

Hybrid Data-Driven and Deep Learning Based Portfolio Optimization

Joy Dip Das¹, Sulalitha Bowala², Ruppa K. Thulasiram¹, Aerambamoorthy Thavaneswaran²

¹Department of Computer Science, University of Manitoba, Winnipeg, Canada

²Department of Statistics, University of Manitoba, Winnipeg, Canada

Email: dasj@myumanitoba.ca, bowalams@myumanitoba.ca, tulsi.thulasiram@umanitoba.ca, Aerambamoorthy.Thavaneswaran@umanitoba.ca

How to cite this paper: Dip Das, J., Bowala, S., Thulasiram, R.K. and Thavaneswaran, A. (2024) Hybrid Data-Driven and Deep Learning Based Portfolio Optimization. *Journal of Mathematical Finance*, 14, 271-310.

<https://doi.org/10.4236/jmf.2024.143016>

Received: February 21, 2024

Accepted: June 21, 2024

Published: June 24, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This research introduces a novel hybrid architecture that combines deep learning, data-driven algorithms, and an affinity propagation-based approach to build robust investment portfolios. This study evaluates the efficacy of BiLSTM and BiGRU in constructing resilient portfolios of stocks from diverse sectors under varying market conditions. The results highlight the superior performance of BiGRU, particularly in dynamic and volatile market scenarios. The research emphasizes the importance of precise stock prediction and effective diversification for building resilient portfolios, leveraging advanced techniques from deep learning and data-driven optimization. Comparative analyses indicate similar performance between portfolios constructed with actual and predicted data using data-driven optimization. The findings offer valuable insights into constructing robust portfolios by employing advanced techniques, thereby enhancing decision-making in financial markets.

Keywords

BiLSTM, BiGRU, Affinity Propagation, Diversification, Data-Driven Algorithm, Portfolio Optimization

1. Introduction

The stock market allows individuals and entities to buy and sell shares or ownership stakes in publicly traded companies. It enables companies to raise funds for expansions, research, and development by selling shares to interested investors. This market provides an avenue for individuals to potentially grow their wealth and plan for their financial future through smart investments. Market dynamics are influenced by diverse elements, including economic indicators,

company performance, geopolitical events, and public sentiment. In simple terms, the stock market is like a giant puzzle with many pieces. Various stock products encompass a wide array of financial instruments available in the market. These products can include stocks of individual companies, exchange-traded funds (ETFs) which bundle multiple stocks, mutual funds comprising a diversified portfolio of stocks managed by professionals, and index funds tracking specific market indices. Portfolio optimization is a crucial concept in finance that involves strategically constructing an investment portfolio to achieve the best possible balance between risk and return. The primary objective is to maximize returns given a specific risk tolerance level or, conversely, reduce risk for a desired amount of return. Traditional portfolio optimization methods utilize mathematical models to assess the historical performance and correlation of assets, aiming to create diversified portfolios that offer optimal risk-return profiles. The process involves selecting a mix of assets that collectively reduce overall portfolio risk while maximizing potential returns. The Markowitz Portfolio Optimization [1] aims to develop investment portfolios through strategic diversification of assets. Its objective is to either maximize anticipated returns for a specified risk level or minimize risk for a targeted return. The resultant efficient frontier illustrates optimal trade-offs between risk and return. Expanding upon the foundations laid by Markowitz, Modern Portfolio Theory (MPT) [2] underscores the importance of diversification in attaining an ideal equilibrium between risk and return, taking into account the interrelation among diverse assets.

The risk-free rate, market risk premium, and the beta associated with the individual asset are incorporated by the Capital Asset Pricing Model (CAPM) [3] [4] [5] [6] [7] for computing the expected yield of an asset. This facilitates the construction of portfolios by assessing the equilibrium between risk and anticipated returns. The CAPM revolutionized asset pricing theory; although widely taught and applied, its simplicity and intuitive appeal are marred by empirical challenges, notably poor performance in real-world applications, possibly stemming from inherent model limitations or inadequacies in market proxies used for testing and application. The Arbitrage Pricing Theory (APT) [8] extends the horizon of portfolio optimization by incorporating diverse factors influencing asset valuations. This allows for a thorough exploration of the interrelationship between risk and return within financial contexts. The Black-Litterman Model [9], in the realm of dynamic management of portfolio emphasizing information utilization, operates within a Bayesian analytic framework, allowing the portfolio manager to express views, subsequently translated into security return forecasts, offering a theoretically and practically appealing tool for portfolio construction despite persisting challenges.

Within the confines of a Markov-based framework characterized by switching regimes, a problem concerning the optimization of risk parity portfolios is formulated and addressed [10]. The aim is to augment the precision of parameter estimation and systematically alleviate the susceptibility of optimal portfolios to discrepancies in estimation. This methodology involves integrating a factor

model associated with the switching of regimes, introducing market dynamics to refine parameter estimation, and subsequently applying this model for the optimization of risk parity. Concentrating on minimizing risk, Mean-Variance Portfolio optimization [11] aims to create a portfolio with the least volatility, regardless of anticipated returns. Monte Carlo simulation for portfolio optimization [12] involves utilizing random sampling to model various potential future scenarios, enabling a comprehensive analysis of portfolio risk and return by generating multiple simulated outcomes based on the specified parameters and assumptions. The conventional method of estimating asset weights includes creating portfolios with equal weights and inverse volatility weights [13]. Various alternative approaches, such as the enhancement of the Sharpe ratio (SRO), optimization of mean-variance (MVO), and portfolio construction through user constraints (CPO) have been explored in this context [13]. Sharpe [14] introduced a mathematical measure to assess portfolio performance, providing a framework for evaluating different portfolio optimization techniques.

These traditional portfolio optimization methods explained above have a myriad of weaknesses in practical scenarios. One significant limitation of MVO and MPT lies in their sensitivity to input data, as reliance on historical information for expected returns and risk parameters can lead to suboptimal portfolios when market conditions deviate from historical patterns. Additionally, the assumption of normality in asset returns and the focus on mean and variance may not accurately capture the non-normal and complex distributions observed in real financial markets, resulting in misestimated risk. These methods often underestimate tail risks, neglecting the potential impact of extreme events on portfolio performance. Furthermore, their single-period analysis overlooks the dynamic nature of financial markets, and the static approach may not adapt well to changing economic environments over time.

In finance, machine learning (ML) and deep learning (DL) algorithms, such as artificial neural networks (ANN), have proven valuable for predicting stock market changes [15]. These advanced approaches outperform traditional statistical methods due to their ability to handle the intricate, non-linear features present in highly dynamic stock market data [16] [17]. Unlike traditional methods relying on assumed data distributions, ML and DL algorithms adapt to changing market conditions, automatically extracting complex patterns for more accurate forecasting¹. Moreover, these technologies extend beyond stock price prediction, finding applications in credit risk assessment, fraud detection, customer sentiment analysis, loan approvals, portfolio optimization, and market trend identification, enhancing decision-making and operational efficiency for financial institutions. In the domain of stock price prediction, a variety of ML-based regression techniques, including decision tree regression, random forest regression, etc. are employed [18]. Additionally, DL algorithms such as ANN, convolutional neural networks (CNN), and recurrent neural networks (RNN), specifically Long short-term memory (LSTM) and Gated Recurrent Unit (GRU), are favored

¹<https://www.gartner.com/en/finance/topics/finance-ai>.

due to having the ability to consider past temporal relationships between data points in stock market data [19]. These DL techniques, known for uncovering subtle patterns in financial data, also play a crucial role in predicting cryptocurrency prices, especially in highly volatile markets [20].

ML based portfolio optimization is garnering attention because of its potential for capturing intricate, non-linear information and vast datasets in financial markets, outperforming traditional methods. Designed to tackle challenges in portfolio optimization related to both mean-variance and mean conditional value at risk (CVaR), ML-based Performance-based regularization (PBR) has been developed [21]. Other than that, ML techniques are fused with MVO for designing portfolio investment strategies [22] [23]. DL models are utilized to directly optimize the Sharpe ratio [24]. Different RNN models such as LSTM and GRU are also heavily explored in portfolio optimization along with ensemble learning approaches [25]. Moreover, recent advances in hybridized ML and DL models have led researchers to explore the efficiency of various hybridized DL models for portfolio optimization [26]. The integration of ML and DL models could potentially lead to a prominent advancement across diverse domains. Hybridized DL algorithms are a prominent field of research in portfolio optimization and the development of investment strategies. Modified Deep Belief Networks and RNNs are heavily utilized in portfolio optimization [27]. Different models such as CNN and RNN are being fused for stock selection and optimization for the formation of profitable risk-averse portfolios [28]. Moreover, various heuristics are also being explored by the amalgamation of these with various deep-learning algorithms for portfolio optimization [29].

A novel data-driven strategy for assessing the risk associated with portfolios, which addresses the limitations of the conventional approaches has been presented in [30]. Based on this concept, an innovative data-dependent approach has been presented in [31] for constructing portfolios consisting of both traditional and cryptocurrencies. Their investigation [31] explored statistical risk metrics and put forth diverse statistical correlations for assessing portfolio composition grounded in these measures. Diversification plays a crucial role in building a robust portfolio. Choueifaty and Coignard [32] devised a mathematical conception of the “diversification ratio” for quantifying the diversity in a portfolio.

Clustering techniques to study the impact of diversification on portfolio optimization techniques have been proposed in [33]. Furthermore, they studied the efficiency of multiple clustering techniques to test the efficacies of distinct clustering methods in increasing the profitability of the portfolios designed by both the traditional and the data-driven algorithms [34].

In the realm of portfolio optimization, traditional techniques face challenges related to the dynamic and non-linear nature of financial markets. Traditional methods often struggle with assumptions of normality, constant parameters, and neglect of transaction costs. Machine Learning approaches solve these issues with traditional portfolio construction techniques by considering the non-normality

of the financial time-series data. General machine learning approaches such as linear regression, support vector regressions, etc. do however, encounter issues of interpretability and overfitting in high-dimensional datasets, which could be handled by the usage of deep learning algorithms in a hybridized manner for portfolio optimization, the focus of this study.

This current study experiments with the novel application of Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) models in the financial domain, specifically for stocks across five distinct financial sectors and cryptocurrencies. The research endeavors to construct a hybrid portfolio integrating both general stocks and cryptocurrencies. This portfolio will be developed through a data-driven approach for portfolio optimization, incorporating diversification strategies. The diversification will be introduced through stock selection facilitated by Affinity Propagation (AP)-based clustering. This research contributes to the understanding of the potential effectiveness of data-driven portfolio optimization with deep learning—BiLSTM and BiGRU models trained under varying market conditions and provide insights into constructing diversified hybrid portfolios that incorporate both general stocks and cryptocurrencies.

Figure 1 illustrates the proposition of a novel integration involving the architecture of BiLSTM and BiGRU. This integration is employed for predicting future financial asset values, followed by diversification in stock selection using AP-based clustering. Subsequently, a data-driven portfolio optimization algorithm is utilized for the construction of robust portfolios. In other words, this research introduces a novel framework that combines DL and a data-driven portfolio optimization algorithm. This innovative approach incorporates the infrequently utilized AP-based clustering for diversification, aiming to create a resilient portfolio optimization technique with a focus on risk aversion.

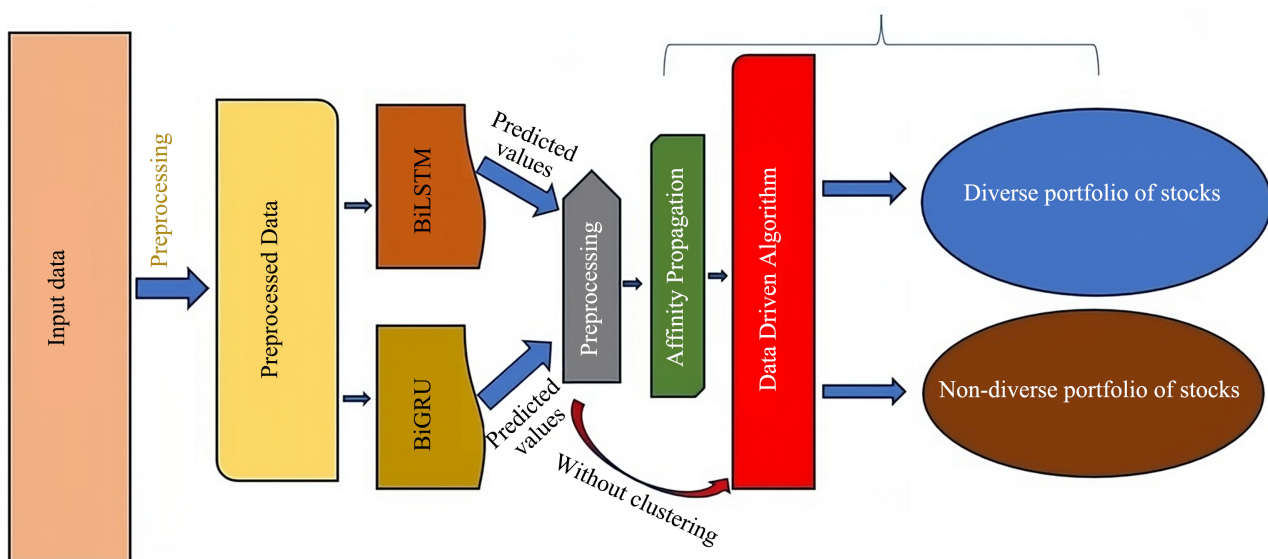


Figure 1. Proposed algorithmic architecture for portfolio optimization.

2. Related Works

Stock market data is notably intricate, characterized by complex interconnections among its components. The dynamics of the stock market are influenced by a multitude of factors such as political, geographical, and socio-economic considerations. The pronounced variability across these diverse factors contributes to fluctuations in stock market trends. Particularly during critical events, accurately predicting stock market performance becomes notably challenging and of utmost importance. Different researchers have used ML and DL algorithms to explore stock market prediction. ML techniques are being heavily explored in the arena of finance and these algorithms are superior to traditional techniques of forecasting such as ARIMA, as the traditional approaches can't identify internal dependencies of the data [35].

2.1. Machine Learning (ML) in Price Prediction and Portfolio Optimization

Machine learning is heavily used in finance for price prediction [36] and clustering of financial assets [34]. Yang and Hospedales [37] have explored self supervised learning (SSL), a class of ML for portfolio diversification in MVO. Their exploration suggested the superiority of SSL over the non-SSL alternatives. To tackle the dimensionality problems, Jaimungal [38] explored various ML and reinforcement learning (RL) for portfolio optimization. Snow [39] considered several weight optimization techniques for several ML algorithms for portfolio optimization. Kaczmarek and Perez [40] have used ML-based stock selection and combined it with the Markowitz mean-variance and Hierarchical Risk Parity (HRP) portfolio construction techniques. Similarly, Jiang *et al.* [41] have used XGboost for portfolio optimization and observed notable efficiency. Similarly, Behera *et al.* [42] explored Support Vector Regression (SVR), XGboost, AdaBoost, K-nearest Neighbours (KNN), and ANN for portfolio optimization. Their research concluded with the findings suggesting the superiority of AdaBoost, a powerful ensemble model.

2.2. Deep Learning (DL) in Price Prediction and Portfolio Optimization

Application of DL models is rising phenomenally in the arena of stock prediction [43]. Jiang [44] has summarized the recent progress in stock market prediction using different DL models and provided a general workflow for the stock prediction domain using DL. Nikou *et al.* [45] have considered the non-linearity and non-stationarity of stock market data when they analyzed different ML and DL algorithms in predicting the daily close price data. Their findings suggest DL models surpass traditional ML algorithms. They also noted that SVR and RF perform well next to the DL models.

DL algorithms are generally seen to be outperforming machine learning algorithms due to having neuron-based learning capabilities. Fernández and Góme

[46] have explored a particular network, the Hopfield network in portfolio optimization. Uysal *et al.* [47] used a risk budgeting model as a layer in the deep neural network which overperformed the traditional techniques. Yu *et al.* [48] used radial basis function (RBF) neural network and have observed sufficient efficiency. Jang and Seong [49] explored the combination of Deep Reinforcement Learning (DRL) and Modern Portfolio Theory (MPT) and observed substantial improvement in profit maximization. Yang [50] further explored DRL in portfolio optimization and reported the superior performance of representation learning-based DRL in portfolio optimization. Ngo *et al.* [51] conducted a comparative study on the performance of DRL with other deep learning and traditional portfolio optimization techniques on the frontier and developed stock market. Their study confirms the better responsiveness of DRL to the market dynamics compared to the other algorithms.

2.3. Hybridizing Machine Learning (ML) and Deep Learning (DL)

The integration of ML and DL models could potentially lead to a prominent advancement across diverse domains. In this amalgamation, ML techniques often serve as a fundamental framework, while DL models, with their complex neural architectures, provide enhanced learning capabilities. Overall, the hybridization of ML and DL broadens the spectrum of applications and augments performance across several domains. LSTM-GRU hybridization is being heavily explored in the financial time series forecasting problem domains. Besides that, LSTM-CNN and GRU-CNN also have been explored for different problem domains [52] [53] including finance [54]. Other than these models, LSTM-GRU-ARIMA [55] is also a popular hybridized model used in different problem domains.

Hybridized DL algorithms are a prominent field of research in the arena of portfolio optimization and the development of investment strategies. Modified Deep Belief Networks and RNNs are heavily utilized in portfolio optimization [27]. Different models such as CNN and RNN are being fused for stock selection and optimization for the formation of profitable risk-averse portfolios [28]. Moreover, various heuristics are also being explored by the amalgamation of these with various deep-learning algorithms for portfolio optimization [29].

BiLSTM is an RNN architecture that utilizes the ability to process sequences of data bi-directionally. Unlike traditional LSTMs, which only consider the past context in a sequence, BiLSTM processes data in both forward and backward directions. This bidirectional processing enables the model to capture dependencies from both preceding and succeeding elements in a sequence, enhancing its ability to understand and learn complex patterns in sequential data. Lu *et al.* [56] have used CNN-BiLSTM using an attention mechanism for efficient price prediction. Pramesti *et al.* [57] have used BiLSTM in stock data prediction for the Indonesian banking sector and found it to be effective. Similar to BiLSTM, BiGRU is bidirectional, meaning it processes information in both forward and

backward directions within a sequence. The GRU is a specific type of RNN cell that is employed in the bidirectional context of BiGRU. Malla *et al.* [58] have explored BiGRU with sentiment data for bitcoin prediction.

This current study aims to devise efficient portfolio optimization technique for maximum profitability by incorporating BiLSTM and BiGRU along with the data-driven portfolio optimization technique improvised with affinity propagation-based clustering for introducing diversity through stock selection.

3. Dataset Consideration

In this research, a diversified portfolio comprising stocks from five distinct sectors-technology, healthcare, energy, banking, and consumer-along with two notable cryptocurrencies, Bitcoin (BTC-USD) and Ethereum (ETH-USD), has been meticulously selected. Each sector is represented by three carefully chosen stocks. The selected stocks include prominent entities such as Apple (AAPL), Microsoft (MSFT), Tesla (TSLA), Johnson & Johnson (JNJ), Pfizer Inc. (PFE), Merck & Co., Inc. (MRK), Exxon Mobil Corporation (XOM), Chevron Corporation (CVX), Enbridge Inc. (ENB), JPMorgan Chase & Co. (JPM), Bank of America Corporation (BAC), HSBC Holdings plc (HSBC), The Procter & Gamble Company (PG), The Coca-Cola Company (KO), and Unilever PLC (UL).

