A Hybrid Approach for a Novel Dynamic Trading System to Produce Robust Cryptocurrency Portfolios

Ehsan Mohammadian Amiri, Akbar Esfahanipour*

Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

ABSTRACT

This study aims to develop a dynamic portfolio trading system for high-risk profiles of cryptocurrencies in two phases: 1) portfolio selection and 2) portfolio construction. In the first phase, we propose a novel algorithmic trading model applying a Convolutional Neural Network (CNN) using a 2-D convolution layer with eight kernels of 3×3 sizes based on the prediction of selected technical indicators to predict buy/sell trading signals. To effectively increase the accuracy of the CNN model, first, the H-step ahead predictions of the selected technical indicators based on Long-short-term-memory (LSTM) along with the indicators themselves have been used to construct input matrices of the CNN model. A new price labeling approach was proposed to determine buying or selling points using the zigzag indicator in our CNN model. Assets with buy signals have been selected to construct the proposed portfolio. In the second phase, we propose a novel robust approach based on Holt-Winters-Multiplicative (HWM) to determine the realized crypto portfolio weights robustly by considering the seasonal effects. The experimental results show that our developed system outperforms the competing models for 30 cryptocurrencies with a high-risk profile in the two phases.

Keywords:

Risky assets, Convolutional neural network (CNN), Zigzag Indicator, Global Minimum Variance Portfolio (GMVP), Holt-Winters-Multiplicative (HWM)

1. Introduction

* Corresponding author Email address: esfahaa@aut.ac.ir

Address: No. 350, Hafez Ave, Valiasr Square, Tehran, Iran 1591634311, Department of Industrial Engineering & Management Systems, Amirkabir University of Technology Tel: +98(21)6454-5369; Mobile: +98(912)347-9906; Fax:+98(21)6695-4569; Website: www.aut.ac.ir/esfahaa

One of the challenges of designing an algorithmic trading system in a fluctuating market such as the cryptocurrency market is the extreme price fluctuations in short time periods. These fluctuations make it difficult to determine the entry (buy) and exit (sell) points of a cryptocurrency. Considering these points as two classes of buying and selling in a classification task, the learning power of a classification model will usually decrease and the model will probably overfit [1]. Therefore, designing an algorithmic trading system with high accuracy in fluctuating conditions has become vital for investors [2].

Therefore, this study has two main goals: 1) designing an efficient hybrid trading system based on CNN to predict buy/sell trading signals in the portfolio selection phase and 2) determining the realized crypto portfolio by developing Golosnoy et al. [3] model in the portfolio construction phase.

We implement several steps to increase the accuracy of our algorithmic trading and reduce the model's complexity in fluctuating conditions. First, we want to evaluate whether the use of predicted indicators in addition to their past values will lead to improve the model accuracy. Based on this idea, we have used the prediction of multi-step ahead of selected indicators using the LSTM model as the input of the classification model. Due to the large quantity of input data, we used a feature selection approach called Eigenvector Centrality Feature Selection (ECFS) to reduce the complexity of the network, remove redundant features, and select appropriate indicators. Also, designing a suitable labeling method for an algorithmic trading system is one of the ways to increase the accuracy of the models significantly [4]. Therefore, in this study, we develop a new labeling approach based on the zigzag approach to reduce noise levels and highlight underlying peaks and valleys in the price trends.

On the other hand, Golosnoy et al. [3] apply an exponential smoothing (ES) to realize covariance matrices and obtain portfolio weights for 30 cryptocurrencies with a high-risk profile. The results show that the ES forecast is well-suited for forecasting the GMVP based on 100 risky assets. However, the ES model is unsuitable for the data with seasonal factors (e.g., day, week, month, year) because they should be adjusted to the smoothing coefficient. The HWM model with three smoothing factors of $\gamma, \beta, \alpha \in (0, 1)$ includes an appealing way to overcome the mentioned problem. Therefore, we adopt the robust HWM model for the determined covariance matrix of cryptocurrencies, and the GMVP composition is computed afterward, which leads to the construction of a portfolio of assets with high-risk profiles.

According to the aforementioned challenges, the main and minor contributions of this study are as follows.

Main contribution:

- Designing an accurate and robust hybrid portfolio trading system with reduced complexity for assets with a high-risk profile.

Minor contributions in the portfolio selection phase:

- Proposing a novel approach based on the Zigzag indicator for labeling the price direction, which reduces the impact of random price fluctuations in the high-volatile cryptocurrency market.
- Adding the predicted values of the technical indicators based on the LSTM as the input of our CNN model, which leads to improve learning power of price direction prediction.
- Ranking and selecting appropriate technical indicators using the ECFS model as a feature selection method.

Minor contributions in the portfolio construction phase:

- Developing Golosnoy et al. [3] model to determine crypto portfolio weights based on the HWM by considering the seasonal effects.
- Obtaining evidence on seasonality patterns of cryptocurrencies by testing for Monday effect among the five highest market capitalization.

The remaining parts of this article are organized as follows: Section 2. Provides the related works of designing trading systems based on deep learning, portfolio trading systems, and labeling methods for deep learning. The proposed dynamic portfolio trading system and the applied methodology are discussed in Section 3. The model evaluation process and data are presented in Section 4. Finally, conclusions are summarized in Section 5.

2. Literature review

This section describes the related studies in the current literature. First, we describe recent developments in trading systems based on deep learning, then provide the primary research available in the literature about portfolio trading systems and labeling methods for deep learning.

2.1 Trading systems based on deep learning

Dakalbab et al. [5] review a comprehensive analysis of 143 research articles that applied AI methodologies in financial trading across 8 different financial markets, using 40 different AI techniques. Among those techniques, deep learning techniques emerged as the most commonly and frequently used in financial trading markets.

In particular, applying deep neural networks based on image processing in finance has been noteworthy. For example, Sezer and Ozbayoglu [6] present a novel algorithmic trading model called CNN-TA, using a 2-D CNN based on image processing for selected ETFs and Dow

Jones 30 Stocks. The results show that their proposed model works better than traditional trading models such as Buy and Hold, LSTM, and MLP, especially when the overall market is not bullish. Luo et al. [7] built a novel Recurrent Reinforcement Learning (RRL) framework based on AI trader and simultaneously adopted two methods: Deep Deterministic Policy Gradient (DDPG) and CNN. Yu and Li [8] suggest a hybrid convolutional recurrent neural network (HCRNN) to forecast the vital trading points (IPs) that are more likely to be followed by a significant stock price rise. Their proposed ITPP-HCRNN achieves an annualized return of 278.46% compared to the market. Wu et al. [9] introduce a novel CNN called SSACNN, which stands for stock sequence array CNN. Using this approach, the SSACNN can accurately analyze and interpret complex stock market data, making it a valuable tool for informed investment decisions. Wu et al. [10] introduce a unique system for classifying and quantifying stock characteristics using fuzzy momentum contrarian uncertainty characteristic system. Khodaee et al. [11] developed a forecasting model for stock price Turning Points (TPs) based on CNN and LSTM models. Their CNN-LSTM-ResNet model outperformed other comparative models in the selected ETFs and the Dow-30 stocks. Buachuen and Kantavat [12] adopt an automated stock trading system using Technical Analysis and The LSTM-CNN hybrid model, which leads to grasping long-term temporal dependencies in stock patterns. The initial experimental study shows the possible advantages of these approaches compared to traditional methodologies. Jing and Kang [13] present a dynamic trading approach using ensemble deep reinforcement learning to generate trading signals based on a state vector that includes embedded candlestick-chart images in the cryptocurrency market. Zou et al. [14] propose a stock trading system based on Deep Reinforcement Learning (DRL). It utilizes the Cascaded Long-Short-Term-Memory (CLSTM-PPO Model) to account for concealed information within daily stock data. The results indicate that their model surpasses the baseline models across critical metrics, including cumulative returns, maximum earning rate, and average profitability per trade. Massahi and Mahootchi [15] suggest a new intraday algorithmic trading system tailored for volatile commodity futures markets. They utilize two methodologies: a Deep Q-network (DQN) algorithm and a robust double-version (DDQN) approach. The results show that their proposed model, which is grounded in real intraday data from gold coin futures contracts, considerably surpasses the benchmarks in terms of return, risk, and risk-adjusted return.

2.2 Portfolio trading systems

Ta et al. [16] propose a dynamic portfolio trading system with ten-year historical stock price data on the daily 500 significant stocks listed on the S&P500. They use the LSTM and RNN models to predict stock movement and asset selection phase. Also, they adopted three methods of equal weight modeling (EW), Monte Carlo simulation (MCS), and the mean-variance optimization model (MVO) for portfolio optimization. Experimental results showed that the returns of the EQ, MCS, and MVO models were almost the same, but due to fewer calculations, the MVO model was recognized as the best-preformed model. Also, Ma et al. [17] build a prediction-based portfolio optimization model using three deep neural networks consisting of deep multilayer perceptron (DMLP), LSTM, and CNN. These models use to forecast the future returns of each stock and asset selection phase. Then, prediction based portfolio optimization models are built by generalizing the frame of mean semiabsolute deviation (MSAD) portfolio model. Ashrafzadeh et al. [18] introduce a novel approach that combines a CNN with fine-tuned hyperparameters using particle swarm optimization (PSO) to stock preselection, along with a mean-variance-forecasting (MVF) model to optimize portfolios consist of 21 stocks in the New York Stock Exchange (NYSE). Alamdari et al. [19] present a dynamic portfolio trading system in two stages. In the asset selection stage, the Pixel Graph Network (PGN) model was applied to determine assets, and after that, a Mean–Conditional Drawdown at Risk (M-CDaR) portfolio optimization model was suggested to choose the best-weighted combination of the selected assets. Table 1 shows the position of the current study in the related literature.

Please insert Table 1 about here.

2.3 Labeling methods for deep learning

Lin et al. [20] propose a hierarchical attention neural semi-Markov conditional random fields (semi-CRF) model as a sequence labeling method. This model employs a hierarchical framework that integrates character and word-level information, utilizing an attention mechanism at each level. As a result, the method can distinguish between more significant and less significant information while constructing the segmental representation. Lu et al. [21] suggest a multi-label neural text classifier named CNN-BiLSTM-Attention that directly derives the precise meaning of labels from the dataset. This classifier is also equipped with a tailored attention mechanism referred to as the multi-label attention mechanism, which is capable of identifying significant text features pertinent to each label. Peng et al. [22] propose a new triple trend labeling method, informed by the analysis of high-frequency data, that aims to decrease the number of trades by potentially impacting the model's training. In addition, an Attention-based CNN–LSTM model for multiple cryptocurrencies (ACLMC) is suggested to enhance model performance by leveraging correlations across various frequencies and currencies. Experimental results indicate that their labeling technique utilizing the ACLMC outperforms traditional baselines regarding financial metrics and reduces the number of transactions.

3. Our proposed dynamic portfolio trading system

This section presents the main processes, including predicting indicators and forming the images as the CNN input, price labeling, predicting trading signals, and realized crypto portfolio weights.

Please insert Figure 1 about here.

Our proposed portfolio trading system, illustrated in Figure 1, is structured in two distinct phases. The first phase, portfolio selection, adopts a hybrid trading system utilizing CNN to predict buy and sell trading signals for each cryptocurrency. In the second phase, portfolio construction, cryptocurrencies with predicted buy trading signals in the following days are fed into the HWM model to determine the realized crypto portfolio weights. Subsections 3.1 to 3.5 describe each component of our dynamic portfolio trading systems. Also, the process of implementing the proposed model along with its setting, including training periods and testing of algorithms is shown in Figure 2.

Please insert Figure 2 about here.

3.1 Ranking and selecting the features using the ECFS method

Before predicting technical indicators, the selected technical indicators for each cryptocurrency have been determined using the ECFS model. For this purpose, the output of the ECFS model is a set of particular values for each indicator, and based on that, the indicators are ranked. Then, input indicators of the prediction model were selected using the scree plots and cumulative eigenvalues. Eventually, we use the ECFS method to rank and select indicators because of the following main reasons:

- Firstly, the ECFS method can reduce the complexity of the network and remove redundant features (Roffo & Melzi [23]). This is an important property of this mehod for the cryptocurrency market with many assets having several related features.

- Secondly, the reliability of the ECFS performance has been demonstrated for seven big datasets, including object recognition, hand-written recognition, biological data, and synthetic testing datasets. Also, the superior performance of the ECFS against seven competing models has been shown as follows. 1) feature selection contain Fisher, 2) FSV (feature selective validation), 3) Inf-FS (infinite feature Selection), 4) MI (mutual information), 5) LS (laplacian score), 6) Relief-F (Relief algorithm feature selection), and 7) RFE (recursive feature elimination) model. (Roffo & Melzi [23]).
- Thirdly, the ECFS model has not been used and evaluated in the algorithmic trading area.

Also, the description of the ECFS method and its corresponding formulas are summarized in Section S.1 of the Supplementary File.

3.2 Predicting technical indicators and forming images as CNN input

After determining the selected indicators, we use predicting indicators to build the input matrix of the CNN as part of asset selection in a portfolio trading system based on image processing, which has not been suggested in the literature. So, we utilize the LSTM deep learning model to predict the indicators as input of our CNN model. The high ability of the LSTM to predict time series, especially financial time series, has been shown in many studies such as Elsworth & Güttel [24].

