Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2024, 2024:116
https://doi.org/10.28919/cmbn/8886
ISSN: 2052-2541

A COMPARATIVE ANALYSIS OF MULTIVARIATE GARCH AND CNN-BILSTM MODELS FOR FORECASTING CONDITIONAL VOLATILITY IN FINANCIAL MARKETS

HAILABE T. ARAYA[1,2,*], JANE ADUDA[3], TESFAHUN BERHANE[4]

[1]Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation, Nairobi 62000, Kenya
[2]Department of Mathematics, Debre Markos University, Debre Markos 269, Ethiopia
[3]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi 62000, Kenya
[4]Department of Mathematics, Bahir Dar University, Bahir Dar 26, Ethiopia

Abstract. In recent years, there has been significant interest in forecasting volatilities within multivariate frameworks for financial assets. Previous research has utilized the VAR-DCC-GARCH model to explore these relationships, offering valuable insights into market dynamics. This paper presents a novel VAR-CNN-BiLSTM model to forecast the conditional correlation between BTC-USD exchange rates and gold prices. The study aimed to improve the accuracy of volatility forecasting for financial assets by introducing this hybrid approach. The hybrid VAR-CNN-BiLSTM model employed the VAR model to capture the linear features and the deep learning network structure that combines the CNN, to capture the hierarchical data structure and BiLSTM layers to capture the long-term dependencies in the data. Results have confirmed that the VAR-CNN-BiLSTM model can achieve better prediction accuracy than the hybrid VAR-DCC GARCH model, in terms of Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) performance measures. The results of this study further indicate a unidirectional causality from the BTC-USD exchange rate to Gold prices. The findings provide valuable insights for traders, financial analysts, and policymakers aiming to understand and anticipate market behaviors involving cryptocurrencies and traditional assets.

*Corresponding author
E-mail address: hayle899@gmail.com
Received September 04, 2024

Keywords: bidirectional long short term memory; convolutional neural network; correlation forecasting; vector auto regression; time series; volatility spillover.

2020 AMS Subject Classification: 68T07.

1. INTRODUCTION

The volatility of stock prices has long been a subject of research interest in financial markets, as it significantly affects investors, traders, and financial institutions. To address this challenge, researchers are employing various statistical and deep learning methods, which have proven to enhance the accuracy of volatility predictions. [1] investigated a novel method of predicting volatility in a univariate setting by utilizing a hybrid model that blends deep learning and GARCH methods. This study demonstrated that by feeding the deep learning models with the residuals from the GARCH model, the new approach outperformed earlier models in terms of predicted accuracy. The prediction of asset return series co-volatility has garnered substantial interest from scholars, professionals, and portfolio managers [2]. Bitcoin and gold are two major financial assets. Bitcoin is a digital asset, whereas gold is a fundamental commodity. The volatilities in these assets are a key concern for policymakers and investors. A substantial rise in commodity and digital exchange prices negatively affects domestic and global economies by increasing inflation and decreasing economic activity.

Since Bitcoin (BTC) has such high volatility and the potential for large profits, it has attracted a lot of attention from investors and researchers in the past several years. At the same time, gold, which is seen as a safe-haven asset due to its stability, has continued to be a mainstay of the financial system. Determining the dynamics and conditional connection between these two assets is essential for managing a portfolio, evaluating risk, and making smart trading choices. The risk level of a portfolio is affected by the simultaneous movements of individual assets across various markets and the inherent risk of assets within a single market. p Analyzing the volatility spillovers between asset and commodity prices has important ramifications for risk management and portfolio optimization on the part of investors and governments [3].

[4] utilized the Baba, Engle, Kraft, and Kroner(BEKK) GARCH model to examine the transmission of shocks and volatility among Bitcoin, Ethereum, and Litecoin. The study found evidence of bidirectional shock spillovers between the Bitcoin-Ethereum and Bitcoin-Litecoin

pairs. [5] employed a rolling window bootstrap method to investigate the causal link between gold and silver returns in the Chinese market. Their findings revealed that across various subperiods, gold exerted both positive and negative influences on silver. [6] investigated the return and volatility transmission among Bitcoin, Ethereum, and Lithecoin during the pre-COVID-19 and COVID-19 periods using the Vector Autoregression-Dynamic Conditional Correlation-Generalized Autoregressive Conditional Heteroskedasticity (VAR-DCC-GARCH) model. It was discovered that the volatility spilovers are bidirectional between Etheruem and Litecoin and unidirectional between Bitcoin and Etheruem. The Dynamic Conditional Correlation-Generalized Autoregressive Conditional Heteroskedasticity (DCC-GARCH) model has two key improvements over earlier models, which are why it has been used in many contemporary studies on stock markets. First off, the DCC model can estimate large correlation coefficient matrices and has obvious computational benefits. Second, the DCC model makes it simple to estimate the parameters required for the correlation process by using a two-step estimation procedure [7, 8].

A VAR model and three Multivariate GARCH models (CCC-GARCH, BEKK-GARCH, and DCC-GARCH) to study the volatility spillover between bitcoin, gold, and crude oil returns [9]. The study also investigated that the data better fits the model and there was a bidirectional spillover between the returns of gold and crude oil and further showed that the DCC-GARCH model provides a better fit than the CCC-GARCH (Constant Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity) model and the BEKK-GARCH model. [10] studied the return and volatility spillovers between the Ghanaian and Nigerian equities markets.The return series for the two markets were analyzed using VARMA-AMGARCH (Vector Autoregressive Moving Average-Asymmetric Multivariate Generalized Autoregressive Conditional Heteroskedasticity). The two stock market return and volatility spillovers showed a strong cross-transmission, according to the authors. Nevertheless, it was also observed that the volatility of the Ghanaian stock market was more susceptible to that of the Nigerian stock market.

