The study focuses on improving the quality of using recurrent neural networks (RNNs) to predict cryptocurrency prices. The formula of the target variable for the model based on the arithmetic mean is developed, which allows us to better take into account the dynamics of cryptocurrency exchanges. The factors affecting this variable were grouped into features based on the volume of daily cryptocurrency trading, the volatility of the relevant prices, and the pre-calculated and selected signals of technical indicators. As part of the study, an algorithm for processing daily data was developed for the model. The results obtained made it possible to create a holistic model for forecasting stock prices. Two recurrent neural networks were trained: one with a long short-term memory (LSTM) and the other with a recurrent gate unit (GRU). To determine the efficiency of the models, the analysis was carried out using two key indicators: the Sortino coefficient, which measures the relative risk/reward for each additional unit of unwanted volatility, and the Sharpe ratio, which measures the return on assets, subtracting the free risk. As a result, it was found that both models have similar results in terms of accuracy (~69 %). Still, the GRU-based model showed significantly better values of the Sortino coefficients (3.13) and Sharpe's coefficient (2.45), which allows us to conclude that it is effective on cryptocurrency exchanges. At the same time, the LSTM model requires more parameters for training than the GRU model with an identical structure, which leads to a longer training time. The obtained scientific and practical results are aimed at more efficient use of recurrent neural networks in price forecasting on cryptocurrency exchanges

Keywords: machine learning, cryptocurrency exchanges, neural networks, deep learning, price prediction

UDC [004.4.6]+336.7
DOI: 10.15587/1729-4061.2023.287094

DEVELOPMENT OF RECURRENT NEURAL NETWORKS FOR PRICE FORECASTING AT CRYPTOCURRENCY EXCHANGES

Victoria Tyshchenko
Doctor of Economic Sciences, Professor*
Svitlana Achkasova
Corresponding author
PhD, Assosiate Professor*
E-mail: svet_achk@meta.ua
Oleksii Naidenko
PhD, Assosiate Professor*
Serhii Kanyhin
Postgraduate Student*
Vlada Karpova
PhD*
*Department of Customs and Financial Services
Simon Kuznets Kharkiv National
University of Economics
Nauky ave., 9-A, Kharkiv, Ukraine, 61166

Received date 03.07.2023
Accepted date 10.10.2023
Published date 30.10.2023

How to Cite: *Tyshchenko, V., Achkasova, S., Naidenko, O., Kanyhin, S., Karpova, V. (2023). Development of recurrent neural networks for price forecasting at cryptocurrency exchanges. Eastern-European Journal of Enterprise Technologies, 5 (4 (125)), 43–54. doi: https://doi.org/10.15587/1729-4061.2023.287094*

1. Introduction

Technological progress, especially in artificial intelligence, drives the exponential growth of information technology. This evolution profoundly transforms social structures and economic ties. A notable manifestation of this trend is the rise of cryptocurrencies. These innovative financial tools have swiftly gained traction. As reported by liga.net in 2023, there are now over 60 million unique cryptocurrency wallets [1], highlighting the significant interest from investors and scholars in this financial market segment.

In this technological revolution, cutting-edge advancements in artificial intelligence, notably recurrent neural networks (RNNs), are pivotal in the cryptocurrency landscape. RNNs are preferred for forecasting tasks due to their efficiency in handling time series data. Their capacity to generalize and discern intricate patterns positions them as a dependable tool.

The inherent volatility of cryptocurrency markets presents significant challenges for investors and traders. Data from CoinMarketCap [2] indicates that many cryptocurrencies experience daily fluctuations by tens of percent, necessitating innovative forecasting methods. Within this scenario, the practical application of RNNs becomes crucial for predicting market directions and formulating investment tactics.

Yet, employing RNNs comes with its set of challenges. A primary hurdle is handling the big data essential for training upcoming models. As per IDC [3], the global data volume is projected to reach 200 zettabytes by 2026. This surge necessitates innovative solutions for data storage and processing.

Given the discussed points, there's a clear need for continued research into using RNNs for cryptocurrency price prediction. Swift advancements in cryptocurrencies mark the current scientific and technological environment. Additionally, the realm of artificial intelligence applications is expanding. These factors highlight the importance of a thorough investigation into the potential of RNNs to enhance operations on cryptocurrency exchanges.

Research into using recurrent neural networks (RNNs) for price forecasting on cryptocurrency exchanges is indeed pertinent. Cryptocurrencies are rapidly evolving, and there's a growing amount of data to process. Within this setting, RNNs can be a robust tool to enhance investment portfolio management in this financial market segment. The rise in cryptocurrency popularity and volatility underscores the demand for innovative market forecasting methods. RNNs can be instrumental in addressing this need. With the surge in data volumes and the broadening applications of artificial intelligence, exploring the capabilities of RNNs on cryptocurrency exchanges becomes a vital avenue for ongoing investigation.

2. Literature review and problem statement

The paper [4] discusses the findings from a study on crafting investment strategies encompassing volatile assets like cryptocurrencies. This research leverages neural networks to suggest a trading strategy that maximizes profit while setting a stop-loss. This is achieved by enhancing the traditional moving average technical indicator. The enhancement involves integrating limited recursion in Elman neural networks with a hybrid neuro-symbolic neural network. This integration ensures that no recursive parts of the network remain trainable. The strategy was modeled using the Eurostoxx50 financial index. Results showcased the strategy's potential to evade negative asset returns, thereby lowering investment risks. Yet, the paper should provide further details on the model's precision. Investment risks in this paper are gauged on a per-trade basis. A future study could evaluate risk tolerance across a series of trades. Moreover, the current neural network only considers a single technical indicator's outcome. A model drawing from multiple indicators could be more insightful, as it might unearth additional nonlinear correlations within price movements.

The paper [5] delves into the attributes of stocks to enhance the accuracy of price forecasts on stock exchanges. The research indicates that public data showcasing prominent funds' investment patterns mirrors their managers' collective views regarding stock characteristics. Like [4], this study doesn't specify the accuracy of the models developed but instead focuses on the combined return of a 50-asset portfolio. Additionally, there's no analysis regarding the model's resilience against risks. For practical application, it's essential to determine if such a model can be reliably used in out-of-sample scenarios, especially under uncertain conditions. Furthermore, when refining the model, it would be beneficial to incorporate insights from the specific characteristics of distinct asset groups.

The paper [6] details research on creating a type 2 chaotic transient-fuzzy deep neuro-oscillatory network with a retrograde signal to predict stock prices. This model underwent training using a dataset comprising 129 assets, of which 9 are cryptocurrencies, spanning 2048 days. The data was represented through ten signals derived from technical indicators. Instead of emphasizing accuracy, the research aimed to reduce the standard deviation while boosting the average cumulative return. While mathematically sound, this method must account for risk tolerance across a series of trades. Drawing from this paper's experiences designing intricate algorithms, it's recommended to consider a model that integrates multiple input features.

