسان جاری)
مدل LSTM (۷) تا (۱۸)	مدل چند متغیره LSTM دو طرفه با سه لایه (متغیرهای ورودی شامل خروجی‌های مدل‌های GARCH بصورت جداگانه به اضافه نوسان جاری)
مدل LSTM (۱۹)	مدل چند متغیره LSTM دو طرفه با سه لایه (متغیرهای ورودی شامل تمام خروجی‌های مدل‌های GARCH همزمان به اضافه نوسان جاری)

منبع: محاسبات تحقیق

متغیرهای اختلاف قیمت باز و بسته‌شدن روزانه (CO_t) و اختلاف بالاترین و پایین‌ترین قیمت روزانه (HL_t) از طریق روابط زیر محاسبه می‌شوند.

$$CO_t = \frac{close_t - open_t}{close_t} \quad (28)$$

$$HL_t = \log \left(\frac{high_t - low_t}{close_t} \right) \quad (29)$$

۱.۴.۵ نتایج مدل‌های تک متغیره

در مدل‌های تک متغیره، تنها متغیر نوسان جاری بعنوان ورودی به مدل LSTM وارد و در هر مرحله با پیچیده کردن شبکه از طریق اضافه نمودن انواع لایه‌ها و افزایش واحدها، مقادیر مربوط به نوسان آتی پیش‌بینی شده‌است. نتایج گزارش شده در جدول شماره (۸) نشان

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۶۵

می‌دهد کلیه معیارهای ارزیابی خطا، از مدل شماره (۱) تا مدل شماره (۳) کاهش داشته است. اضافه کردن یک لایه پیچشی به مدل شماره (۳) و طراحی مدل شماره (۴)، تنها باعث کاهش محسوس در معیار RMSPE شده و معیارهای MAE و RMSE را افزایش داده است. لذا اضافه کردن لایه پیچشی، باعث بهبود مدل نشده و در مراحل بعدی از این لایه صرف‌نظر شده است. با این وجود براساس کلیه معیارهای خطا، مدل شماره (۴) از کلیه مدل‌های GARCH عملکرد بهتری در پیش‌بینی نوسان بیت‌کوین دارد.

جدول ۸. معیارهای ارزیابی مدل‌های LSTM تک متغیره

مدل	RMSPE	RMSE	MAE
مدل شماره ۱	۰/۳۲۴۲۷۷	۰/۰۵۶۰۹۷	۰/۰۳۷۳۱۴
مدل شماره ۲	۰/۳۰۸۵۰۷	۰/۰۵۴۴۲۲	۰/۰۳۵۹۷۷
مدل شماره ۳	۰/۲۷۶۷۴۵	۰/۰۵۱۶۸۲*	۰/۰۳۵۶۵۵*
مدل شماره ۴	۰/۲۲۶۱۱۵*	۰/۰۵۴۲۰۸	۰/۰۳۷۱۵۲

منبع: محاسبات تحقیق

۲.۴.۵ نتایج مدل چند متغیره با متغیرهای کمکی

در اولین مدل چندمتغیره، متغیر نوسان جاری به همراه متغیرهای کمکی (CO_t) ، (HL_t) و $(Volume_t)$ بعنوان ورودی به مدل وارد و مقادیر نوسان آتی برای بازه زمانی ارزیابی، پیش‌بینی شده است. ورود متغیرها مذکور بعنوان متغیر کمکی به مدل باعث بهبود قدرت پیش‌بینی مدل در تمام معیارهای خطا، نسبت به کلیه مدل‌های خانواده GARCH شده است. نتایج گزارش شده در جدول شماره (۹) نشان می‌دهد، اضافه کردن لایه چهارم در مدل شماره (۶) نسبت به مدل شماره (۵)، علاوه بر عدم ایجاد بهبود قابل توجه در معیار RMSPE، باعث افزایش معیارهای MAE و RMSE گردیده است. لذا مدل LSTM چند متغیره با سه لایه، نسبت به ۴ لایه عملکرد بهتری را ارائه می‌کند.

جدول ۹. معیارهای ارزیابی مدل‌های LSTM چند متغیره با متغیرهای کمکی

مدل	RMSPE	RMSE	MAE
مدل شماره ۵	۰/۲۱۷۲۳۳	۰/۰۵۱۶۹۷*	۰/۰۳۷۶۱۴*
مدل شماره ۶	۰/۲۱۳۸۷۲*	۰/۰۵۳۶۰۷	۰/۰۳۷۷۱۹

منبع: محاسبات تحقیق

بررسی دو مدل چند متغیره در این بخش، نسبت به مدل‌های LSTM تک متغیره نشان می‌دهد در مجموع (از نظر شاخص‌های RMSE و MAE) بهبود محسوسی در عملکرد مدل، اتفاق نیفتاده و استفاده از متغیرهای کمکی صرفاً باعث بهبود عملکرد مدل‌های LSTM نسبت به مدل‌های خانواده GARCH شده است.