The impact of volatility spillover on the returns of the East African securities markets in Nairobi, Uganda, and Dar el-Sahara was investigated [11]. The Vector Autoregressive model was utilized in the study to model the evolution of return series. Furthermore, possible co-integration was examined using the Johansen co-integration test. To investigate the dynamics of conditional variances, a further dynamic conditional correlation model was used. The results of the study showed that there is a causal relationship that is bidirectional between the Nairobi Securities Exchange and the Dar es Salaam Securities Exchange. The fact that correlations between stock returns change over time is widely known. [12, 13] have observed that there is a tendency for correlations among market returns to decrease during bull markets and increase during bear markets. Furthermore, it is now well acknowledged that there is a substantial increase in correlation between foreign stock markets at times of market volatility, or stock market crises . [14] evaluated the volatility interconnectedness between oil and coffee markets using multivariate GARCH models, focusing on covolatility forecasting with high-frequency data. The study concluded that the varying conditional correlation (VCC) model with Student's t-distributed innovation terms provides the most accurate forecasts. [15] studied the hybrid DCC-GARCH models with Deep Learning models to check the forecasting performance of correlations. improves the latter's DCC-GARCH models. These studies have concentrated on the relationships between the stock markets in China, Hong Kong, and the US. The results imply that the DCC-GARCH models' ability to predict market dynamics is much enhanced by the use of deep learning techniques. [15] investigated the integration of hybrid DCC-GARCH models with Deep Learning techniques to improve the forecasting performance of DCC-GARCH model. Their study demonstrates that the incorporation of Deep Learning methods enhances the performance of traditional DCC-GARCH models.

[16] introduced the VAR-CNN-Long Short Term Memory (VACL) model to address the challenge of forecasting volatility. However, there remains a research gap in extending these findings to include additional forecasting of volatility and conditional correlations, particularly in comparison to the DCC-GARCH model. This paper aimed to fill this gap and do a comparative analysis of two advanced forecasting models: the Vector Autoregressive - Convolutional Neural Network - Bidirectional Long Short-Term Memory (VAR-CNN-BiLSTM) and the Vector Autoregressive Dynamic Conditional Correlation Generalized Autoregressive Conditional

Heteroskedasticity (VAR-DCC-GARCH) model. The primary objective is to predict the conditional correlation and volatility of BTC-USD exchange rates and gold prices. The VAR-CNN-BiLSTM model integrates traditional time-series analysis with deep learning techniques, leveraging the strengths of CNN for feature extraction and BiLSTM for capturing long-term dependencies in the data. In contrast, the VAR-DCC-GARCH model employs a well-established econometric approach to model time-varying correlations and volatilities, making it a robust benchmark for comparison.

By analyzing the predictive performance of these models, this study aimed to provide insights into their effectiveness and applicability in financial markets, contributing to the broader field of econometric and machine learning-based forecasting methods. The results of this research have significant implications for traders, financial analysts, and policymakers who seek to understand and anticipate market behaviors involving cryptocurrencies and traditional assets like gold.

2. Methodology

The study used VAR , CNN and BiLSTM as components to create a new hybrid models and compare with VAR- DCC-GARCH. Each of these components has unique traits that may be extracted from historical data.

2.1. VAR model.

[17] popularized the Vector Autoregression (VAR) model, which is a prominent econometric technique for examining the dynamic interactions between several time series variables. According to [17], all variables within a VAR system are considered endogenous. The rationale behind constructing a VAR model is based on the assumption that all variables being investigated are endogenous, with typically none being treated as exogenous [18]. Vector auto-regression is a useful technique when two or more time series interact with one another. This model is autoregressive, meaning that each variable is defined as a function of the variables' historical values [19]. The study stated that the VAR Model cannot be applied unless the time series data is stationary [16].

Mathematical model for VAR model. Let $Y_t = (Y_{1,t}, Y_{2,t}, \ldots, Y_{n,t})^\top$ represent an $n \times 1$ vector comprising time series variables. The fundamental vector autoregressive model of order p is denoted as VAR(p), is

$$Y_t = c + \sum_{i=1}^{p} \pi_i Y_{t-i} + \varepsilon_t$$
$$\varepsilon_{t|F_{t-1}} \sim N(0, H_t) \tag{1}$$

In this context, π_i refers to $n \times n$ coefficient matrices, c represents an $n \times 1$ vector of constants, and ε_t denotes an $n \times 1$ unobservable white noise vector process with a mean of zero. H_t is a positive definite covariance matrix, and F_{t-1} is a set of past information. Determining the lag length of the VAR model involves finding the value of p that minimizes specific model selection criteria by fitting VAR(p) models of different orders, such as $p = 0, \ldots, p_{max}$. The Akaike Information Criterion (AIC) was used in this study to identify the optimal lag length (p) for the model.

2.2. DCC-GARCH model. Once the VAR (p) model has been estimated, the residuals have been gathered for subsequent DCC- GARCH modeling. The Dynamic Conditional Correlation (DCC) model, developed by [20] and [21], builds upon [22] Constant Conditional Correlation (CCC) model by introducing time-varying correlation matrices that are parameterized using a limited number of variables. This model represented one of the most adaptable iterations of MGARCH models. The conditional variance-covariance matrix H_t is determined based on the information available up to Ω_t and is broken down using the Cholesky method in the following manner:

$$H_t = \begin{pmatrix} h_{11,t}^2 & \cdots & h_{1n,t} \\ h_{12,t} & \cdots & h_{2n,t} \\ \vdots & \ddots & \vdots \\ h_{1n,t} & \cdots & h_{nn,t}^2 \end{pmatrix} = \begin{pmatrix} \sqrt{h_{11,t}^2} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{22,t}^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & & \sqrt{h_{nn,t}^2} \end{pmatrix} \begin{pmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1n,t} \\ \rho_{12,t} & 1 & \cdots & \rho_{2n,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n,t} & \cdots & & 1 \end{pmatrix}$$
$$\begin{pmatrix} \sqrt{h_{11,t}^2} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{22,t}^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & & \sqrt{h_{nn,t}^2} \end{pmatrix} \tag{2}$$

Furthermore, Eq. (2) can be expressed as $H_t = D_t R_t D_t$

Hence, H_t is the conditional covariance matrix with elements $[H_t]_{ij} = \sqrt{h_{iit} h_{jjt}}\,\rho_{ij}$, where $\rho_{ii} =$

1. D_t represents a $k \times k$ diagonal matrix having conditional variance $\sqrt{h_{iit}}$ on its diagonal, while R_t signifies a correlation matrix that varies with time. Every individual element h_{iit} can be represented using a univariate GARCH model in the following manner:

$$h_{iit} = \alpha_{i0} + \sum_{i=1}^{p} \alpha_{ip} \varepsilon_{i,t-p}^2 + \sum_{j=1}^{q} \beta_{jq} h_{j,t-q}^2, \quad \text{for } i = 1, 2, \ldots, m \tag{3}$$