The paper [7] introduces research on designing a neural network for predicting stock market prices. The authors suggest a structure that combines a bilinear projection with an attention mechanism. While the model's accuracy touches 80 %, there's no cumulative return or investment risk level assessment. Drawing on this paper, considering input features as time series might be beneficial in ensuring a reliable accuracy level.

The paper [8] discusses the creation of a neural network for predicting prices on stock exchanges. The authors introduce a novel deep-learning approach that melds wavelet transforms with stacked autoencoders. The model showcases satisfactory MAPE, R, and Theil U metrics values. Yet, much like in [6], it doesn't account for risk tolerance across consecutive trading actions. Incorporating insights from recurrent neural networks in handling time series might be beneficial.

The paper [9] discusses the advantages of using modeling technology in business risk management. It underscores the value of employing recurrent neural networks for price predictions, leading to a decrease in risky ventures.

Although the models presented are trained on big datasets, enhancements could benefit their accuracy. The paper [10] highlighted a shortcoming of RNNs in handling ongoing input streams. An adaptive forget gate was introduced, enabling the network to refresh its state at specific intervals. Testing revealed that this new method outperforms traditional RNN techniques. This solidifies the merit of developing models using this innovative approach.

A review of these sources points towards the potential benefits of undertaking a research initiative to refine an RNN model for predicting prices on cryptocurrency exchanges.

3. The aim and objectives of the study

The study aims to develop a model for forecasting prices on cryptocurrency exchanges based on recurrent neural networks.

To achieve this aim, the following objectives were accomplished:

– to categorize model features into groups based on cryptocurrency clusters and technical indicators-based signals. Clusters should consider daily trading volume and price volatility;

– to define a target variable for the model;

– to build and evaluate the model.

4. Materials and methods of research

4. 1. Research object

The object of this research is the development of a model that predicts cryptocurrency exchange prices using recurrent neural networks. This encompasses the identification of model input features and the subsequent training process.

This study's central hypothesis posits that price prediction on cryptocurrency exchanges can be enhanced by developing a recurrent neural network with intricate input features and a training strategy aligned with the model's future application.

This research operates on the premise that there exist discernible patterns within price dynamics, which the model can identify given sufficient training data. However, certain limitations exist. The model does not account for external influences like global economic shifts, political interventions, or technological advancements.

4. 2. Data and tools

This study's models were trained and validated using data from the Binance API [11] from June 2021 to June 2023. This dataset captures daily metrics such as close, open, high, lowest prices, and trading volumes. Our analysis targeted cryptocurrencies exchanged with BUSD, which serves as Binance's counterpart to the US dollar. The dataset comprises 289 distinct cryptocurrencies. A brief overview of the data is provided in Table 1. The full dataset can be accessed at [12].

Table 1

A sample of input data collected using the Binance API for the period from June 2021 to June 23

Crypto-currency symbol	End of the period	Start of the period	Open price	High price	Low price	Close price	Trading volume	Trading volume (BUSD)
BTC	2021-06-01	2021-06-01	37266.5	37918.54	35683.88	36685.87	19265.08	704653105.2
BTC	2021-06-02	2021-06-02	36685.87	38235.22	35909.54	37568.67	14636.82	546204536.3
BTC	2021-06-03	2021-06-03	37568.66	39475.45	37168.01	39250	17391.70	671676510.4
BTC	2021-06-04	2021-06-04	39249.99	39285.96	35578.23	36853.91	19950.22	738544042.6
BTC	2021-06-05	2021-06-05	36848.29	37924.61	34823.78	35521.12	16838.29	612038548.6
BTC	2021-06-06	2021-06-06	35521.12	36475.78	35230	35807.49	10837.52	389495323.9
BTC	2021-06-07	2021-06-07	35810.93	36809.14	33313.77	33573.37	18293.17	650225499.6
BTC	2021-06-08	2021-06-08	33570	34085.63	31050.01	33396.46	31935.94	1045356523
BTC	2021-06-09	2021-06-09	33403.31	37574.48	32418.21	37408.94	32116.48	1126081299

Source: Compiled by the authors based on [11]

Python was chosen as the primary tool for data analysis, model training, and visualization [13]. Essential libraries such as NumPy, Pandas, Scikit-learn, TensorFlow, Matplotlib, Seaborn, and Keras were employed. These libraries offer extensive resources for data processing, feature engineering, model training, validation, and graphical representation. The utilization of these libraries simplifies both the development and evaluation phases. Due to its vast library ecosystem, intuitive syntax, scalability, high performance, and easy integration capabilities, Python was deemed the most suitable for the study's objectives.

4. 3. Neural network type

RNNs (Recurrent Neural Networks) were selected for forecasting stock prices due to their proficient handling of time series data. They process the entire sequence context and identify nonlinear relationships within the series. Given the intricate nonlinear patterns frequently appearing in financial time series, this capability is vital for forecasting stock prices. At its core, RNNs are a subset of neural networks where connections form a directed sequence-aligned graph [10]. This design endows them with memory loops, enabling information retention within the network.

However, early RNN iterations encountered the 'vanishing gradient' issue, leading to protracted training durations for gradient methods when working with RNNs [14, 15]. This challenge arose because the error gradient, vital for gradient methods, vanished as it retrogressively moved through the network. Consequently, the RNN's initial layers ceased learning. In situations with extended sequences, RNNs struggled to relay information from preceding steps forward, exhibiting a short-term memory deficit.

Fig. 1 illustrates a common issue with earlier RNN versions: unit 1 experienced a vanishing error gradient because of damped backflow, impeding the accurate transfer of information to block 4. A new RNN variant, the long short-term memory (LSTM), emerged to address the short-term memory challenge.

Fig. 2, derived from [10], displays an LSTM unit. This unit comprises three gates: input, output, and forget. These gates regulate information flow and determine the unit's status. In sequence, LSTMs link these units, each functioning as a memory module.

As shown in Fig. 2, the forget gate (1) tells the unit what information should be «forgotten» or discarded from the internal state. The input gate (2) indicates what new information to store in the state. The output gate (3), (4) is what the unit outputs to the outside, a filtered version of the internal state:

$$f_t = \sigma\left(W_t * \left[h_{t-1}, X_t\right] + b_t\right), \tag{1}$$

$$i_t = \sigma\left(W_i * \left[h_{t-1}, X_t\right] + b_i\right), \tag{2}$$

$$o_t = \sigma\left(W_o * \left[h_{t-1}, X_t\right] + b_o\right), \tag{3}$$

$$\hat{C}_t = \tanh\left(W_c * \left[h_{t-1}, X_t\right] + b_c\right). \tag{4}$$

Then the internal state of the cell is calculated by the formula (5):

$$C_t = i_t * \hat{C}_t + f_t * C_{t-1}. \tag{5}$$

The final output from the block, or h_t, is then filtered by the formula (6):

$$h_t = o_t * \tanh\left(C_t\right). \tag{6}$$

As with every neural network, weights are associated with each input. These weight matrices are combined with gradient optimization to make the feedforward unit learn. The weight matrices can be seen in the formulas above as W_f, b_f, W_i, b_i, W_o, b_o, W_c, b_c, respectively.