۳.۴.۵ مدل چند متغیره ترکیبی با مدل‌های گارچ

در مدل‌های شماره (۷) تا (۱۸) خروجی هر یک از مدل‌های GARCH به‌همراه متغیر نوسان جاری بعنوان ورودی مدل LSTM در نظر گرفته شده و به صورت یک مدل ترکیبی، مقدار نوسان آتی برای بازه زمانی ارزیابی پیش‌بینی گردیده است. نتایج معیارهای ارزیابی خطای پیش‌بینی در جدول شماره (۱۰) آورده شده است. براساس نتایج جدول شماره (۱۰)، ترکیب هر یک از خروجی‌های مدل‌های GARCH با LSTM، از هر یک از نتایج بدست آمده در مراحل قبل، عملکرد بهتری در پیش‌بینی نوسان قیمت بیت‌کوین دارد. لذا مدل‌های ترکیبی GARCH-LSTM هم نسبت به مدل‌های GARCH و هم مدل‌های LSTM تک‌متغیره و LSTM چندمتغیره با استفاده از متغیرهای کمکی، برتری قابل توجهی دارند. همچنین با توجه به معیار RMSPE، مدل ترکیبی GARCH-LSTM با توزیع خطای نرمال و از نظر معیارهای خطای RMSE و MAE مدل EGARCH-LSTM، با توزیع‌های خطای به ترتیب تعمیم‌یافته و تی‌استیودنت، بهترین عملکرد در پیش‌بینی نوسان بیت‌کوین را دارند.

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۶۷

جدول ۱۰. معیارهای ارزیابی مدل‌های ترکیبی LSTM و خانواده GARCH

مدل	RMSPE	RMSE	MAE
LSTM-(توزیع نرمال)GARCH: مدل شماره (۷)	۰/۱۸۳۷*	۰/۰۴۷۷	۰/۰۳۲۲
LSTM-(توزیع تی‌استیودنت)GARCH: مدل شماره (۸)	۰/۱۸۵۲	۰/۰۴۶۵	۰/۰۳۲۱
LSTM-(توزیع خطای تعمیم یافته)GARCH: مدل شماره (۹)	۰/۱۹۰۱	۰/۰۴۶۵	۰/۰۳۱۳
LSTM-(توزیع نرمال)EGARCH: مدل شماره (۱۰)	۰/۱۸۹۳	۰/۰۴۶۰	۰/۰۳۱۷
LSTM-(توزیع تی‌استیودنت)EGARCH: مدل شماره (۱۱)	۰/۱۸۴۴	۰/۰۴۶۴	۰/۰۳۱*
LSTM-(توزیع خطای تعمیم یافته)EGARCH: مدل شماره (۱۲)	۰/۱۸۹۳	۰/۰۴۵*	۰/۰۳۱۵
LSTM-(توزیع نرمال)GJR-GARCH: مدل شماره (۱۳)	۰/۱۸۵۷	۰/۰۴۶۷	۰/۰۳۲۲
LSTM-(توزیع تی‌استیودنت)GJR-GARCH: مدل شماره (۱۴)	۰/۱۹۲۳	۰/۰۴۷۲	۰/۰۳۱۹
LSTM-(توزیع خطای تعمیم یافته)GJR-GARCH: مدل شماره (۱۵)	۰/۱۸۶۴	۰/۰۴۶۸	۰/۰۳۲۰
LSTM-(توزیع نرمال)TARCH: مدل شماره (۱۶)	۰/۱۸۹۷	۰/۰۴۶۶	۰/۰۳۱
LSTM-(توزیع تی‌استیودنت)TARCH: مدل شماره (۱۷)	۰/۱۹۶۵	۰/۰۴۸۳	۰/۰۳۳۱
LSTM-(توزیع خطای تعمیم یافته)TARCH: مدل شماره (۱۸)	۰/۱۹۳۹	۰/۰۴۷۲	۰/۰۳۲۸

منبع: محاسبات تحقیق

در مرحله بعدی کلیه خروجی‌های مدل‌های GARCH به صورت همزمان، به‌همراه متغیر نوسان جاری بعنوان ورودی به مدل LSTM وارد و نوسان آتی برای بازه زمانی ارزیابی پیش‌بینی می‌شود. نتایج جدول شماره (۱۱) نشان می‌دهد در هیچ یک از معیارهای خطا، مدل بدست آمده نسبت به بهترین مدل در مرحله قبل برتری ندارد.