The dataset for stock analysis encompasses information spanning the years 1998 to 2022, serving as the training data. This temporal range encapsulates pivotal financial events, including the www crash of 2000, the global financial crisis of 2008, and the 2020 COVID-19-induced financial crisis. Additionally, minor yet impactful financial events within this timeframe are also considered. Leveraging this extensive training data, the study formulates predictions for the year 2023, subsequently utilizing the predicted data for portfolio optimization.

For Bitcoin, the training data spans from 2014 to 2022, while Ethereum's training dataset covers the period from 2017 to 2022. This comprehensive approach enables the study to incorporate critical financial events, fostering a robust analysis for predictive modeling and portfolio optimization in the cryptocurrency domain.

4. Algorithm Design

This section discusses the algorithmic construction of BiLSTM, BiGRU, Affinity Propagation, and Data-Driven Portfolio Optimization algorithm.

4.1. Bidirectional Long Short-Term Memory (BiLSTM)

BiLSTM is an advancement of the conventional LSTM architecture, specifically crafted to consider dependencies in both forward and backward directions within sequential data. LSTM, a variant of RNN is specifically efficient in considering dependencies of long-term time series. BiLSTM enhances the capabilities of traditional LSTMs by processing input sequences in both forward and backward directions simultaneously.

Figure 2 shows a basic chain architecture of LSTM². This chain architecture of LSTM represents how sequential data is processed, where the information passes in a single direction.

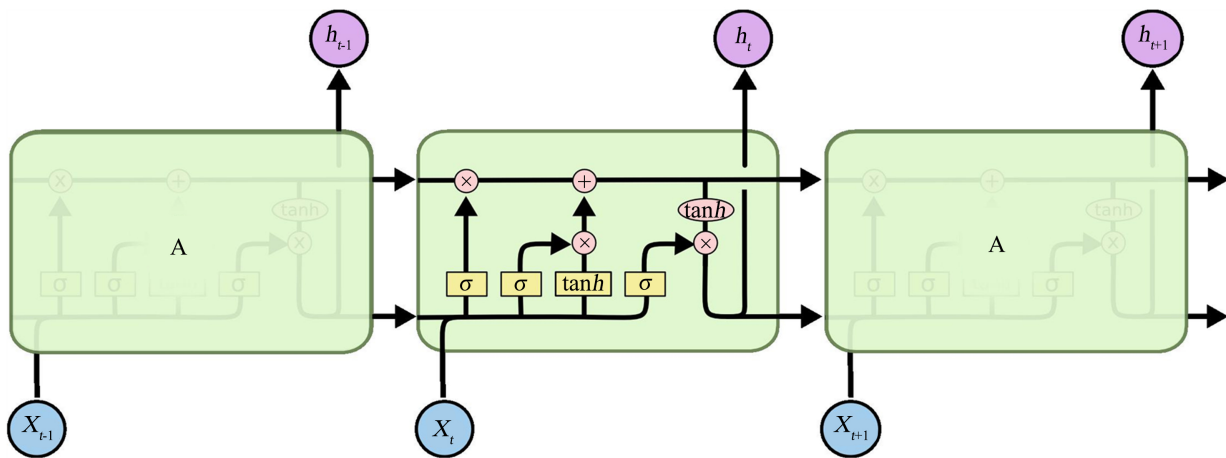


Figure 2. A basic LSTM chain architecture.

In **Figure 2**, the forget gate, determines the information to be excluded from the cell state. The input gate decides which values to update in the cell state. The cell state is responsible for storing long-term information. The output gate controls the information to be output from the cell state. h_t represents the hidden state, encompassing short-term information for output.

Figure 3 demonstrates a BiLSTM architecture³. Within an LSTM unit, key components encompass a cell, an input gate, a forget gate, and an output gate. These components work together to selectively retain and update information over sequential steps. The key innovation in BiLSTM is its bidirectional architecture. Instead of processing the input sequence only from left to right (as in traditional LSTMs), BiLSTM processes it in both directions simultaneously. In the forward pass, the input sequence is handled from the start to the finish. The hidden states at each step capture the information up to that point, considering the context from the past. Simultaneously, in the backward pass, the input sequence is processed from the end to the beginning. The hidden states in this direction capture information considering the context of the future. At each time step, the concatenation of hidden states from both the forward and backward passes takes place. This results in a representation that encodes information from both the past and the future, providing a more comprehensive context for each element in the sequence. By considering bidirectional dependencies, BiLSTM has an enhanced ability to capture long-range dependencies and understand the context in which each element appears within the sequence. BiLSTM is trained using *backpropagation through time* (BPTT) and optimized using techniques like *gradient clipping* to address vanishing or exploding gradient issues. **Algorithm 1** captures the steps in BiLSTM model.

²<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.

³<https://8f430952.rocketcdn.me/wp-content/uploads/2021/07/image-5.jpeg>.

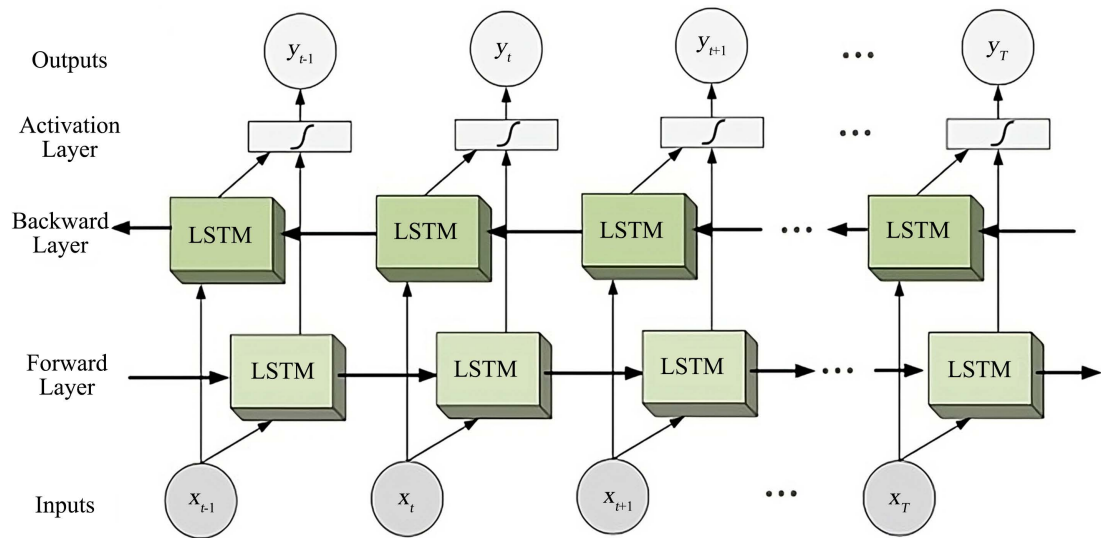


Figure 3. A BiLSTM architecture.

Algorithm 1. Bidirectional Long Short-Term Memory (BiLSTM).

Input: Input sequence X of length T

Output: Output sequence h

- 1 Initialize forward and backward LSTM networks;
- 2 Initialize parameters, weights, and biases;
- 3 **for** $t \leftarrow 1$ **to** T **do**
- 4 Forward pass through the forward LSTM network: $h_t^f \leftarrow \text{LSTM}_{\text{forward}}(x_t, h_{t-1}^f, c_{t-1}^f)$;
- 5 Backward pass through the backward LSTM network: $h_t^b \leftarrow \text{LSTM}_{\text{backward}}(x_t, h_{t+1}^b, c_{t+1}^b)$;
- 6 Concatenate forward and backward hidden states: $h_t \leftarrow [h_t^f, h_t^b]$;
- 7 Perform further processing or use the concatenated states for downstream tasks;
- 8 **return** h

4.2. Bidirectional Gated Recurrent Unit (BiGRU)

BiGRU is an advancement of the traditional GRU neural network architecture designed to capture bidirectional dependencies in sequential data. GRU is a variant of RNN that, like LSTM, handles sequential time series. BiGRU enhances the capabilities of GRU by processing input sequences in both forward and backward directions simultaneously.

Figure 4 illustrates a basic GRU architecture⁴. It consists of a reset gate and an update gate, allowing it to selectively retain and update information over sequential steps. The reset gate forgets specific past information based on weighted calculation, similarly, the storage of relevant information is decided by weighted calculation by the update gate. This mechanism helps GRU capture long-term dependencies in sequential data while maintaining computational efficiency.

In the GRU model framework as mentioned in **Figure 4**, r_t acts as the reset gate, and z_t operates as the update gate. The vector $\tilde{h}_t$ signifies a candidate activation vector. The functions σ and $\tanh$ denote the sigmoid and hyperbolic tangent functions, respectively.

⁴https://www.researchgate.net/figure/GRU-structure-diagram_fig2_355705028.

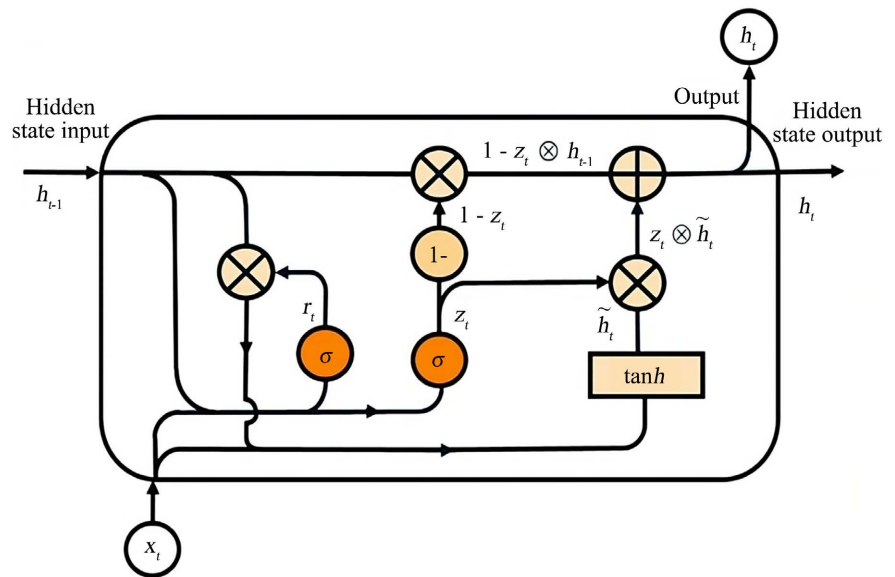


Figure 4. A GRU architecture.

Figure 5 illustrates the architecture of BiGRU⁵. BiGRU is employed for tasks involving sequential data analysis, where understanding the context of information requires considering both preceding and succeeding information. The key innovation in BiGRU is its bidirectional architecture. Instead of processing the input sequence only from left to right, as in traditional GRUs, BiGRU processes it in both directions simultaneously similar to the process explained for BiLSTM model earlier. By considering bidirectional dependencies, BiGRU has an enhanced ability to capture long-range dependencies and understand the context in which each element appears within the sequence. **Algorithm 2** captures the steps in BiGRU model.

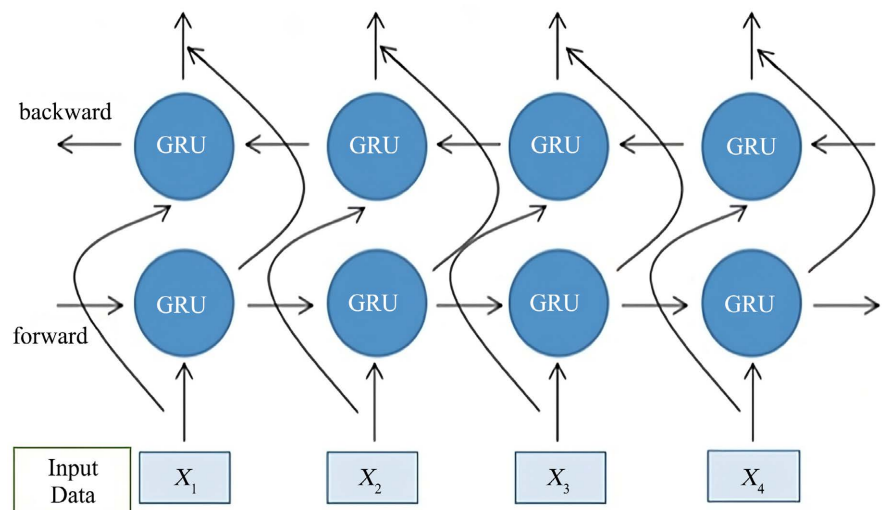


Figure 5. BiGRU architecture.

⁵<https://www.researchgate.net/publication/366162819/figure/fig1/AS:11431281106354346@1670632077172/BiGRU-model-structure-diagram.jpg>.

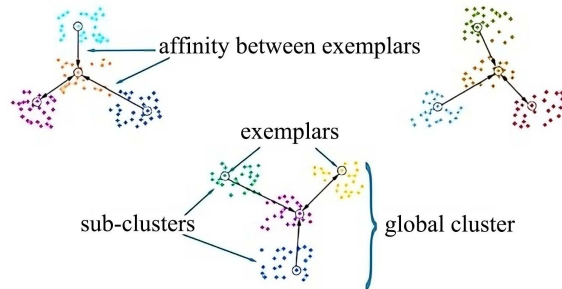
Algorithm 2. Bidirectional GRU.

Data: Input sequence $\mathbf{X} = (x_1, x_2, \dots, x_T)$
Result: Bidirectional GRU output $\mathbf{H} = (\vec{h}_1, \vec{h}_2, \dots, \vec{h}_T, \overleftarrow{h}_T, \dots, \overleftarrow{h}_1)$

- 1 Initialize GRU forward and backward hidden states: $\vec{h}_0 = \overleftarrow{h}_{T+1} = \mathbf{0}$;
- 2 **for** $t = 1$ **to** T **do**
- 3 Perform forward GRU step;
- 4 $\vec{h}_t = \text{GRU_Step}(\vec{h}_{t-1}, x_t)$;
- 5 Perform backward GRU step;
- 6 $\overleftarrow{h}_t = \text{GRU_Step}(\overleftarrow{h}_{t+1}, x_t)$;
- 7 Concatenate forward and backward hidden states;
- 8 $\mathbf{H}[t] = [\vec{h}_t, \overleftarrow{h}_t]$;
- 9 **return** $\mathbf{H}$;

4.3. Clustering: Affinity Propagation

It is a type of clustering algorithm that autonomously identifies clusters and representative data points, known as exemplars, within a dataset. It operates by iteratively exchanging messages between data points based on their pairwise similarities, represented in a similarity matrix. The algorithm maintains responsibility and availability matrices, updating them to assess the suitability of each data point to act as an exemplar for others. Exemplars are chosen by maximizing the sum of responsibility and availability, facilitating the formation of clusters around these influential points. **Figure 6** illustrates the Affinity Propagation⁶. Notably, Affinity Propagation dynamically determines the number of clusters, making it advantageous when prior knowledge of cluster count is unavailable. The algorithm converges when the matrices stabilize, ensuring consistent exemplar selection and cluster assignment.

**Figure 6.** Affinity propagation.

This clustering algorithm operates by iteratively updating two key matrices, responsibility (R) and availability (A), to determine representative data points, called exemplars, within a dataset. **Figure 6** depicts the Affinity Propagation technique for clustering. The similarity matrix S quantifies pairwise similarities between data points. The responsibility matrix R represents each point's suitability to serve as an exemplar for others, considering alternative exemplars. Simultaneously, the availability matrix A accumulates evidence for a point to choose

⁶https://www.frontiersin.org/files/Articles/10792/fninf-05-00018-r1/image_m/fninf-05-00018-g003.jpg.

another as its exemplar. The iterative updates are influenced by the damping factor Λ , which prevents oscillations and controls the impact of new values on existing ones. Exemplars are chosen based on maximizing the sum of responsibility and availability. This dynamic process of message passing continues until convergence, resulting in stable matrices and consistent exemplar selection. Affinity Propagation's strength lies in its ability to automatically determine the number of clusters and exemplars without requiring prior information.

The key equations for the affinity propagation are:

Responsibility (R) update:

$$R(i, k) \leftarrow S(i, k) - \max_{k' \neq k} \{A(i, k') + S(i, k')\}$$

Availability (A) Update:

$$A(i, k) \leftarrow \min \left\{ 0, R(k, k) + \sum_{i' \neq i, i' \neq k} \max \{0, R(i', k)\} \right\}$$

Damping: Both the responsibility and availability matrices are updated with a damping factor (Λ) to prevent oscillations:

$$R(i, k) \leftarrow \Lambda \cdot \text{old}R(i, k) + (1 - \Lambda) \cdot R(i, k)$$

$$A(i, k) \leftarrow \Lambda \cdot \text{old}A(i, k) + (1 - \Lambda) \cdot A(i, k)$$

Exemplar Assignment:

$$\text{Exemplar for } i \text{ is } \arg \max_k \{R(i, k) + A(i, k)\}$$

$S(i, k)$ represents the similarity between data points i and k .

$R(i, k)$ is the responsibility of point i to be the exemplar for point k .

$A(i, k)$ is the availability of point k for being the exemplar for point i .

Λ is the damping factor that controls the influence of the new values on the existing ones.

Algorithm 3 provides an overview of the affinity propagation algorithm.

Algorithm 3. Affinity propagation.

Data: Similarity matrix S
Result: Exemplars and clusters

- 1 Initialize availability matrix A and responsibility matrix R to zeros;
- 2 Set damping factor $\text{damping} = 0.5$;
- 3 **while** *not converged* **do**
- 4 **for each data point** i **do**
- 5 **for each exemplar** k **do**
- 6 $R(i, k) = S(i, k) - \max_{k' \neq k} (A(i, k') + S(i, k'))$ for all $k' \neq k$;
- 7 **for each data point** i **do**
- 8 **for each exemplar** k **do**
- 9 $A(i, k) = \min(0, R(k, k) + \sum \max(0, R(i', k)))$ for all $i' \neq i$;
- 10 $R = \text{damping} \cdot \text{old}R + (1 - \text{damping}) \cdot R$;
- 11 $A = \text{damping} \cdot \text{old}A + (1 - \text{damping}) \cdot A$;
- 12 **for each data point** i **do**
- 13 Set i as an exemplar if $A(i, i) + R(i, i) > 0$;
- 14 **for each data point** i **do**
- 15 Assign i to the exemplar with the highest net responsibility ($R + A$);

4.4. Evaluation Metrics

For quantifying model performance, this study has explored four popular evaluation metrics. These metrics are used to quantify how well the deep learning models are in forecasting future values.

Mean Squared Error (MSE): It assesses the average squared discrepancy between the observed values and the true values in a dataset. It is given by

$$\text{MSE} = \frac{1}{m} \sum_{j=1}^m (b_j - \hat{b}_j)^2;$$

here b_j is the original price and the $\hat{b}_j$ is the forecasted value.

Mean Absolute Error (MAE): It quantifies the mean extent of errors between forecasted and genuine values in a dataset. Mathematically,

$$\text{MAE} = \frac{1}{m} \sum_{j=1}^m |b_j - \hat{b}_j|$$

where the number of observations in the dataset is represented by m . The true value for observation j is represented by b_j . The predicted value for observation j is represented by $\hat{b}_j$.

R-Square: In a regression model, the R-squared metric quantifies the ratio of the variance in the dependent variable that the independent variables account for. Mathematically,

$$R^2 = 1 - \frac{\text{SSR}}{\text{SST}}$$

where SSR is the sum of squared discrepancies between the observed values and the predicted values and SST is the total sum of squared discrepancies between the observed values and the average of the observed values.