These units are then connected, as shown in Fig. 3. This allows the LSTM network to store information from past steps and make time-series predictions. Using the LSTM cell architecture, the network can eliminate the vanishing gradient problem. This problem prevented older RNN architectures from achieving good time-series predictions.

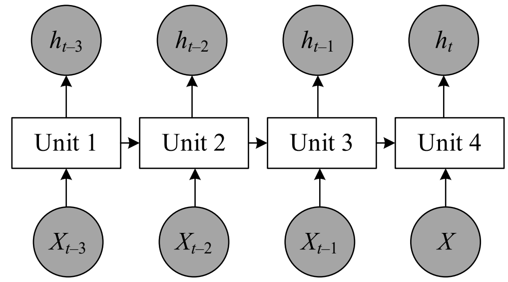

Fig. 1. A typical structure of a recurrent neural network: X_t – input sequence, h_t – output sequence

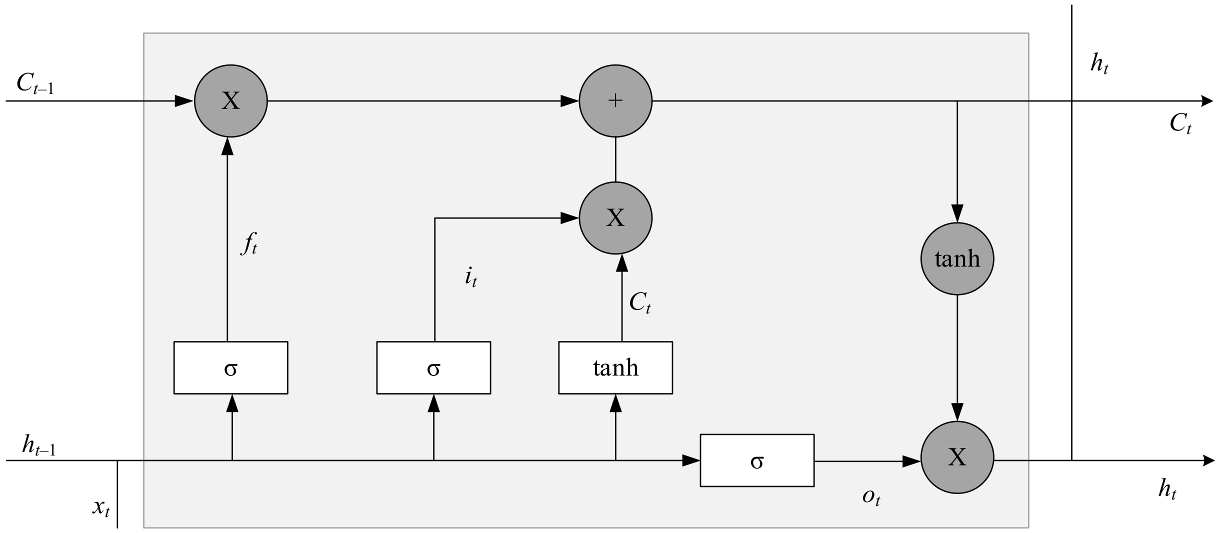

Fig. 2. Unit of long short-term memory: X_t – input data sequence, h_t – output data sequence, C_t – unit status, f_t – forget gate, i_t – input gate, *to* – output gate, $\hat{C}_t$ – internal unit's state

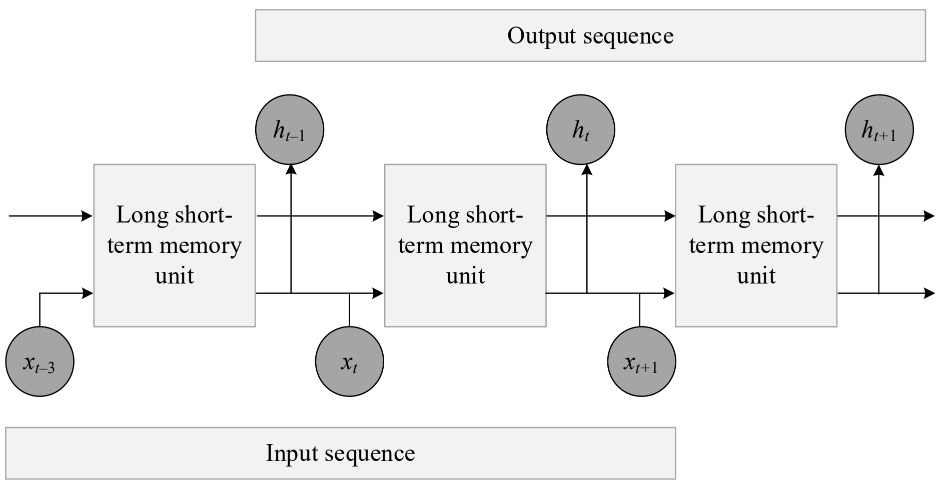

Fig. 3. Long short-term memory unit within a recurrent neural network

As shown in Fig. 3, the architecture is a standard version of the LSTM unit. Researchers are constantly improving and modifying the block architecture to make the LSTM network more efficient and reliable for different tasks. An example is the used architecture of the LSTM unit, where connections are added to each valve, which allows them to look at the internal state of the C_{t-1}.

Another common modification of the RNN is the gated recurrent unit or GRU, shown in Fig. 4.

Fig. 4 shows the main distinction between the GRU and the LSTM. The GRU integrates the input and forget gates into a singular update gate. Furthermore, the internal state of the unit and the hidden state are merged. As a result, a GRU block is more straightforward than a conventional LSTM. An advantage of the RBF lies in its rapid learning capability. Later in this paper, the effectiveness of both the GRU and LSTM methods will be tested and compared. Notably, the GRU and LSTM can identify long-term patterns in time series, a crucial aspect for forecasting market trends. Their adaptability enables them to fit various scenarios and effectively address the nonlinear dependencies frequently seen in financial time series. However, while recognizing their merits, it's vital to understand that employing RNNs for stock price prediction necessitates a comprehensive knowledge of the data and a systematic approach to its pre-processing, which constitutes the core of this research.