جدول ۱۱. نتیجه برآورد مدل LSTM با کلیه خروجی‌های GARCH

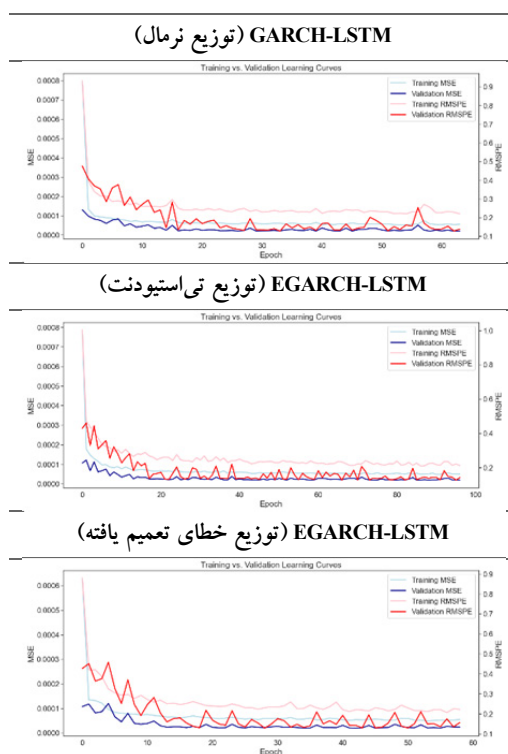
مدل	RMSPE	RMSE	MAE
LSTM-(خروجی کلیه مدل‌های)GARCH: مدل شماره (۱۹)	۰/۱۸۶۹۵	۰/۰۴۸۴۴	۰/۰۳۳۰۶

منبع: محاسبات تحقیق

۶. مدل نهایی

با توجه به نتایج بدست آمده در جدول شماره (۱۰) سه مدل GARCH-LSTM (با توزیع خطا نرمال)، EGARCH-LSTM (با توزیع‌های خطای تعمیم‌یافته و توزیع تی‌استیودنت) از

نظر معیارهای خطا، بعنوان مدل‌های بهینه معرفی شده‌است. لذا در این مرحله سه مدل مذکور، براساس داده‌های آموزش و ارزیابی برآورد و مقادیر نوسان آتی برای بازه زمانی پیش‌بینی (۱۲۰ داده انتهایی) استخراج شده و معیارهای خطا برای این بازه زمانی محاسبه می‌شود. شکل شماره (۷) تغییرات معیارهای RMSE و RMSE در فرآیند مدل‌سازی LSTM را نشان می‌دهد. کاهش معیارهای خطا در بازه ارزیابی نسبت به بازه آموزش، بهینه‌بودن فرآیند یادگیری و برآورد مدل را نشان می‌دهد.



شکل ۷. تغییرات معیارهای خطا برای مدل‌های منتخب

منبع: یافته‌های تحقیق

نتایج معیارهای خطا برای پیش‌بینی بخش انتهایی داده‌ها براساس سه مدل منتخب که در جدول شماره (۱۲) گزارش شده، نشان می‌دهد در هر سه معیار مدل EGARCH-LSTM با توزیع خطای تی استیودنت، برتری دارد.

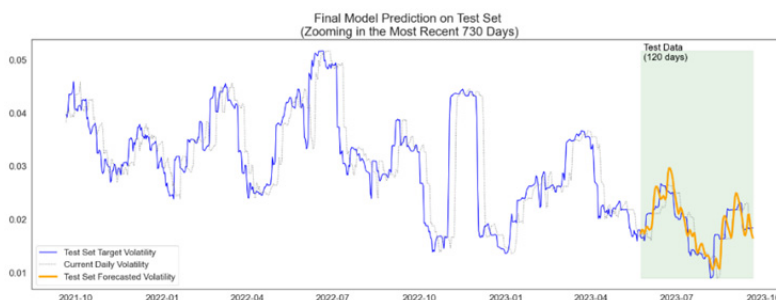
پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۶۹

جدول ۱۲. نتایج سه مدل ترکیبی GARCH-LSTM نهایی

مدل	RMSPE	RMSE	MAE
GARCH - LSTM (توزیع نرمال)	۰.۱۸۳۷۶	۰.۰۰۲۹۳	۰.۰۰۲۶۰
EGARCH - LSTM (توزیع خطای تعمیم یافته)	۰.۱۷۳۸۷	۰.۰۰۲۹۹	۰.۰۰۲۵۶
EGARCH - LSTM (توزیع تی استیودنت)	۰.۱۵۰۹۱*	۰.۰۰۲۷۰*	۰.۰۰۲۲۴*

منبع: محاسبات تحقیق

شکل شماره (۸) پیش‌بینی مدل نهایی برای بازه زمانی مذکور را نشان می‌دهد.