Assume that N cryptocurrencies are selected for trading in the portfolio. The data required to run the proposed model includes the OHLC values of these cryptocurrencies. After data gathering, the value of M selected technical indicators and their predicted values for each cryptocurrency are used to form images as input to the CNN model to determine the buy/sell trading signals. The input of the CNN model is considered a two-dimensional image in which rows represent each indicator, and the columns represent time steps according to the (daily) trading frequency. The M first rows of the matrix contain the past values of the selected technical indicators from period $t\text{-}h$ to t. The M next rows contain the predicted values of the selected technical indicators from period $t+1$ to $t + h$. Eq. (1) shows the input matrix (image) of the classification model where $Indicator_{M,t}$ is Mth indicator at time t, and $\widehat{Indicator}_{m,t+h}$ is h steps ahead forecasting of Mth indicator. The LSTM network architecture and related formulas are described in Section S.2 of the Supplementary File.

$$
\begin{bmatrix}
Indicator_{1,t-h} & \cdots & Indicator_{1,t-2} & Indicator_{1,t-1} & Indicator_{1,t} \\
Indicator_{2,t-h} & \cdots & Indicator_{2,t-2} & Indicator_{2,t-1} & Indicator_{2,t} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
Indicator_{m,t-h} & \cdots & Indicator_{m,t-2} & Indicator_{m,t-1} & Indicator_{m,t} \\
\widehat{Indicator}_{1,t+1} & \widehat{Indicator}_{1,t+2} & \widehat{Indicator}_{1,t+1} & \cdots & \widehat{Indicator}_{1,t+h} \\
\widehat{Indicator}_{2,t+1} & \widehat{Indicator}_{2,t+2} \, . & \widehat{Indicator}_{2,t+1} & \cdots & \widehat{Indicator}_{2,t+h} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\widehat{Indicator}_{m,t+1} & \widehat{Indicator}_{m,t+2} & \widehat{Indicator}_{m,t+1} & \cdots & \widehat{Indicator}_{m,t+h}
\end{bmatrix}
\tag{1}
$$

3.3 Labeling (determining the target of the classification model)

We propose a novel approach based on the Zigzag indicator for labeling the price direction and determining the target of the classification model, which reduces the impact of random price fluctuations in the high-volatile cryptocurrency market. Also, using the Zigzag as a prediction target in the CNN model, compared to the well-known up-down labeling (for example, based on return), prevents the deep learning model from getting confused to some extent, and small changes in the pattern of indicators do not lead to a change in the value of the target variable [25]. The Zigzag indicator is usually used for a two-class labeling problem, which is used in conjunction with Elliot Wave Theory to determine the positioning of each wave in the prevalent cycle [26]. The indicator makes it easier to identify directions across all time frames by filtering out minor price movements, as shown in Figure 3.

Please insert Figure 3 about here.

In the process of calculating the zigzag indicator, an important parameter that leads to fundamental changes in the identification of pivot points is the Percentage of Price Change (PPC) parameter. The larger the PPC value, the smaller the fluctuations are removed from the price trend, and, thus the fewer the number of pivot points. This is shown in Figure 4. Also, the calculation process of the zigzag indicator is shown in Figure 5 [27]. In Table S.1 of the Supplementary File, we will show that changes in PPC and, as a result, changes in the smoothness of the zigzag indicator can significantly affect the accuracy of the proposed CNN-LSTM model.

After calculating the Zigzag indicator for the close price, we specify the value of the target (label) for time t as Eq. (2).

Please insert Figure 4 about here.

Please insert Figure 5 about here.

$$Label_t = \begin{cases} 1 & if \quad zigzag_{t+1} - zigzag_t \geq 0 \\ -1 & if \quad zigzag_{t+1} - zigzag_t < 0 \end{cases} \tag{2}$$

Where value 1 indicates the buy signal and value -1 indicates the sell signal.

Eventually, we use the zigzag approach for three main reasons:

- Firstly, the Zigzag indicator reduces the effect of random price fluctuations and is used to identify price directions and their changes [28].
- Secondly, this indicator reduces noise levels in price time series so that a notable decrease in noise during the labeling process reduces the frequency of the buying and selling signals.
- Thirdly, as shown in Figure 4, the fluctuations of the ZZ indicator change significantly over time and directly affect the model's accuracy. In addition, the cryptocurrency market is still in the early stage and it is highly volatile, so finding the optimal parameter of the ZZ indicator is necessary to maximize the model's accuracy in a fluctuating market.

3.4 Classification model for determining trading signals

After determining the input matrix and the target labels, a CNN model is used to classify each of the N-selected assets. As shown in Figure 6, layers of this network include: ***i)*** a 2-dimensional convolution layer with eight kernels 3×3 sizes, ***ii)*** batch normalization layer, ***iii)*** ReLU (Rectified Linear Unit) layer, ***iv)*** a 2-D fully connected layer, ***v***) softmax layer, and ***vi)*** classification layer. Also, the general recommendation in designing the architecture of a CNN network is to use increasing size for convolution layers and also to use kernel size with sizes 1, 3, and 5 [29]. The architectural framework used in this network is similar to the Alexnet network [30]. with a learning rate value of 0.01.

Please insert Figure 6 about here.

3.5 Global Minimum Variance Portfolio

After the classification model for determining trading signals of each asset using the CNN model, it is determined which assets are selected to be included in the portfolio. So, we adopt an HWM prediction to estimate the realized GMVP proportions to smooth actual covariance matrices, and the GMVP composition is subsequently calculated for assets with a high-risk profile. One of the current characteristics of cryptocurrencies is their high price volatility, which creates a risky investment environment. In the realized crypto portfolio weights of the proposed method, the GMVP based on HWM is used, which can model extreme risks to prevent severe losses [31].

To calculate the GMVP weights, the realized covariance matrix has to be estimated first. The realized covariance matrix can be described as shown in Eq. (3) [3].

$$R_t = \sum_{i=1}^{n} x_{t,i} x_{t,i}' \tag{3}$$

Where $x_{i,t} = p(t-1+\frac{i}{n}) - (t-1+\frac{i-1}{n})$ *with* $i=1,\ldots,n$, $p()$ is a Brownian stochastic, n is uniformly spaced intraday log-price vectors during day t, and $X_{t,i}$ is ith intraday return vector at day t. To determine the appropriate allocation of funds to risky assets, investors must establish the optimal portfolio proportions. The GMVP holds significance within the framework of Markowitz portfolio theory, serving as the initial reference point for the mean-variance efficient frontier that its structure is acquired through the selection of weights w_t that minimizes the variance of the portfolio as shown in Eq. (4).

$$W_t = \arg\min_{\omega_t} \omega_t'^{T} \sum_t \omega_t \text{, subject to } \omega_t' l = 1 \text{ where } l \text{ is a vector of ones} \tag{4}$$

Now, GMVP weights and the realized GMVP proportions are shown in Eqs. (5 and 6).

$$w_t^{GMVP} = \frac{\Sigma^{-1}_t l}{l' \Sigma^{-1}_t l} \tag{5}$$

$$\omega_t^{GMVP} = \frac{R^{-1}_t l}{l' R^{-1}_t l} \tag{6}$$

Golosnoy et al. [32] have demonstrated that the realized measures R_t and ω_t can serve as accurate point estimators of Σ_t and w_t, respectively. In order to make optimal decisions regarding minimum-variance portfolios, it is necessary to predict the composition of the GMVP, which involves forecasting w_{t+1} based on the information available at time t. Golosnoy et al. [3] have utilized the ES model for forecasting GMVP compositions. However, the ES model may not be suitable for datasets that exhibit seasonal patterns (e.g., day, week, month, year) as these seasonal factors need to be adjusted for the smoothing coefficient.

3.5.1 Statistical significance tests

This study focuses on investigating the Monday effect (a type of seasonal effect) in the cryptocurrency market, assuming a simple linear regression model:

$$d_{idt} = \alpha + \beta D_{idt} + \varepsilon_{idt} \tag{7}$$

Where *i* represents the various crypto assets, *t* denotes the specific time period being examined, *d* signifies the day of the week (d=Monday), and D_{idt} is the different daily dummy variables utilized. A dummy variable has been generated for each day of the week. However, our focus will primarily be on the days when anomalies have been identified in the existing literature such as Monday. The null hypothesis H_0 is denoted by β=0, indicating no statistical disparity between the daily return associated and the remaining days of the week. The alternative hypothesis H_1 is represented by $\beta \neq 0$, indicating that the difference in the daily return under consideration is statistically significant at a 0.1 significance level. Based on the aforementioned literature, we expect to observe positive average returns on Monday and the Monday [33,34].

3.5.2 Holt Winters Multiplicative (HWM) predictions

Bolarinwa et al. [35] revealed the HWM model as the most robust model based on out-of-sample forecast accuracy, among other models. Although extensive research has been carried out on HWM, no single study exists which use the HWM models for construction of a robust portfolio. Hence, we adopt the HWM to determine the realized covariance matrices based on the presence of seasonal effect, and the GMVP composition is computed afterward. In this regard, we incorporated our suggested equations into the Golosnoy et al. [3] model that addresses the seasonal effects, thereby rectifying its deficiencies. In other words, we recommend the inclusion of Eq. (8)-(12) in their model.

$$I_t = \gamma \frac{\Sigma_t}{S_{t-l}} + (1-\gamma)(I_{t-1} + T_{t-1}), \quad \gamma \in (0,1) \tag{8}$$

$$T_t = \beta(I_t - I_{t-1}) + (1-\beta)T_{t-1}, \qquad \beta \in (0,1) \tag{9}$$

$$S_t = \alpha \frac{\Sigma_t}{I_t} + (1-\alpha)S_{t-1}, \quad \alpha \in (0,1) \tag{10}$$

$$\hat{\Sigma}_{t+m} = (I_t + mT_t)S_{t-l+m} \tag{11}$$

Where Σ_t is the covariance matrices at time t, $\hat{\Sigma}_{t+m}$ is the prediction of the covariance matrix for m-step-ahead, I_t is the smoothing amount of level at time t, T_t is the smoothing amount of trend at time t, S_t is the smoothing amount of season at time t, l is season length, m is the number of period after time t and γ, β, α are level, trend, and season smoothing coefficients, respectively. Finally, the GMVP weights is calculated using Eq. (12).

$$w_t = \frac{\hat{\Sigma}_{t+1}^{-1} l}{l' \hat{\Sigma}_{t+1}^{-1} l} \tag{12}$$

4. Model evaluation

This section is dedicated to presenting data and results of our proposed dynamic portfolio trading system in the cryptocurrency market.

4.1 Data

In this study, the daily prices of thirty risky assets (as shown in Table 2) selected by the highest cryptocurrency market capitalization are obtained from finance.yahoo.com for training and testing of the proposed model between 7/1/2018 and 7/1/2023. Due to the increasing depth of the cryptocurrency market over the past year, the last year of the mentioned period is for testing the model, and the rest is for training our deep learning models. The training period is also considered using a sliding time window approach with a time step of 1 month (30 days) to update the learning models every month. Therefore, the testing period includes 12 sub-periods.

Please insert Table 2 about here.

In our study, 19 technical indicators and their predictions are used as the selected features. These indicators cover a variety of categories, including trend, mean reversion, relative strength, volume, and momentum for price and trading volume. This set of indicators has previously been used in studies such as Monsalve et al. [36], Hsu et al. [37], Kara et al. [38].

The utilized indicators include the A/D oscillator, Chaikin oscillator, MACD, Stochastic oscillator, Momentum, Chaikin volatility, Negative volume index, Positive volume index, RSI, A/D line, Bollinger band, Highest high, Lowest low, On balance volume, Price rate of change, PV trend, Typical price, the volume rate of change, Weighted closing price, and William A/D line. To evaluate the performance of the proposed LSTM-CNN-FS-ZZ model of this research, it has been compared with the CNN-based model, the LSTM-CNN-based model, and the LSTM-CNN-FS-based model, as explained next. To evaluate the LSTM model, Root Mean Square Error as $RMSE = \sqrt{\sum_{t=1}^{n} \frac{(\hat{y}_t - y_t)^2}{n}}$ is used,where y_t is the actual value of an indicator at time *t*, $\hat{y}_t$ is the predicted value of that indicator at time *t*, and *n* is the number of observations in the testing period in the LSTM model. The accuracy criterion is used to evaluate the performance of the CNN model, which indicates the ratio of the number of correct predictions of trading signals to all the predictions of the model. Finally, the compounded return of the trading systems compared to the Buy & Hold (BaH) strategy is utilized to evaluate the overall performance of the portfolio selection phase.

In the portfolio construction phase, the performance of models was evaluated using the Sharpe ratio, Sortino ratio along with statistical tests to see their statistical significance. The equations for the Sharpe and Sortino ratios are given by Eqs. (13) and (14), respectively [39].

$$sh_R = \frac{(R_p - R_f)}{\delta_p} \tag{13}$$

$$s_R = \frac{(R_p - R_f)}{\delta_{p-dowm}} \tag{14}$$

Where R_p is the return of the portfolio, R_F is the risk-free rate, and σ_p is the portfolio standard deviation. The modeling of the evaluation process of the proposed models was run in MATLAB 2022 software and in a system with features: Core i7 CPU and 16 GB RAM.

4.2 Results

Our dynamic portfolio trading system's overall performance is assessed in two phases: trading system performance and portfolio trading system performance. The trading system performance evaluates the effectiveness of our proposed model in the asset selection phase for each cryptocurrency, as well as the classifier's ability to differentiate between BaH returns. On the other hand, the portfolio trading system performance measures the performance of the proposed system by constructing a portfolio of cryptocurrencies.

4.2.1 Results of trading system performance (Portfolio selection phase)

In the results section, due to the high number of assets in the portfolio, the results of four cryptocurrencies with the highest market capitalization and popularity, which include Bitcoin (BTC), Ethereum (ETH), Binance Coin (BNB), and Cardano (ADA) are mentioned as examples. In addition, the average results of thirty cryptocurrencies (Avg. 30 Cryptos) are also evaluated. In multivariate statistics, a scree plot is a line plot of the eigenvalues of factors or principal components in an analysis [40]. The scree plot is used to determine the number of factors to retain in an exploratory factor analysis (FA) or principal components to keep in a principal component analysis (PCA). The procedure of finding statistically significant factors (components) using a scree plot is known as a scree test.