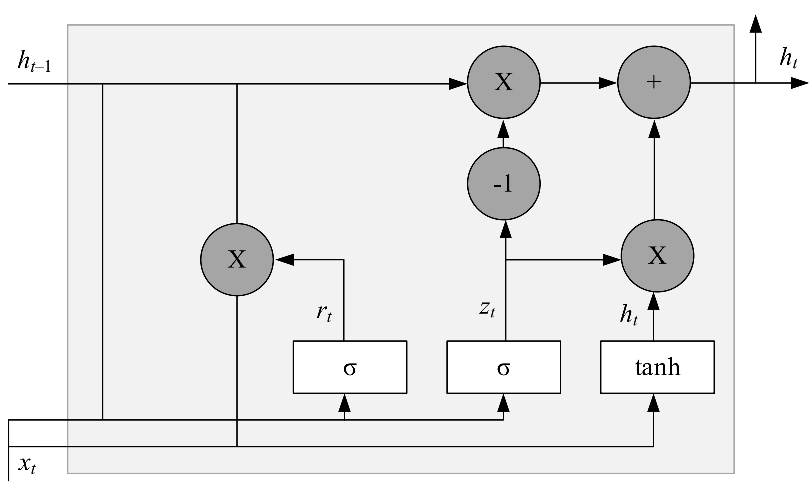

Fig. 4. Gated recurrent unit

4. 4. Evaluation of the obtained models

The accuracy of the models was assessed in this study. Yet, it's important to note that the accuracy measure only evaluates the capability to anticipate price shifts at specific moments without accounting for associated trading risks. For investors, the ideal models are those that forecast accurately and balance risk and return favorably. To this end, the Sharpe [16] and Sortino [17] ratios were employed to examine the efficacy of the derived trading strategies.

The significance of these ratios in model analysis cannot be overstated. The first reason is their intrinsic risk adjustment, which is pivotal for gauging performance in practical scenarios. By factoring in risk, it becomes feasible to juxtapose a model's profit potential against the inherent risks. Secondly, the standardization offered by these ratios provides a straightforward method to gauge model performance. This facilitates more informed determinations of the model's continued deployment or potential refinements based on the findings.

The Sharpe ratio, in essence, contrasts the performance of investments, including cryptocurrencies, to a risk-free asset, taking risk adjustments into account. It's computed as the differential between the returns of the investment and that of the risk-free asset, all divided by the standard deviation of the investment's return. This ratio conveys the additional yield an investor garners for each incremental risk unit. The Sharpe ratio can be expressed using formula:

$$S=\frac{E\left[R-R_f\right]}{\sigma}, \tag{7}$$

where R – the return on the portfolio (asset), R_f – the return on an alternative investment (usually a risk-free interest rate is taken), $E[R-R_f]$ – the risk premium (the mathematical expectation of the excess of asset return over the return on an alternative investment), and, σ – the standard deviation of the portfolio (asset) return.

The Sharpe ratio determines how well an asset's return compensates for the risk an investor is taking. Investing in the asset with the higher Sharpe ratio will be less risky when comparing two assets with the same expected return.

The Sortino coefficient assesses the profitability and risk of an investment tool, portfolio, or strategy. Though it bears similarities to the Sharpe ratio in its calculation, it diverges by utilizing the «downside volatility» instead of the general portfolio volatility. Here, the volatility is determined based on returns that fall beneath the minimum acceptable portfolio return (MAR) [18]. The methodology for its calculation is articulated in formulas (8) and (9):

$$S=\frac{R-T}{DR}, \tag{8}$$

where R – the average portfolio return, T – the minimum acceptable level of portfolio profitability, and DR – a downward deviation or «downward volatility»:

$$DR=\sqrt{\int_{-\infty}^{t}\left(T-r\right)^2 f(r)dr}, \tag{9}$$

where T – the annual target return, originally called the minimum acceptable return MAR, r – a random variable representing the return for the distribution of annual returns, and $f(r)$ – a distribution for annual returns, for example, a log-normal distribution.

According to the sources [16, 17], Sharpe and Sortino values greater than one indicate acceptable and above 2.5 – excellent model quality.

5. Results of developing a recurrent neural network for price forecasting at cryptocurrency exchanges

5. 1. Formation of models features groups

Three feature groups were derived from the available data: classification by trading volume, classification by price volatility, and technical indicator signals. Each of these groups is elaborated upon below.

Trading volume is indicative of the liquidity on an exchange. High liquidity can translate to more fluid price movements. Moreover, a surge in trading volume may represent heightened interest and trading activity, signaling investor confidence or doubt. When there's an uptick in trading volume accompanying a price rise or fall, it may suggest a more robust trend. Such correlations are considered by the model, enhancing its forecasting precision.

Daily trading volume values were restricted to 0 to 0.07 billion BUSD to address extreme deviations. Any values beyond this bracket were treated as outliers and subsequently replaced. The data was then clustered into 4 distinct groups using the K-means method [18]. Fig. 5 offers a visual representation of the distribution of cryptocurrency trading volumes, depicted through a histogram.

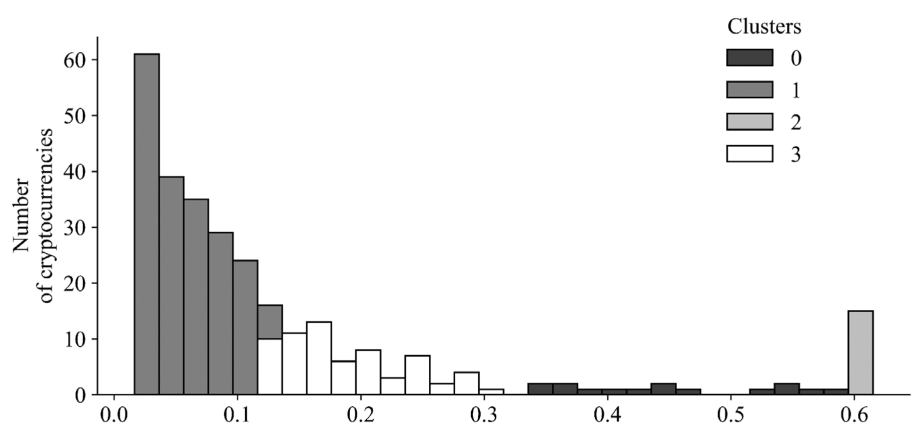

Fig. 5. Histogram of the distribution of cryptocurrency clusters calculated by the K-means method based on daily trading volume as of June 01, 2023, for 289 cryptocurrencies. Source data and calculations: [12]

Fig. 5 reveals that most cryptocurrencies experience daily trades of up to 200 million BUSD. However, a distinct group comprises significant cryptocurrencies like bitcoin, ether, and several stablecoins, registering daily trading volumes surpassing 600 million BUSD. Incorporating the one-hot encoded clusters as features in the RNN model is expected to elucidate specific nonlinear relationships for each group, potentially enhancing forecast precision.