Where, $a_{i0} > 0$, $\alpha_{ii} > 0$, and $\beta_{ii} > 0$ are non-negative, and $\sum_{q=1}^{q_i} \alpha_{ii,p} + \sum_{p=1}^{p_i} \beta_{ii,q} < 1$. According to [23], the selection of a lag order of (1,1) within the GARCH family model is deemed sufficient to encapsulate all observed volatility clustering present in the data. This assertion underscores the effectiveness of the GARCH(1,1) model in capturing the essential characteristics of financial data, particularly in terms of volatility dynamics. When constructing the conditional correlation matrix R_t, it is necessary to confirm that the covariance matrix H_t is positive definite, and that all elements of R_t are less than or equal to one. In order to determine whether these requirements are met, [24, 25] recommend the following method:

$$\mathbf{R}_t = \mathbf{Q}_t^{*^{-1}} \mathbf{Q}_t \mathbf{Q}_t^{*^{-1}} \tag{4}$$

$$\mathbf{Q}_t^* = \begin{bmatrix} q_{11} & 0 & 0 & \cdots & 0 \\ 0 & q_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & q_{mm} \end{bmatrix} \tag{5}$$

where Q_t^* is a diagonal matrix consisting of the diagonal elements specified in Eq. (5). A symmetric positive definite conditional covariance matrix, $Q_t = (q_{ij,t})$ can be expressed as follows:

$$\mathbf{Q}_t = (1 - \alpha - \beta)\bar{Q} + \alpha \varepsilon_{t-1} \varepsilon_{t-1}' + \beta \mathbf{Q}_{t-1} \tag{6}$$

$\bar{\mathbf{Q}} = \text{cov}(\varepsilon_t \varepsilon_t') = \mathbb{E}(\varepsilon_t \varepsilon_t')$ is an $N \times N$ unconditional covariance of the standadized residual of univariate GARCH model. Parameters α and β are nonnegative scalars, ensuring that $\alpha + \beta < 1$ to maintain positivity of the covariance matrix. The conditional correlation $r_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t} q_{jj,t}}}$ can be represented in the standard correlation form by defining $Q_t = (q_{ij,t})$ as shown:

$$r_{ij,t} = \frac{(1 - \alpha - \beta)\bar{Q} + \alpha \varepsilon_{t-1} \varepsilon_{t-1}' + bQ_{t-1}}{\sqrt{(1 - \alpha - \beta)\bar{Q} + \alpha \varepsilon_{t-1} \varepsilon_{t-1}' + bQ_{t-1}} \cdot \sqrt{(1 - \alpha - \beta)\bar{Q} + \alpha \varepsilon_{t-1} \varepsilon_{t-1}' + bQ_{t-1}}} \tag{7}$$

2.3. CNN model. [26] pioneered the Convolutional Neural Network (CNN), which has since become one of the most popular and widely adopted neural network architectures. The core structure of the CNN model includes a convolutional layer, a pooling layer, and a fully connected layer [27].Convolutional layer is CNN's fundamental structural component. It utilitzed to extract features such as temporal patterns, trends, anomalies, and local relationships by applying a series of filters to the incoming data. The feature map's dimensions are decreased and the overfitting issue is lessened with the use of the pooling layer. In order to emphasize patterns and notable spikes in time series data, the study used a max-pooling layer, which tends to maintain the most salient aspects of the input.

If the input vector is $\mathbf{x}_0 \in \{x_1, x_2, \ldots, x_n\}$, the output $\mathbf{y}_{ij}^1$ from the initial convolutional layer can be described as follows [28]:

$$\mathbf{y}_{ij}^1 = Relu\left(b_j^1 + \sum_{m=1}^{M} w_{m,j}^1 x_{i+m-1,j}^0\right) \tag{8}$$

In this context, the output $\mathbf{y}_{ij}^1$ is determined using the input $\mathbf{x}_{ij}^0$ from the preceding layer, with b_j^1 denoting the bias associated with the jth feature map. The kernel weights are denoted by w, and the ReLU activation function was employed in the investigation [29]. In a similar manner, the output vector of the k^{th} convolutional layer is expressed as follows:

$$\mathbf{y}_{ij}^l = Relu\left(b_j^l + \sum_{m=1}^{M} w_{m,j}^l x_{i+m-1,j}^{l-1}\right) \tag{9}$$

Following the convolutional layer, the max pooling layer reduces the spatial dimensions of the feature maps the convolutional layer generates, hence reducing the number of parameters and computing expense.

2.4. BiLSTM model. The study will initially addressed a unidirectional LSTM model in order to provide a better understanding of the design and functionality of a BiLSTM network. LSTM models excel in handling long-term dependencies and processing long-sequence data. The gradient disappearance and explosion issues in modeling training were successfully resolved by the LSTM model, which enhances the hidden layer structure of RNN by adding a set of gating units made up of input gates, forgetting gates, and output gates [30]. The structure of the LSTM network is shown in Fig. 1. In order to anticipate time-series outcomes, the LSTM architecture processes data using a particular algorithm in the forget gates. The forget gate is first fed the input for the current time step as well as the output from the previous time step.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{10}$$

Here, the input value at the current time is x_t, the output from the previous time step is h_{t-1}, the bias applied to the forget gate is b_f, the forget gate's weight is W_f, and the range of f_t is (0,1). Additionally, the input gate receives inputs from both the past and current times as well as their output values. The state of the candidate cell at the input gate as well as the output value are calculated using the following formula:

$$i_t = \sigma(W_i[\cdot h_{t-1}, x_t] + b_i) \tag{11}$$

$$\widetilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{12}$$

where W_c is the weight of the candidate gate, b_c is the bias value of the candidate gate , W_i is the weight of the input gate, and the value range of i_t is (0, 1). The process of changing cell values or model parameters at this point is the next step in the LSTM model, and it is done as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{13}$$

The range of values for C_t is (0, 1). At processing time t, the output value h_{t-1} and the input value x_t serve as inputs for the output gate, and the output from this gate is determined using the following formula:

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{14}$$

where b_o is the output gate's bias value, W_o is the output gate's weight, and the value range of O_t is (0.1). Ultimately, the output gate produces the LSTM's ultimate output value, which is the outcome of a computation made with the following formula:

$$h_t = O_t \cdot \tanh(C_t) \tag{15}$$

After elucidating the activity of the LSTM network, the operation of BiLSTM networks can be explained. BiLSTM networks represent an advancement of bidirectional RNN-based LSTM networks. By integrating a forward LSTM layer with a backward LSTM layer, BiLSTM networks fully capture both past and future information, unlike traditional LSTM networks, which predict the next output based solely on past time series data. This dual input of forward and backward sequence information enhances the model's robustness [31]. This paper utilized the BiLSTM neural network to capture bidirectional sequential features from the information extracted by the CNN layer, effectively leveraging the long-term dependencies in the sample data for learning, and ultimately producing the output. The fully connected layer produces the volatility and correlation prediction results. Fig. 1 illustrates the BiLSTM structure, constructed with LSTM blocks. BiLSTM, comprising both forward and backward LSTM components, necessitates a reversal of the computation.