In Figure 7, the results of running the ECFS model for the first 4 cryptocurrencies are shown in the form of scree plots. The indicators that are on the upper side of the horizontal line (5%)

are the indicators that are selected. For ADA, BNB, and ETH, two indicators of A/D line and OBV have been selected as indicators. For BTC, in addition to these two indicators, A/D oscillator indicator has also been selected. An interesting point to note in the feature selection results is the selection of an indicator based on crowd sentiment and trading volume as the OBV. These results somehow indicate the low depth of the cryptocurrency market because decisions based on volume and market sentiments can be significant in the trading process.

Please insert Figure 7 about here.

After ranking and selecting appropriate technical indicators, the values of the selected indicators are predicted based on LSTM. The RMSE values for fitting the selected technical indicators of each cryptocurrency for each of the twelve training periods in the LSTM-CNN-FS models are presented in Table S.2 of the Supplementary File.

Tables 3 and 4 present the model training and test accuracy results for the CNN model, the LSTM-CNN model, the LSTM-CNN-FS model, and our proposed LSTM-CNN-FS-ZZ model for the 12 training and testing periods, respectively. As seen in Figure 8, the average testing accuracy of the 12 models with different training/testing periods of the LSTM-CNN-FS model is higher than the average learning accuracy of the CNN-based model of all cryptocurrencies (%64.262 accuracies). In other words, using the selected indicator prediction as input to the CNN learning model and the Zigzag indicator as a labeling method will improve the accuracy of our proposed trading system.

Please insert Table 3 about here.

Please insert Table 4 about here.

Please insert Figure 8 about here.

It can be seen that in Figure 8, the use of indicator predictions as input in the CNN model has increased the accuracy of the model by about 5%. By implementing ECFS model in our LSTM-CNN-FS model, the accuracy of the model has not changed significantly. Therefore, the goal of using feature selection has been achieved, which in addition to reducing the dimensions of the issue, also leads to an increase in the speed of solving the model (especially in this study, each indicator is predicted by an LSTM) and can prevent the possibility of overfitting or confusion of the model. The cryptocurrency market is still in a nascent stage of following the static curve and is highly volatile.

Please insert Figure 9 about here.

Therefore, it is necessary to find the optimal parameter that maximizes the accuracy the model in a fluctuating market. The variation in the accuracy of the CNN model with the changes of the PPC parameter in the training periods for different cryptocurrencies is shown in Figure 9 and Table S.1 in the Supplementary File. The evaluation range is [0.02, 0.18] based on trial and error. The noteworthy point is the difference between the maximum accuracy values in different training periods. In addition, the maximum accuracy value has occurred almost in the second half of the considered interval for the PPC parameter. After evaluating the prediction and classification models, our proposed trading system (Regardless of portfolio and for all cryptocurrencies one by one) is also compared with the BaH strategy, CNN, LSTM-CNN, and LSTM-CNN-FS methods. Each method, model and strategy has been implemented, and relevant financial calculation scenarios have been analyzed. In the BaH strategy, the stock is bought at the beginning of the test period and sold at the end of the test period. Figure 10 shows the daily compounded returns of Bitcoin over 30 trading days and 12 testing periods. As can be seen, the compounded returns of the proposed model are significantly better than the compounded returns of the competing models in all testing periods, as shown in Figure 11.

Please insert Figure 10 about here.

Please insert Figure 11 about here.

Also, the results of the compounded return of each cryptocurrency are shown in Table 5. The average annualized return for our proposed trading system (LSTM-CNN-FS-ZZ) is 35.13%, whereas BaH's average annualized return is 8.91%, CNN's average annualized return is 18.83, LSTM-CNN's average annualized return is 24.40%, and LSTM-CNN-FS average annualized return is 26.47%. Our proposed method's average annualized return is almost four times better than BaH's average annualized returns. At the same time, our proposed model is the only model with positive annualized returns for all cryptocurrencies because we adopt a novel approach based on the Zigzag indicator for labeling the price direction, which reduces the impact of random price fluctuations in the high-volatile cryptocurrency market.

Please insert Table 5 about here.

4.2.2 Results of portfolio trading system performance (Portfolio construction phase)

In the first step, we examine the Monday effect in the cryptocurrency market. The Monday effect can be classified as seasonal effect. The Monday effect refers to the tendency of Monday returns to be negative or lower compared to the rest of the week. Table 6 shows significant positive and larger average returns at a 10 percent significance level that occurred on Monday for ETH, BNB, ADA, and SOL. On the other hand, a reason for the not significantly positive Monday return of Bitcoin might be its maturity.

Please insert Table 6 about here.

After confirming the existence of seasonality effect in the cryptocurrency market, we adopt the HWM model to determine the realized covariance matrices and subsequently calculate the GMVP composition. To compare our proposed portfolio trading system performance with the benchmark models, the average testing periods of the Sharpe ratio and the Sortino ratio for the proposed model were compared with the equal-weighted (EW) approach. The equal weight approach is a method of proportional measurement that assigns equal significance to every asset within a portfolio. Therefore, this approach avoids concentrating too much of the weight on a few large assets and gives more weight to assets at the lower end of the market cap range. According to Figures 12 to 14, the GMVP based on the HWM showed a better performance compared to the equal-weighted approach in the two Sharpe and Sortino ratios because the GMVP is a portfolio of assets with a high-risk profile that has minimal volatility or variance than all other optimal portfolios.

Please insert Figure 12 about here.

Please insert Figure 13 about here.

Please insert Figure 14 about here.

Therefore, the risk-to-reward portfolio is very favorable for investors, especially risk-averse investors. However, in the periods 1 and 7 coincided with the sharp growth of cryptocurrencies in the middle of 2022 and the beginning of 2023, our proposed LSTM-CNN-FS-ZZ based on Predicted weights by the HWH method had a poorer performance in handling Large positive fluctuations despite a higher ability to handle small volatility.

Lastly, Table 7 summarizes each model's performance for diverse portfolio trading systems, utilizing descriptive statistics to compare further. Panels A and B of Table 7 display the descriptive statistics of daily returns and the results of statistical tests, respectively.

As indicated in Panel A of Table 7, our proposed trading system combined with the equally-weighted (EW) method yielded the highest daily mean return of 0.0047. Following this, the LSTM-CNN-FS+EW returns 0.0039, while the LSTM-CNN+EW records a return of 0.0037.

Also, the analysis indicates that the LSTM-CNN-FS-ZZ (our proposed trading system to predict buy/sell trading signals in the portfolio selection phase) combined with HWM (our proposed approach to calculate the GMVP composition in the portfolio construction phase) has the lowest level of risk, with a standard deviation of 0.0102 and a downside deviation of 0.0029. In comparison, the LSTM-CNN-FS+HWM ranks the second in terms of risk metrics.

This notable decrease in risk metrics of the GMVP approach based on the HWM has achieved the highest Sharpe and Sortino ratios in comparison to the EW approach, even though EW has the highest return.

Please insert Table 7 about here.

In Panel B of Table 7, we employ a one-tailed T-test for the assessment of means, an ANOVA-test for the analysis of variance, and two significance tests to evaluate the Sharpe and Sortino ratios, where the Markowitz portfolio optimization model (mean-variance) is considered as the baseline model for this analysis without including any prediction models. To satisfy the minimum sample size of the parametric T-test, we use daily returns instead of this study's test periods (monthly returns). Also, the null hypothesis H_0 is that the obtained empirical Sharpe or Sortino ratio is of zero mean. In mathematical terms, it says $E[Sh_R]$ or $E[S_R] = 0$. If we do not reject the null hypothesis, it indicates that the Sharpe or Sortino ratios lack statistical significance. On the other hand, if we reject the null hypothesis, we determine that the Sharpe ratio is statistically significant.

By observing the probability of p-values, it can be stated that the Sharpe and Sortino ratios of all portfolio trading systems are statistically significant, with a very high confidence level of 99%. Also, the T-test and ANOVA-test reveal that the proposed trading systems +EW and +HWM models are identified as this study's most influential cryptocurrencies portfolio trading systems compared to the base system (i.e., mean-variance), achieving higher profits and minimizing drawdowns, respectively.

5. Conclusion

This section describes the summary of the current study and the main results in the subsection 5.1. The research limitations and directions for future research are illustrated in subsection 5.2.

5.1 Concluding remarks

In this study, first, we propose the LSTM-CNN-FS-ZZ model to develop a trading system in crypto selection phase, where we utilize the LSTM, a type of recurrent neural network, to H-step ahead forecasting of the selected technical indicators, the Zigzag indicator to the labeling of the CNN model. The experimental results of the ECFS model have shown that although the feature selection model has not significantly increased the model's accuracy, it has led to a significant dimension reduction. This means that the number of indicators for prediction by the LSTM model is reduced which is a time-consuming operation with high computational complexity. On the other hand, it has reduced the input image dimensions of the CNN model. Also, he ZigZag approach can solve one of the challenges in labeling the networks of financial time series. We know that financial markets are highly non-linear and fluctuate widely due to the large number of events that can directly or indirectly affect market trends. One way to partially correct short-term adverse price movements that are problematic in swing trading is to use price smoothing techniques to identify trends. Reducing these short-term and consecutive fluctuations could facilitate correctly labeling a classification machine learning model. The Zigzag indicator effectively filters out noises and identifies significant price movements [41]. We also claim that the use of selected technical indicators with their H-step ahead forecasting can lead to increased accuracy of our proposed model. Therefore, in the proposed model, this issue has been considered, and multi-step ahead forecasting of indicators has been used Figure 11 and Table 5 show that the average annualized return for our proposed LSTM-CNN-FS-ZZ is 35.13%, whereas the BaH's average annualized return is 8.91%, and our claim was confirmed. In addition, our proposed model is the only model with positive annualized returns for all cryptocurrencies compared to other models. In realized crypto portfolio weights phase, we proposed an HWM model with three smoothing parameters $\gamma, \beta, \alpha \in (0, 1)$ to determine the best-weighted combination of risky assets as an extended version of the proposed model by Golosnoy et al. [3]. This approach is built to achieve the minimum possible portfolio variance for a given set of risky assets (cryptocurrencies) by selecting asset weights that minimize the variance, which leads to the construction of a robust portfolio. Figures 12-14 indicate that our developed portfolio trading system based on robust HWM forecasts outperforms the competing models for assets with a high-risk profile.

Also, the analysis of Table 7 shows that our proposed LSTM-CNN-FS-ZZ+HWM exhibits the lowest risk level, characterized by a standard deviation of 0.0102 and a downside deviation of 0.0029, which leads to the achievement of the highest Sharpe and Sortino ratios. Since our proposed model is the most favorable combination of high-risk assets among all efficient portfolios, which represents the minimum variance or risk in light of expected returns, we recommend it to risk-averse investors in high-risk markets such as cryptocurrencies.

5.2 Limitations and future research

Although we offered a new approach to portfolio trading systems with a significantly compounded return, the study has limitations that provide opportunities for further research. Examining the performance of the proposed model in other markets, such as the stock market, oil price, gold price, and interest rate, may be interesting for further research. In general, the noteworthy ideas that can influence future studies are as follows: 1) investigating our models with respect to the uncertain environments by incorporating fuzzy information for investor preferences, 2) constructing a portfolio with considering practical constraints such as transaction cost [42], 3) comparing the performance of our labeling approach with alternative approaches such as semi-Markov conditional random fields (semi-CRF) model [20] and triple trend labeling method [22] 4) tuning parameters of the CNN in our proposed portfolio trading system based on some metaheuristic algorithms such as genetic algorithm, 5) as an alternative to the suggested trading strategies, genetic programming can be used to generate trading rules [43], and 6) comparing the performance of our proposed GMVP approach based on HWM to state-of-the-art models in portfolio optimization such as mean-CDaR [19], data-driven risk measures [44], and EvoFolio [45].