Volatility is pivotal in stock exchanges, indicating the extent of price variations and associated uncertainty. The indicator for volatility is derived from the standard deviation of the daily closing price's percentage alteration, as depicted in formula (10):

$$S_x = \sqrt{\frac{\sum_i^n \left(x_i - \bar{x}\right)^2}{n-1}}, \tag{10}$$

where n – number of days, $\bar{x}$ – average close price, x_i – close price on a particular day.

Abnormal deviations led to constraining the values between 0.04 and 0.13. Any values beyond this range were replaced with these extreme limits. The data were categorized into four clusters utilizing the K-means clustering method [18]. The findings were illustrated in a histogram (Fig. 6).

As illustrated in Fig. 6, most cryptocurrencies exhibit volatility values ranging from 0.05 to 0.1. This range signifies a moderate degree of price variability. The spread of these volatility metrics suggests their potential as neural network input data once transformed into a uniform code. By introducing volatility as a feature, deep learning models can be better equipped to understand the intricacies of price movement dynamics. These dynamics influence stock market sentiment, diverse trading behaviors, and external determinants. Incorporating volatility enhances the model's predictive accuracy regarding price direction.

Technical indicators are analytical tools in stock exchanges, forecasting price shifts based on past data [19]. Incorporating buy and sell cues derived from these indicators as features enhances the capacity of deep learning models to identify price movement trends. A compilation of 26 prevalent stock market indicators was gathered for this study, drawing from references [19, 20], as detailed in Table 2.

The list of technical indicators presented in Table 2 is effective for identifying trends in the stock exchange. Calculations were made for each indicator, the results of which are shown in Fig. 7.

As depicted in Fig. 7, both the Sortino and Sharpe coefficients exhibit relatively low average and minimum values. Such an outcome aligns with expectations, given that signals from technical indicators are typically more effective when amalgamated with other signals [13]. In the realm of neural networks, a synergistic effect emerges during the model training phase.

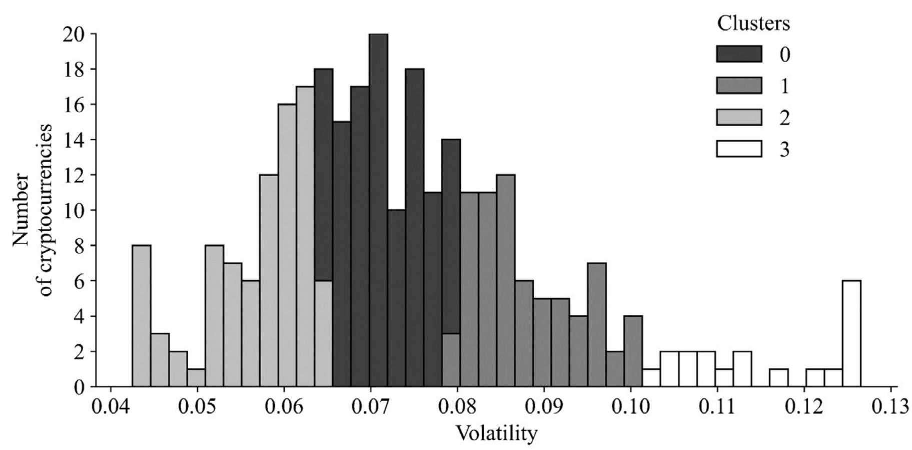

Fig. 6. Histogram of the distribution of cryptocurrency clusters calculated by the K-means method based on the price volatility for the period 06.2021 – 06.2023 for 289 cryptocurrencies. Source data and calculations: [12]

Table 2

List of common technical indicators

Short title	Full title	Short title	Full title
CCI	Commodity Channel Index	RSI	Relative Strength Index
WPR	Williams Percentage Range	STOCH	Stochastic Oscillator
BB	Bollinger Bands	TSI	True Strength Index
FI	Force Index	UO	Ultimate Oscillator
ATR	Average True Range	ADI	Accumulation and Distribution Index
VI	Vortex Index	EOM	Ease Of Movement
TEMA	Triple Exponential Moving Average	OBV	On-Balance Volume
MI	Mass Index	VPT	Volume-Price Trend
MACD	Moving Average Convergence/Divergence	DC	Donchian Channel
KST	Know Sure Thing Oscillator	KC	Keltner Channel
ADX	Average Direction Index	CC	Coppock Curve
DPO	Detrend Price Oscillator	UI	Ulcer Index
MFI	Money Flow Index	ICH	Ichimoku Kinko Hyo

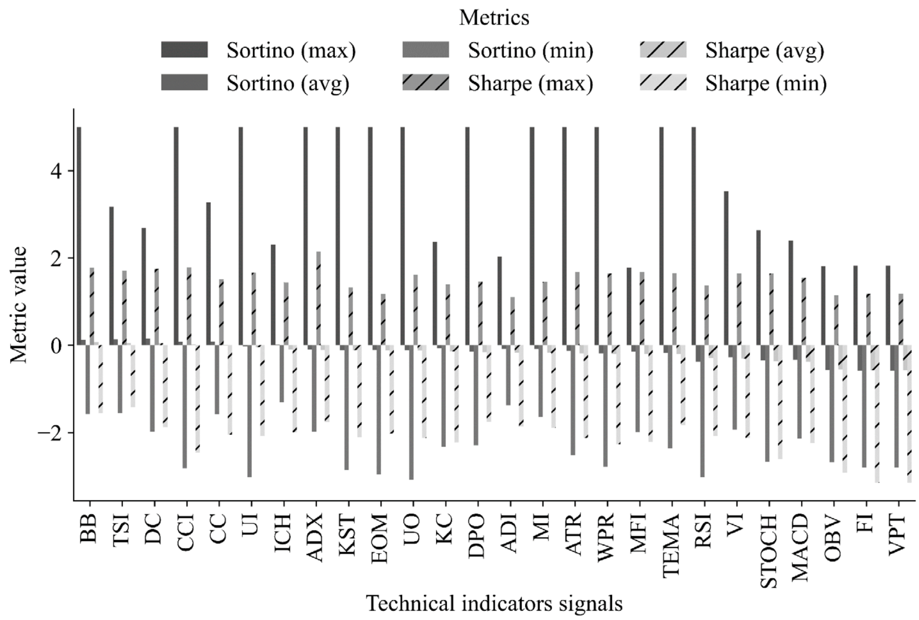

Fig. 7. Calculated metric values for 26 signals based on technical indicators from 06.2021 to 06.2023 for 289 cryptocurrencies. Dataset and code: [12]

Incorporating these indicators into a neural network as binary signals to forecast price direction is crucial. It fosters a notable enhancement in analysis efficiency and simplicity. With these indicators, the neural network can more readily discern behavioral patterns.