شکل ۸. پیش‌بینی مدل EGARCH-LSTM (توزیع تی استیودنت)

منبع: یافته‌های تحقیق

۷. نتیجه‌گیری

در این پژوهش مدل‌سازی نوسان روزانه بیت‌کوین با استفاده از مدل‌های ترکیبی خانواده گارچ و حافظه کوتاه مدت طولانی مورد بررسی قرار گرفت. بررسی مجموعه کاملی از مدل‌های منفرد در هر یک از حوزه‌های ذکر شده و ترکیب آن‌ها نشان داد طراحی یک مدل ترکیبی GARCH-LSTM باعث بهبود نتایج پیش‌بینی نسبت به هر یک از مدل‌های خانواده GARCH و مدل‌های تک متغیره LSTM خواهد شد. با توجه به معرفی رویکرد مدل‌سازی ترکیبی نتایج کلی این پژوهش به صورت زیر دسته بندی شده است.

- استفاده از خانواده مدل‌های GARCH و توزیع‌های مختلف خطا می‌تواند پیش‌بینی‌های مناسبی برای نوسان ارائه نماید. در این پژوهش از نظر معیار RMSPE،

- مدل EGARCH (1,1) با توزیع خطای تی-استیودنت، بهترین عملکرد را در میان ۱۲ مدل برآورد شد از خانواده GARCH داشت.
- استفاده از مدل LSTM تک متغیره، در صورت طراحی مناسب لایه‌ها و ویژگی‌ها، نسبت به تمام مدل‌های GARCH از نظر پیش‌بینی نوسان و معیارهای خطا برتری محسوس دارد.
 - افزایش پیچیدگی‌های مدل‌سازی LSTM مانند افزایش لایه‌ها، واحدهای هر لایه و ... لزوماً باعث بهبود عملکرد مدل از نظر معیارهای خطا، در پیش‌بینی مقادیر آتی نخواهد شد. لذا طراحی بهینه در مدل‌سازی یادگیری ماشین بسیار بااهمیت است.
 - استفاده از ویژگی‌های یک متغیر بعنوان متغیرهای کمکی احتمالاً باعث بهبود نتایج مدل‌سازی LSTM نسبت به مدل‌های GARCH خواهد شد، که این موضوع می‌تواند بعلت اطلاعات اضافه در مدل باشد اما این متغیرها که موجب پیچیده‌تر شدن مدل LSTM می‌شوند، همواره نتایج یک مدل ساده LSTM را بهبود نخواهند داد.
 - مدل ترکیبی GARCH-LSTM که از خروجی‌های GARCH بعنوان ورودی مدل LSTM استفاده می‌کند، نه تنها نسبت به کلیه مدل‌های GARCH، بلکه نسبت به مدل‌های LSTM ساده تک متغیره و حتی مدل‌های چندمتغیره با متغیرهای کمکی ناشی از ویژگی‌های اضافی، عملکرد بهتری در پیش‌بینی نوسان دارد.
 - استفاده از خروجی مدل‌های GARCH به صورت جداگانه و بعنوان ورودی مدل LSTM چندمتغیره، نتایج بهتری نسبت به استفاده همزمان از این خروجی‌ها در مدل LSTM چندمتغیره دارد.
 - در این پژوهش مدل LSTM - EGARCH با توزیع خطای تی-استیودنت بهترین عملکرد از نظر پیش‌بینی نوسان قیمت بیت‌کوین را نشان داد.
 - در این مطالعه براساس بررسی‌های انجام شده که در متن پژوهش آورده شده، پنجره زمانی یک ماهه (۳۰ روز) برای محاسبه نوسان جاری در نظر گرفته شده است. پیشنهاد می‌شود مدل‌سازی انجام شده در پنجره‌های زمانی مختلف اجرا و نتایج مورد مقایسه قرار گیرد.

پیش‌بینی نوسان با استفاده از ترکیب ... (سعید صدرزاده مقدم و دیگران) ۱۷۱

- در پایان بعنوان تحقیقات آتی، طراحی مدل MIDAS-LSTM و همچنین مدل‌های GARCH-MIDAS-LSTM برای بررسی و ارزیابی دقت پیش‌بینی مقادیر آتی پیشنهاد می‌گردد.