The supplementary is available at:
file:///C:/Users/pc/Downloads/Supplementary%20File-SCI-2407-9205-1.pdf

References

1. Wang, W., Li, W., Zhang, N., et al. "Portfolio formation with preselection using deep learning from long-term financial data", Expert Systems with Applications, 143, p.113042 (2020). DOI: 10.1016/j.eswa.2019.113042
2. Kovačević, T., Merćep, A., Begušić, S., et al. "Optimal trend labeling in financial time series", IEEE access, pp. 83822 – 83832 (2023). DOI: 10.1109/ACCESS.2023.3303283
3. Golosnoy, V., Gribisch, B., and Seifert, M. I. "Exponential smoothing of realized portfolio weights", Journal of Empirical Finance, 53, pp. 222-237 (2019). DOI: 10.1016/j.jempfin.2019.07.00

4. Han, Y., Kim, J., and Enke, D. “A machine learning trading system for the stock market based on n-period min-max labeling using xgboost”, Expert Systems with Applications, 211, p. 118581 (2023). DOI:10.1016/j.eswa.2022.118581
5. Dakalbab, F., Talib, M. A., Nassir, Q., et al. “Artificial intelligence techniques in financial trading: a systematic literature review”, Journal of King Saud University-Computer and Information Sciences, p. 102015 (2024). DOI: 10.1016/j.jksuci.2024.102015
6. Sezer, O. B., and Ozbayoglu, A. M. “Algorithmic financial trading with deep convolutional neural networks: time series to image conversion approach”, Applied Soft Computing, 70, pp.525-538 (2018). DOI: 10.1016/j.asoc.2018.04.024
7. Luo, S., Lin, X., and Zheng, Z. “A novel cnn-ddpg based ai-trader: performance and roles in business operations”, Transportation Research Part E: Logistics and Transportation Review, 131, pp. 68-79 (2019). DOI: 10.1016/j.tre.2019.09.013
8. Yu, X., and Li, D. “Important trading point prediction using a hybrid convolutional recurrent neural network”, Applied Sciences, 11(9), p.3984. (2021). DOI: 10.3390/app11093984
9. Wu, J. M. T., Li, Z., Srivastava, G., et al. “A graph-based convolutional neural network stock price prediction with leading indicators”, Software: Practice and Experience, 51(3), pp.628-644 (2021a). DOI: doi.org/10.1002/spe.2915
10. Wu, M. E., Syu, J. H., Lin, J. C. W., et al. “Effective fuzzy system for qualifying the characteristics of stocks by random trading”, IEEE transactions on fuzzy systems, 30(8), pp.3152-3165 (2021b). DOI: 10.1109/TFUZZ.2021.3105192
11. Khodaee, p., Esfahanipour, A., and Mehtari Taheri.,H. “Forecasting turning points in stock price by applying a novel hybrid cnn-lstm-resnet model fed by 2d segmented images”, Engineering Applications of Artificial Intelligence 116, p.105464 (2022). DOI: 10.1016/j.engappai.2022.105464
12. Buachuen, W., and Kantavat, P. “Automated stock trading system using technical analysis and deep learning models”, In Proceedings of the 13th International Conference on Advances in Information Technology, pp. 1-9 (2023, December). DOI: 10.1145/3628454.3631670
13. Jing, L., and Kang, Y. “Automated cryptocurrency trading approach using ensemble deep reinforcement learning: learn to understand candlesticks”, Expert Systems with Applications, 237, p. 121373 (2024). DOI: 10.1016/j.eswa.2023.121373
14. Zou, J., Lou, J., Wang, B., et al. “A novel deep reinforcement learning based automated stock trading system using cascaded lstm networks”, Expert Systems with Applications, 242, p. 122801 (2024). DOI:10.1016/j.eswa.2023.122801
15. Massahi, M., and Mahootchi, M. “A deep q-learning based algorithmic trading system for commodity futures markets”, Expert Systems with Applications, 237, p.121711 (2024). DOI:10.1016/j.eswa.2023.121711
16. Ta, V. D., Liu, C. M., and Tadesse, D. A. “Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading”, Applied Sciences, 10(2), p. 437 (2020). DOI:10.3390/app10020437
17. Ma, Y., Han, R., and Wang, W. “Prediction-based portfolio optimization models using deep neural networks”, Ieee Access, 8, pp. 115393-115405 (2020). DOI: 10.1109/ACCESS.2020.3003819

18. Ashrafzadeh, M., Taheri, H. M., Gharehgozlou, M., et al. “Clustering-based return prediction model for stock pre-selection in portfolio optimization using pso-cnn+ mvp”, Journal of King Saud University-Computer and Information Sciences, 35(9), p. 101737 (2023). DOI:10.1016/j.jksuci.2023.101737
19. Alamdari, M. K., Esfahanipour, A., and Dastkhan, H. “A portfolio trading system using a novel pixel graph network for stock selection and a mean-cdar optimization for portfolio rebalancing”, Applied Soft Computing, 152, p. 111213 (2024). DOI: 10.1016/j.asoc.2023.111213
20. Lin, J. C. W., Shao, Y., Djenouri, Y., et al. “A recurrent neural network with an attention model for sequence labeling”, knowledge-based systems, 212, p. 106548 (2021). DOI: 10.1016/j.knosys.2020.106548
21. Lu, G., Liu, Y., Wang, J., et al. “CNN-BiLSTM-Attention: A multi-label neural classifier for short texts with a small set of labels”, Information Processing & Management, 60(3), p. 103320 (2023). DOI:10.1016/j.ipm.2023.103320
22. Peng, P., Chen, Y., Lin, W., et al. “Attention-based cnn–lstm for high-frequency multiple cryptocurrency trend prediction”, Expert systems with applications, 237, p. 121520 (2024). DOI:10.1016/j.eswa.2023.121520
23. Roffo, G., and Melzi, S. “Feature selection via eigenvector centrality”, In The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Springer, pp. 1-12. (2016). DOI: 10.48550/arXiv.1704.05409
24. Elsworth, S., and Güttel, S. “Time series forecasting using lstm networks: a symbolic approach”, arXiv preprint arXiv:2003, P. 05672 (2020). DOI: 10.48550/arXiv.2003.05672
25. Qi, L., Khushi, M., and Poon, J. “Event-driven lstm for forex price prediction”, In 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), pp. 1-6 (2020). DOI: 10.1109/CSDE50874.2020.9411540
26. Seo, Y. H. “Profitable chart patterns in forex and sock market: fibonacci analysis, harmonic pattern, elliott wave, and x3 chart pattern”,(2019). ISBN: 9781393764137, 1393764134
27. Özorhan, M. O., Toroslu, İ. H., and Şehitoğlu, O. T. “Short-term trend prediction in financial time series data”, Knowledge and Information Systems, 61, pp. 397-429 (2019). DOI: 10.1007/s10115-018-1303-x
28. Wang, X., Zhang, H., Zhang, Y., et al. “Learning nonstationary time-series with dynamic pattern extractions”, IEEE Transactions on Artificial Intelligence, 3(5), pp. 778-787 (2021). DOI: 10.1109/TAI.2021.3130529
29. Alzubaidi, L., Zhang, J., Humaidi, A. J., et al. “Review of deep learning: concepts, cnn architectures, challenges, applications, future directions”, Journal of Big Data, 8(1), pp. 1-74 (2021). DOI: 10.1109/TKDE.2020.2981333
30. Alom, M. Z., Taha, T. M., Yakopcic, C., et al. “The history began from alexnet: a comprehensive survey on deep learning approaches”, arXiv preprint arXiv:1803, p. 1164 (2018). DOI: 10.48550/arXiv.1803.01164
31. Bodnar, T., and Okhrin, Y. “Boundaries of the risk aversion coefficient: should we invest in the global minimum variance portfolio”, Applied Mathematics and Computation, 219, pp. 5440–5448 (2013). DOI: 10.1016/j.amc.2012.11.049
32. Golosnoy, V., Schmid, W., Seifert, M. I., et al. “Statistical inferences for realized portfolio weights”, Econometrics and Statistics, 14, pp. 49-62. (2020). DOI: 10.1016/j.ecosta.2018.08.003
33. Kaiser, L. “Seasonality in cryptocurrencies”, Finance Research Letters, 31, (2019). DOI: 10.1016/j.frl.2018.11.007

34. Baur, D. G., Cahill, D., Godfrey, K., et al. "Bitcoin time-of-day, day-of-week and month-of-year effects in returns and trading volume", Finance Research Letters, 31, pp. 78-92 (2019). DOI: 10.1016/j.frl.2019.04.023
35. Bolarinwa, I. A., Abutu, B., and Ugbedeojo, O. I. "On robustness of three linear time series models to an outlier", African Journal of Advances in Science and Technology Research, 12(1), pp. 35-42 (2023). DOI: 10.27620/afrop.1511.45858
36. Monsalve, S., Suárez-Cetrulo, A. L., Cervantes, A., et al. "Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators", Expert Systems with Applications, 149, p. 113250 (2020). DOI: 10.1016/j.eswa.2020.113250
37. Hsu, M.-W., Lessmann, S., Sung, M.-C., et al. "Bridging the divide in financial market forecasting: machine learners vs. financial economists", Expert Systems with Applications, 61, pp. 215-234 (2016). DOI: 10.1016/j.eswa.2016.05.033
38. Kara, Y., Boyacioglu, M. A., and Baykan, Ö. K. "Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul stock exchange", Expert Systems with Applications, 38(5), pp. 5311-5319 (2011). DOI: 10.1016/j.eswa.2010.10.027
39. Sortino, F. A. "The sortino framework for constructing portfolios: focusing on desired target return to optimize upside potential relative to downside risk", Elsevier, pp. 23-49 (2009). ISBN: 9780080961682
40. Holland, S. M. "Principal components analysis (pca) ", Department of Geology, University of Georgia, Athens, GA, p. 2501 (2008). DOI:10.1038/s43586-022-00184-w
41. Wang, D., and Shang, Y. "A new active labeling method for deep learning", In 2014 International joint conference on neural networks (IJCNN), pp. 112-119 (2014). DOI: 10.1109/IJCNN.2014.6889457
42. Babazadeh, H., and Esfahanipour, A. "A novel multi period mean-var portfolio optimization model considering practical constraints and transaction cost", Journal of Computational and Applied Mathematics, 361, pp. 313-342 (2019). DOI: 10.1016/j.cam.2018.10.039
43. Mousavi, S., Esfahanipour, A., and Zarandi, "Multitree genetic programming-based learning of interpretable and accurate tsk systems for dynamic portfolio trading", Applied Soft Computing, 34, pp. 449-462 (2015). DOI: 10.1016/j.asoc.2015.05.021
44. Bowala, S., and Singh, J. "Optimizing portfolio risk of cryptocurrencies using data-driven risk measures", Journal of Risk and Financial Management, 15(10), p. 427 (2022). DOI: 10.3390/jrfm15100427
45. Guarino, A., Santoro, D., Grilli, L., et al. "Evofolio: a portfolio optimization method based on multi-objective evolutionary algorithms", Neural Computing and Applications, 36(13), pp. 7221-7243 (2024). DOI:10.1007/s00521-024-09456-w

Biographies:

Ehsan Mohammadian Amiri is currently a Ph.D candidate at the Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran. He received his MSc degree in Industrial Engineering from Khajeh Nasir Toosi University of Technology, Tehran, Iran, in 2017. His current research interests are in the fields of artificial intelligence, machine learning fuzzy sets theory, risk management, and related applications. He has published 10 research papers in relevant scientific journals and presented 5 papers at international conferences.

Akbar Esfahanipour is an associate professor at the Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran. He received his PhD in Industrial Engineering from Tarbiat Modares University in the field of intelligent decision making in the stock market. He also pursued a postdoctoral fellowship in management information systems at McMaster University. His research interests include financial resilience, behavioral finance, financial risk analysis, and developing intelligent systems for forecasting, decision making, and risk analysis for managers and investors.

List of Figure captions:

Figure 1. Structure of the proposed portfolio trading system for cryptocurrencies

Figure 2. Overview of the proposed dynamic portfolio trading model

Figure 3. Example of the ZigZag indicator for Bitcoin

Figure 4. Indicator changes and standard deviation of Zigzag indicator with PPC parameter for the Bitcoin

Figure 5. Overview of the process of implementing the Zigzag indicator as a pseudo-code style

Figure 6. A CNN architecture in our proposed model

Figure 7. Selected features for the first 4 cryptocurrencies using the ECFS method

Figure 8. Average training and testing period accuracy (%) of all cryptocurrencies for different models

Figure 9. Sample of variations of LSTM-CNN-FS-ZZ accuracy(%) with changes of the PPC parameter

Figure 10. Compounded return of Bitcoin in testing periods for different models

List of Table captions:

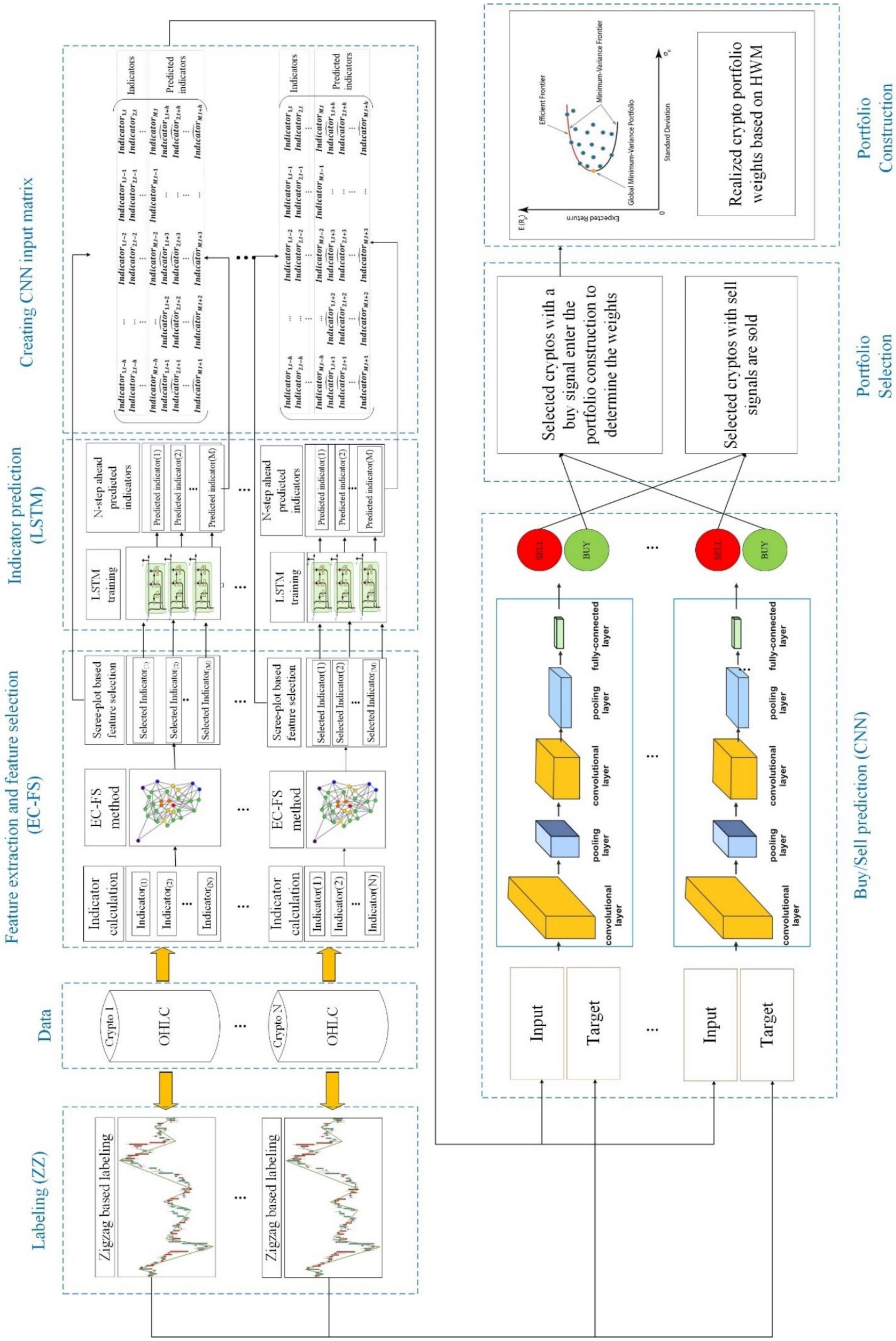


Figure 1. Structure of the proposed portfolio trading system for cryptocurrencies

```
BEGIN
  READ prices of selected assets (open, high, low, close)
  DEFINE number of testing periods (12)
  DEFINE training period days (30)
  DEFINE CNN training horizon (1000)
  DEFINE CNN test horizon (30)
  DEFINE CNN image length (5)
  DEFINE LSTM training horizon (1000)

  FOR each Crypto DO
    CALCULATE technical indicators
    RANK indicators using ECFS
    SELECT indicators using scree-plot
  END FOR

  FOR each testing period DO
    FOR each Crypto DO
      LSTM training phase
      CALCULATE selected indicators
      LSTM training for each indicator in the LSTM training horizon
      SAVE LSTM networks

      PREDICT indicators in CNN image length for CNN training horizon by trained LSTM network


      CNN training phase
      CREATE images by indicators and predictions (CNN input)
      CALCULATE Zigzag indicator for CNN training horizon
      CNN training

      CNN testing phase
      PREDICT indicators in CNN image length for CNN testing horizon
      CREATE images by indicators and their predictions (CNN input)
      CALCULATE Zigzag indicator for CNN testing horizon
      COMPUTE accuracy of the CNN network for the CNN testing period
      SAVE CNN networks
    END FOR

    FOR each day of CNN testing Horizon DO
      PREDICT label of each crypto by related CNN network (crypto selection)
      HWM approach of realized GMVP (weight determination)
      REBALANCE portfolio weights based on new calculated weights.
    END FOR

    CALCULATE portfolio return based on Cryptos weights
    SAVE results
  END FOR
END
```

Figure 2. Overview of the proposed dynamic portfolio trading model

Figure 3. Example of the ZigZag indicator for Bitcoin

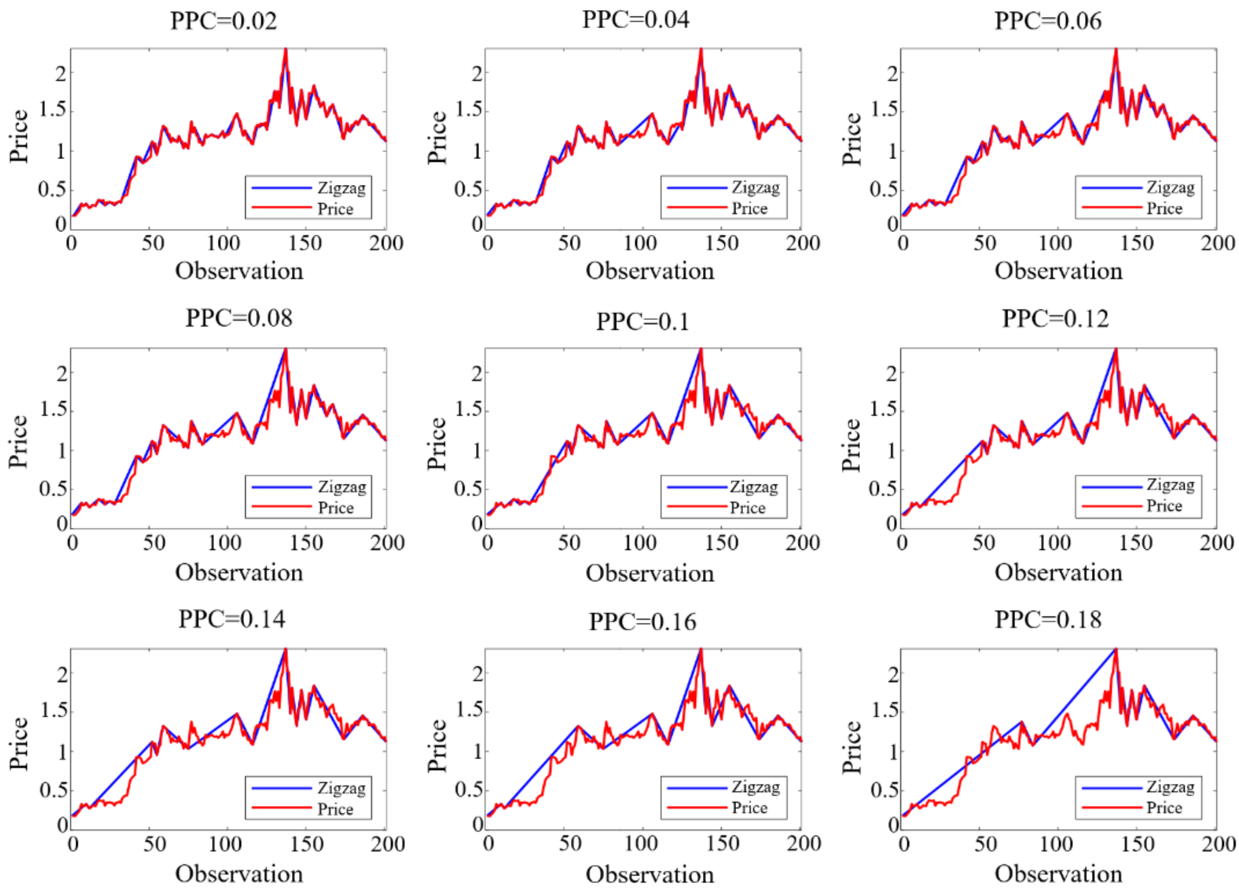


Figure 4. Indicator changes and standard deviation of Zigzag indicator with PPC parameter for the Bitcoin

```
BEGIN
  high_point = 0
  low_point = 0
  FOR each price_point in data_set DO
    IF price_point > previous_high_point THEN
      high_point = price_point
    END IF
    IF price_point < previous_low_point THEN
      low_point = price_point
    END IF
    distance = high_point - low_point
    IF distance > ppc THEN
      mark pivot points as high_point and low_point
      IF the pivot point is marked THEN
        set opposite point as the new high or low point
      END IF
    END IF
  END FOR
  OUTPUT pivot points as Zigzag indicator
END
```

Figure 5. Overview of the process of implementing the Zigzag indicator as a pseudo-code style

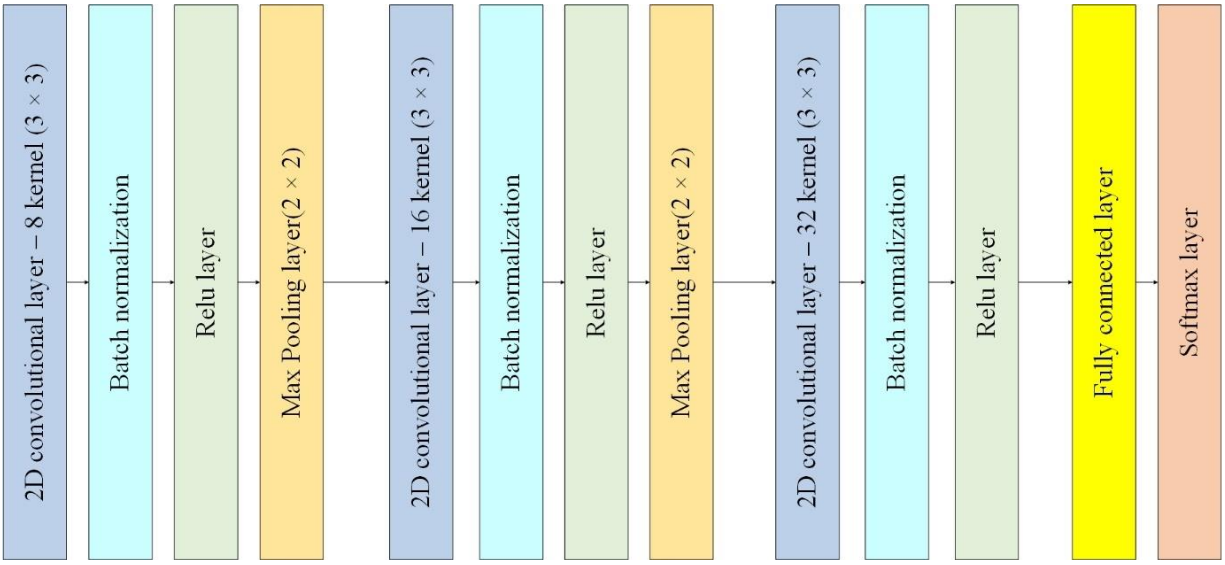
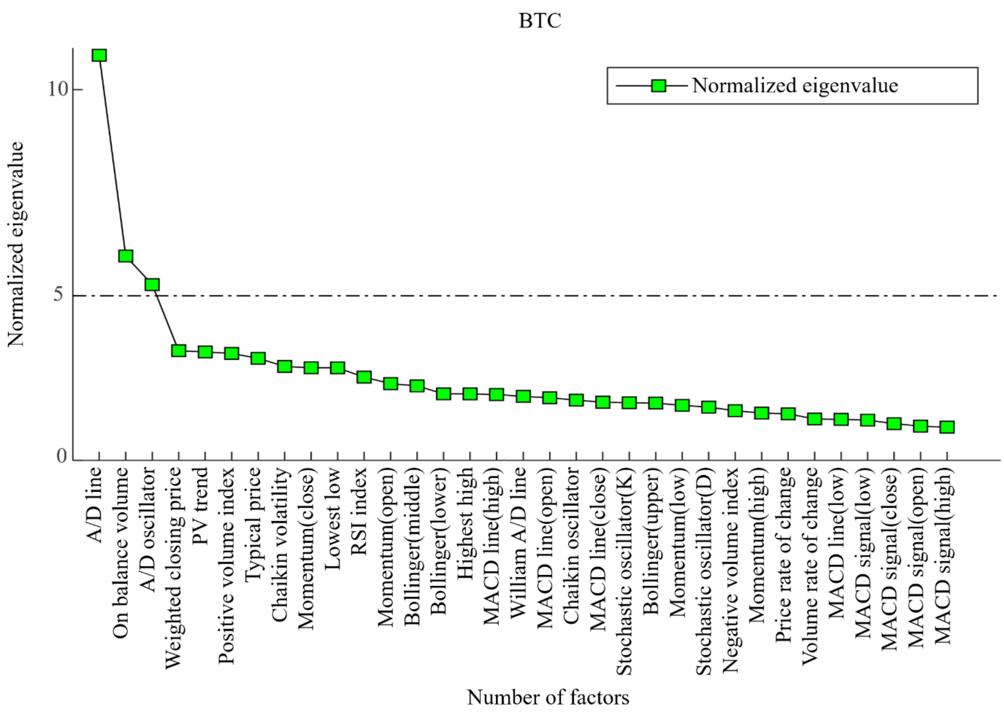
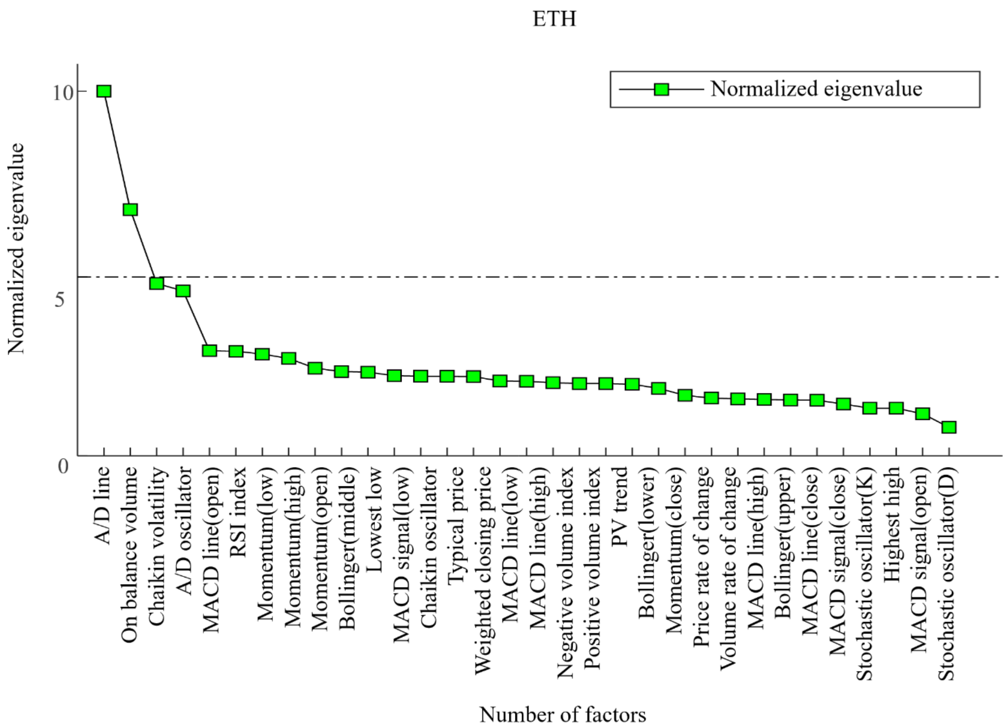
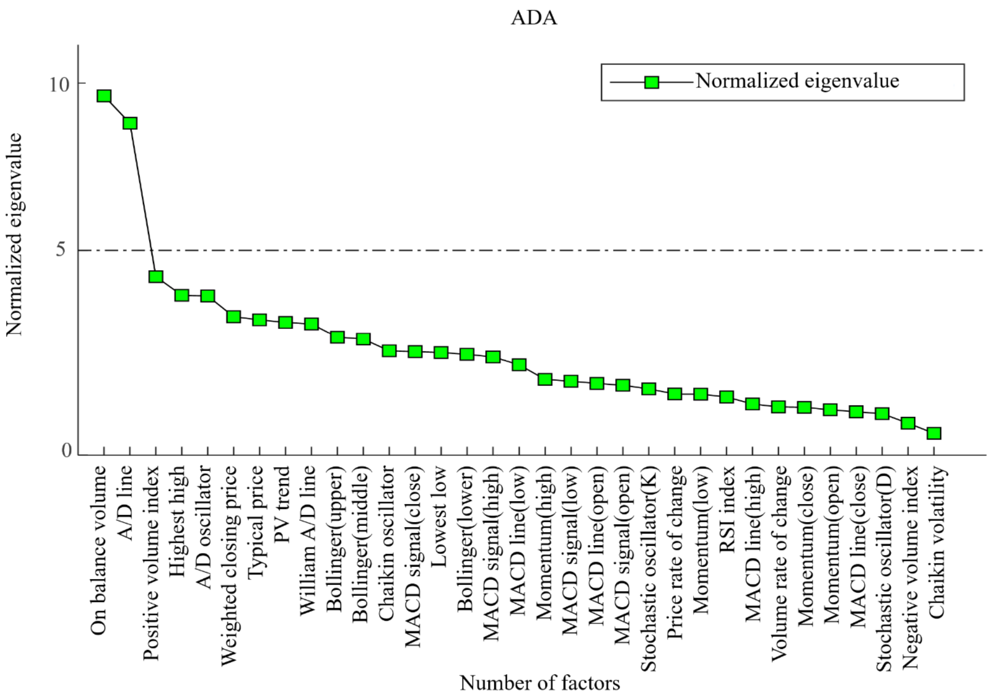