5. 2. Developing a target variable for the model

In training neural networks for stock price predictions using time-series data, selecting the target variable (Y) is as critical as selecting features. This target variable must encapsulate the information essential for realizing the specified investment objective. Factoring in the investment objective and the forecast horizon is crucial when pinpointing the target variable. Distinct target variables can yield varying outcomes. Some conventional target variables for forecasting stock price behavior are provided in formulas (11)–(15) as referenced in [14, 15]. Here, P is the price, t – is the period, and n – is the number of days.

The price change is calculated as the difference between the price of an asset at a future and current time. This simple approach allows you to learn about the expected price movement over a certain period:

$$Y_t = P_{t+n} - P_t. \tag{11}$$

The return is calculated as a relative price change. It allows you to find out how much the price is expected to rise or fall in percentage terms:

$$Y_t = \frac{P_{t+n} - P_t}{P_t}. \tag{12}$$

The logarithmic return is the natural logarithm of the ratio between future and current prices, which has convenient mathematical properties such as additivity over time:

$$Y_t = \log\left(\frac{P_{t+n} - P_t}{P_t}\right). \tag{13}$$

Price direction is a binary classification of price direction indicating whether the price of an asset will rise or fall:

$$Y_t = \begin{cases} 1, \text{ if } P_{t+n} > P_t, \\ 0, \text{ if } P_{t+n} \le P_t. \end{cases} \tag{14}$$

Excess return is the difference between the asset's return and a benchmark return, such as a stock index or a risk-free rate. This allows you to predict the value of an asset relative to other assets on the exchange:

$$Y_t = \left(\frac{P_{t+n} - P_t}{P_t}\right) - risk\ free\ rate. \tag{15}$$

The formulas (16), (17) are based on the arithmetic mean:

$$W_t = \frac{\sum_t^n \left(\frac{P_{t+n} - P_{t+n-1}}{P_{t+n-1}}\right)}{n}. \tag{16}$$

$$Y_t = \begin{cases} 1, \text{ if } W_t > 0, \\ -1, \text{ if } W_t < 0, \\ 0, \text{ if } W_t = 0. \end{cases} \tag{17}$$

Thus, the target variable (17) is formed, which is calculated based on the arithmetic mean of the future change in the closing price (Table 1) over n periods in percentage terms. The target variable is a signal to trade and takes the value 1 – buy, 0 – hold, or –1 – sell. In other words, if the price increases over the subsequent n periods, the target variable will signal a buy and vice versa. Such a target variable can offset the negative impact of noise price fluctuations during training.

5. 3. Building and testing the model

Based on the previous calculations, three groups of features were formed: involvement in clusters by trading volume, involvement in clusters by price volatility, and technical indicators-based signals. The target variable was also created based on the arithmetic mean (17). Fig. 8 shows the developed algorithm for generating model features in the time series format for the model.

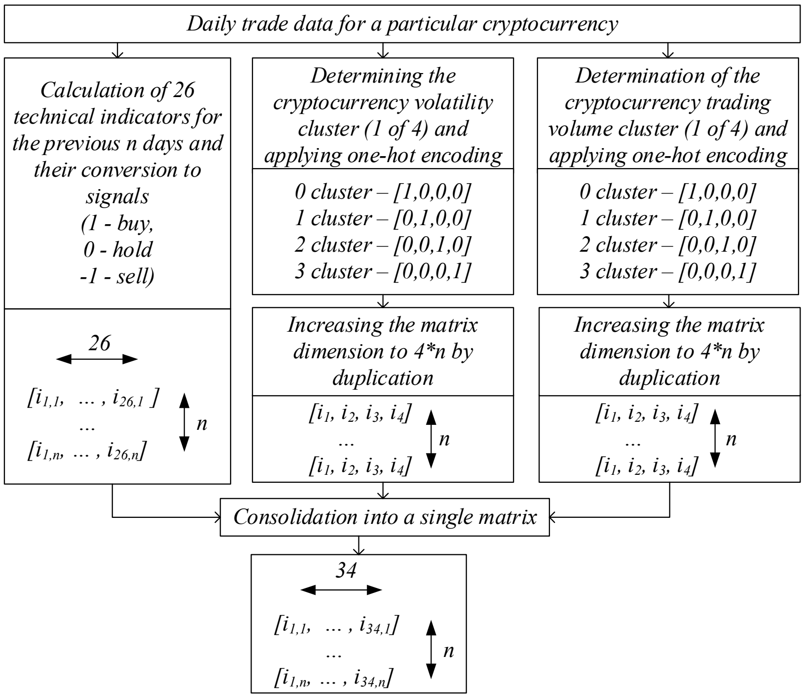

Fig. 8. Algorithm for generating model features in the format of time series for a single cryptocurrency: n — the dynamic parameter of the time series duration (in days)

As shown in Fig. 8, the data is transformed into a matrix format before being transferred to the neural network. It is worth noting that the algorithm is not applied to the first n days of the sample due to the lack of relevant input data. A diagram of the model training and validation process is shown in Fig. 9, with an extended explanation of the algorithm for calculating the Sharpe and Sortino coefficients in Fig. 10.

The maximum values of the time series size and the threshold for predicting the target variable in Fig. 9 are due to the lack of significant computing capabilities and time to test more combinations. Fig. 11 shows the results of evaluating the obtained 54 RNN models (27 GRU and 27 LSTM).

As shown in Fig. 11, the model trains better when the target value window is increased. However, this reduces the values of the Sharpe and Sortino ratios. Expanding the time series window increases the accuracy, but there is no apparent effect on the values of the ratios. At the same time, the best results of the coefficients were obtained at 30 days for the LSTM and 40 days for the GRU. Thus, it was decided to use the RNN models presented in Fig. 12.

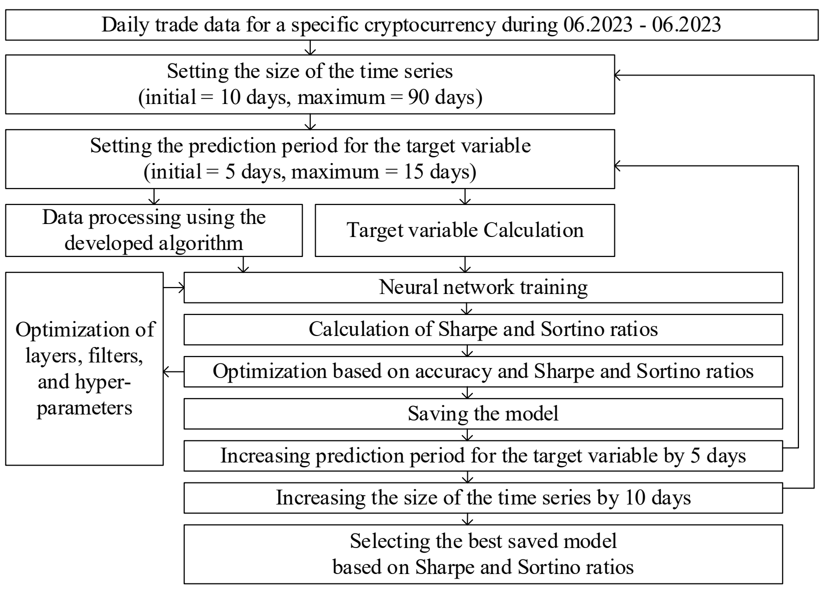

Fig. 9. Model training algorithm for price forecasting on a cryptocurrency exchange