Mean Absolute Percentage Error (MAPE): It serves as a measure to evaluate the precision of a forecasting model by determining the mean percentage disparity between anticipated and real values. Its significance lies in appraising model performance across varied datasets, furnishing a standardized gauge of predictive accuracy. Mathematically,

$$\text{MAPE} = \frac{1}{m} \sum_{j=1}^m \left| \frac{A_j - P_j}{A_j} \right| \times 100.$$

The number of observations in the dataset is represented by m . A_j represents the actual value for observation j . P_j represents the predicted value for observation j .

4.5. Data Driven Portfolio Optimization

Two innovative estimators were introduced in [30]. These functions, characterized by smaller variances, are employed to formulate a novel risk measure for a portfolio that takes into account moments of higher order. The sign correlation ($\rho_{P,\text{sgn}}$) and volatility correlation ($\rho_{P,\text{vol}}$) are expressed as:

$$\rho_{P,\text{sgn}} = \text{Corr}\left(\text{Sgn}\left(R_p - \mu_p\right), R_p - \mu_p\right)$$

$$\rho_{P,\text{vol}} = \text{Corr}\left(\left|R_P - \mu_p\right|, \left(R_P - \mu_p\right)^2\right)$$

Here, μ_p represents the expected return of the portfolio. These are used in [31] to introduce four distinct risk measures. The portfolio's expected return is indicated by μ_p . R_P represents the portfolio return. The inverse of the cumulative distribution function (CDF) is denoted by $F(x)$ of portfolio return. $\tilde{R}_p$ and $\hat{\sigma}_p$ serve as estimates of the expected return and standard deviation (SD) of a financial portfolio, respectively, derived from the past l observations.

The mean absolute deviation (MAD) assessment for the portfolio is computed using:

$$\text{MAD}_P = 2\hat{\rho}_{P,\text{sign}}\hat{\sigma}_P\sqrt{F(\tilde{R}_p)(1-F(\tilde{R}_p))}$$

The volatility estimate of the portfolio, employing volatility correlation reduction (VEV), is quantified through:

$$\text{VEV}_P = \sqrt{1 - \hat{\rho}_{P,\text{vol}}^2}\hat{\sigma}_P$$

The volatility assessment of the portfolio utilizing sign correlation reduction (VES) is given below:

$$\text{VES}_P = \sqrt{1 - \hat{\rho}_{P,\text{sign}}^2}\hat{\sigma}_P$$

For the volatility estimate of the portfolio considering both volatility and sign correlation reduction (VESV), the formula is as follows:

$$\text{VESV}_P = \sqrt{(1 - \hat{\rho}_{P,\text{sign}}^2)(1 - \hat{\rho}_{P,\text{vol}}^2)}\hat{\sigma}_P$$

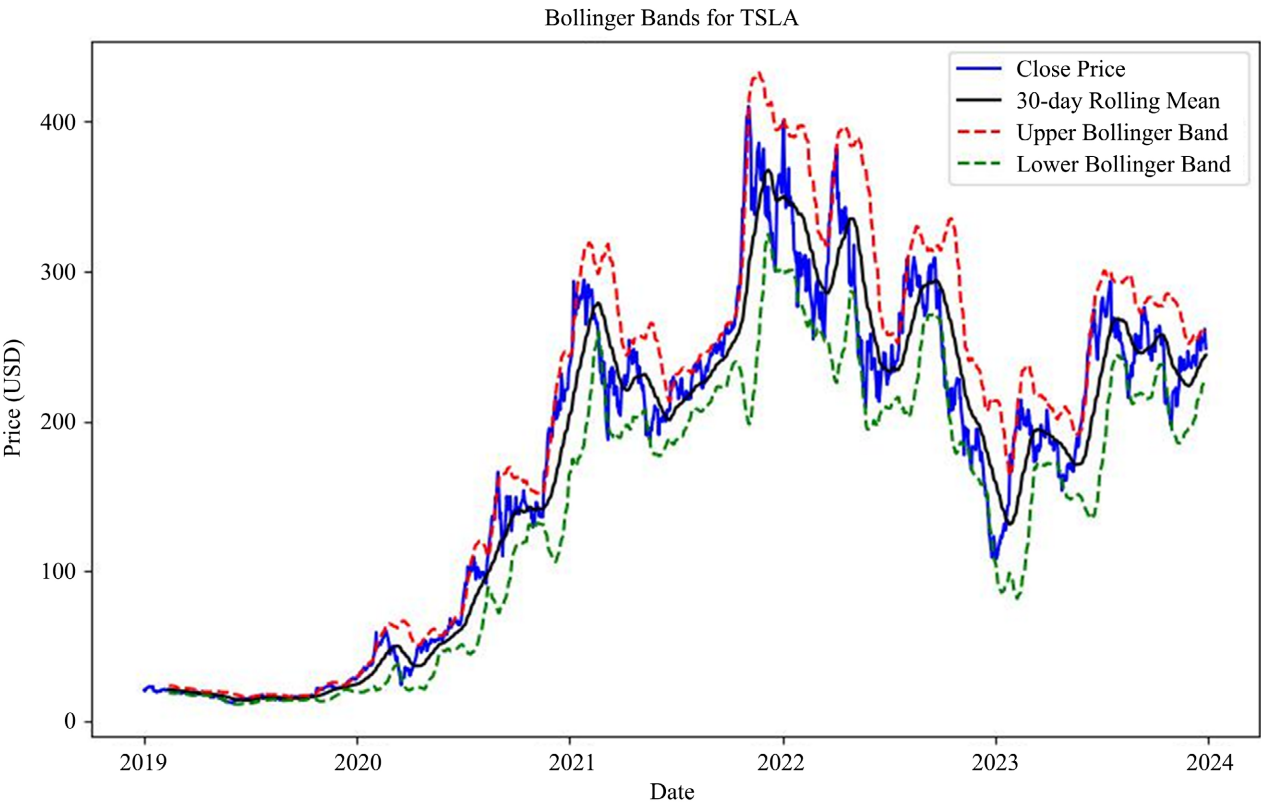
The data-driven portfolio optimization strategy considers these four distinct data-driven risk measures. Unlike other existing traditional portfolio optimization techniques, the data-driven portfolio optimization algorithm considers the underlying non-normality of the financial time series.

5. Results and Discussions

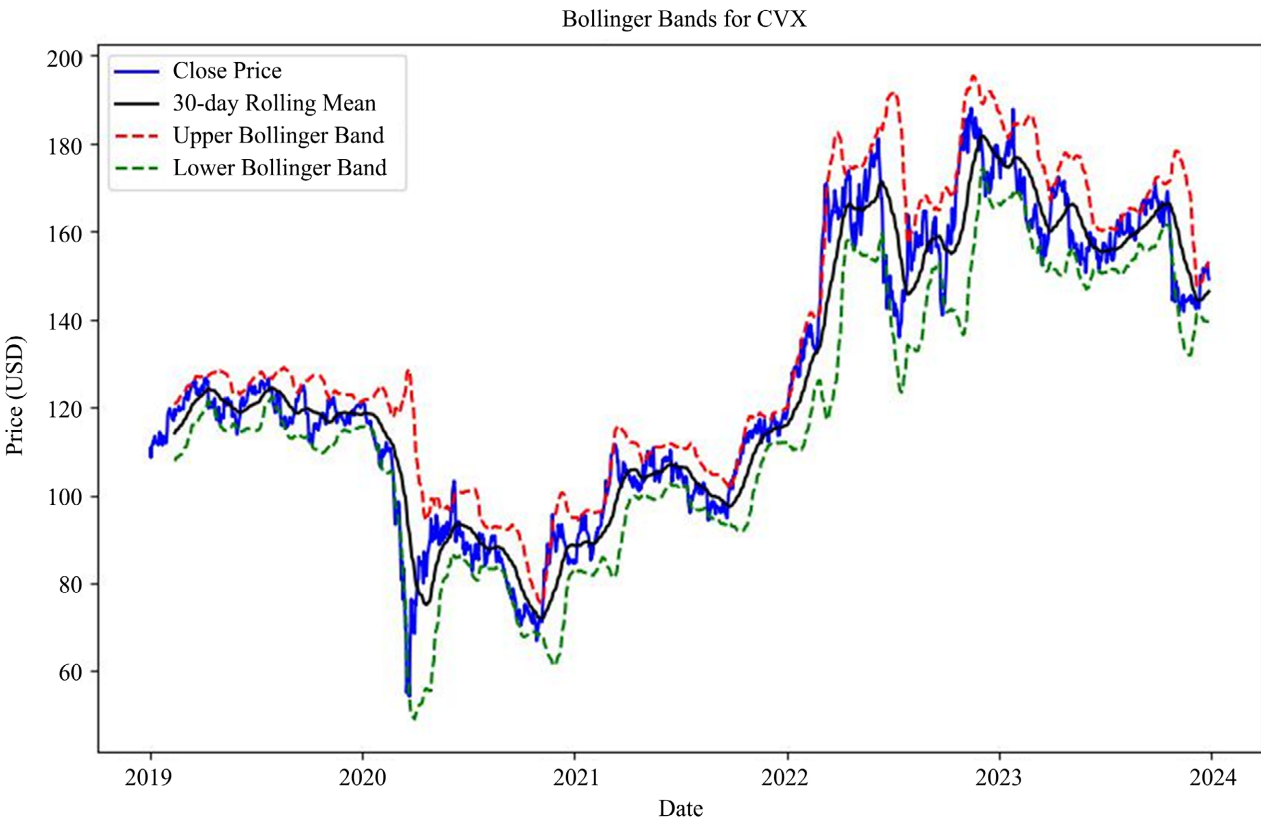
5.1. Exploratory Data Analysis

The dataset considered for this research is from 1998 to 2023. This dataset captures all major and minor events in this timeframe. Thus the BiLSTM and BuGRU models learn from the patterns of the drastic market events. This section explores interesting market intricacies especially from 2019 to 2023.

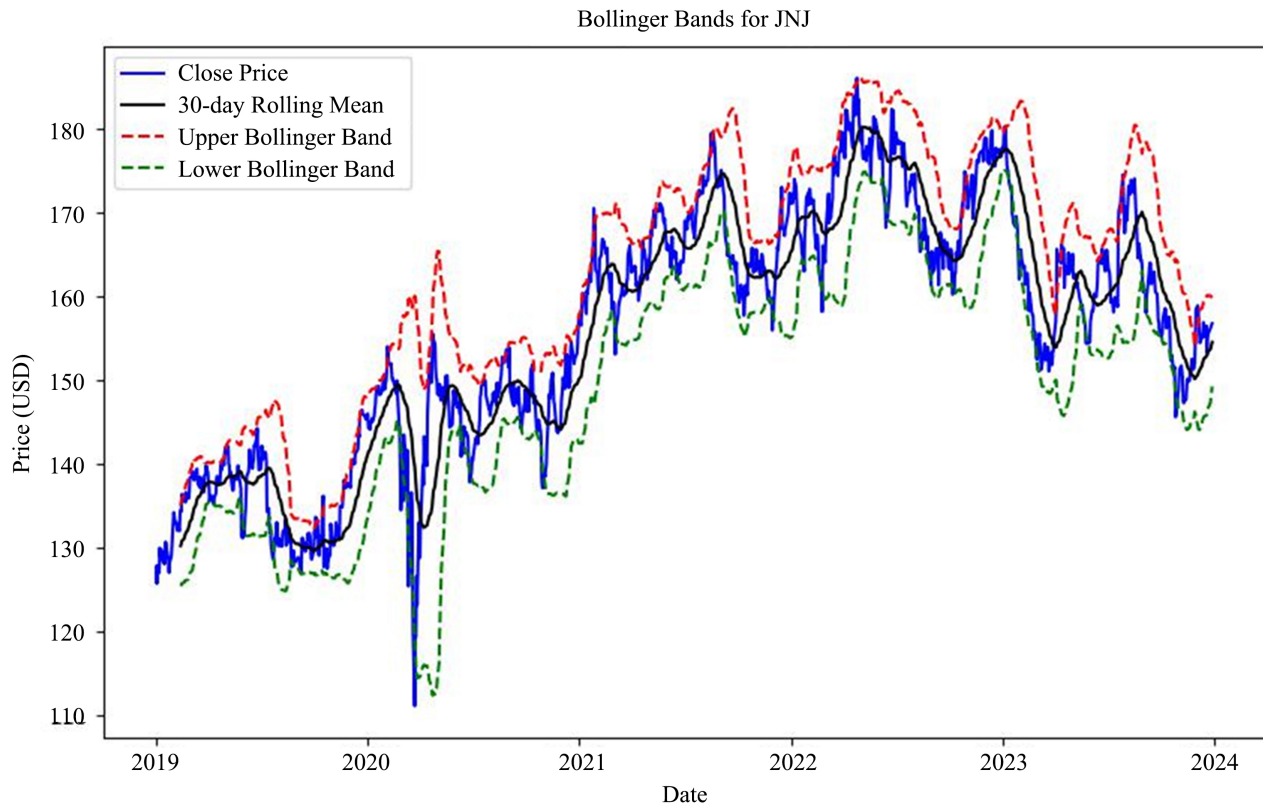
Examining **Figure 7** and **Table 1**, it becomes evident that Tesla (TSLA) exhibits the highest level of volatility. In addition to Tesla (TSLA), Chevron (CVX) demonstrates notable volatility, albeit not to the extent observed in Tesla. Beyond traditional stocks, Bitcoin (BTC) manifests the utmost volatility, attributed to the inherently unpredictable nature of cryptocurrencies. Forecasting the performance of fluctuating assets proves challenging, primarily due to their pronounced volatility. Consequently, the formulation of portfolios incorporating highly unpredictable cryptocurrencies becomes a formidable task.



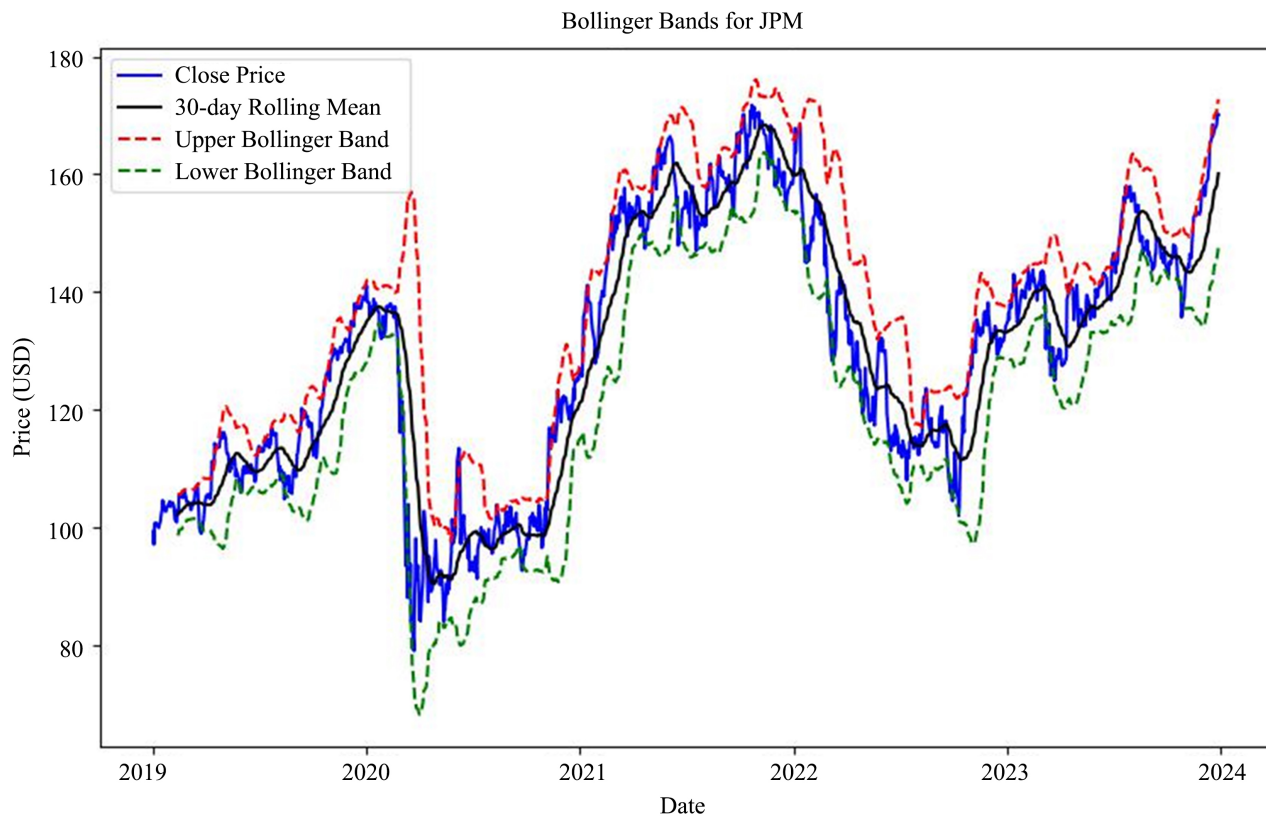
(a)



(b)



(c)



(d)