Figure 6. A CNN architecture in our proposed model


BTC
Normalized eigenvalue
10
5
0
Normalized eigenvalue
A/D line
On balance volume
A/D oscillator
Weighted closing price
PV trend
Positive volume index
Typical price
Chaikin volatility
Momentum(close)
Lowest low
RSI index
Momentum(open)
Bollinger(middle)
Bollinger(lower)
Highest high
MACD line(high)
William A/D line
MACD line(open)
Chaikin oscillator
MACD line(close)
Stochastic oscillator(K)
Bollinger(upper)
Momentum(low)
Stochastic oscillator(D)
Negative volume index
Momentum(high)
Price rate of change
Volume rate of change
MACD line(low)
MACD signal(low)
MACD signal(close)
MACD signal(open)
MACD signal(high)
Number of factors


ETH
Normalized eigenvalue
10
5
0
Normalized eigenvalue
A/D line
On balance volume
Chaikin volatility
A/D oscillator
MACD line(open)
RSI index
Momentum(low)
Momentum(high)
Momentum(open)
Bollinger(middle)
Lowest low
MACD signal(low)
Chaikin oscillator
Typical price
Weighted closing price
MACD line(low)
MACD line(high)
Negative volume index
Positive volume index
PV trend
Bollinger(lower)
Momentum(close)
Price rate of change
Volume rate of change
MACD line(high)
Bollinger(upper)
MACD line(close)
MACD signal(close)
Stochastic oscillator(K)
Highest high
MACD signal(open)
Stochastic oscillator(D)
Number of factors



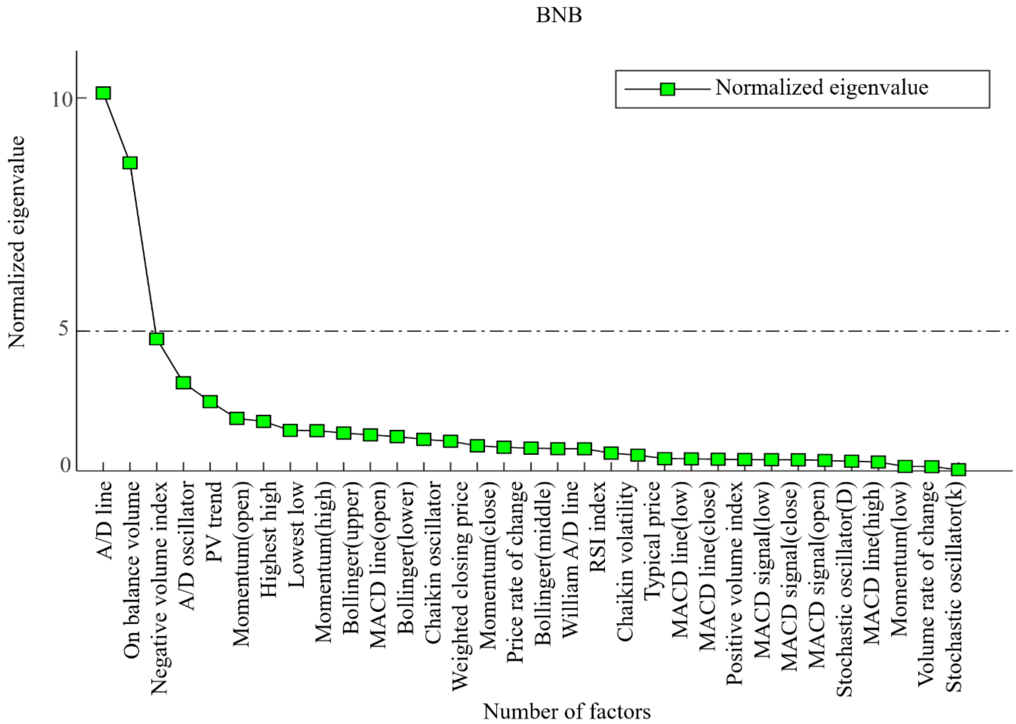


Figure 7. Selected features for the first 4 cryptocurrencies using the ECFS method

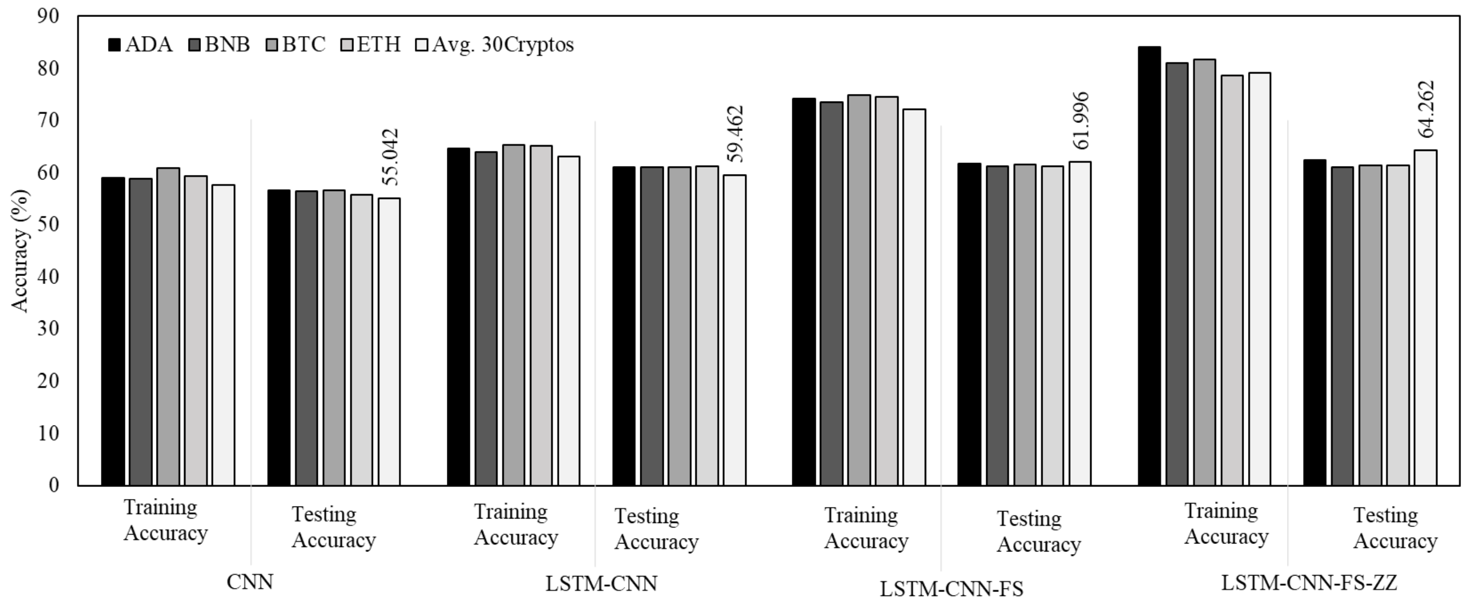


Figure 8. Average training and testing period accuracy (%) of all cryptocurrencies for different models

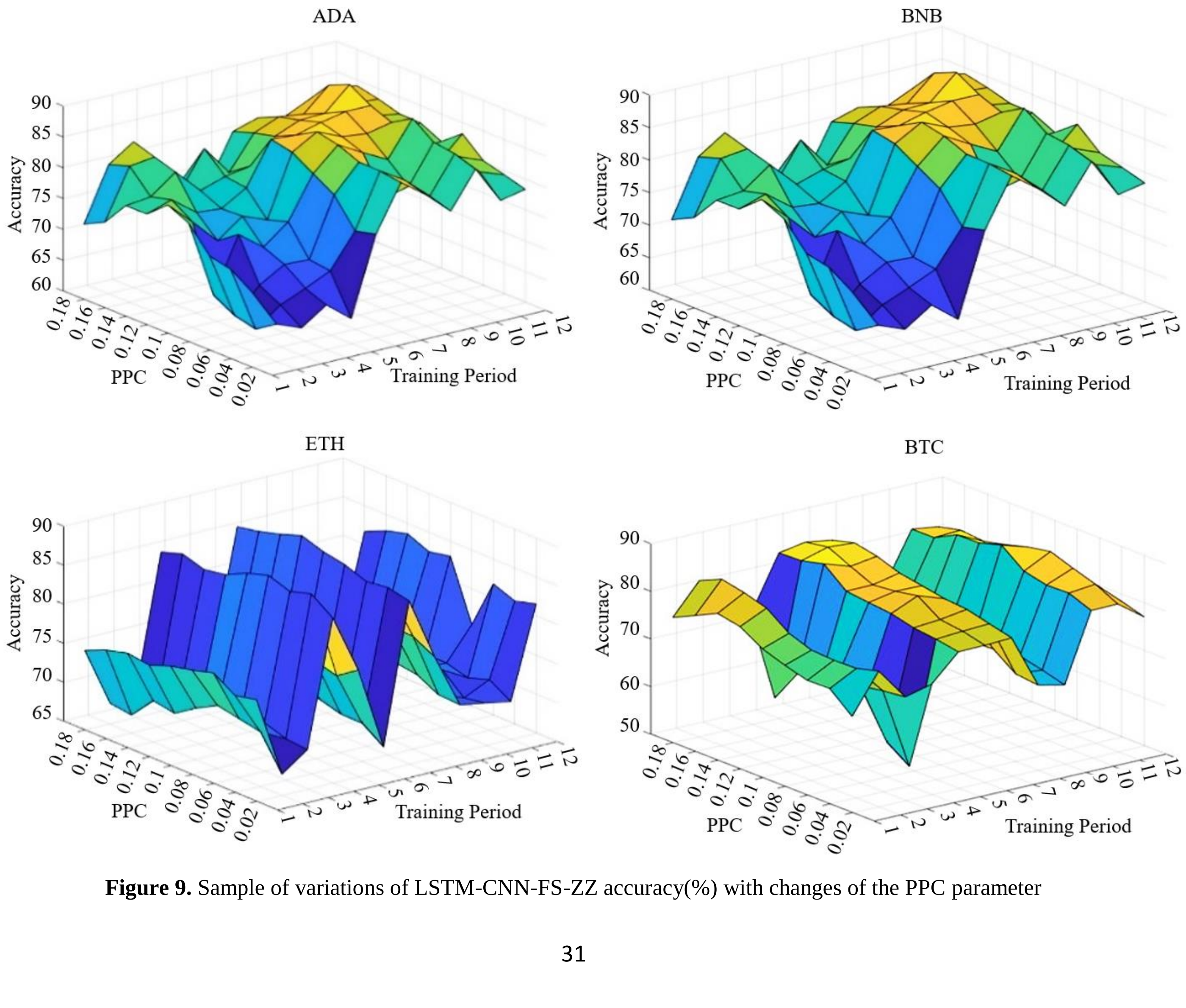


Figure 9. Sample of variations of LSTM-CNN-FS-ZZ accuracy(%) with changes of the PPC parameter