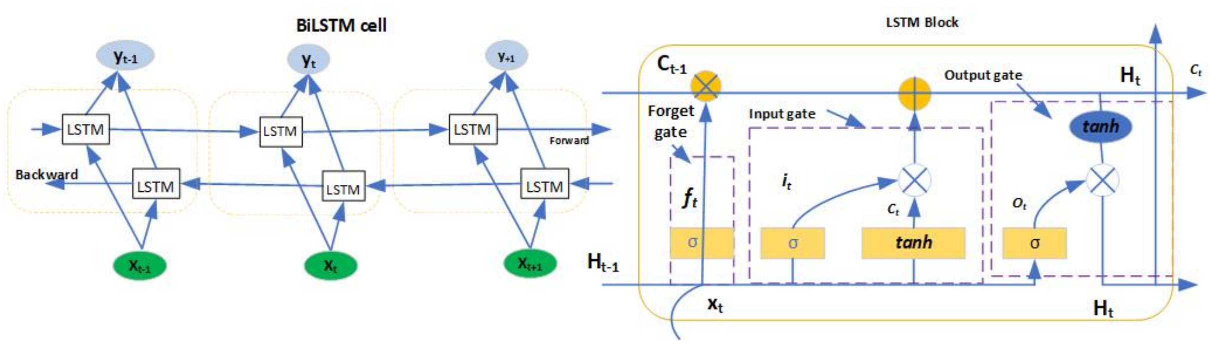


Fig. 1. Internal architecture of the BiLSTM model.

2.5. VAR-CNN-BiLSTM model. The study proposes a novel methodology for forecasting the conditional correlation and volatility of assets by employing a hybrid multivariate model integrating Vector autoregressive model(VAR), Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) architectures. First, the mean model was built using VAR , which allowed simultaneous analysis of multiple time series variables, capturing the interdependencies and dynamic relationships between them. VAR models facilitate Granger causality testing, allowing researchers to assess the causal relationships between variables. The parameters of the VAR model can be estimated using various techniques, such as ordinary least squares (OLS) or maximum likelihood estimation (MLE). After doing mean model, the standard residuals will be an input for the hybrid CNN-BiLSTM model. CNN-LSTM model can indeed be utilized for investigating dynamic relationships and dependencies between multiple time series variables, including how their correlations and volatilities evolve in response to changing market conditions and economic factors.

Lastly, CNN and BiLSTM architectures receive the predicted output of the GARCH family models. CNN effectively captures spatial dependencies, while BiLSTM excels at capturing long-term dependencies, leveraging both temporal and spatial features for improved forecasting.

Data Preprocessing

The missing values were addresed through an imputation of interpolation. Additionally, identify outliers in the data that may skew the analysis and consider removing or correcting them using statistical methods or domain knowledge. Data points that are located at the outermost limits of a dataset are referred to be outliers. In order to ensure that the values contributed equally and support the effectiveness of the training process, the study normalized the residual data series as described in [32].

$$y_i^j = \frac{x_i^j - \bar{x}^j}{s^j} \tag{16}$$

In this context, y_i^j denotes the standardized value of the j-th series at index i, x_i^j refers to the original input data value for series j, $\bar{x}^j$ represents the mean of the input data values for series j, and s^j indicates the standard deviation of the input data for series j. Each economic and financial time series dataset was divided into two subsets: a training set and a test set, with 80% of the data allocated for training and the remaining 20% reserved for testing model accuracy. After this division, the data must be reshaped into a 3D format to be used as input for the hybrid CNN-BiLSTM model. Thus, input dimensions are samples, time steps, and features. The number of time steps (window size) was a hyperparameter that represents the number of previous lags used as input to predict the next time steps. The study used empirical testing to fix an optimal value of the windows hyperparameter. The sequence of observations for asset i is $\{x_{i,t_1}, x_{i,t_2}, \ldots, x_{i,t_n}\}$, where i is the number of assets. This observations can be rearranged in a supervised learning form as shown in the matrics below

$$\begin{bmatrix} [a_{1,t_1}, a_{2,t_1}] & [a_{1,t_2}, a_{2,t_2}] & \cdots & [a_{1,t_p}, a_{2,t_p}] \\ [a_{1,t_2}, a_{2,t_2}] & [a_{1,t_3}, a_{2,t_3}] & \cdots & [a_{1,t(p+1)}, a_{2,t(p+1)}] \\ \vdots & \vdots & \ddots & \vdots \\ [a_{1,t(m-1)}, a_{2,t(m-1)}] & [a_{1,t_m}, a_{2,t_m}] & \cdots & [a_{1,t(m+p-1)}, a_{2,t(m+p-1)}] \end{bmatrix}$$

Once the assets residuals are prepared as inputs for the CNN, they are further processed. The output for the hybrid CNN-BiLSTM model will be in the form of 2x2 covariance matrices, where the diagonal elements represent the variances and the off-diagonal elements represent the correlations of the asset residuals.

$$\left\{ \begin{array}{c} \begin{pmatrix} h_{11,t_1} & h_{12,t_1} \\ h_{21,t_1} & h_{22,t_1} \end{pmatrix} \\ \cdots \\ \begin{pmatrix} h_{11,t_m} & h_{12,t_m} \\ h_{21,t_m} & h_{22,t_m} \end{pmatrix} \end{array} \right\}$$

This observations have to changed to multiple examples (samples) by developing a matrix X which served as independent variable of the model and y as dependent variable of the model of which the model can learn. Then, divide the time series to examples where each sample has size equal to the number of time steps (lagged variables) that is p and the size of learning samples is m. The obtained size of the independent and predicted matrix will have size $(m-1)\times p$ and $m\times 1$. The Deep learning models described in this paper are outlined in Table 1. The goal is to pick out features from the input dataset using CNN layers. Subsequently, the outputs of these CNN layers are fed into layers of BiLSTM and an output dense layer to aid with sequence prediction.

Table 1. Deep learning models' internal hybrid structure.