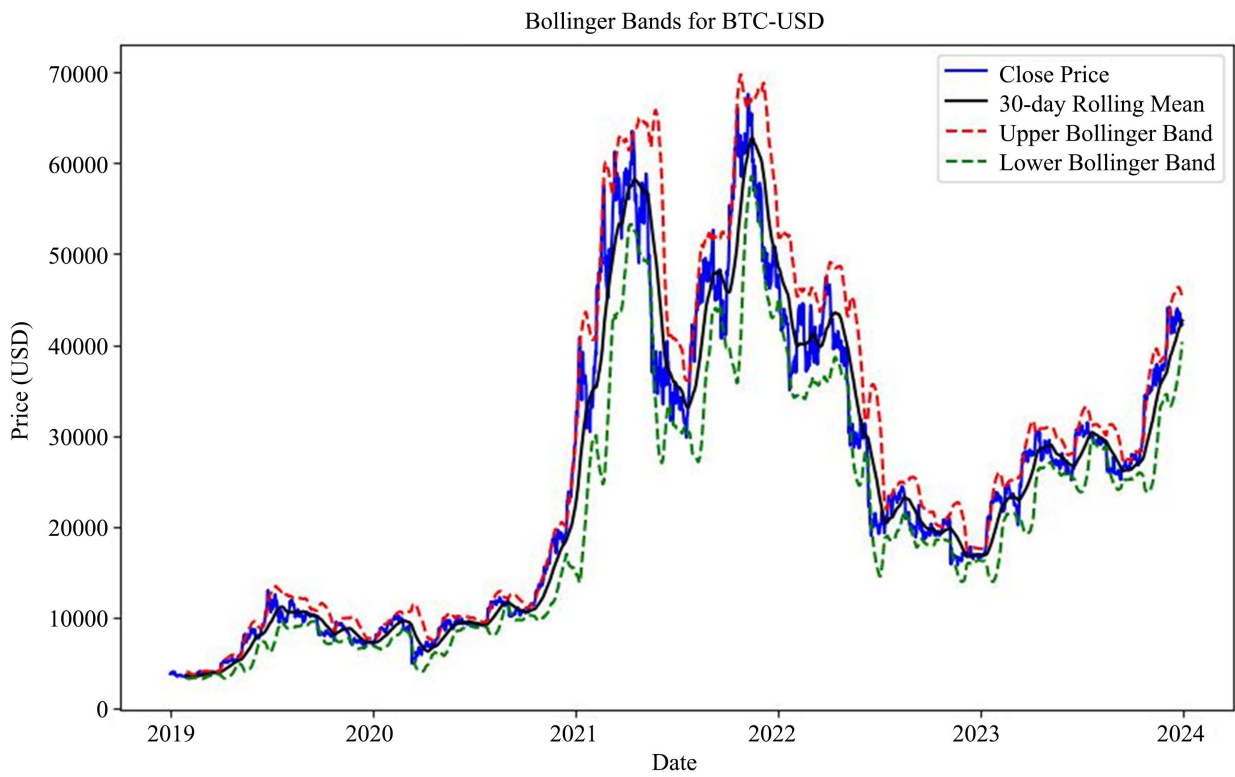
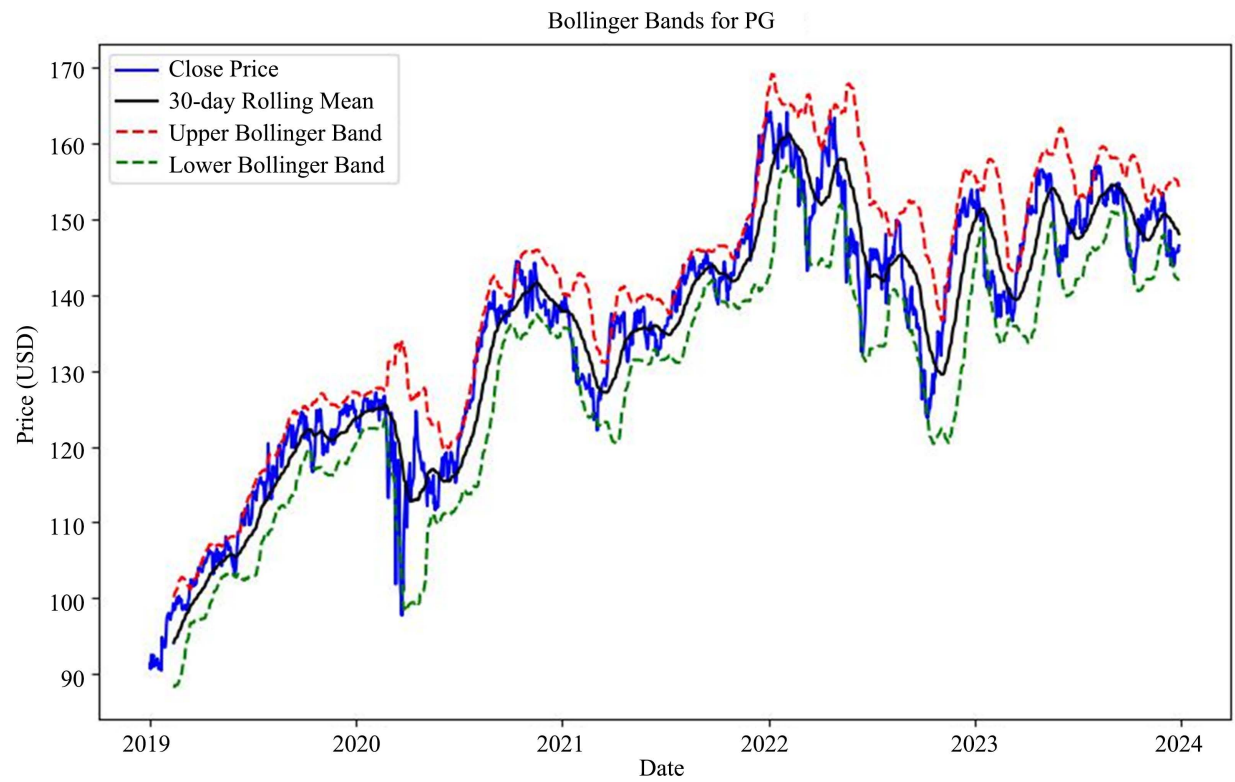


Figure 7. Bollinger bands for stocks and cryptos from 2019 to 2023. (a) Tesla, (b) Chevron, (c) Johnson & Johnson, (d) JP Morgan & Chase, (e) The Procter & Gamble Company, (f) Bitcoin.

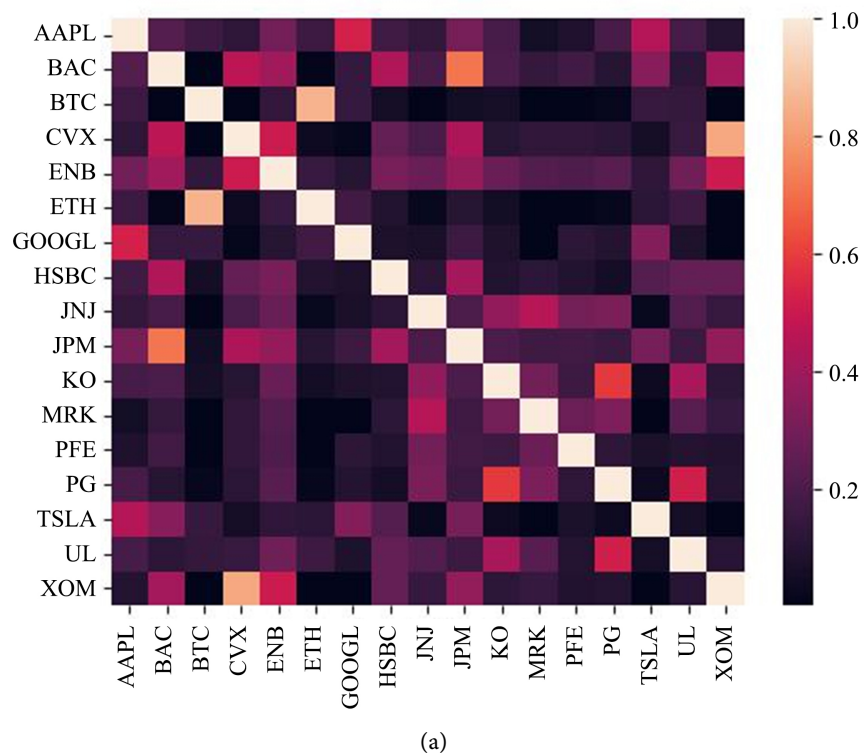
Table 1. Volatility with 20 days rolling window.

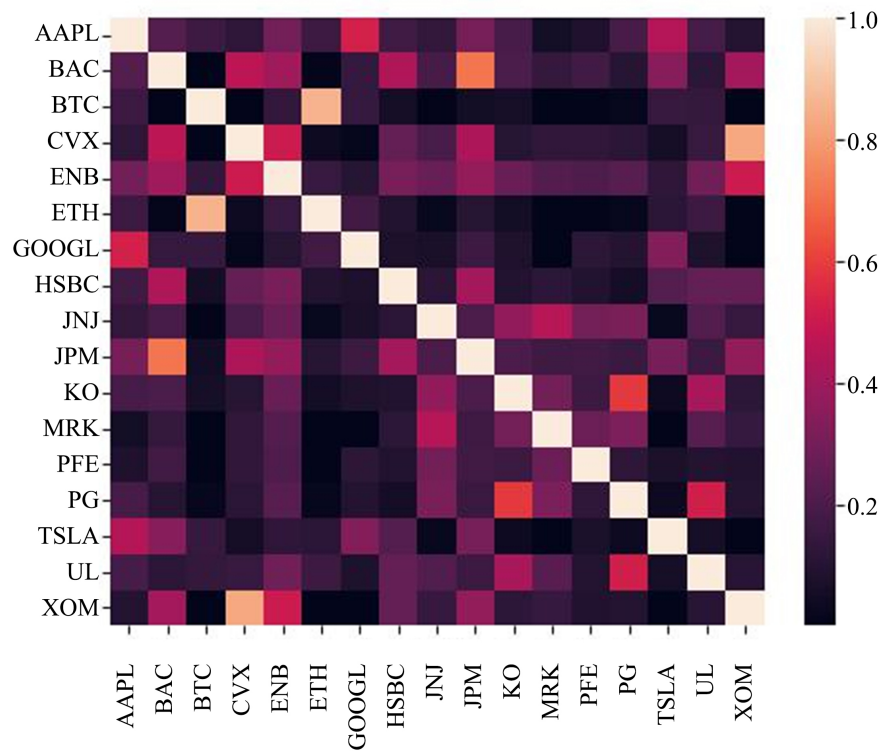
Components	Volatility
TSLA	108.30
JNJ	14.89
JPM	20.01
CVX	30.96
PG	16.65
BTC	16064.68

The visual representation in **Figure 7** indicates that throughout the COVID-19 pandemic in 2020, the general market performance experienced a downturn, excluding certain technology stocks such as TESLA. Both Tesla and Bitcoin also displayed a gradual increase in their prices during this period. After the conclusion of the COVID-19 era, the lingering impact of the financial crisis persisted, leading to noticeable price fluctuations in 2022.

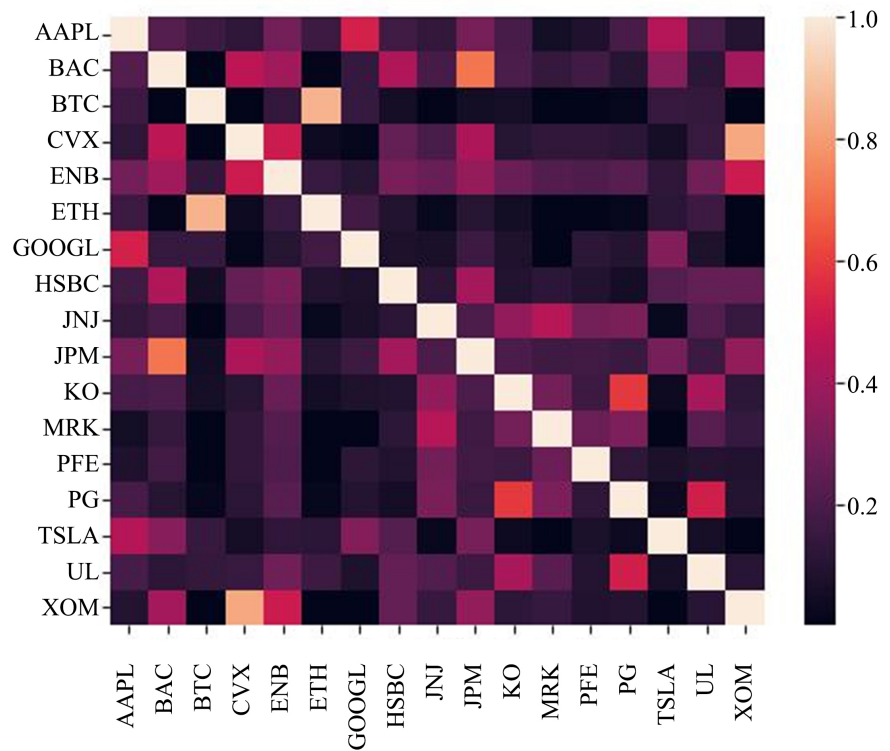
5.2. Performance Analysis of BiLSTM and BiGRU

The correlation among various financial assets analyzed in this study for the year 2023 is depicted as a heat map in **Figure 8**. The BiLSTM and BiGRU DL models were trained to forecast asset prices for the same year. In this examination, the correlation between assets is assessed using both the predicted data from the models and the actual data.





(b)



(c)

Figure 8. Correlation analysis using actual data and forecasted data of 2023. (a) Correlation with actual data; (b) Correlation with BiLSTM predicted data; (c) Correlation with BiGRU predicted data.

The correlation matrix derived from the actual data of 2023 **Figure 8(a)** reveals a robust correlation between stocks in the energy sector and those in the banking sector, particularly Bank of America (BAC). Additionally, BTC demonstrates a significant association with ETH. Furthermore, stocks within the same sector consistently exhibit elevated levels of correlation.

The correlation analysis in **Figure 8(b)** and **Figure 8(c)** indicates that the predicted data successfully incorporates all the significant correlations observed among financial assets in the actual data. This suggests that the DL models have effectively captured the correlations among assets essential for constructing portfolios.

Furthermore, upon reviewing **Table 2**, it becomes evident that BiGRU exhibits its superiority over BiLSTM in forecasting the yearly prices of financial assets. The four evaluation metrics examined in this research consistently endorse the enhanced performance of BiGRU compared to BiLSTM in the prediction of financial asset prices.

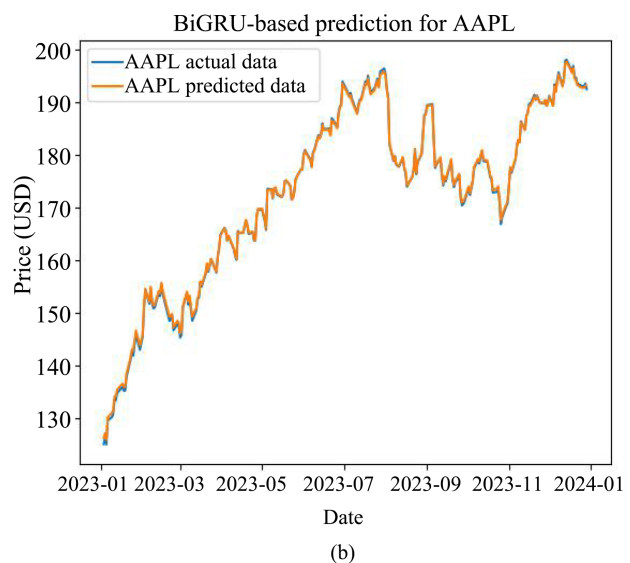
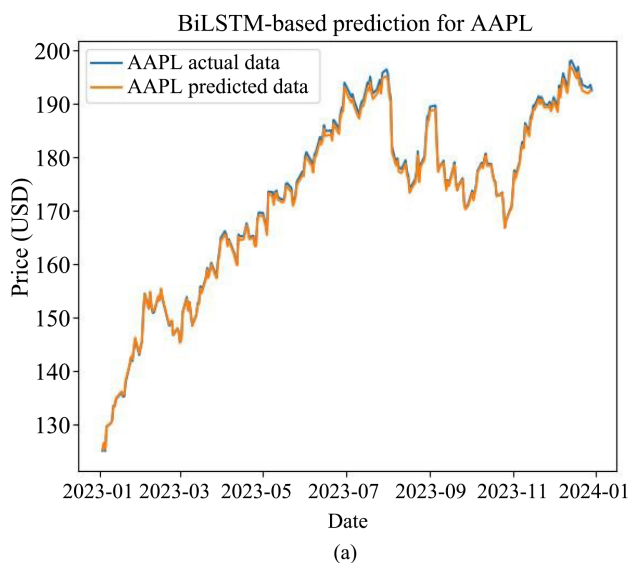
Table 2. Performance comparison of BiLSTM and BiGRU.

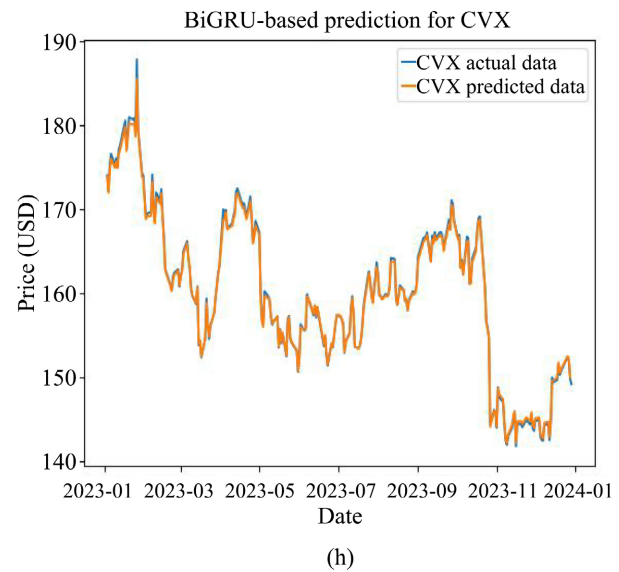
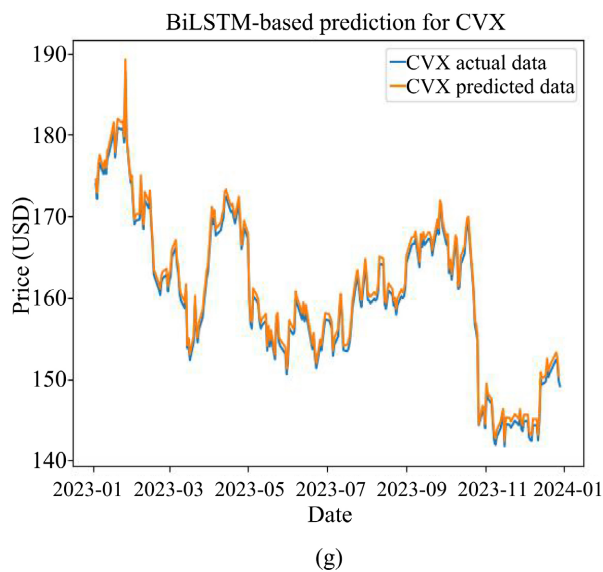
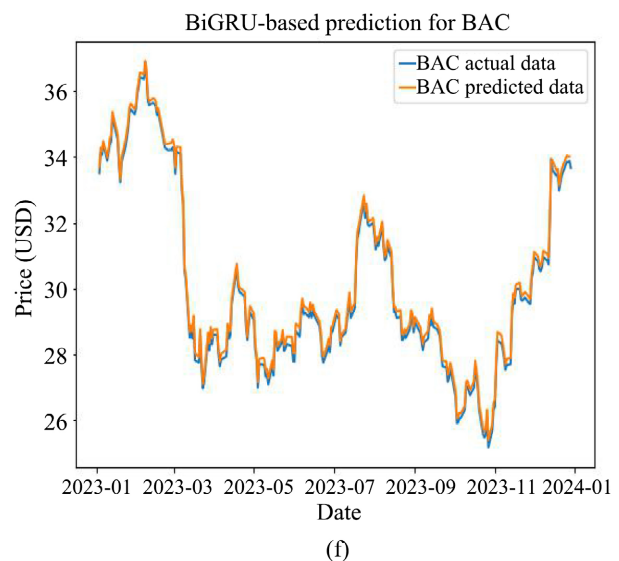
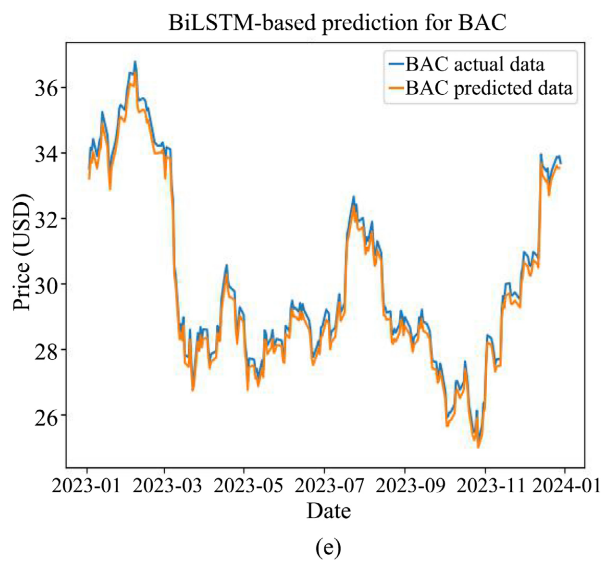
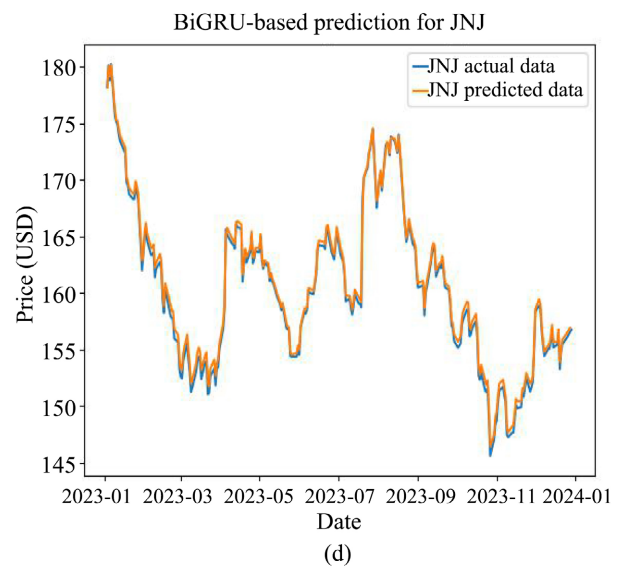
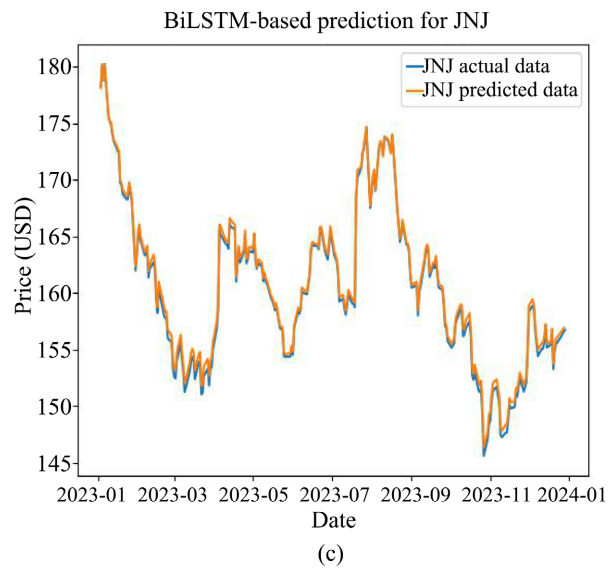
Data Series	Metric	Algorithms	
		BiLSTM	BiGRU
AAPL	MSE	0.40	0.14
	MAE	0.54	0.30
	R-Squared	0.998	0.999
	MAPE	0.30%	0.18%
BAC	MSE	0.08	0.03
	MAE	0.27	0.17
	R-Squared	0.9901	0.9959
	MAPE	0.90%	0.58%
BTC-USD	MSE	0.0006	0.0003
	MAE	0.019	0.014
	R-Squared	0.9994	0.9997
	MAPE	14.7%	11.1%
CVX	MSE	0.63	0.11
	MAE	0.78	0.25
	R-Squared	0.993	0.999
	MAPE	0.49%	0.16%
JNJ	MSE	0.2	0.2
	MAE	0.4	0.4
	R-Squared	0.99	0.99
	MAPE	0.3%	0.3%