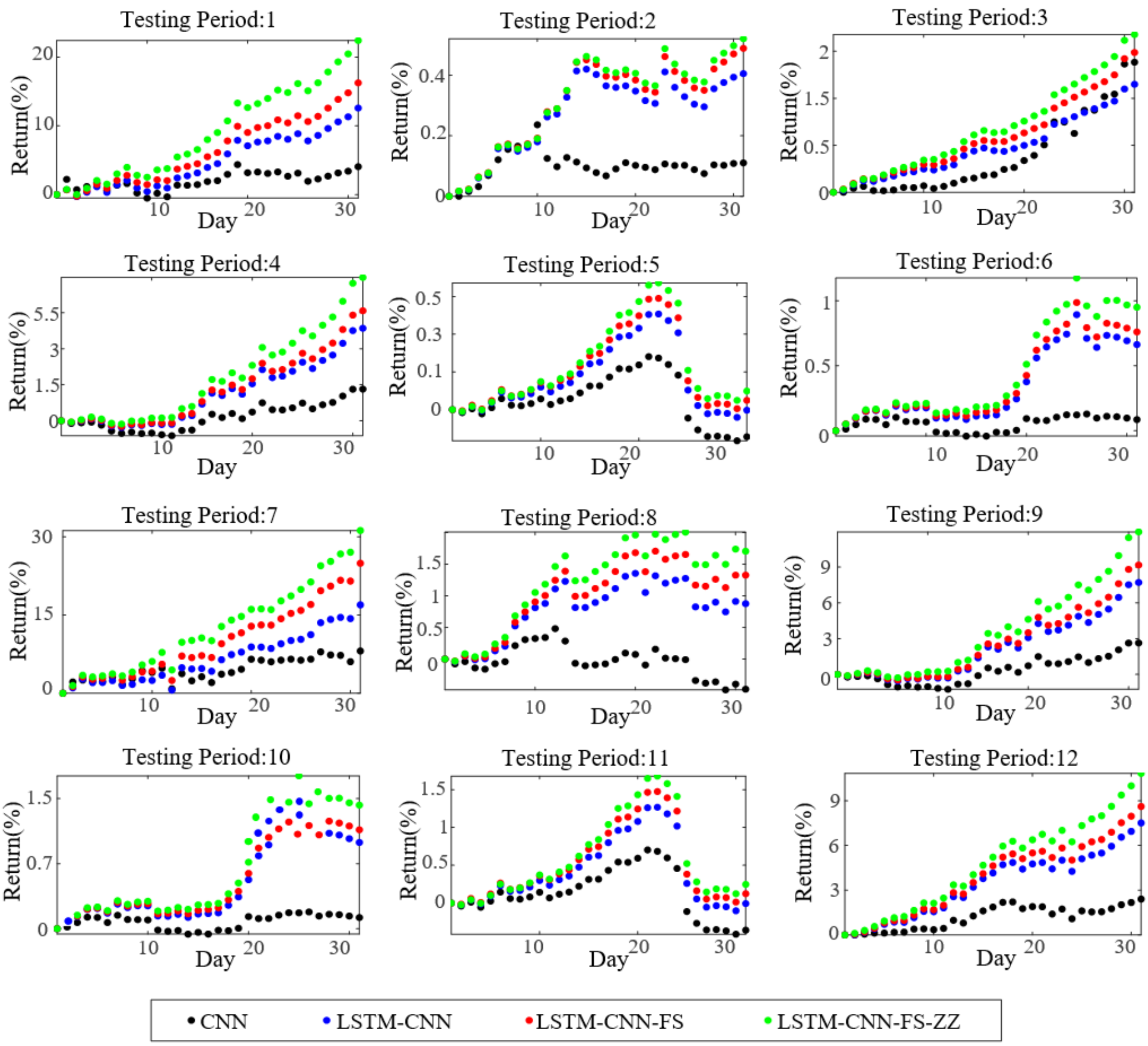


Figure 10. Compounded return of Bitcoin in testing periods for different models

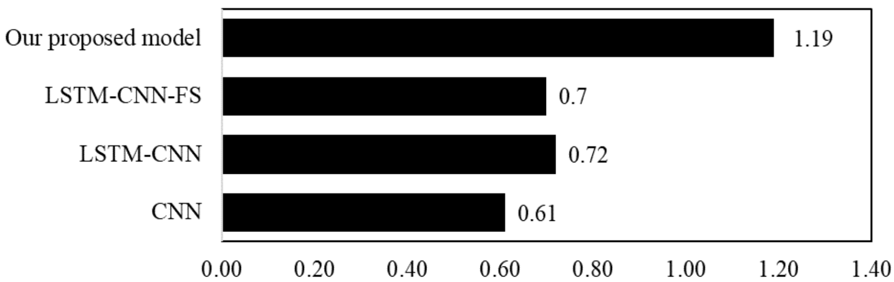


Figure 11. Average end-of-period compounded return of Bitcoin

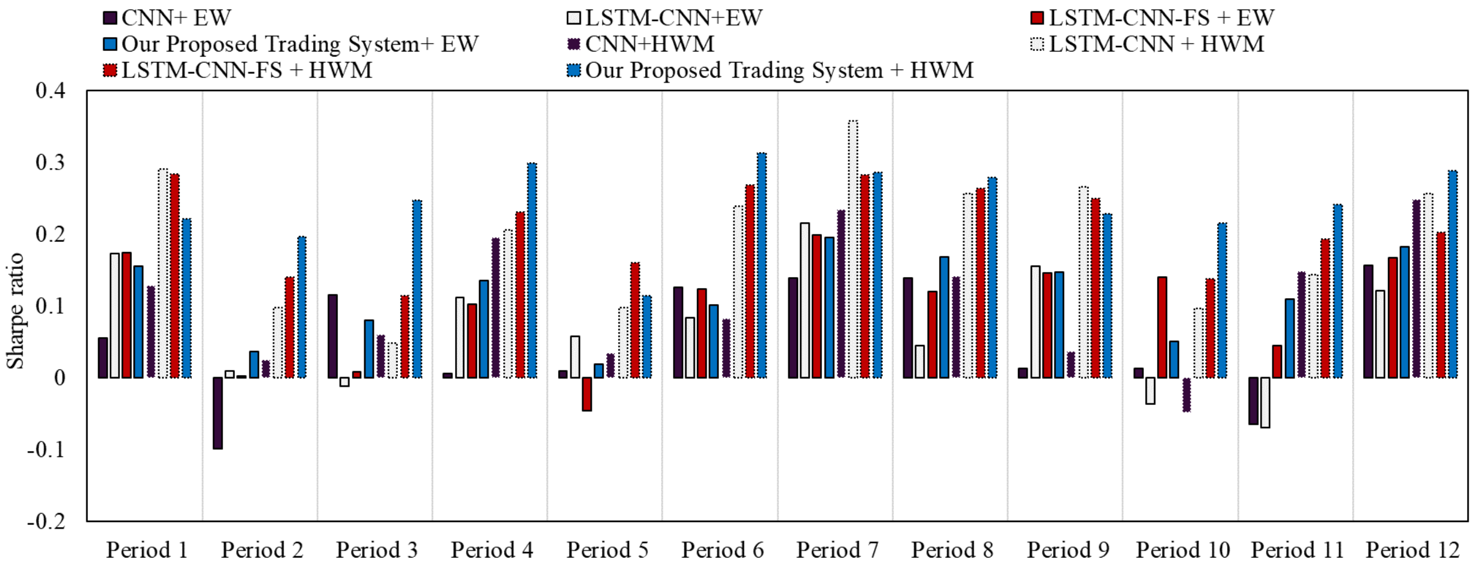


Figure 12. Average daily Sharpe ratio for each testing period

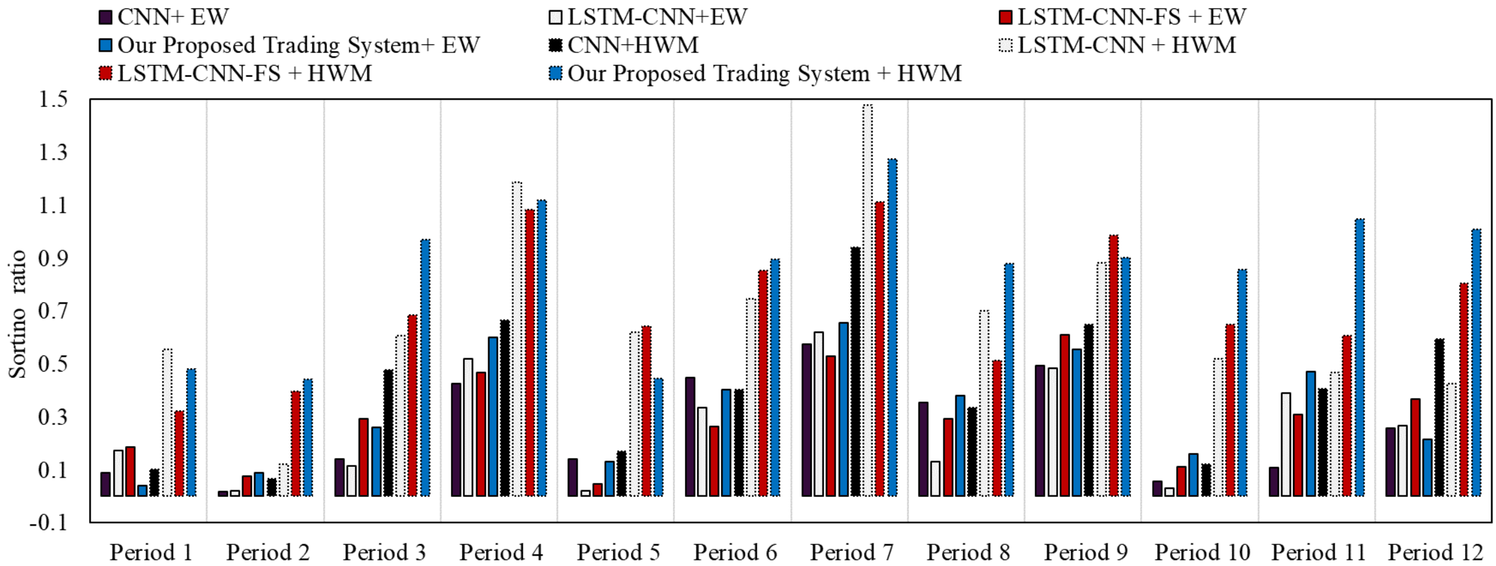


Figure 13. Average daily Sortino ratio for each testing period

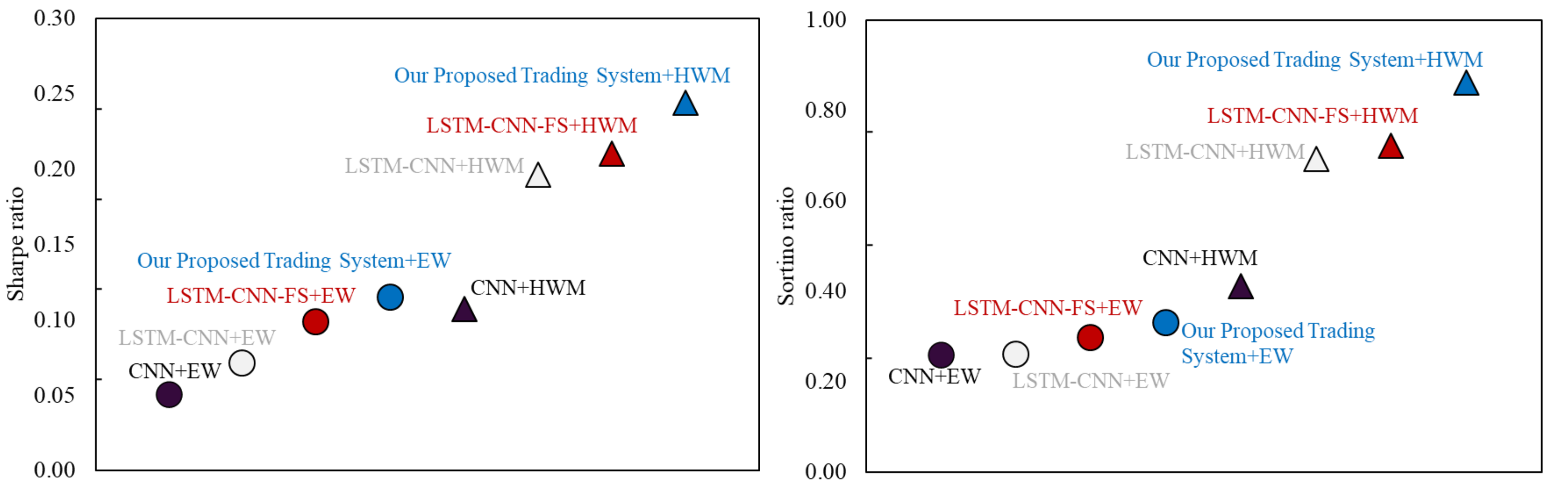


Figure 14. Average testing period Sharpe ratio and Sortino ratio of each approach for all periods

Table 1. Related studies in the literature

Author(s) & Year	Research Area				
	Algorithmic trading				Portfolio construction /opt
	Input Type	Models	Application		
			Trading system	Portfolio trading system	
Sezer and Ozbayoglu (2018)	28 stocks of Dow30 and ETFs	CNN-TA	✓	×	×
Luo et al. (2019)	OLHCV of stock-index future and 18 indicators	CNN-DDPG	✓	×	×
Ma et al. (2020)	Stocks of China Securities 100 index in the Chinese stock market	DMLP, LSTM, and CNN	×	Predicting stock returns as a portfolio selection phase	MSAD and EW
Ta et al. (2020)	500 large-cap stocks listed on the American Stock Exchange S&P 500	LSTM and RNN	×	Predicting stock movement as a portfolio selection phase	EQ, MCS and MVO
Yu and Li (2021)	50 American stocks	ITPP-HCRNN	✓	×	×
Khodaee et al. (2022)	Two markets of Exchange-traded funds (ETFs) and Dow	CNN-LSTM-ResNet	✓	×	×
Ashrafzadeh et al. (2023)	OLHCV of 21 stocks in the New York Stock Exchange	CNN- PSO	✓	Predicted stock returns as a portfolio selection phase	MVF
Jing and Kang (2024)	OLHCV of cryptocurrency market	Ensemble deep reinforcement learning	✓	×	×
Zou et al. (2024)	30 stocks in the SSE 50 in China, 30 stocks in SENSEX in India, and 30 stocks in FTSE 100 in UK	DRL and Cascaded LSTM	✓	×	×
Alamdari et al. (2024)	OLHCV of 18 stocks from the New York Stock Exchange	PGN	×	Predicting buy trading signals as a portfolio selection phase	M-CDaR
Massahi and Mahootchi (2024)	Commodity futures markets	DQN and DDQN	✓	×	×
Current Study	**30 cryptocurrencies, 19 technical indicators + forecasting**	**LSTM-CNN- ZZ-FS**	×	**Predicting buy trading signals as a portfolio selection phase**	**HWM to determine the realized portfolio weights**

Table 2. The list of 30 cryptocurrencies selected by highest market capitalization.