پیوست ۱

تعیین مرتبه‌های p و q مدل‌های GJR-GARCH و TARCH

	توزیع نرمال TARCH(1, 2, 0)		توزیع نرمال GARCH (1,1,0)
	توزیع تی استیودنت TARCH(1, 2, 1)		توزیع تی استیودنت Gjr-GARCH(1,1,2)
	توزیع خطای تعمیم یافته TARCH(1, 2, 1)		توزیع خطای تعمیم یافته GARCH (1, 1, 1)

کتابنامه

- Assaf, A. M., Haron, H., Abdull Hamed, H. N., Ghaleb, F. A., Qasem, S. N., & Albarrak, A. M. (2023). A Review on Neural Network Based Models for Short Term Solar Irradiance Forecasting. *Applied Sciences*, 13(14), 8332.
- Banik, S., Sharma, N., Mangla, M., Mohanty, S. N., & Shitharth, S. (2022). LSTM-based decision support system for swing trading in the stock market. *Knowledge-Based Systems*, 239, 107994.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2017). GARCH modeling of cryptocurrencies. *Journal of Risk and Financial Management*, 10(4), 17.
- Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4), 157–168.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339–350.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669.
- García-Medina, A., & Luu Duc Huynh, T. (2021). What drives Bitcoin? An approach from continuous local transfer entropy and deep learning classification models. *Entropy*, 23(12), 1582.
- Hajizadeh, E., Seifi, A., Fazel Zarandi, M. H., & Turksen, I. B. (2012). A hybrid modeling approach for forecasting the volatility of the S&P 500 index return. *Expert Systems with Applications*, 39(1), 431–436.
- Khalidi, R., el Afia, A., & Chiheb, R. (2019). Forecasting BTC volatility: A comparative study between parametric and nonparametric models. *Progress in Artificial Intelligence*, 8(4), 511–523.
- Kim, 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, 25–37.
- Kristjanpoller, W., Fadic, A., Minutolo, M. C. (2014). Volatility forecast using hybrid neural network models. *Expert Systems with Applications*, 41(5), 2437–2442.
- Kristjanpoller, W., Minutolo, M. C. (2016). Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Systems with Applications*, 65, 233–241.
- Kristjanpoller, W., Minutolo, M. C. (2018). A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis, and principal components analysis. *Expert Systems with Applications*, 109, 1–11.
- Lamoureux, C. G., & Lastrapes, W. D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *The Journal of Finance*, 45(1), 221–229.

- Liu, Y. (2019). Novel volatility forecasting using deep learning–Long ShortTerm Memory Recurrent Neural Networks. *Expert Systems with Applications*, 132, 99–109.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PloS One*, 13(3), 0194889.
- Peng, Y., Albuquerque, P. H. M., Camboim de Sá, J. M., Padula, A. J. A., & Montenegro, M. R. (2018). The best of two worlds: Forecasting high-frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. *Expert Systems with Applications*, 97, 177–192.
- Quintero Valencia, D. E., et al. (2019). Pronóstico de volatilidad de la TRM mediante un modelo híbrido LSTM–GARCH. PhD thesis, Universidad del Rosario.
- Ramos-Pérez, E., Alonso-González, P. J., & Núñez Velázquez, J. J. (2019). Forecasting volatility with a stacked model based on a hybridized Artificial Neural Network. *Expert Systems with Applications*, 129, 1–9.
- Rayadurgam, Vikram Chandramouli, & Mangalagiri, Jayasree. (2023). Does inclusion of GARCH variance in deep learning models improve financial contagion prediction? *Finance Research Letters*, 54.
- Seo, M., & Kim, G. (2020). Hybrid forecasting models based on neural networks for the volatility of Bitcoin. *Applied Sciences* (Switzerland), 10(14), 00.
- Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, 157(00).
- Vijayalakshmi, V. (2024). Implementation of sentiment analysis in stock market prediction using variants of GARCH models. In *Computational Intelligence Methods for Sentiment Analysis in Natural Language Processing Applications* (pp. 227–249). Morgan Kaufmann.
- Zeng, H., Shao, B., Dai, H., Yan, Y., & Tian, N. (2023). Prediction of fluctuation loads based on GARCH family-CatBoost-CNNLSTM. *Energy*, 126125.
- Zolfaghari, M., Sahabi, B., & Bakhtyaran, M. (2020). Designing a model for forecasting the stock exchange total index returns (Emphasizing on combined deep learning network models and GARCH family models). *Financial Engineering and Securities Management* (Portfolio Management), 11(42), 138–171.