Continued

PG	MSE	0.5	0.3
	MAE	0.7	0.5
	R-Squared	0.98	0.99
	MAPE	0.5%	0.3%

Figure 9 presents the performance of BiGRU and BiLSTM models in various financial sectors. Notably, the predictions generated by the BiGRU model exhibit greater efficiency compared to those produced by the BiLSTM model. As illustrated in **Table 2**, a comparative analysis reveals the superior performance of BiGRU-based predictions over BiLSTM for both AAPL and the highly volatile BTC-USD. Specifically, for AAPL, the evaluation metrics indicate a lower MSE of 0.14 for BiGRU in contrast to 0.40 for BiLSTM, signifying the enhanced accuracy of BiGRU predictions. This suggests that the predicted values generated by BiGRU closely align with the actual values. Moreover, BiGRU exhibits a reduced MAE compared to BiLSTM, indicating a lower error in BiGRU predictions. The error percentage for BiGRU is notably lower at 0.18%, while for BiLSTM, it stands at 0.30% in the case of AAPL. This observation is evident in **Figure 9(a)** and **Figure 9(b)**, where the predictions for AAPL using BiGRU are notably sharp and accurate in contrast to those based on BiLSTM. In the context of highly volatile cryptocurrencies, such as BTC-USD, BiGRU outperforms BiLSTM, demonstrating superior performance with a minimal Mean Squared Error (MSE) of 0.0003. Moreover, when considering Mean Absolute Error (MAE), BiGRU exhibits lower values. In terms of error percentage, predictions based on BiGRU exhibit a 0.58% error, whereas BiLSTM-based predictions show a higher error rate of 0.90%. This observation is evident in **Figure 9(k)** and **Figure 9(l)**, where the predictions for BTC generated by BiGRU are characterized by sharp precision, contrasting with the less precise predictions produced by BiLSTM.





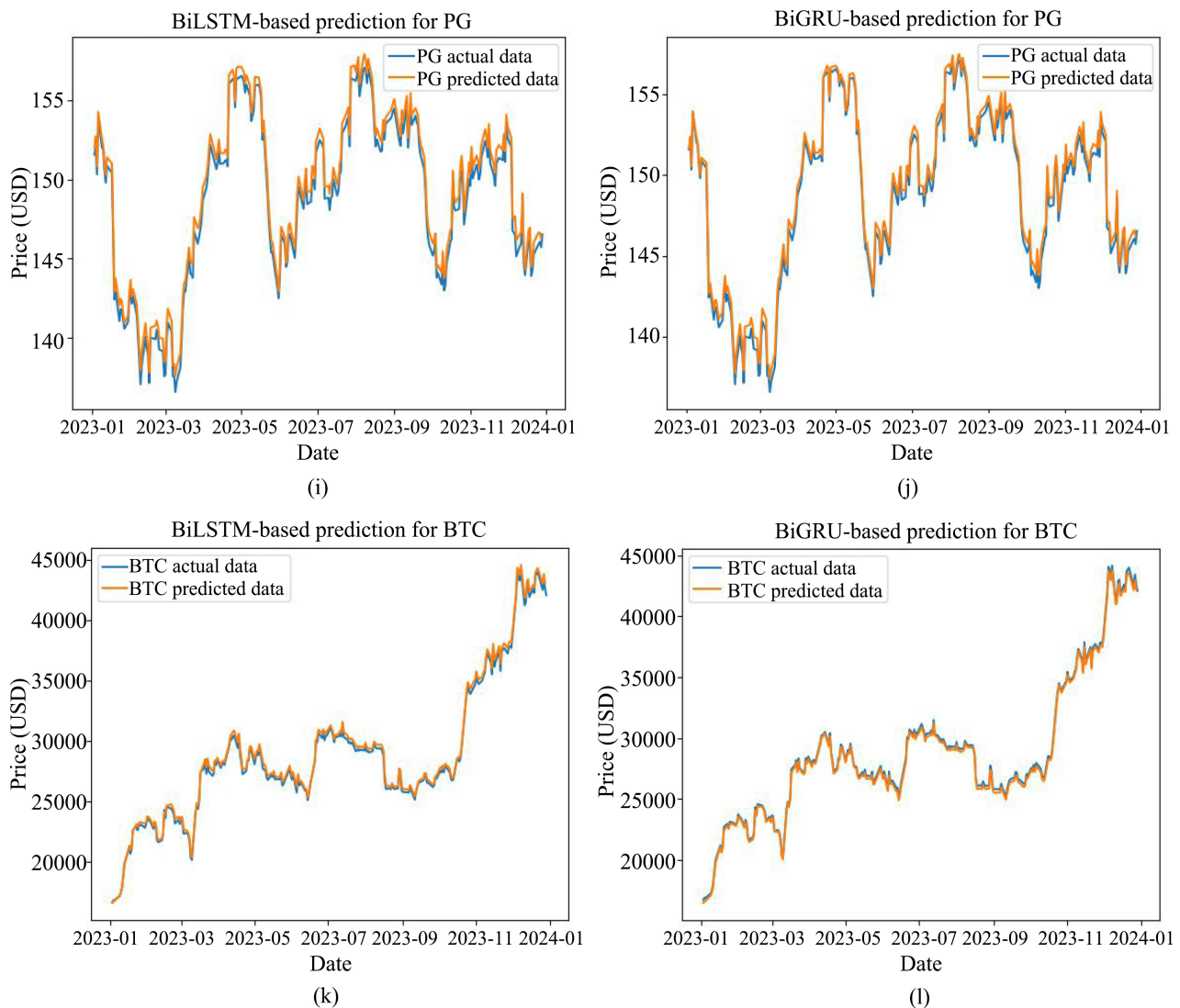


Figure 9. BiLSTM and BiGRU predictions. (a) BiLSTM-based AAPL prediction; (b) BiGRU-based AAPL prediction; (c) BiLSTM-based JNJ prediction; (d) BiGRU-based JNJ prediction; (e) BiLSTM-based BAC prediction; (f) BiGRU-based BAC prediction; (g) BiLSTM-based CVX prediction; (h) BiGRU-based CVX prediction; (i) BiLSTM-based PG prediction; (j) BiGRU-based PG prediction; (k) BiLSTM-based BTC prediction; (l) BiGRU-based BTC prediction.

Similar observations can be made for the financial assets of the other sectors as well. Therefore, While both BiLSTM and BiGRU demonstrate accurate price predictions and closely align with the actual data, the BiGRU algorithm outperforms the BiLSTM algorithm, particularly for highly unpredictable assets, showcasing greater accuracy.

5.3. Portfolio Performance

In the creation of financial portfolios, this study explores the Data-Driven approach, taking into account essential features of financial time series, such as non-normality, which are often overlooked by traditional methods. These features play a crucial role in constructing robust portfolios. Additionally, the study

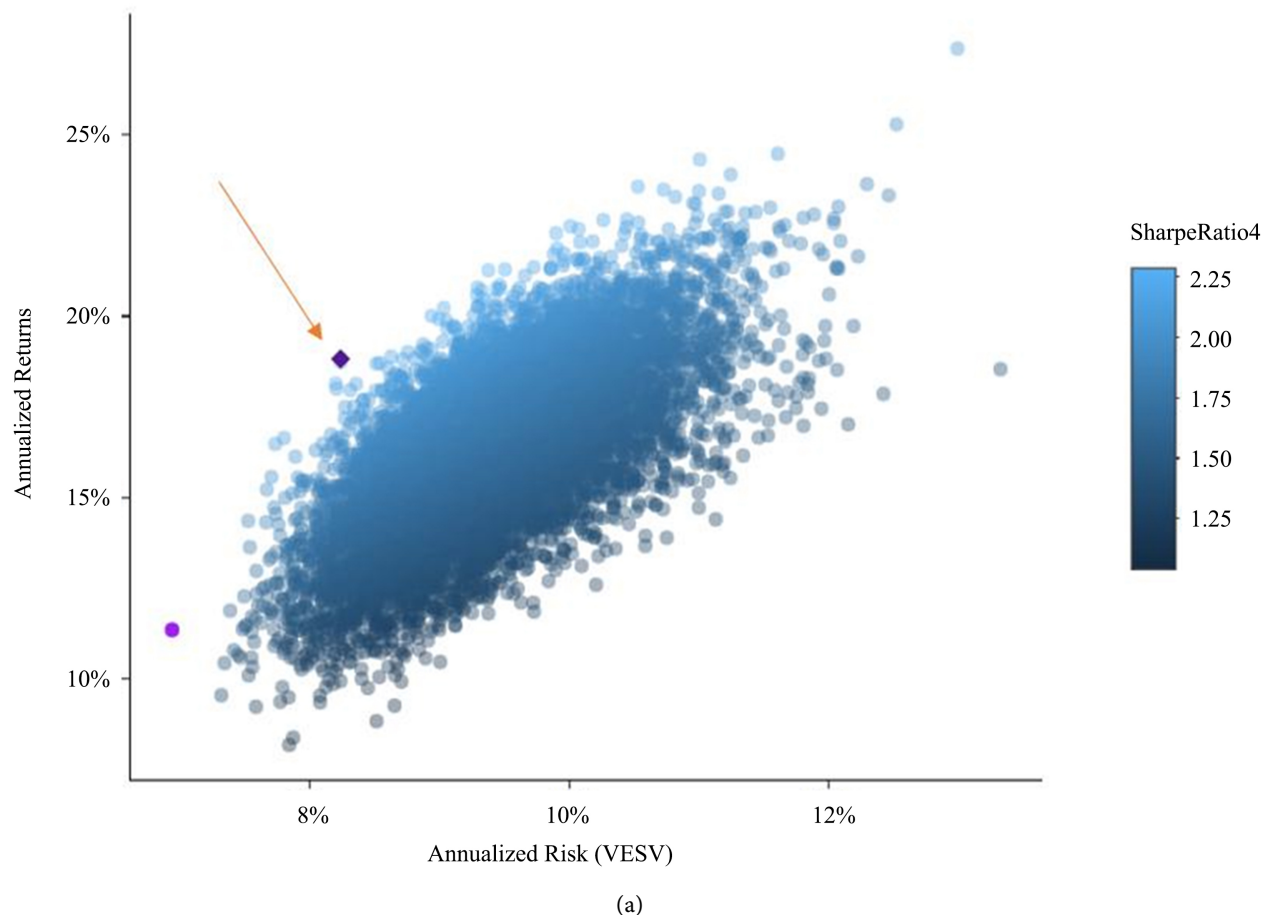
incorporates the concept of diversification in stock selection. The performance of portfolios is evaluated under two scenarios: 1) where all stocks are included without specific selection (no diversification), and 2) where the affinity propagation clustering technique is employed to introduce diversity in stock selection. Consequently, the study assesses the performance of data-driven portfolios with and without clustering, providing insights into the impact of diversification on portfolio performance.

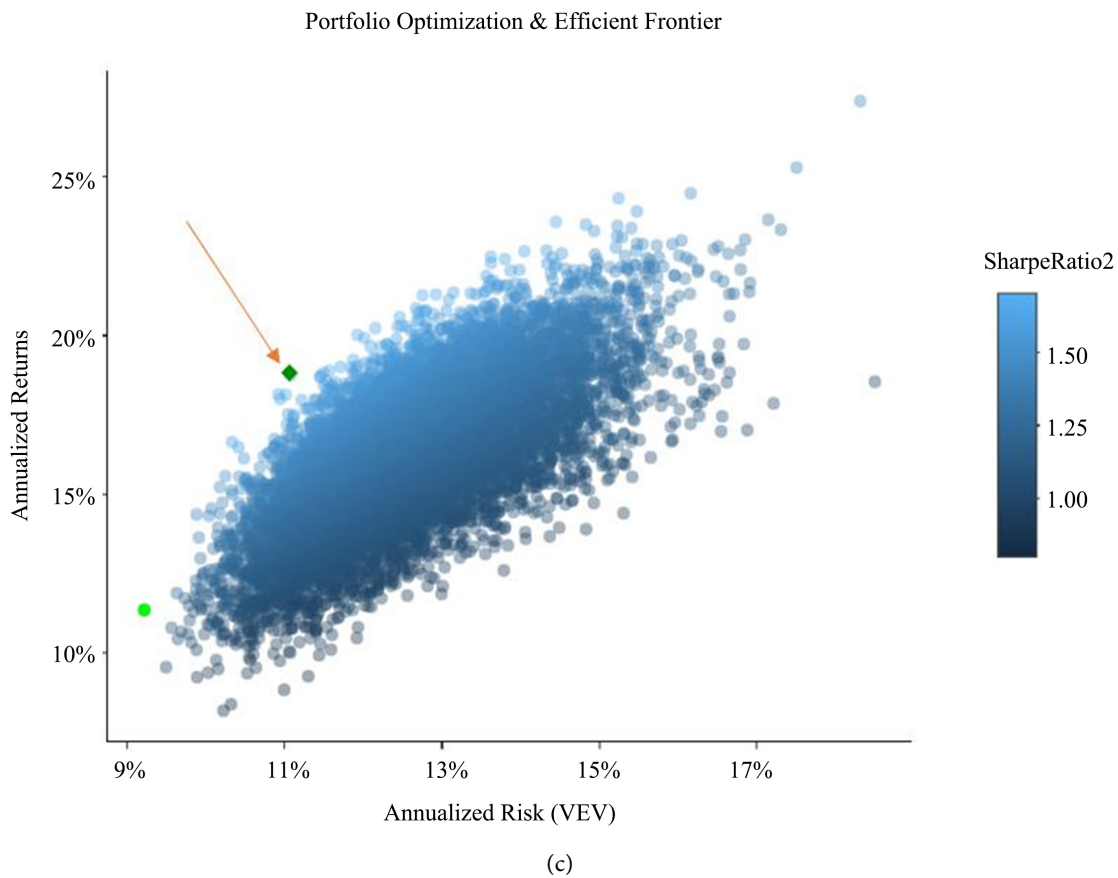
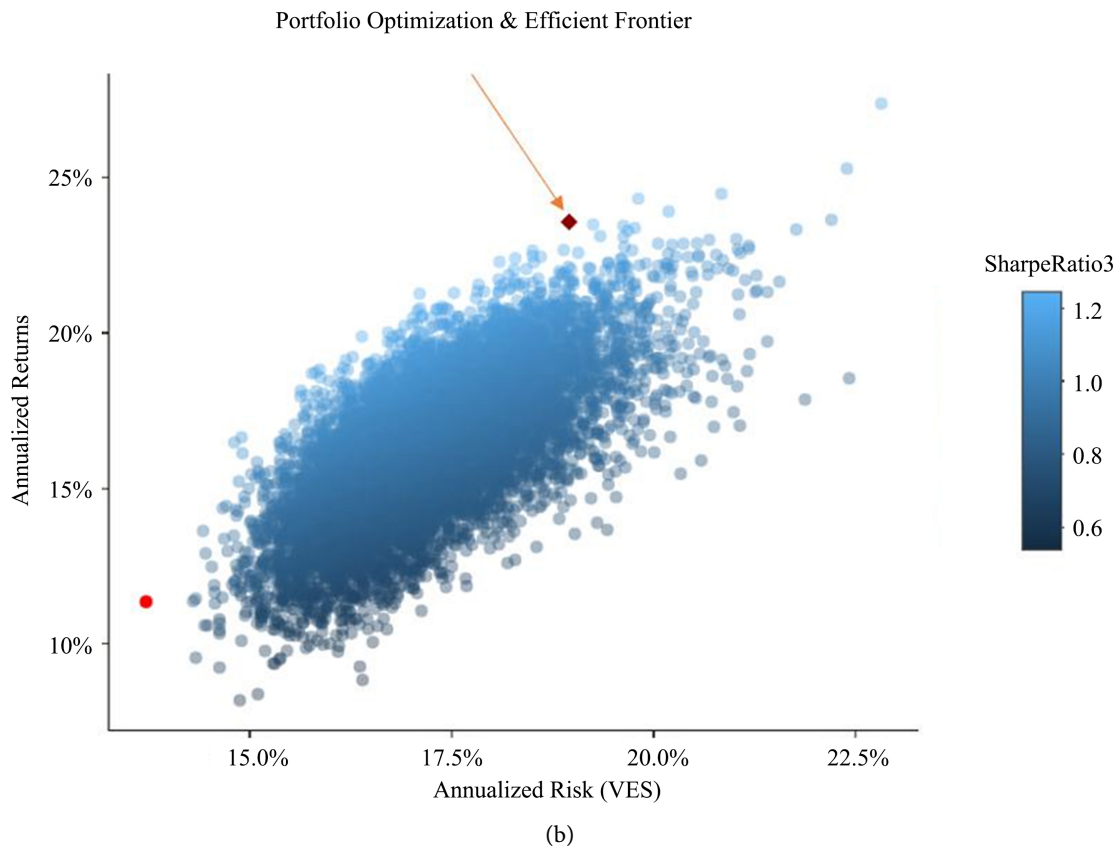
5.3.1. Without Diversification

In this research, we build portfolios encompassing 15 stocks and 2 cryptocurrencies, incorporating both real and forecasted data.