HDL model	**Structure of Layer**
CNN-BiLSTM	conv1D layer (filters: 32, filter size:3, relu activation) +maxpooling1D(Pooling size:1,padding:same) + conv1D layer (filters: 128, filter size: 2, relu activation) + max-pooling1D (polling size:1, padding: same) + flatten layer + BiLSTM layer (neurons: 64, relu activation) + BiLSTM layer (neurons: 32;relu activation) + dense layer (neuron: 4, linear activation)

2.6. Forecasting performance evaluation. The assessment of forecasting efficacy in multivariate DCC-GARCH and CNN-LSTM models was undertaken utilizing the root mean square error (RMSE) and mean squared error (MSE) as primary estimation criteria.

$$\text{MAE} = \frac{1}{N}\sum_{k=1}^{N}|Y_k - \hat{Y}_k| \tag{17}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(Y_k - \hat{Y}_k)^2} \tag{18}$$

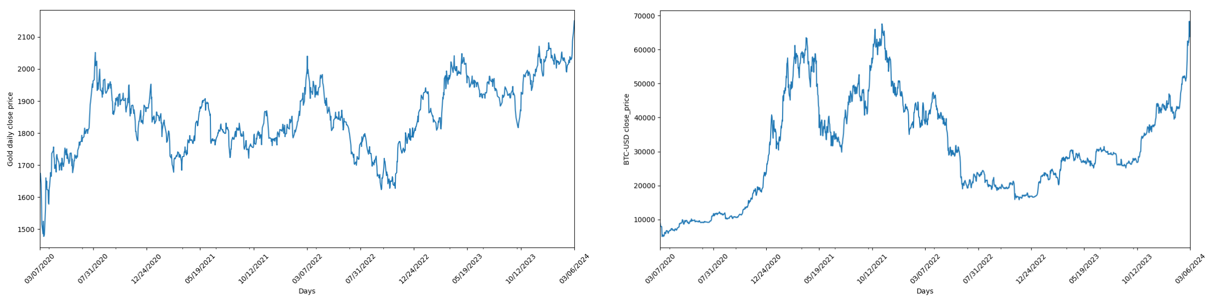
Where N is the number of observations, Y_t are the actual values at time t, and $\hat{Y}_k$ are the predicted values of the model at time t.

3. Data

This paper utilized daily time series data of Bitcoin and Gold closing prices, expressed in US dollars. Both series contain data spanning from 16 May 2019 to 15 May 2024, extracted from the yahoo Finance.com website. A total of 1828 daily observations were recorded. The missed values were filled using linear interpolation. The dataset was split into three portions: the training set, which accounted for 60% of the data, and the validation set, which constituted the remaining 20% to assess the performance of the trained model. The remaining 20% was used for testing to evaluate the final performance of the model. Fig. 2 below depicted the historical trends of Bitcoin USD (BTC-USD) exchange rate and gold prices. It was evident that both variables exhibited similar directional movements. Additionally, both price series appeared to be non-stationary. [33] noted that most financial research favored the use of asset returns over asset prices. This preference was due to the fact that returns provided investors with a comprehensive and scale-independent overview of the investment's performance. Furthermore, return series exhibited beneficial statistical characteristics. For the sake of interpretation, financial returns are commonly calculated in percentage as:

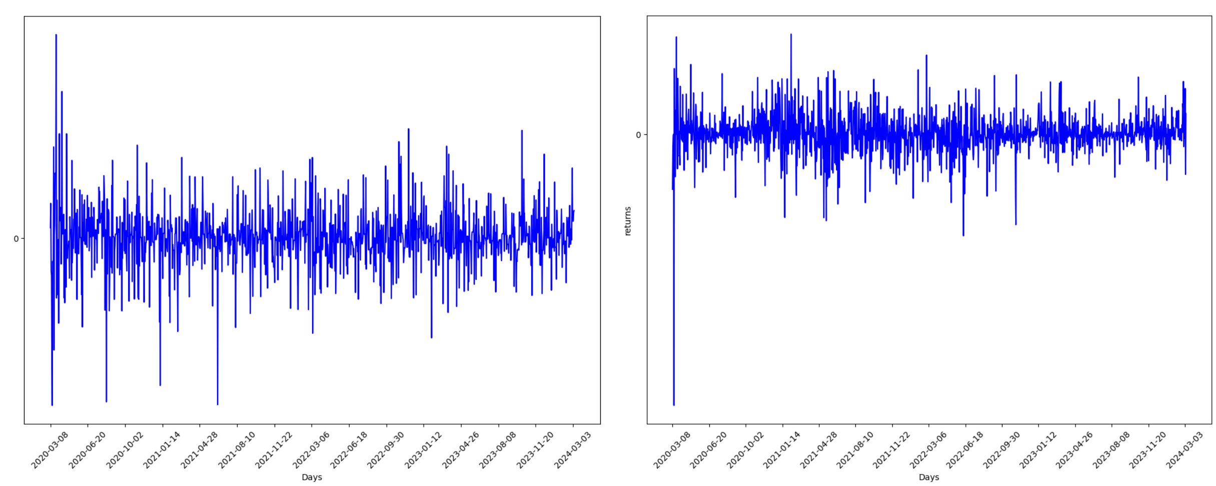
$$r_{i,t} = \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \times 100 \tag{19}$$

where $P_{i,t}$ denotes the current closing prices of the i^{th} asset and $P_{i,t-1}$ is the previous trading day of the i^{th} assets. The descriptive statistics and timing diagrams for the returns series are presented in Fig. 3 and Table 2. These visuals reveal several abnormal peaks in the return sequence data for both series. Daily returns oscillate around zero, visually suggesting the stationarity of the return series. The daily return series also revealed periods of high volatility, which is stylized to represent returns, frequently followed by other high volatility, and small volatility, which is typically followed by other small volatility.



(a) Gold prices **(b)** BTC-USD prices

Fig. 2. Time series graphs of BTC-USD and Gold prices.



(a) Gold return **(b)** BTC-USD return

Fig. 3. Time series graphs of BTC-USD and Gold returns.

Table 2. Descriptive statistics of BTC-USD and Gold returns.

Statistic	**BTC-USD**	**Gold**
Mean	0.0011	0.0003
Median	0.0004	0.0003
Variance	0.0012	5.86*10-5
Skewness	-1.278	-0.217
Kurtosis	18.526	7.539
Standard deviation	0.0354	0.0076
Jarque-Bera	26626.95	4341.257
	(0.00)	(0.00)

** Values below the Jarque–Bera statistics represent p-values at 5% significance level.

In addition to visually examining the time plot of the log return series for the BTC-USD exhange rate and Gold prices, the study conducted stationarity tests using the Augmented Dickey-Fuller test. Table 3, revealed p-values that provide sufficient evidence to reject the null hypothesis of non-stationarity. Consequently, the log return series were found to be stationary at the 5% significance level. Given that stationarity is an essential characteristic for time series analysis, we proceeded with further analysis of the return series.