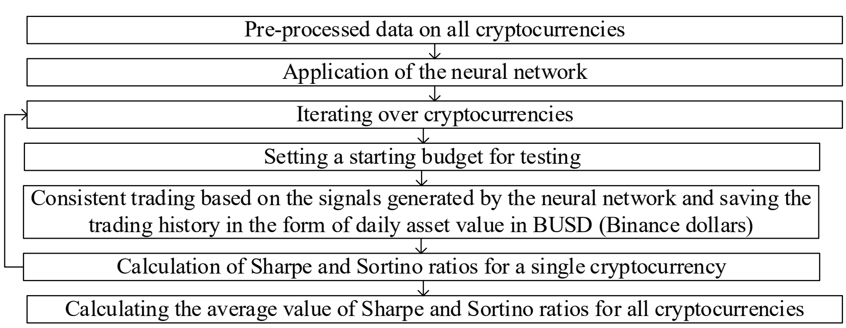

Fig. 10. Algorithm for calculating the Sharpe and Sortino ratios as part of the model training algorithm for forecasting prices on a cryptocurrency exchange

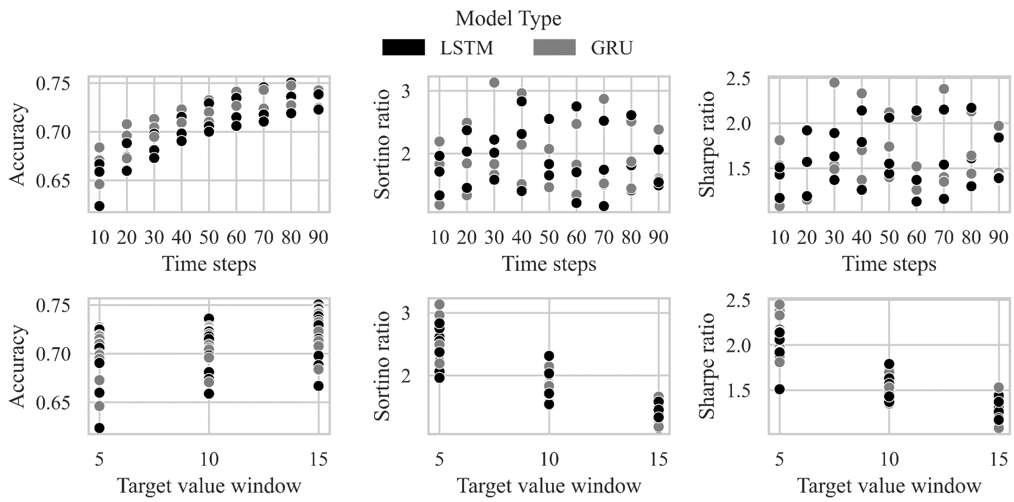

Fig. 11. Results of the evaluation of recurrent neural networks trained based on data for June 2021 to June 2023. Source data, code and models: [12]

Architecture of the developed recurrent neural network with long short-term memory

Long short-term memory recurrent layer (input)
parameters: 188,000
shape: 40:200

Dropout layer
parameters: 0
shape: 40:200

Long short-term memory recurrent layer
parameters: 120,400
shape: 40:100

Dropout layer
parameters: 0
shape: 40:100

Long short-term memory recurrent layer
parameters: 30,200
форма: 50

Dropout layer
parameters : 0
shape: 50

Dense layer (output)
parameters: 153
shape: 3

Architecture of the developed recurrent neural network with a gated reccurent unit

Layer with gated reccurent unit (input)
parameters: 141,600
shape: 30:200

Dropout layer
parameters: 0
shape: 30:200

Layer with gated reccurent unit
parameters: 90,600
shape: 30:100

Dropout layer
parameters: 0
shape: 30:100

Layer with gated reccurent unit
parameters: 22,800
shape: 50

Dropout layer
parameters: 0
shape: 50

Dense layer (output)
parameters: 153
shape: 3

Fig. 12. Architecture of the developed recurrent neural networks. Source data, code, and models: [12]

As can be seen in Fig. 12, there are 338 ths. parameters in the LTSM model and 255 ths. parameters in the GRU model. The training history is shown in Fig. 13, 14.

As can be seen from Fig. 13, 14, the accuracy for the GRU is 69.47 %, and for the LTSM 69.04 %. The Sortino and Sharpe ratios are 3.13 and 2.45 for the GRU and 2.83 and 2.14 for the LTSM. Accuracy could be improved by covering more previous days, but this would reduce the values of the Sharpe and Sortino coefficients. It is worth noting that it is difficult to achieve high accuracy due to the high degree of uncertainty in cryptocurrency markets.

The study used data for 289 cryptocurrencies, but it needs to be clearer to visualize the history of model testing on each due to the results' similarity. Therefore, it is necessary to form a representative list of cryptocurrencies.

A matrix of intersections derived from the previously obtained cryptocurrency clusters by volatility and daily trading volume is presented in Fig. 15.

As shown in Fig. 15, 249 cryptocurrencies (86 % of the total) are located within 6 cluster intersections. A representative list is formed from cryptocurrencies with average trading volumes within their cluster intersection. Fig. 16, 17 demonstrate the test results and the Sharpe and Sortino ratios values for the created models for each cryptocurrency.