The efficient frontiers of the constructed portfolios are alike for both predicted and actual data, given the close resemblance between the actual values and the predicted values generated by both BiLSTM and BiGRU. The efficient frontiers constructed with four data-driven risk measures are illustrated in **Figure 10**. This figure presents the efficient frontiers for portfolios formed with four data-driven risk measures. In **Figure 10(b)**, it can be observed that the portfolio constructed with VES-based risk measure showed maximum profitability.

Portfolio Optimization & Efficient Frontier





Portfolio Optimization & Efficient Frontier

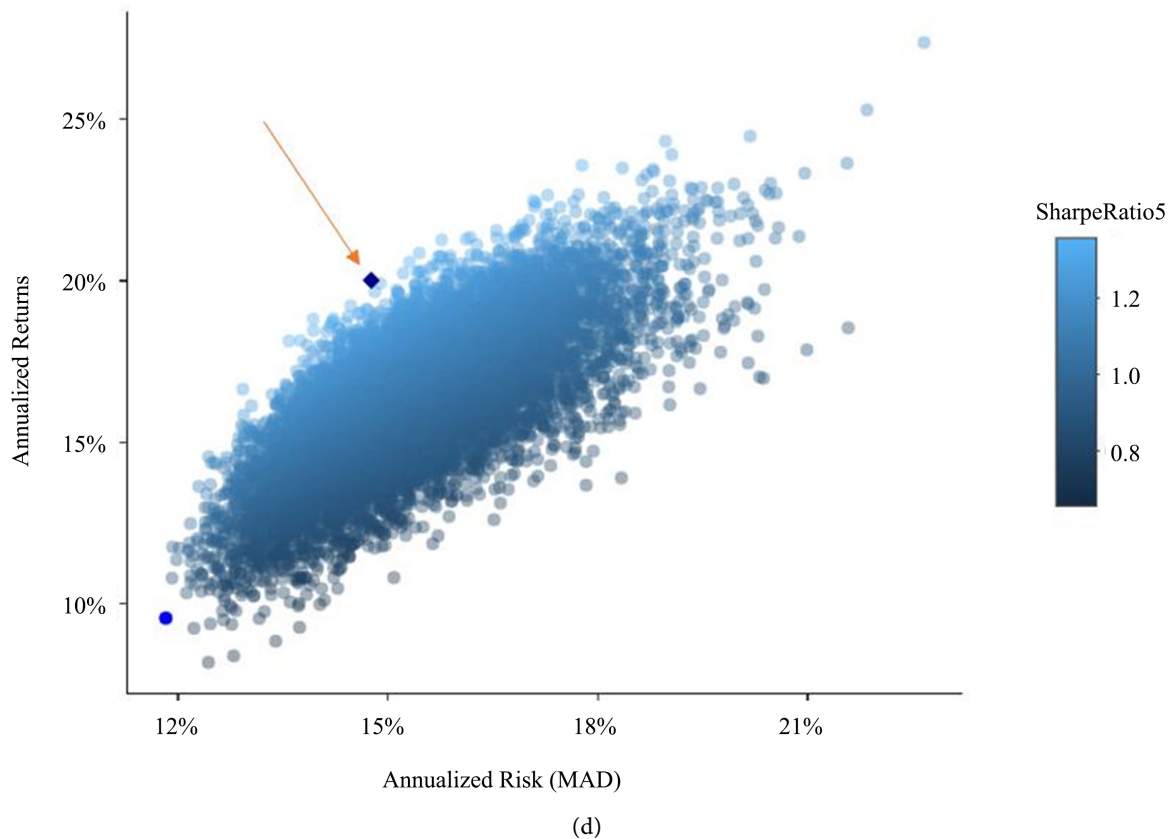


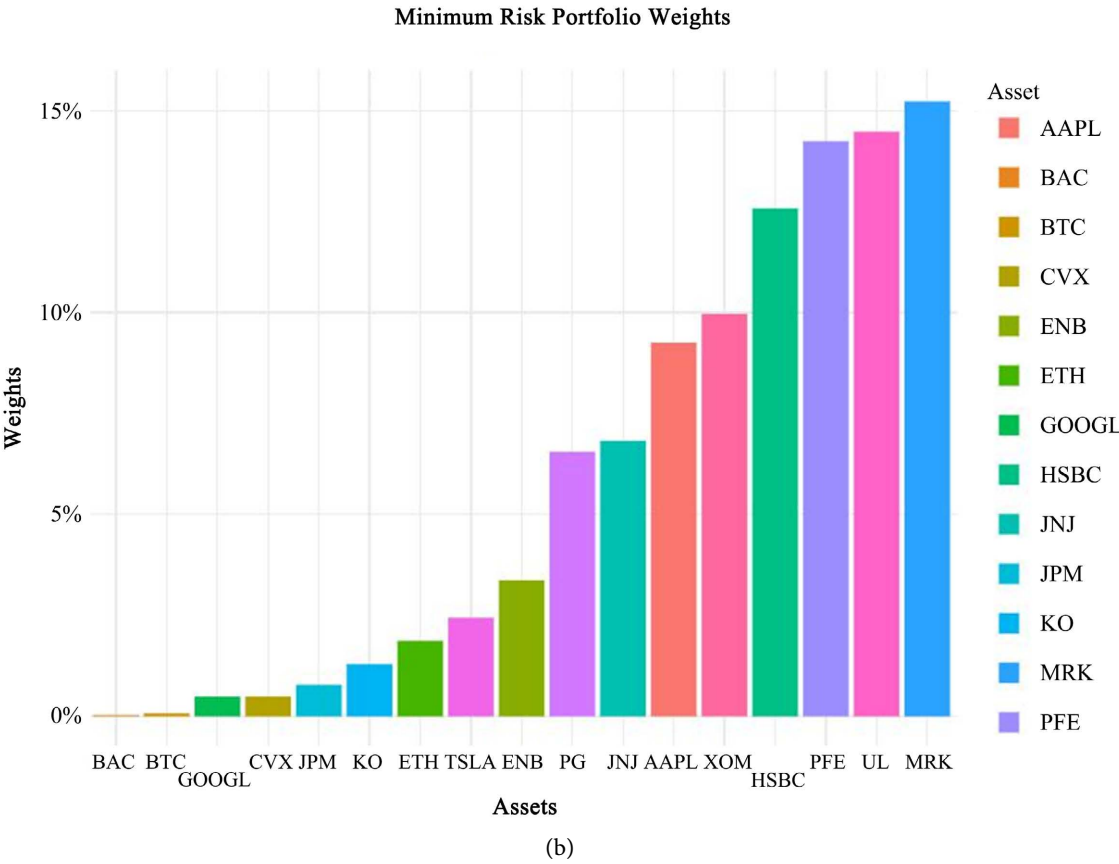
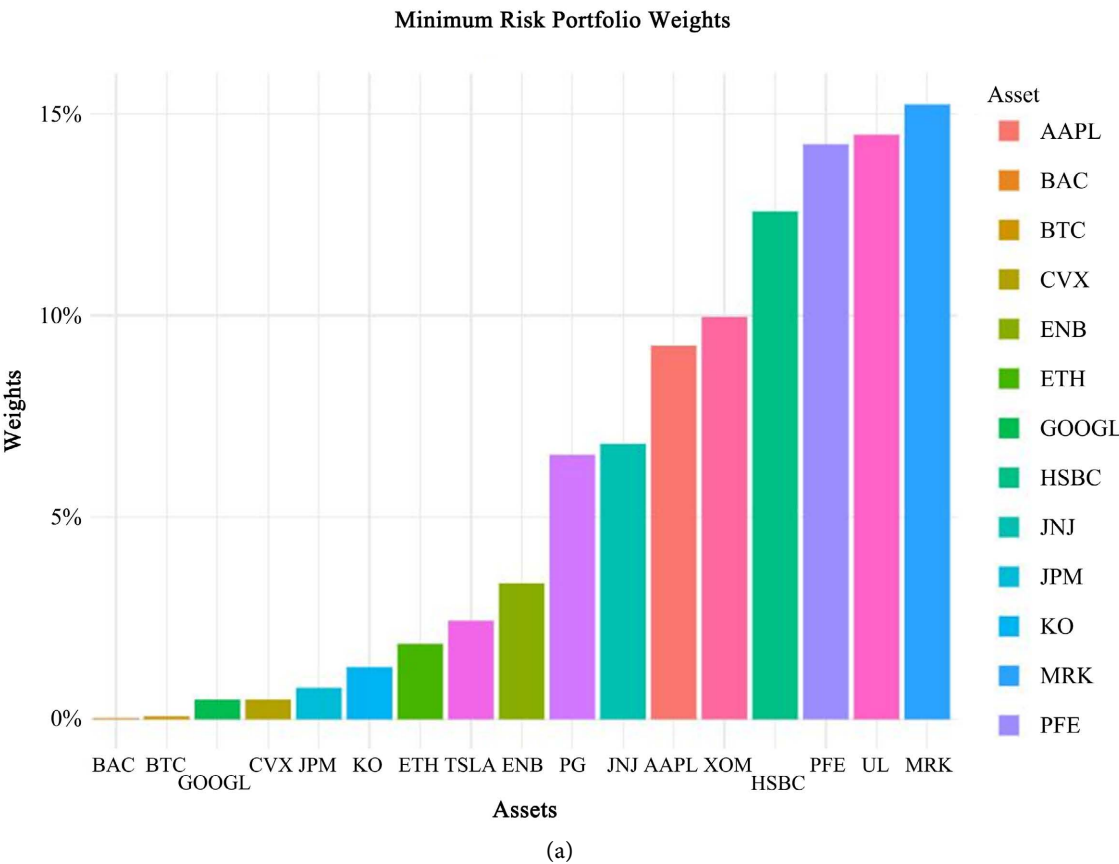
Figure 10. Performance of portfolio based on data-driven risk measures. (a) Efficient Frontier based on VESV risk measure; (b) Efficient Frontier based on VES risk measure; (c) Efficient Frontier based on VEV risk measure; (d) Efficient Frontier based on MAD risk measure.

Figure 11 illustrates the optimization of portfolio weights under two paradigms: minimum risk and maximum Sharpe ratio for both actual and predicted values. **Figure 11(b)** and **Figure 11(d)** indicate a striking resemblance between portfolio weight optimization based on predicted price data and that based on actual data.

5.3.2. With Diversification through Affinity Propagation

Utilizing affinity propagation for stock selection revealed the presence of six clusters. To enhance diversification in the stock selection process, this research identifies the financial asset with the highest mean returns within each cluster. Consequently, the stocks chosen from each cluster are incorporated to construct a robust portfolio. This approach ensures diversification in the portfolio, a critical factor in bolstering its resilience.

Figure 12 presents the performance of portfolios constructed with diversification under four data-driven risk measures. Here, portfolios constructed with both actual and predicted data showed similar efficient frontiers. In **Figure 12(c)**, it could be observed that the portfolio constructed with VEV-based risk measure showed maximum profitability compared to the risk.



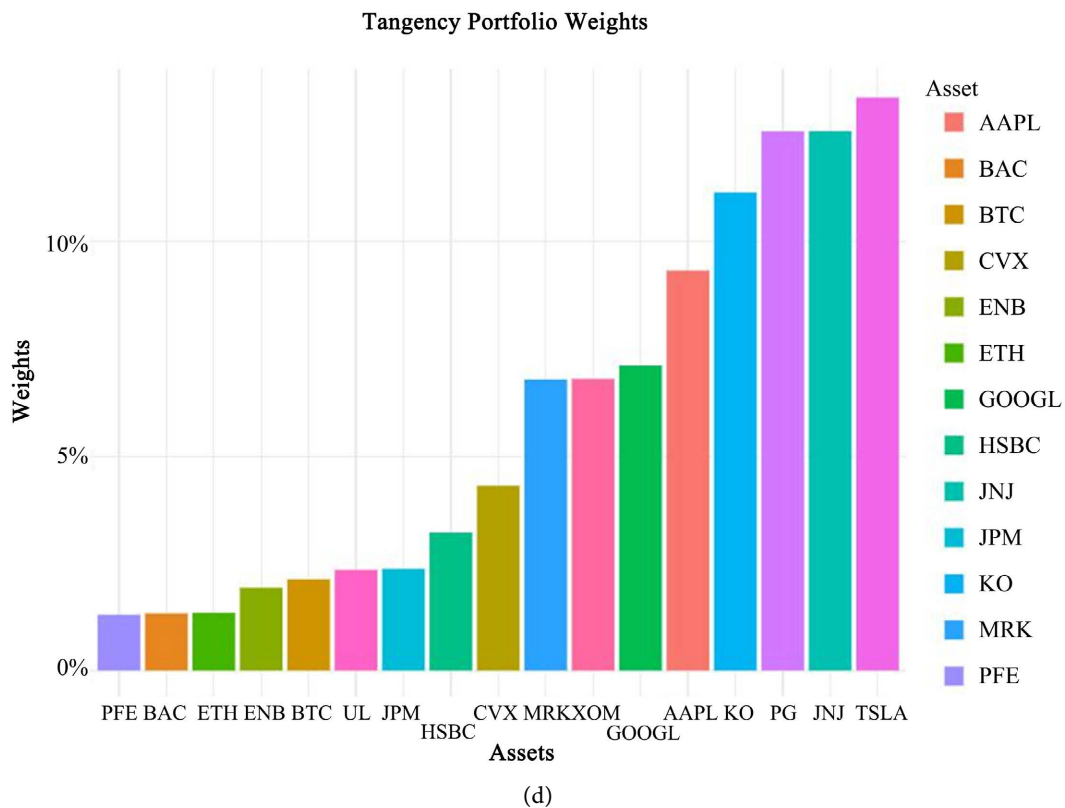
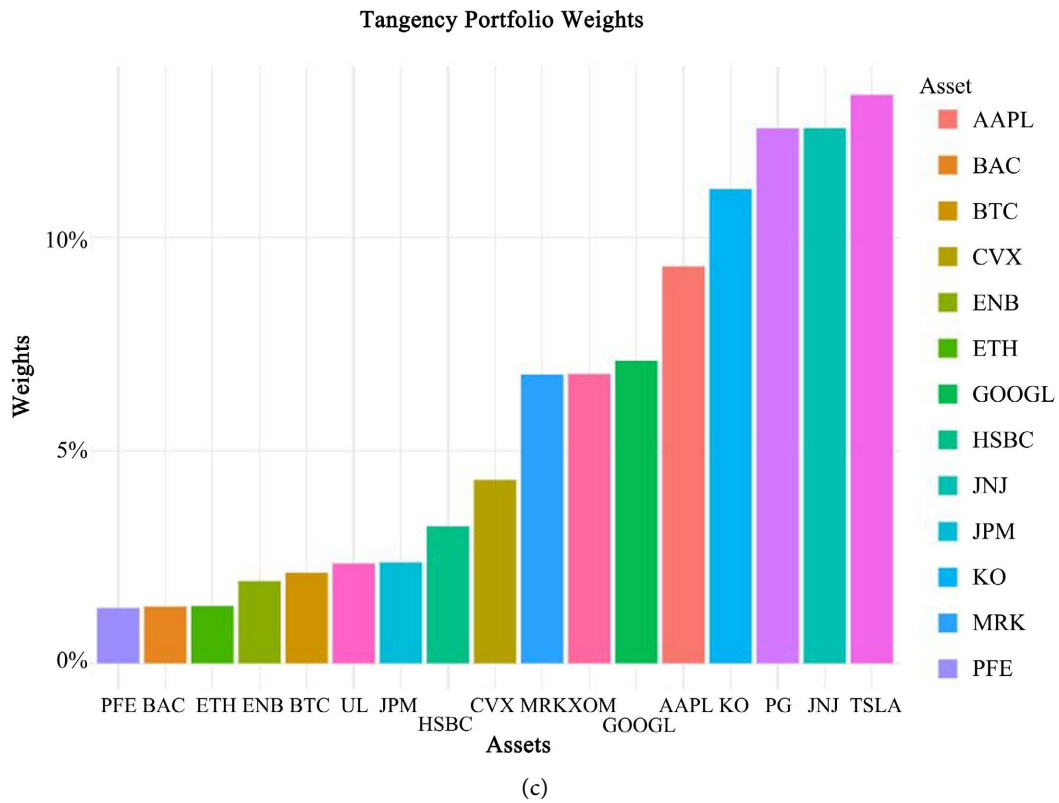
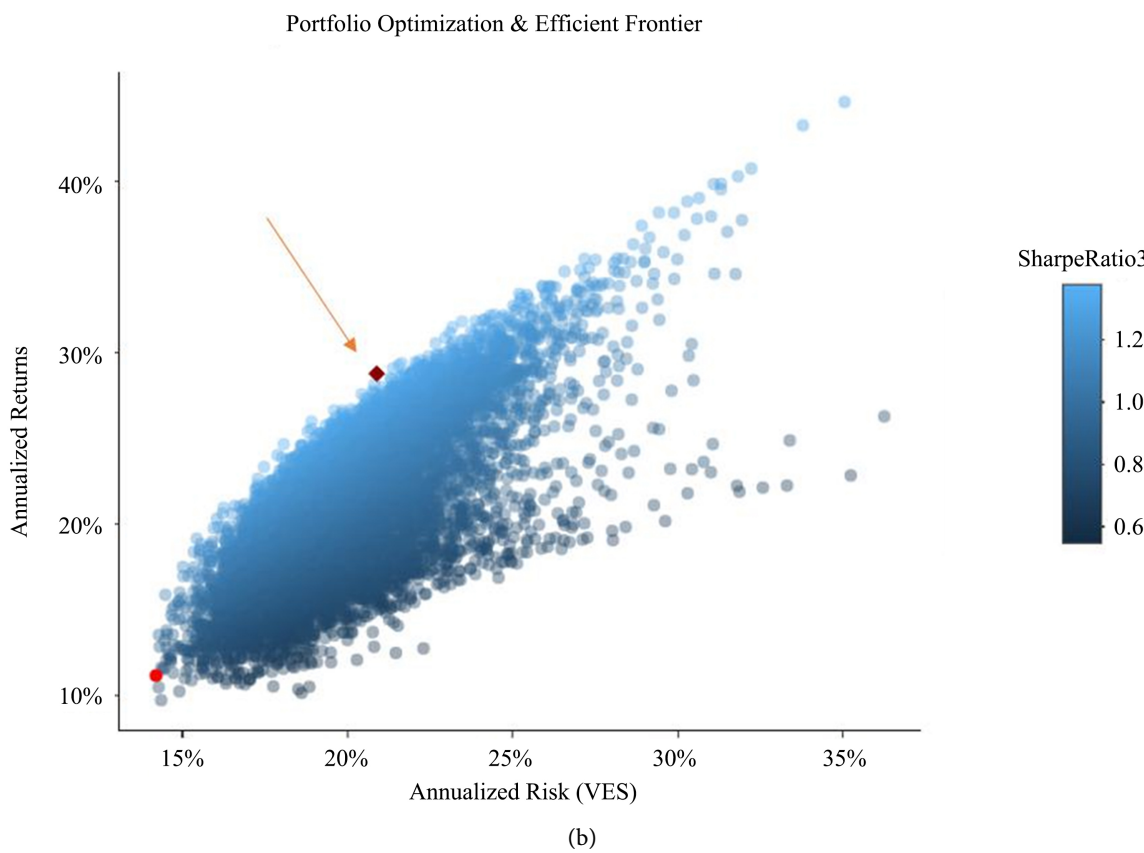
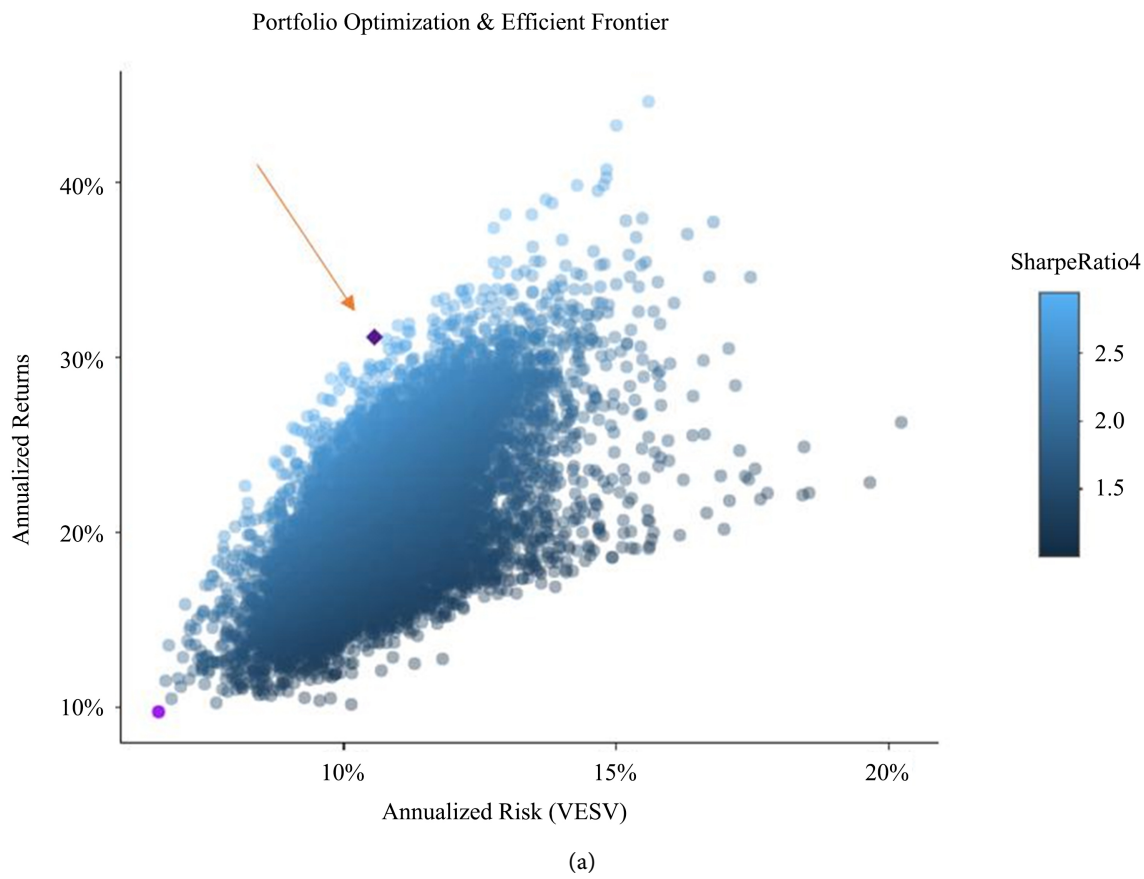