Table 3. The ADF unit root test for log returns stationarity.

Series	t-statistic	MacKinnon approximate p-value
BTC-USD	-16.23	-2.86
GOLD	-21.26	-2.12

Table 4 presents the results of ARCH-LM test, F statistic and TR2 statistic, which indicates that all time series have ARCH effect, meanwhile, Ljung-Box statistic is also different from zero, so it is suitable for building GARCH model.

Table 4. ARCH LM test summary statistics.

return series	LM-statistic	p-value
BTC-USD	2.13	0.0071
GOLD	24.2	0.006

Table 4 presents the results of the ARCH LM test for the two return series. The p-values suggest rejecting the null hypothesis of “no ARCH effect” at the 5% significance level. These findings imply that the price log return series exhibit volatility, necessitating the use of GARCH models for accurate modeling.

4. Results and Discussion

VAR(2) was constructed using the lag order test of the VAR model. Furthermore, the Granger causality test was employed to examine the correlation between the prices of gold and BTC-USD. In this study, past returns of exchange rates Bitcoin-USD are useful in forecasting the returns of Gold reported in Table 5. This implies a unidirectional causal relationship from Bitcoin-USD to Gold.

Table 5. The return series' Granger causality(F) test results.

	BTC-USD	GOLD
BTC-USD		6.186(0.01293)
GOLD	0.5244(0.469)	

Note that the parenthesis indicates the p-value.

Various criteria were employed to determine the optimal lag length of the VAR model with the results presented in Table 6. The performance of each criterion in the simulation study depends on the sample size to which it is applied [34]. The study showed that the Schwarz information criterion and Hannan-Quinn criterion perform better in large samples, whereas the Akaike information criterion and final prediction error perform better in small samples.This finding has been further corroborated by [11]. Because the Schwarz information criterion performs better than other criteria in bigger samples, it was used in this investigation. As a result, the investigation determined that two was the ideal lag time for the VAR model. The residuals from the VAR(2) model were estimated and subsequently examined for serial and cross-correlation using the multivariate Portmanteau test. The findings, presented in Table 7, indicate that the test did not provide enough evidence to reject the null hypothesis, suggesting that there are no significant serial or cross-correlations in the residuals of the VAR(2) model.

Table 6. Findings from the VAR(2) model.

Parameter	Estimate	Standard Error	t-value	Pr(>| t |)
r1.l1	-0.05805	0.02631	-2.206	0.0275 *
r2.l1	0.0684	0.0189	2.575	0.04651
r1.l2	0.0443	0.0263	1.682	0.032
r2.l2	0.1603	0.1186	1.352	0.1767
r1.l1	0.015	0.0058	2.578	0.0100 *
r2.l1	0.0001	0.02631	0.005	0.9957
r1.l2	0.004	0.0058	0.801	0.4234
r2.l2	0.053	0.026	2.037	0.0418 *

Note. r1 represents the BTC-USD exchange return; r2 corresponds to the Gold return. The terms .l1 and .l2 indicate the first and second lags, respectively.

4.1. Volatility Transmission Analyzed Using the DCC Model. The results obtained from the DCC-GARCH model indicate that the conditional variance of the standardized residuals is affected by the previous volatility of the two residual series as well as their correlations. Using the DCC-GARCH model, the study discussed the dynamic process of the co-movement between the BTC-USD exchanges and Gold indices. The parameter of the univariate GARCH model is estimated, and the results are shown in Table 8. Since $\sigma_{i,t}$ is assumed to follow a univariate GARCH(1,1) process, the coefficient $A_{i,i}$ represents ARCH effects, while $B_{i,i}$ (i=1,2) represents GARCH effects. All parameters of the univariate GARCH(1,1) model for the two assets are statistically significant at 5% level of significance. The study found that the coefficients of lagged conditional variance $B_{i,i}$ (i=1,2) in the GARCH model are greater than the coefficients of past return errors $A_{i,i}$. This suggests a fundamental stability effect in the volatilities of BTC-USD exchange and gold returns. The statistically significant parameter of the univariate GARCH model $B_{i,i}$ in a DCC-GARCH model not only showed us past volatilities have a strong and persistent effect on current volatilities for the given time series of an individual indices but also the presence of bi-directional volatility transmission among the indices. The significance of all ARCH and GARCH coefficients α and β respectively showed as the DCC-GARCH model can effectively incorporate and respond to external information. This results that the impact of external information in the model was long-lasting and persistent. Furthermore, the estimated parameters of the DCC -GARCH model are greater than zero. This indicated that the conditional correlations are time-varying. This challenges the assumption of constant conditional correlations, which was often unrealistic in empirical studies, as discussed by [35]. Since $\alpha+\beta<1$, the DCC-GARCH models were mean-reverting but did so gradually.

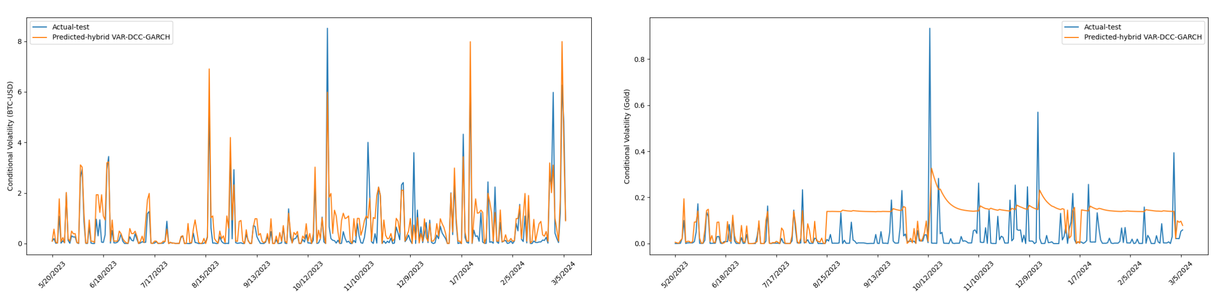
Table 7. Results of Multivariate Portmanteau test for autocorrelation in the residuals of VAR(2).

Variable	m	Q(m)	df	p-value
BTC-USD	1	0.03484757	2	0.12493
GOLD	2	9.36262447	4	0.05264784

Table 8. Results obtained from the DCC-GARCH model.