Symbol	Name	Symbol	Name	Symbol	Name
BTC	Bitcoin	TRX	TRON	ICON	ICX
ETH	Ethereum	XMR	Monero	Lisk	LSK
XRP	XRP	Neo	NEO	Zilliqa	ZIL
BCH	Bitcoin Cash	Dash	DASH	Decred	DCR
EOS	EOS	Binance Coin	BNB	Bitcoin Gold	BTG
LTC	Litecoin	Ethereum Classic	ETC	0x	ZRX
XLM	Stellar Lumens	NEM	XEM	Siacoin	SC
ADA	Cardano	Ontology	ONT	Maker	MKR
IOTA	MIOTA	Qtum	QTUM	Dogecoin	DOGE
USDT	Tether	Zcash	ZEC	Waves	WAVES

Table 3. The accuracies (%) of the CNN and LSTM-CNN models for each of the cryptocurrency

Testing Period	Training Accuracy					Testing Accuracy				
	ADA	**BNB**	**BTC**	**ETH**	**Avg. 30Cryptos**	**ADA**	**BNB**	**BTC**	**ETH**	**Avg. 30Cryptos**
1	61.6327* 60.4082**	60.0000* 64.0816**	62.0408* 69.3878**	60.4082* 67.7551**	55.1020* 59.5918**	43.6667* 51.6667**	51.3333* 58.3333**	64.6667* 71.6667**	43.6667* 51.6667**	51.6667* 61.6667**
2	57.9592* 61.6327**	62.8571* 64.0816**	61.2245* 64.4898**	60.0000* 65.7143**	59.5918* 60.0000**	56.6667* 57.6667**	61.6667* 71.6667**	72.6667* 80.6667**	46.0000* 50.0000**	64.6667* 69.6667**
3	50.6122* 63.2653**	61.2245* 65.3061**	62.0408* 68.1633**	61.6327* 64.8980**	59.1837* 68.5714**	63.3333* 64.3333**	45.6667* 50.6667**	53.6667* 57.6667**	45.6667* 46.6667**	69.0000* 70.0000**
4	57.5510* 63.2653**	52.6531* 59.5918**	60.0000* 65.7143**	61.6327* 67.7551**	59.1837* 66.1224**	53.6667* 60.6667**	63.3333* 65.3333**	50.3333* 56.3333**	62.3333* 63.3333**	54.6667* 58.6667**
5	58.3673* 66.1224**	55.5102* 60.8163**	63.2653* 65.7143**	63.6735* 68.9796**	62.4490* 66.5306**	61.3333* 66.3333**	46.0000* 51.0000**	62.3333* 65.3333**	45.0000* 53.0000**	53.6667* 58.6667**
6	59.5918* 64.8980**	60.4082* 65.7143**	63.2653* 66.1224**	59.1837* 64.8980**	58.7755* 62.4490**	57.0000* 65.0000**	60.0000* 63.0000**	60.0000* 61.0000**	66.0000* 72.0000**	51.6667* 56.6667**
7	61.6327* 66.5306**	59.1837* 64.4898**	58.7755* 65.7143**	59.5918* 68.1633**	57.5510* 66.5306**	66.6667* 66.6667**	64.6667* 70.6667**	41.3333* 43.3333**	69.0000* 74.0000**	61.3333* 66.3333**
8	57.5510* 65.7143**	60.0000* 64.4898**	59.1837* 64.8980**	55.5102* 62.8571**	59.5918* 61.2245**	42.6667* 48.6667**	51.3333* 57.3333**	46.0000* 50.0000**	55.0000* 62.0000**	48.0000* 54.0000**
9	63.6735* 66.5306**	57.1429* 64.0816**	57.1429* 61.6327**	56.3265* 62.8571**	57.1429* 62.4490**	60.3333* 64.3333**	62.3333* 63.3333**	54.6667* 57.6667**	58.3333* 65.3333**	70.0000* 74.0000**
10	63.2653* 68.5714**	61.6327* 64.4898**	62.4490* 66.9388**	59.1837* 63.2653**	59.5918* 64.4898**	56.6667* 57.6667**	62.3333* 63.3333**	54.6667* 61.6667**	50.3333* 58.3333**	55.6667* 61.6667**
11	58.7755* 64.4898**	58.3673* 64.0816**	60.8163* 66.1224**	57.1429* 60.8163**	60.0000* 64.4898**	55.0000* 65.0000**	62.3333* 63.3333**	70.0000* 73.0000**	72.6667* 81.6667**	54.6667* 57.6667**
12	57.5510* 63.6735**	57.9592* 66.1224**	59.5918* 59.1837**	58.3673* 63.2653**	57.9592* 63.6735**	63.3333* 65.3333**	47.0000* 55.0000**	48.3333* 54.3333**	55.6667* 57.6667**	65.6667* 70.6667**

The accuracies with an asterisk (*) and (**) are accuracies (%) of the CNN model and LSTM-CNN model, respectively.

Table 4. The accuracies (%) of the LSTM-CNN-FS and our proposed models for each of the cryptocurrency

Testing Period	Training Accuracy					Testing Accuracy				
	ADA	**BNB**	**BTC**	**ETH**	**Avg. 30Cryptos**	**ADA**	**BNB**	**BTC**	**ETH**	**Avg. 30Cryptos**
1	70.2040* 80.8163**	73.8775* 84.4898**	79.1836* 80.8163**	77.5510* 79.1837**	69.3877* 71.0204**	51.6666* 48.6667**	54.3333* 54.3333**	67.6666* 70.6667**	50.6666* 48.6667**	58.6666* 56.6667**
2	71.4285* 82.0408**	73.8775* 84.4898**	74.2857* 84.8980**	75.5102* 73.0612**	69.7959* 80.4082**	59.6666* 58.6667**	71.6666* 71.6667**	81.6666* 77.6667**	54.0000* 52.0000**	67.6666* 67.6667**
3	73.0612* 83.6735**	75.1020* 85.7143**	76.7346* 78.3674**	71.3709* 73.0612**	78.3673* 80.0000**	64.3333* 68.3333**	51.6666* 48.6667**	56.6666* 60.6667**	47.6666* 49.6667**	71.0000* 73.0000**
4	71.4285* 77.5510**	66.1224* 66.1225**	73.8775* 73.8776**	77.5510* 88.1633**	75.9183* 77.5510**	60.6666* 61.6667**	68.3333* 66.3333**	58.3333* 55.3333**	67.3333* 68.3333**	61.6666* 58.6667**
5	75.9183* 82.0408**	70.6122* 81.2245**	75.5102* 86.1225**	78.7755* 80.4082**	76.3265* 86.9388**	65.3333* 68.3333**	55.0000* 55.0000**	64.3333* 63.3333**	51.0000* 51.0000**	59.6666* 57.6667**
6	74.6938* 85.3061**	75.5102* 77.1429**	75.9183* 86.5306**	74.6938* 72.2449**	72.2449* 82.8571**	64.0000* 62.0000**	64.0000* 60.0000**	61.0000* 63.0000**	75.0000* 74.0000**	56.6666* 61.6667**
7	76.3265* 86.9388**	74.2857* 84.8980**	75.5102* 86.1225**	77.9591* 88.5714**	76.3265* 86.9388**	68.6666* 71.6667**	67.6666* 70.6667**	46.3333* 46.3333**	71.0000* 73.0000**	67.3333* 68.3333**
8	75.5102* 86.1225**	74.2857* 84.8980**	74.6938* 76.3265**	72.6530* 83.2653**	71.0204* 81.6327**	51.6666* 50.6667**	53.3333* 55.3333**	55.0000* 55.0000**	63.0000* 60.0000**	51.0000* 53.0000**
9	76.3265* 86.9388**	73.8775* 75.5102**	71.4285* 73.0612**	72.6530* 74.2857**	72.2449* 82.8571**	67.3333* 67.3333**	63.3333* 68.3333**	57.6666* 58.6667**	63.3333* 64.3333**	75.0000* 74.0000**
10	78.3673* 88.9796**	74.2857* 75.9184**	76.7346* 87.3469**	73.0612* 74.6939**	74.2857* 84.8980**	59.6666* 59.6667**	66.3333* 64.3333**	58.6666* 58.6667**	55.3333* 58.3333**	61.6666* 58.6667**
11	74.2857* 84.8980**	73.8775* 84.4898**	75.9183* 86.5306**	70.6122* 72.2449**	74.2857* 84.8980**	64.0000* 65.0000**	64.3333* 65.3333**	75.0000* 71.0000**	77.6666* 81.6667**	58.6666* 59.6667**
12	73.4693* 84.0816**	75.9183* 86.5306**	68.9795* 79.5918**	73.0612* 83.6735**	73.4694* 84.0816**	64.3333* 66.3333**	55.0000* 52.0000**	56.3333* 56.3333**	58.6666* 56.6667**	67.6666* 68.6667**

*The accuracies with an asterisk (*) and (**) are accuracies (%) of the LSTM-CNN-FS model and our proposed model, respectively.

Table 5. Comparison of annualized returns of the proposed system (LSTM-CNN-FS-ZZ) with BaH, CNN, and LSTM-CNN Models (Cryptocurrencies– test period: 2022–2023)

Crypto	BaH	CNN	LSTM-CNN	LSTM-CNN-FS	Our proposed model	Crypto	BaH	CNN	LSTM-CNN	LSTM-CNN-FS	Our proposed model
BTC	0.59	0.61	0.72	0.70	1.19	**ETC**	0.40	0.66	0.69	0.69	0.77
ETH	0.83	0.66	0.89	0.91	1.03	**XEM**	-0.23	-0.11	-0.11	0.08	0.14
XRP	0.54	0.57	0.68	0.69	0.66	**ONT**	-0.14	0.08	0.10	0.11	0.13
BCH	1.95	1.21	1.37	1.20	1.54	**QTUM**	0.04	0.21	0.29	0.20	0.29
EOS	-0.16	0.11	0.17	0.17	0.29	**ZEC**	-0.38	-0.10	-0.11	0.02	0.01
LTC	1.21	0.58	0.59	0.68	0.61	**ICX**	-0.21	0.07	0.07	0.07	0.08
XLM	-0.02	0.05	0.01	0.02	0.09	**LSK**	-0.20	-0.38	-0.21	-0.04	0.02
ADA	-0.35	0.31	0.38	0.36	0.69	**ZIL**	-0.47	0.10	0.12	0.12	0.12
MIOTA	-0.30	0.05	0.09	0.09	0.09	**DCR**	-0.29	0.07	0.08	0.14	0.19
USDT	0.00	0.00	0.00	0.00	0.00	**BTG**	0.04	0.22	0.34	0.28	0.30
TRX	0.17	-0.17	-0.01	0.06	0.11	**ZRX**	-0.29	-0.39	-0.27	-0.22	0.02
XMR	0.47	0.33	0.49	0.55	0.51	**SC**	-0.12	0.31	0.37	0.19	0.42
NEO	0.18	0.09	0.18	0.17	0.19	**MKR**	-0.04	-0.11	-0.10	-0.10	0.02
DASH	-0.08	-0.11	-0.10	-0.01	0.01	**DOGE**	0.03	0.41	0.51	0.40	0.53
BNB	0.14	0.22	0.28	0.21	0.26	**WAVES**	-0.64	0.10	0.11	0.20	0.23

Model	BaH	CNN	LSTM-CNN	LSTM-CNN-FS	Our proposed model
Average	**0.09**	**0.19**	**0.25**	**0.26**	**0.35**

Table 6. The results of statistical significance test

VARIABLES/Crypto	BTC	ETH	BNB	ADA	SOL
Mean	0.0101	0.0143	0.0222	0.0145	0.0455
Variance	0.0100	0.0171	0.0221	0.0244	0.0550
Observations	261	261	261	261	154
T_Stat	1.3327	1.6742	1.7611	1.7099	1.6677
P-value	0.1838	0.9024	0.0794	0.0885	0.0934

Table 7. Comparison of the performance of portfolio trading systems based on statistical tests (from 7/1/2022 to 7/1/2023)

Models	CNN+ EW	LSTM-CNN+ EW	LSTM-CNN-FS+ EW	Proposed trading systems +EW	CNN+ HWM	LSTM-CNN+ HWM	LSTM-CNN-FS+ HWM	proposed trading systems + HWM
Panel A: Daily return descriptive statistics								
Mean	0.0029	0.0037	0.0039	0.0047	0.0016	0.0023	0.0022	0.0025
Standard dev.	0.0518	0.0541	0.0433	0.0410	0.0148	0.0119	0.0104	0.0102
Downside dev.	0.0113	0.0149	0.0132	0.0143	0.0039	0.0033	0.0031	0.0029
Panel B: Statistical tests for daily returns								
T-test	3.5610	4.1582	4.2540	4.8221	2.2140	2.5051	2.7976	2.9881
ANOVA-test	12.8871	16.6630	18.4182	19.4321	30.5789	31.1190	35.6887	37.5460
P-value $_{Sharpe}$	1.06e-06	2.15e-06	7.77e-05	3.17e-06	1.99 e-07	9.31e-06	3.13e-07	5.25e-07
P-value $_{Sortino}$	7.06e-04	2.15e-05	5.36e-06	1.99e-06	8.87 e-06	2.43e-07	2.88e-07	3.09e-07