Figure 11. Portfolio weight distribution based on risk and Sharpe ratio: actual and predicted data. (a) Minimum Risk Portfolio with actual data; (b) Minimum Risk Portfolio with predicted data; (c) Tangency Portfolio: max Sharpe ratio with actual data; (d) Tangency Portfolio: max Sharpe ratio with predicted data.



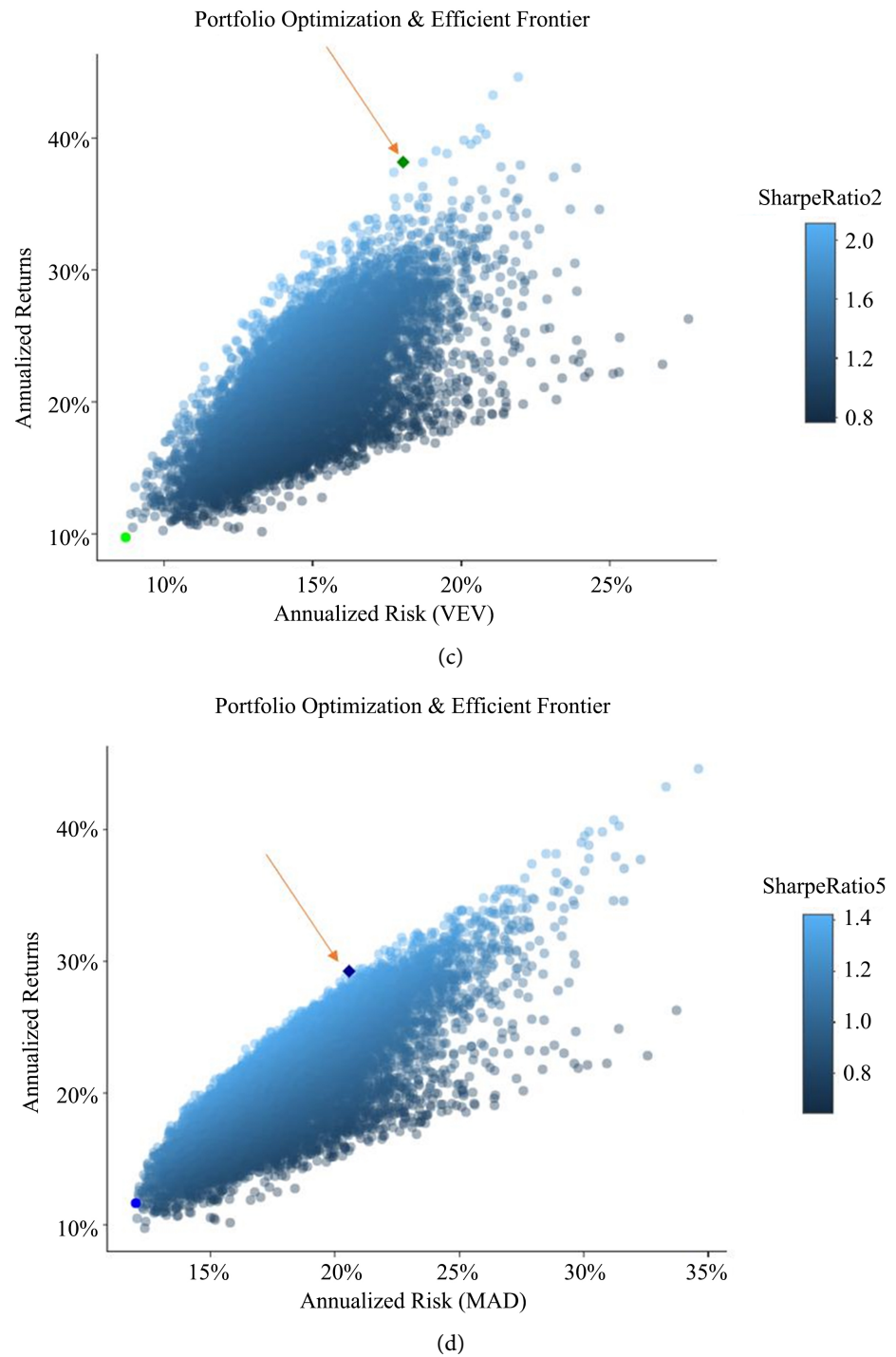
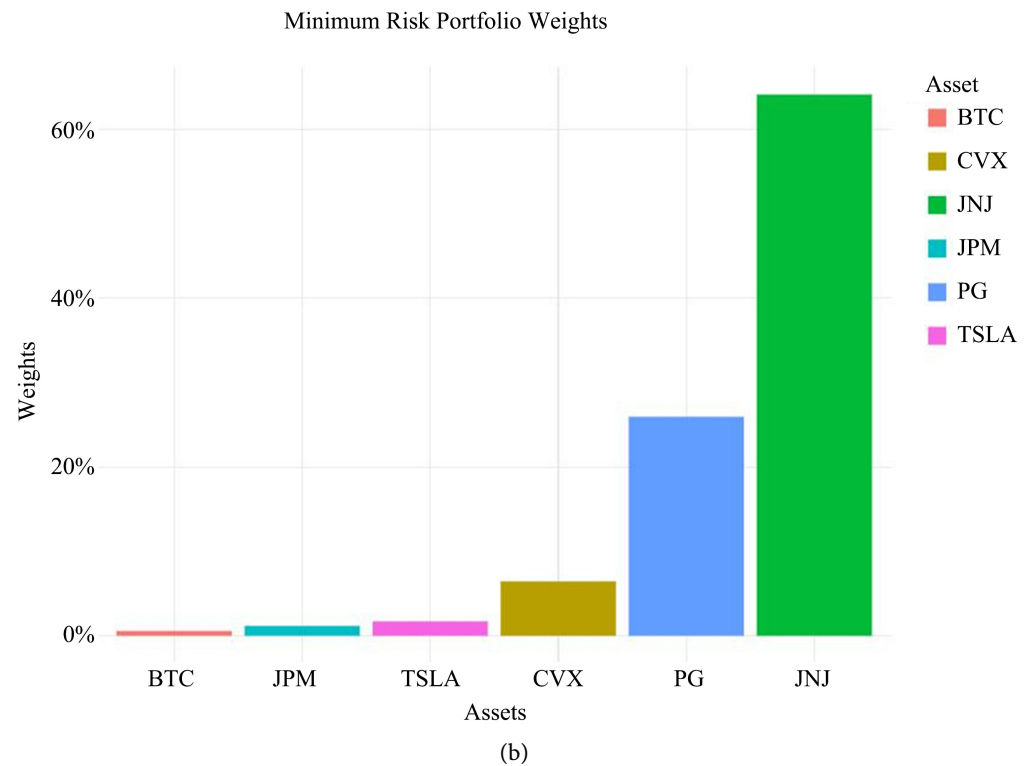
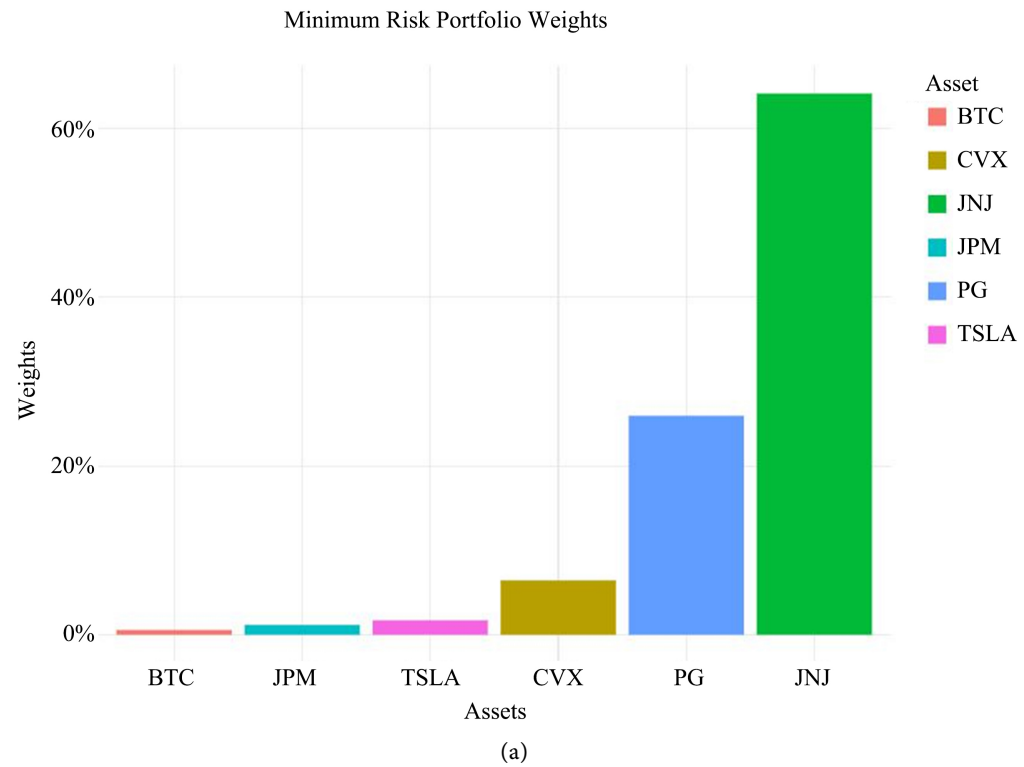


Figure 12. Performance of portfolio based on data-driven risk measures. (a) Efficient Frontier based on VESV risk measure; (b) Efficient Frontier based on VES risk measure; (c) Efficient Frontier based on VEV risk measure; (d) Efficient Frontier based on MAD risk measure.

Upon comparing **Figure 10** and **Figure 12**, it is observed that the portfolios' performance doubled across all data-driven risk measures following the implementation of diversification through Affinity Propagation clustering for stock selection.

Figure 13 depicts the optimization of portfolio weights employing two approaches: minimizing risk and maximizing the Sharpe ratio for both actual and predicted values. A notable similarity in portfolio weight optimization between predicted and actual data is observed in **Figure 13(b)** and **Figure 13(d)**.



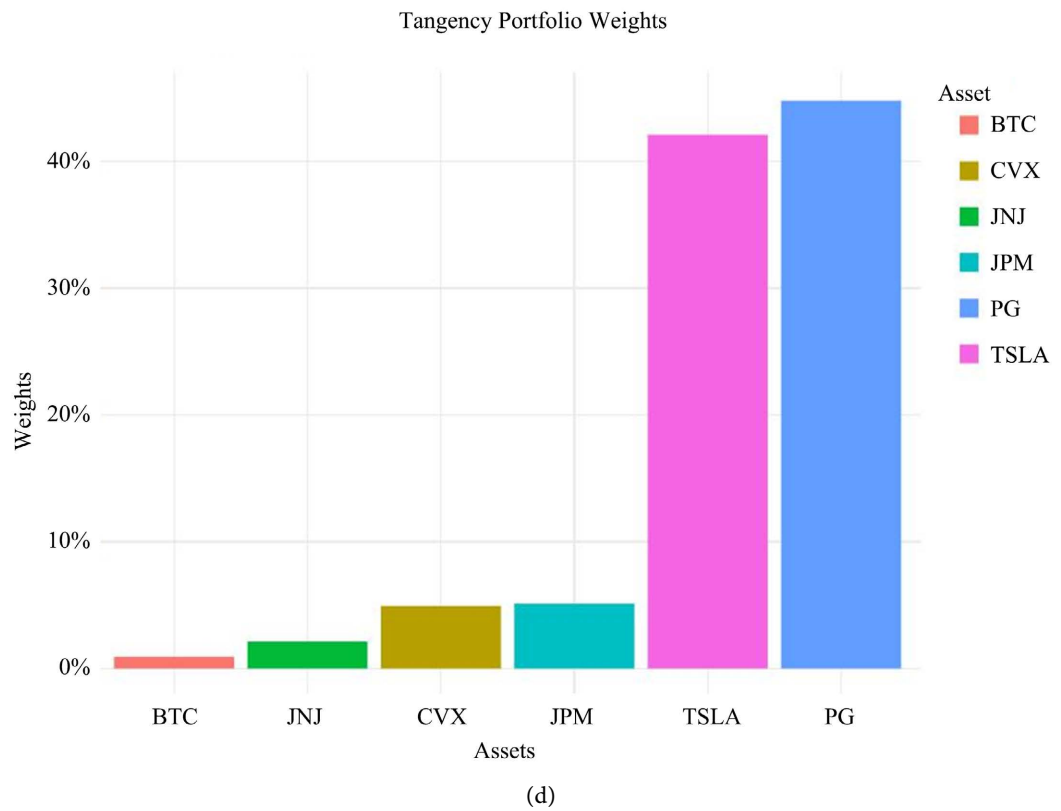
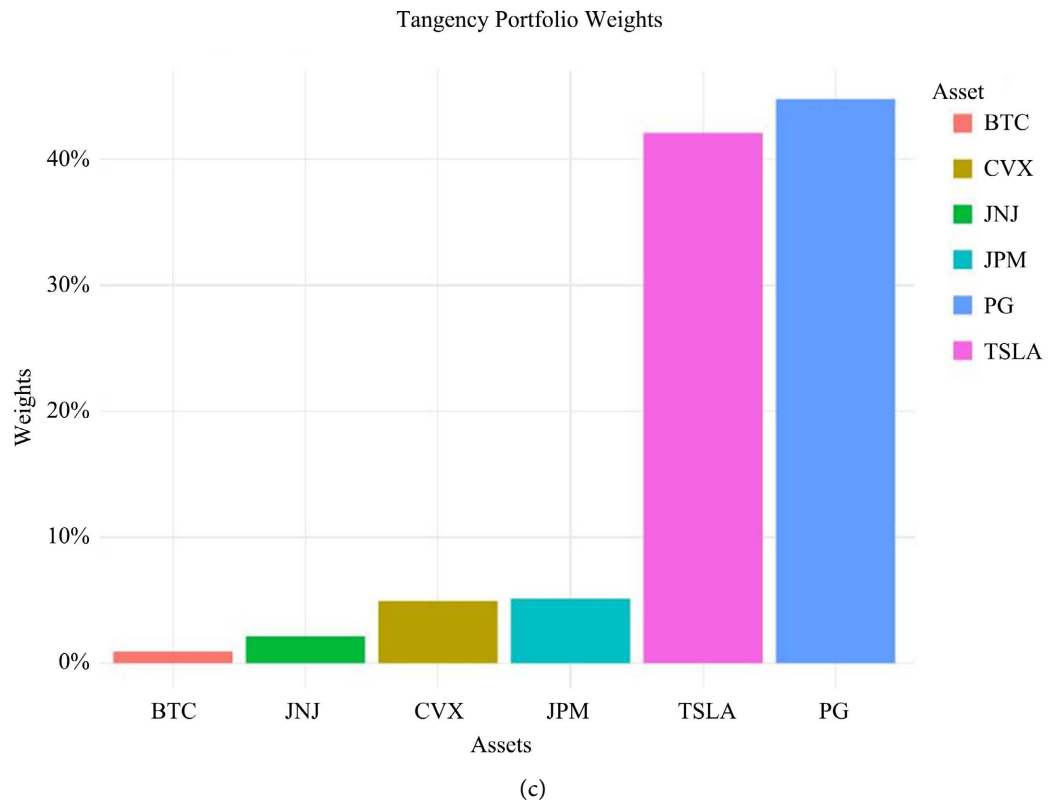


Figure 13. Portfolio weight distribution based on risk and Sharpe ratio with actual and predicted data. (a) Minimum Risk Portfolio with actual data; (b) Minimum Risk Portfolio with predicted data; (c) Tangency Portfolio: max Sharpe ratio with actual data; (d) Tangency Portfolio: max Sharpe ratio with predicted data.

The proficient performance of the BiLSTM and BiGRU algorithms in predicting financial asset prices across diverse market conditions leads to negligible discrepancies between actual and forecasted prices. Consequently, the data-driven algorithm exhibits comparable efficiency in portfolio weight optimization using both actual and data-driven approaches. The evaluation of these two algorithms suggests that the innovative integration of BiLSTM/BiGRU with the data-driven portfolio optimization algorithm constitutes a lucrative technique. This approach demonstrates maximum profitability under diverse market conditions and with stocks from various sectors. In certain instances, in the domain of time series predictions, it has been observed that the fundamental GRU architecture outperforms the LSTM model, as documented by Yamak *et al.* [59]. Drawing a parallel from this observation, a similar trend is discerned in the comparison between BiGRU and BiLSTM architectures. The rationale behind this superiority of GRU-based architectures may be attributed to their inherent design characteristics, such as a simplified structure with fewer parameters and the absence of an explicit memory cell. These attributes enable GRU models to capture and retain relevant information more efficiently in certain contexts, potentially leading to enhanced performance in specific time series prediction scenarios.