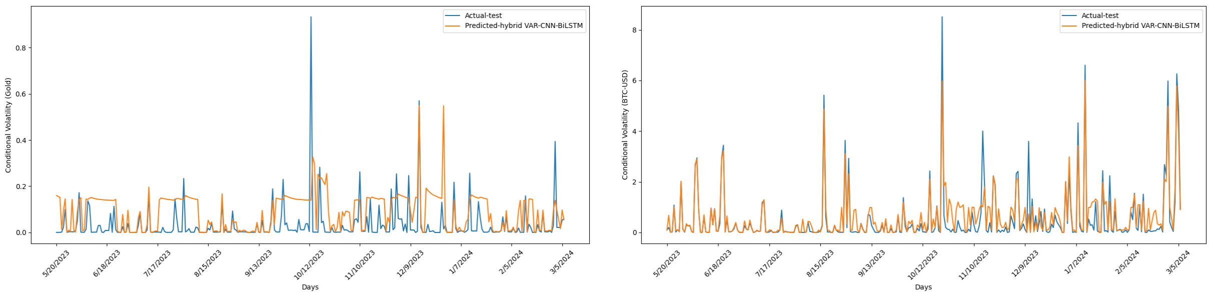
Parameter	Estimate	Standard Error	t-value	Pr(>| t |)
μ_1	-0.009445	0.031433	-0.30049	0.763806
ω_1	0.033092	0.018284	1.80990	0.070311
A_{11}	0.032966	0.019812	1.66393	0.036127
B_{11}	0.931179	0.033146	28.09308	0.00000
μ_2	-0.008644	0.029881	-0.28927	0.772372
ω_2	0.089103	0.067123	1.32746	0.184358
A_{22}	0.045781	0.033632	1.36124	0.0273437
B_{22}	0.870627	0.087084	1.36124	0.00000
α	0.006869	0.003809	1.80351	0.071308
β	0.987513	0.004418	223.53295	0.00000

4.2. Prediction results of conditional volatility on hybrid models. The study introduced a novel hybrid model, denoted as the hybrid VAR-CNN-BiLSTM, for forecasting conditional volatility and correlations. After forecasting correlation and covariance using the proposed model, the study compared it with an existing hybrid econometric model, the VAR-DCC-GARCH model. Fig. 4 and Fig. 5 illustrated the conditional volatility of BTC-USD exchanges and Gold using the hybrid VAR-DCC-GARCH model and the hybrid VAR-CNN-BiLSTM model, respectively.



(a) (b)

Fig. 4. Volatility forecasts for a hybrid of VAR-DCC-GARCH models. (a) Conditional volatility forecast for BTC. (b) Conditional volatility forecast for GOLD.



(a) (b)

Fig. 5. Volatility forecasts for a hybrid of VAR-CNN-BiLSTM models. (a) Conditional volatility forecast for BTC. (b) Conditional volatility forecast for GOLD.

The evaluation results, presented in Table 9, demonstrated that the hybrid VAR-CNN-BiLSTM models outperform the hybrid VAR-DCC-GARCH models in forecasting the volatility of BTC-USD exchange rates. The prediction results of the new proposed models were shown in Fig. 5(a)–(b). In comparison to hybrid VAR-DCC-GARCH model, the hybrid VAR-CNN-BiLSTM exhibited notable improvements, with a decrease of 28.57% in MAE and 36.58% in MSE for BTC-USD exchanges and a decrease of 22.9% in MAE and 14.29% in MSE for Gold. Consequently, the hybrid VAR-CNN-BiLSTM model outperforms optimal prediction performance.

TABLE 9. Bivariate model comparison under normal distribution for volatility of BTC-USD exchanges and Gold.

Model	**MAE**		**MSE**	
	BTC-USD	**GOLD**	**BTC-USD**	**GOLD**
Hybrid VAR-DCC-GARCH	0.3416	0.093	0.3516	0.014
Hybrid VAR-CNN-BiLSTM	0.244	0.0717	0.223	0.012

The evaualtion results of the the forecasted covolatility of BTC-USD and Gold return sereis are shown on Table 10. The MSE and RMSE results obtained from the models VAR-DCC-GARCH is larger than VAR-CNN-Bilstm mode under the normal distrbution. Therefore, the hybrid VAR-CNN-BiLSTM model is more effective for estimating and forecasting the variance-covariance of the BTC-USD and Gold return series. In addition to predicting the conditional correlation, the correlation was found to be positive and significant [14]. This suggests that the two return series tend to increase or decrease in tandem.

Table 10. Comparison of the bivariate hybrid model under normal distribution for the covolatility of gold and BTC-USD exchanges.

Statistics	VAR-DCC	VAR-CNN-BiLSTM
MSE	0.05	0.03484757
MAE	0.127	0.00689

5. Conclusion

In conclusion, the key findings are as follows: The study used the dynamic relationships between BTC-USD exchanges and Gold using historical data. it also examined spillover effects in variances between them. The mean model were studied using VAR model and the conditional covarinaces were further investigated using DCC-GARCH and hybrid CNN-BiLSTM models. In addition to this the causality test between the series were checked using the returns. There is a unidirectional effect across the series. Further more, the research analysis the predictive power accuracy between DCC-GARCH and CNN-BiLSTM in forecasting the conditional voaltility in each indeces and correlation between them. The result indicated that the hybrid VAR-CNN-BiLSTM model had higher accuracy as compared to VAR-DCC-GARCH model.

6. Recommendation

The study was limited to investigating the co-volatility prediction of BTC-USD exchanges with Gold using the two hybrid models. As a result , the study recommended the following ideas for future works. Future study aiming at the co-movement of securites and forcasting it using both asymmetric and symmetric DCC-GARCH models and further will compare it with deep learning models like Gated neural networks. Furthermore, the study can give a way to constracd a new hybrid DCC-GARCH-LSTM model in order to model conditional correlation between assets.

Funding Statement

This work is supported by the Pan African University Institute for Basic Sciences, Technology and Innovation.

Acknowledgements

The authors would like to thank Pan African University for providing the necessary support.

Data Availability Statement

The data of this study can be obtained by contacting the corresponding author.

Conflict of Interests

The authors declare that there is no conflict of interests.

References

[1] H.T. Araya, J. Aduda, T. Berhane, A hybrid garch and deep learning method for volatility prediction, J. Appl. Math. 2024 (2024), 6305525. https://doi.org/10.1155/2024/6305525.

[2] L. Bauwens, C. Hafner, S. Laurent, Handbook of volatility models and their applications, Wiley, Hoboken, 2012.

[3] A.A. Salisu, T.F. Oloko, Modeling oil price–US stock nexus: A VARMA–BEKK–AGARCH approach, Energy Econ. 50 (2015), 1–12. https://doi.org/10.1016/j.eneco.2015.03.031.

[4] P. Katsiampa, S. Corbet, B. Lucey, Volatility spillover effects in leading cryptocurrencies: A Bekk-Mgarch analysis, Finance Res. Lett. 29 (2019), 68–74. https://doi.org/10.1016/j.frl.2019.03.009.

[5] G.-D. Liu, C.W. Su, The dynamic causality between gold and silver prices in China market: a rolling window bootstrap approach, Finance Res. Lett. 28 (2019), 101–106. https://doi.org/10.1016/j.frl.2018.04.007.

[6] I. Yousaf, S. Ali, The COVID-19 outbreak and high frequency information transmission between major cryptocurrencies: evidence from the VAR-DCC-GARCH approach, Borsa Istanbul Rev. 20 (2020), S1–S10. https://doi.org/10.1016/j.bir.2020.10.003.

[7] R.F. Engle, O. Ledoit, M. Wolf, Large dynamic covariance matrices, J. Bus. Econ. Stat. 37 (2019), 363–375. https://doi.org/10.1080/07350015.2017.1345683.

[8] C. Wu, Measuring and comparing the value-at-risk using garch and carsr models for csi 300 index, Theor. Econ. Lett. 08 (2018), 1179–1187. https://doi.org/10.4236/tel.2018.86078.

[9] F. Gharbi, Time-varying volatility spillovers among bitcoin and commodity currencies, J. Stat. Econ. Meth. 8 (2019), 25–39.

[10] M.A. Shettima, A.I. Gambo, O. Abraham, et al. Return and volatility spillovers between Ghanaian and Nigerian equity markets, Int. J. Econ. Bus. Manage. Res. 3 (2019), 1–19.

[11] N. Yunvirusaba, J. Aduda, A. Kube, Volatility spillover effects among securities exchanges in east Africa, Int. J. Econ. Finance, 11 (2019), 32–41.

[12] A. Ang, G. Bekaert, International asset allocation with time-varying correlations, National Bureau of Economic Research, Cambridge, MA, 1999. https://doi.org/10.3386/w7056.

[13] F. Longin, B. Solnik, Extreme correlation of international equity markets, J. Finance 56 (2001), 649–676. https://doi.org/10.1111/0022-1082.00340.

[14] D. Yeshiwas, Y. Berelie, Forecasting the covolatility of coffee arabica and crude oil prices: a multivariate garch approach with high-frequency data, J. Prob. Stat. 2020 (2020), 1424020. https://doi.org/10.1155/2020/1424020.

[15] J. Ni, Y. Xu, Forecasting the dynamic correlation of stock indices based on deep learning method, Comput. Econ. 61 (2023), 35–55. https://doi.org/10.1007/s10614-021-10198-3.

[16] A. Sinha, R. Tayal, A. Vyas, et al. Forecasting electricity load with hybrid scalable model based on stacked non linear residual approach, Front. Energy Res. 9 (2021), 720406. https://doi.org/10.3389/fenrg.2021.720406.

[17] C.A. Sims, Macroeconomics and reality, Econometrica 48 (1980), 1. https://doi.org/10.2307/1912017.

[18] D.N. Gujarati, Basic econometrics, 4th ed., McGraw-Hill, 2005.

[19] Machine Learning +, Vector autoregression (VAR) – Comprehensive guide with examples in python, https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python.

[20] R. Engle, K. Sheppard, Theoretical and empirical properties of dynamic conditional correlation multivariate garch, National Bureau of Economic Research, Cambridge, MA, 2001. https://doi.org/10.3386/w8554.

[21] R. Engle, Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models, J. Bus. Econ. Stat. 20 (2002), 339–350. https://doi.org/10.1198/073500102288618487.

[22] T. Bollerslev, Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model, Rev. Econ. Stat. 72 (1990), 498–505. https://doi.org/10.2307/2109358.

[23] C. Brooks, S.P. Burke, Information criteria for GARCH model selection, Eur. J. Finance 9 (2003), 557–580. https://doi.org/10.1080/1351847021000029188.

[24] E. Orskaug, Multivariate dcc-garch model: -with various error distributions, Thesis, Institutt for Matematiske Fag, 2009. http://hdl.handle.net/11250/259296.

[25] S.J. Lim, M. Masih, Exploring portfolio diversification opportunities in Islamic capital markets through bitcoin: evidence from MGARCH-DCC and wavelet approaches, MPRA Paper No. 79752, 2017. https://mpra.ub.uni-muenchen.de/79752.

[26] Y. Lecun, L. Bottou, Y. Bengio, et al. Gradient-based learning applied to document recognition, Proc. IEEE 86 (1998), 2278–2324. https://doi.org/10.1109/5.726791.

[27] Y. Li, W. Dai, Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model, J. Eng. 2020 (2020), 344–347. https://doi.org/10.1049/joe.2019.1203.

[28] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: a comprehensive review, Neural Comput. 29 (2017), 2352–2449. https://doi.org/10.1162/neco_a_00990.

[29] V. Nair, G.E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: 27th International Conference on Machine Learning, Haifa, Israel, 2010.

[30] J. Zhang, L. Ye, Y. Lai, Stock price prediction using cnn-bilstm-attention model, Mathematics 11 (2023), 1985. https://doi.org/10.3390/math11091985.

[31] J. Eapen, D. Bein, A. Verma, Novel deep learning model with cnn and bi-directional lstm for improved stock market index prediction, in: 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), IEEE, Las Vegas, NV, USA, 2019: pp. 0264–0270. https://doi.org/10.1109/CCWC.2019.8666592.

[32] H. Widiputra, A. Mailangkay, E. Gautama, Multivariate cnn-lstm model for multiple parallel financial time-series prediction, Complexity 2021 (2021), 9903518. https://doi.org/10.1155/2021/9903518.

[33] R.S. Tsay, Analysis of financial time series, Wiley, 2005.

[34] V.K.S. Liew, Which lag length selection criteria should we employ? Econ. Bull. 3 (2004), 1–9. https://ssrn.com/abstract=885505.

[35] L. Bauwens, S. Laurent, J.V.K. Rombouts, Multivariate GARCH models: a survey, J. Appl. Econometrics 21 (2006), 79–109. https://doi.org/10.1002/jae.842.