6. Key Findings

From the results section, this study identifies the following key findings from this research.

- 1) Based on the observed results from **Table 2** and **Figure 9**, superior performance of BiGRU-based predictions is observed compared to BiLSTM predictions.
- 2) Comparing **Figure 10** and **Figure 12**, it is observed introduction of diversification through Affinity Propagation enhances portfolio performance by two-fold compared to portfolios without diversification.
- 3) In **Figure 10(b)**, where portfolio construction lacks diversification, portfolios utilizing the VES-based risk measure exhibit superior performance compared to those constructed with alternative data-driven risk measures. Conversely, in **Figure 12(c)**, with diversification incorporated into portfolio construction, portfolios employing the VEV-based risk measure demonstrate out-performance over portfolios constructed with other data-driven risk measures.
- 4) Utilizing a data-driven approach for portfolio weight optimization, coupled with diversification through clustering techniques proves to be effective when implementing BiLSTM and BiGRU-based predictions.
- 5) BiGRU demonstrates significant effectiveness in predicting prices compared to BiLSTM. Therefore, it is recommended to prefer BiGRU, particularly when dealing with dynamic market conditions and highly volatile assets, when constructing robust portfolios.

7. Practicality, Limitations and Future Plan

The methodology proposed involves training the model across diverse market conditions spanning the period from 1998 to 2022. Subsequently, the algorithm undergoes rigorous testing using data from the year 2023, employing a suite of significant evaluation metrics. The algorithm, as proposed, adeptly assimilates patterns learned from historical observations, retaining contextual insights from pivotal crisis events (2001 dot-com burst, 2008 global crisis). Consequently, the optimized technique demonstrates notable efficacy in accurate prediction and the judicious optimization of portfolios to enhance profitability. Diversifying the portfolio almost doubled its ability to withstand challenges. The algorithm proposed is ready for practical use in industry after a few adjustments. These changes would involve considering transaction costs, liquidity of the assets, interest rates, inflation, and other macroeconomic factors. While the suggested method has been employed by training with significant market crises, there is an opportunity to subject it to further testing across diverse market scenarios, including both bullish and bearish conditions. Additionally, incorporating a range of economic factors beyond these scenarios would enhance the applicability of the approach, making it a more robust tool for portfolio managers. Exploring these avenues constitutes the forthcoming direction of our research.

8. Conclusions

In this study, the focal point has been resilient portfolio construction by incorporating deep learning and data-driven portfolio optimization techniques. Achieving resilience in portfolios involves not only accurate stock prediction but also effective stock selection for diversification. The core of efficient portfolio construction lies in the accurate forecast of stock prices, ensuring profitability across diverse market conditions. This study has investigated the effectiveness of two bidirectional recurrent neural networks, BiLSTM and BiGRU, in predicting stock prices under varying market conditions with financial assets from different sectors. The comparative analysis in this paper revealed the superior efficacy of BiGRU over BiLSTM in forecasting prices under challenging circumstances for financial assets across sectors.

To construct portfolios this study has employed forecasted data alongside actual data using a data-driven approach for the year 2023. The research indicates that employing deep learning-based predicted data for data-driven portfolio optimization yielded results that are quite comparable with actual data. This approach was tested both with and without diversification in stock selection. Diversified portfolios outperformed non-diversified ones, demonstrating almost twice the performance. For non-diversified portfolios, those utilizing the sign-correlation reduction based (that is, VES-based) risk measure showed favorable performance compared to portfolios built with other data-driven risk measures (VEV, VESV and MAD). In the case of diversified portfolios, those employing the volatility correlation reduction based (that is, VEV-based) risk

measure outperformed portfolios with other data-driven risk measures (VES, VESV and MAD).

In essence, the proposed novel deep learning-based data-driven portfolio optimization technique ensures maximum profitability under diverse market conditions. The incorporation of diversification with Affinity propagation enhances profitability, and selecting appropriate risk measures further improves overall portfolio performance. This study contributes valuable insights for robust portfolio construction, leveraging advanced techniques for enhanced decision-making in financial markets. Further improvements in results could be achieved by considering recent advances in quantum machine learning concepts.

Acknowledgements

The first two authors acknowledge financial support from the University of Manitoba Graduate Fellowship (UMGF) and Graduate Enhancement of Tri-Council Stipends (GETS), University of Manitoba. The last two authors acknowledge the Discovery Grants from the Natural Sciences and Engineering Research Council (NSERC) Canada.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Markowitz, H. (1952) Portfolio Selection. *The Journal of Finance*, **7**, 77-91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- [2] Fabozzi, F.J., Gupta, F. and Markowitz, H.M. (2002) The Legacy of Modern Portfolio Theory. *The Journal of Investing*, **11**, 7-22. <https://doi.org/10.3905/joi.2002.319510>
- [3] Mossin, J. (1966) Equilibrium in a Capital Asset Market. *Econometrica*, **34**, 768-783. <https://doi.org/10.2307/1910098>
- [4] Treynor, J.L. (1962) Toward a Theory of Market Value of Risky Assets. *Risk Books*, 15-22.
- [5] Lintner, J. (1965) The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, **47**, 13-37. <https://doi.org/10.2307/1924119>
- [6] Sharpe, W.F. (1964) Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, **19**, 425-442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- [7] Fama, E.F. and French, K.R. (2004) The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, **18**, 25-46. <https://doi.org/10.1257/0895330042162430>
- [8] Solnik, B. (1983) International Arbitrage Pricing Theory. *The Journal of Finance*, **38**, 449-457. <https://doi.org/10.1111/j.1540-6261.1983.tb02251.x>
- [9] Cheung, W. (2010) The Black-Litterman Model Explained. *Journal of Asset Management*, **11**, 229-243. <https://doi.org/10.1057/jam.2009.28>

- [10] Costa, G. and Kwon, R.H. (2018) Risk Parity Portfolio Optimization under a Markov Regime-Switching Framework. *Quantitative Finance*, **19**, 453-471. <https://doi.org/10.1080/14697688.2018.1486036>
- [11] Broadie, M. (1993) Computing Efficient Frontiers Using Estimated Parameters. *Annals of Operations Research*, **45**, 21-58. <https://doi.org/10.1007/bf02282040>
- [12] Detemple, J.B., Garcia, R. and Rindisbacher, M. (2003) A Monte Carlo Method for Optimal Portfolios. *The Journal of Finance*, **58**, 401-446. <https://doi.org/10.1111/1540-6261.00529>
- [13] Pai, G.A.V. (2017) Metaheuristics for Portfolio Optimization. Wiley. <https://doi.org/10.1002/9781119482840>
- [14] Sharpe, W.F. (1966) Mutual Fund Performance. *The Journal of Business*, **39**, 119-138. <https://doi.org/10.1086/294846>
- [15] Zhao, Y. and Yang, G. (2023) Deep Learning-Based Integrated Framework for Stock Price Movement Prediction. *Applied Soft Computing*, **133**, Article ID: 109921. <https://doi.org/10.1016/j.asoc.2022.109921>
- [16] M, H., E.A., G., Menon, V.K. and K.P., S. (2018) NSE Stock Market Prediction Using Deep-Learning Models. *Procedia Computer Science*, **132**, 1351-1362. <https://doi.org/10.1016/j.procs.2018.05.050>
- [17] Sirisha, U.M., Belavagi, M.C. and Attigeri, G. (2022) Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison. *IEEE Access*, **10**, 124715-124727. <https://doi.org/10.1109/access.2022.3224938>
- [18] Kumbure, M.M., Lohrmann, C., Luukka, P. and Porras, J. (2022) Machine Learning Techniques and Data for Stock Market Forecasting: A Literature Review. *Expert Systems with Applications*, **197**, Article ID: 116659. <https://doi.org/10.1016/j.eswa.2022.116659>
- [19] Shahi, T.B., Shrestha, A., Neupane, A. and Guo, W. (2020) Stock Price Forecasting with Deep Learning: A Comparative Study. *Mathematics*, **8**, Article 1441. <https://doi.org/10.3390/math8091441>
- [20] Pintelas, E., Livieris, I.E., Stavroyiannis, S., Kotsilieris, T. and Pintelas, P. (2020) Investigating the Problem of Cryptocurrency Price Prediction: A Deep Learning Approach. In: Maglogiannis, I., Iliadis, L. and Pimenidis, E., Eds., *Artificial Intelligence Applications and Innovations*, Springer, 99-110. https://doi.org/10.1007/978-3-030-49186-4_9
- [21] Ban, G., El Karoui, N. and Lim, A.E.B. (2018) Machine Learning and Portfolio Optimization. *Management Science*, **64**, 1136-1154. <https://doi.org/10.1287/mnsc.2016.2644>
- [22] Chen, W., Zhang, H., Mehlawat, M.K. and Jia, L. (2021) Mean-Variance Portfolio Optimization Using Machine Learning-Based Stock Price Prediction. *Applied Soft Computing*, **100**, Article ID: 106943. <https://doi.org/10.1016/j.asoc.2020.106943>
- [23] Paiva, F.D., Cardoso, R.T.N., Hanaoka, G.P. and Duarte, W.M. (2019) Decision-making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. *Expert Systems with Applications*, **115**, 635-655. <https://doi.org/10.1016/j.eswa.2018.08.003>
- [24] Zhang, Z., Zohren, S. and Roberts, S. (2020) Deep Learning for Portfolio Optimization. *The Journal of Financial Data Science*, **2**, 8-20. <https://doi.org/10.3905/jfds.2020.1.042>
- [25] Ma, Y., Han, R. and Wang, W. (2021) Portfolio Optimization with Return Prediction Using Deep Learning and Machine Learning. *Expert Systems with Applications*,

- 165, Article ID: 113973. <https://doi.org/10.1016/j.eswa.2020.113973>
- [26] Cao, H.K., Cao, H.K. and Nguyen, B.T. (2020) Delafo: An Efficient Portfolio Optimization Using Deep Neural Networks. In: Lauw, H., Wong, R.W., Ntoulas, A., Lim, E.P., Ng, S.K. and Pan, S., Eds., *Advances in Knowledge Discovery and Data Mining*, Springer, 623-635. https://doi.org/10.1007/978-3-030-47426-3_48
- [27] Sharma, M. and Shekhawat, H.S. (2022) Portfolio Optimization and Return Prediction by Integrating Modified Deep Belief Network and Recurrent Neural Network. *Knowledge-Based Systems*, **250**, Article ID: 109024. <https://doi.org/10.1016/j.knosys.2022.109024>
- [28] Singh, P., Jha, M., Sharaf, M., El-Meligy, M.A. and Gadekallu, T.R. (2023) Harnessing a Hybrid CNN-LSTM Model for Portfolio Performance: A Case Study on Stock Selection and Optimization. *IEEE Access*, **11**, 104000-104015. <https://doi.org/10.1109/access.2023.3317953>
- [29] Cui, T., Du, N., Yang, X. and Ding, S. (2024) Multi-period Portfolio Optimization Using a Deep Reinforcement Learning Hyper-Heuristic Approach. *Technological Forecasting and Social Change*, **198**, Article ID: 122944. <https://doi.org/10.1016/j.techfore.2023.122944>
- [30] Thavaneswaran, A., Liang, Y., Yu, N., Paseka, A. and Thulasiram, R.K. (2021) Novel Data-Driven Resilient Portfolio Risk Measures Using Sign and Volatility Correlations. 2021 *IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, Madrid, 12-16 July 2021, 1742-1747. <https://doi.org/10.1109/compsac51774.2021.00260>
- [31] Bowala, S. and Singh, J. (2022) Optimizing Portfolio Risk of Cryptocurrencies Using Data-Driven Risk Measures. *Journal of Risk and Financial Management*, **15**, Article 427. <https://doi.org/10.3390/jrfm15100427>
- [32] Choueifaty, Y. and Coignard, Y. (2008) Toward Maximum Diversification. *The Journal of Portfolio Management*, **35**, 40-51. <https://doi.org/10.3905/jpm.2008.35.1.40>
- [33] Das, J.D., Bowala, S., Thulasiram, R.K. and Thavaneswaran, A. (2023) Resilient Portfolio Optimization Using Traditional and Data-Driven Models for Cryptocurrencies and Stocks. 2023 *IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*, Torino, 26-30 June 2023, 1343-1348. <https://doi.org/10.1109/compsac57700.2023.00204>
- [34] Das, J.D., Bowala, S., Thulasiram, R.K. and Thavaneswaran, A. (2023) Portfolio Diversification with Clustering Techniques. 2023 *IEEE Symposium Series on Computational Intelligence (SSCI)*, Mexico City, 5-8 December 2023, 97-102. <https://doi.org/10.1109/ssci52147.2023.10371938>
- [35] Rundo, F., Trenta, F., di Stallo, A.L. and Battiato, S. (2019) Machine Learning for Quantitative Finance Applications: A Survey. *Applied Sciences*, **9**, Article 5574. <https://doi.org/10.3390/app9245574>
- [36] Vijh, M., Chandola, D., Tikkiwal, V.A. and Kumar, A. (2020) Stock Closing Price Prediction Using Machine Learning Techniques. *Procedia Computer Science*, **167**, 599-606. <https://doi.org/10.1016/j.procs.2020.03.326>
- [37] Yang, Y. and Hospedales, T.M. (2023) An Evaluation of Self-Supervised Learning for Portfolio Diversification. In: Iliadis, L., Papaleonidas, A., Angelov, P. and Jayne, C., Eds., *Artificial Neural Networks and Machine Learning—ICANN 2023*, Springer, 283-294. https://doi.org/10.1007/978-3-031-44213-1_24
- [38] Jaimungal, S. (2021) Reinforcement Learning and Stochastic Optimisation. *Finance and Stochastics*, **26**, 103-129. <https://doi.org/10.1007/s00780-021-00467-2>

- [39] Snow, D. (2020) Machine Learning in Asset Management—Part 2: Portfolio Construction—Weight Optimization. *The Journal of Financial Data Science*, **2**, 17-24. <https://doi.org/10.3905/jfds.2020.1.029>
- [40] Kaczmarek, T. and Perez, K. (2021) Building Portfolios Based on Machine Learning Predictions. *Economic Research-Ekonomska Istraživanja*, **35**, 19-37. <https://doi.org/10.1080/1331677x.2021.1875865>
- [41] Jiang, Z., Ji, R. and Chang, K. (2020) A Machine Learning Integrated Portfolio Re-balance Framework with Risk-Aversion Adjustment. *Journal of Risk and Financial Management*, **13**, Article 155. <https://doi.org/10.3390/jrfm13070155>
- [42] Behera, J., Pasayat, A.K., Behera, H. and Kumar, P. (2023) Prediction Based Mean-Value-at-Risk Portfolio Optimization Using Machine Learning Regression Algorithms for Multi-National Stock Markets. *Engineering Applications of Artificial Intelligence*, **120**, Article ID: 105843. <https://doi.org/10.1016/j.engappai.2023.105843>
- [43] Hu, Z., Zhao, Y. and Khushi, M. (2021) A Survey of Forex and Stock Price Prediction Using Deep Learning. *Applied System Innovation*, **4**, Article 9. <https://doi.org/10.3390/asi4010009>
- [44] Jiang, W. (2021) Applications of Deep Learning in Stock Market Prediction: Recent Progress. *Expert Systems with Applications*, **184**, Article ID: 115537. <https://doi.org/10.1016/j.eswa.2021.115537>
- [45] Nikou, M., Mansourfar, G. and Bagherzadeh, J. (2019) Stock Price Prediction Using DEEP Learning Algorithm and Its Comparison with Machine Learning Algorithms. *Intelligent Systems in Accounting, Finance and Management*, **26**, 164-174. <https://doi.org/10.1002/isaf.1459>
- [46] Fernández, A. and Gómez, S. (2007) Portfolio Selection Using Neural Networks. *Computers & Operations Research*, **34**, 1177-1191. <https://doi.org/10.1016/j.cor.2005.06.017>
- [47] Uysal, A.S., Li, X. and Mulvey, J.M. (2023) End-to-End Risk Budgeting Portfolio Optimization with Neural Networks. *Annals of Operations Research*, **339**, 397-426. <https://doi.org/10.1007/s10479-023-05539-4>
- [48] Yu, L., Wang, S. and Lai, K.K. (2008) Neural Network-Based Mean-Variance-Skewness Model for Portfolio Selection. *Computers & Operations Research*, **35**, 34-46. <https://doi.org/10.1016/j.cor.2006.02.012>
- [49] Jang, J. and Seong, N. (2023) Deep Reinforcement Learning for Stock Portfolio Optimization by Connecting with Modern Portfolio Theory. *Expert Systems with Applications*, **218**, Article ID: 119556. <https://doi.org/10.1016/j.eswa.2023.119556>
- [50] Yang, S. (2023) Deep Reinforcement Learning for Portfolio Management. *Knowledge-Based Systems*, **278**, Article ID: 110905. <https://doi.org/10.1016/j.knosys.2023.110905>
- [51] Ngo, V.M., Nguyen, H.H. and Van Nguyen, P. (2023) Does Reinforcement Learning Outperform Deep Learning and Traditional Portfolio Optimization Models in Frontier and Developed Financial Markets? *Research in International Business and Finance*, **65**, Article ID: 101936. <https://doi.org/10.1016/j.ribaf.2023.101936>
- [52] Xia, K., Huang, J. and Wang, H. (2020) LSTM-CNN Architecture for Human Activity Recognition. *IEEE Access*, **8**, 56855-56866. <https://doi.org/10.1109/access.2020.2982225>
- [53] Luo, L. (2018) Network Text Sentiment Analysis Method Combining LDA Text Representation and GRU-CNN. *Personal and Ubiquitous Computing*, **23**, 405-412. <https://doi.org/10.1007/s00779-018-1183-9>

- [54] Kim, T. and Kim, H.Y. (2019) Forecasting Stock Prices with a Feature Fusion LSTM-CNN Model Using Different Representations of the Same Data. *PLOS ONE*, **14**, e0212320. <https://doi.org/10.1371/journal.pone.0212320>
- [55] Pierre, A.A., Akim, S.A., Semenyio, A.K. and Babiga, B. (2023) Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches. *Energies*, **16**, Article 4739. <https://doi.org/10.3390/en16124739>
- [56] Lu, W., Li, J., Wang, J. and Qin, L. (2020) A CNN-BILSTM-AM Method for Stock Price Prediction. *Neural Computing and Applications*, **33**, 4741-4753. <https://doi.org/10.1007/s00521-020-05532-z>
- [57] Pramesti, M.I., Indikawati, F.I. and Prahara, A. (2022) Multivariate Time Series Stock Price Data Prediction in the Banking Sector in Indonesia Using Bidirectional Long Short-Term Memory (biLSTM). *Signal and Image Processing Letters*, **4**, 28-37.
- [58] Malla, J., Lavanya, C., Jayashree, J. and Vijayashree, J. (2022) Bidirectional Gated Recurrent Unit (BiGRU)-Based Bitcoin Price Prediction by News Sentiment Analysis. In: Reddy, V.S., Prasad, V.K., Wang, J. and Reddy, K.T.V., Eds., *Soft Computing and Signal Processing. ICSCSP 2022*, Springer, 31-40. https://doi.org/10.1007/978-981-19-8669-7_4
- [59] Yamak, P.T., Yujian, L. and Gadosey, P.K. (2019) A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting. *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, Sanya, 20-22 December 2019, 49-55. <https://doi.org/10.1145/3377713.3377722>