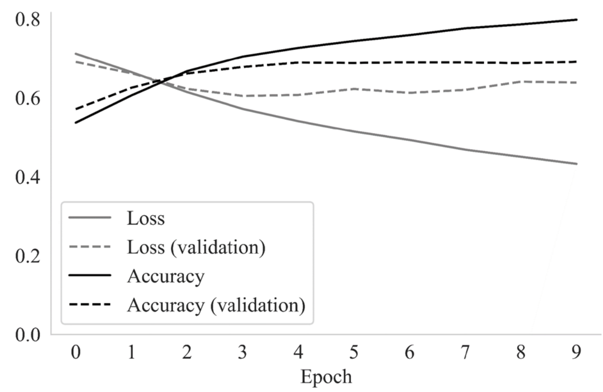

Fig. 13. Training history of a recurrent neural network with a long short-term memory based on data for June 2021 to June 2023. Source data, code, and models: [12]

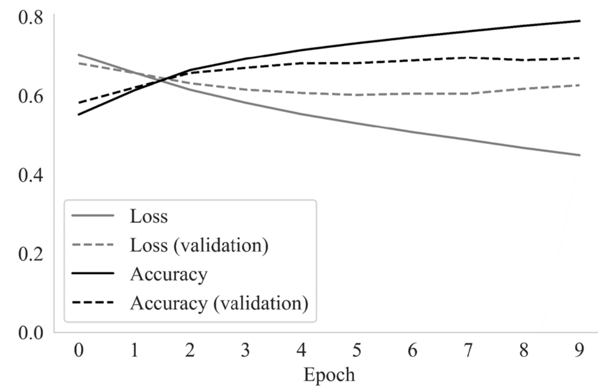

Fig. 14. Training history of a recurrent neural network with a gated recurrent unit based on data for June 2021 to June 2023. Source data, code, and models: [12]

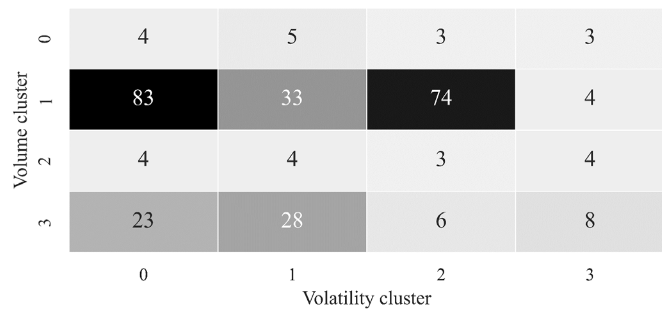

Fig. 15. Matrix of cryptocurrency distribution between clusters of daily trading volumes and volatility. Source data and code: [12]

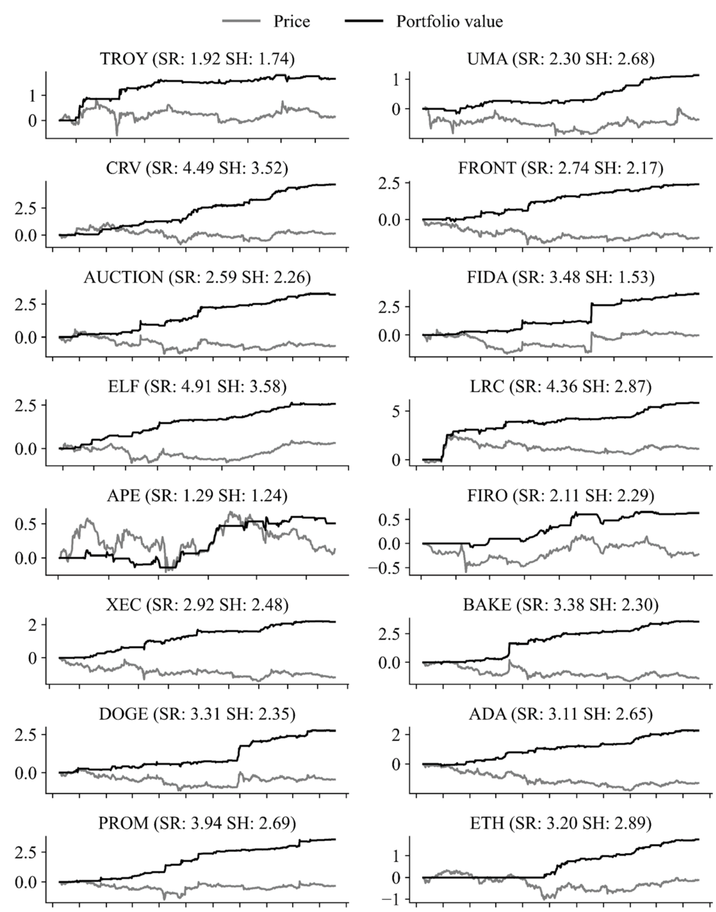

Fig. 16. Results of testing the model based on the recurrent neural network with a gated recurrent unit on data for June 2021 to June 2023. Source data, code, and models: [12]

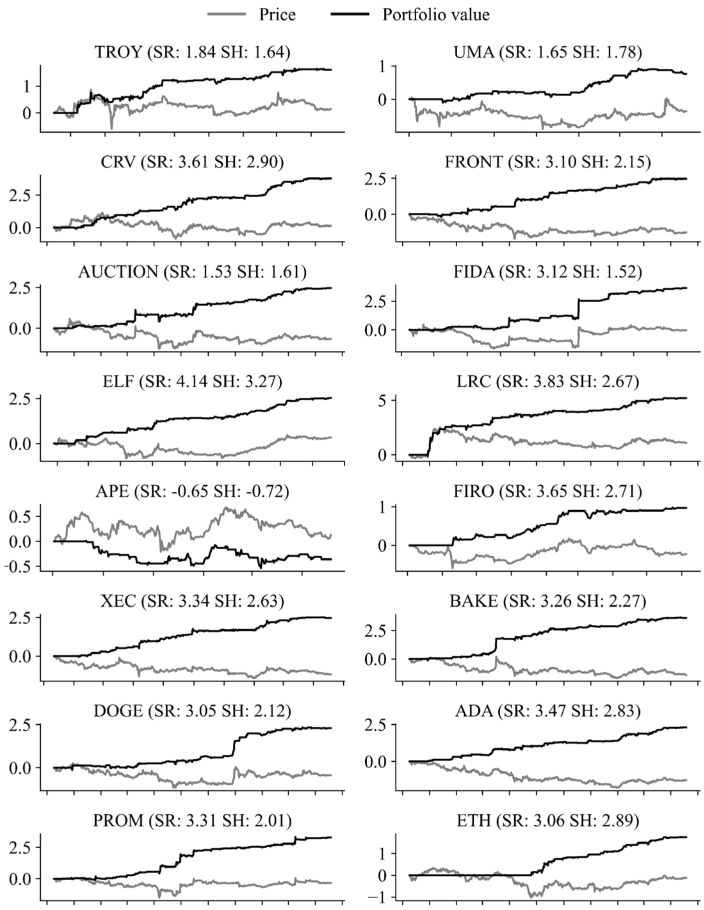

Fig. 17. Results of testing the model based on the recurrent neural network with a long short-term memory on data for June 2021 to June 2023. Source data, code, and models: [12]

As shown in Fig. 16, 17, the increase in asset value occurred for most of the tests even though the price of the respective assets fell during the second half of the test period. This indicates the potential of the models for practical work at cryptocurrency exchanges.

6. Discussion of the results of the development of a recurrent neural network for price forecasting on cryptocurrency exchanges

Significant advancements have been observed in leveraging neural networks for stock exchange forecasting, as highlighted in [4]. The unique characteristics of the proposed models have been underscored through data collection from the international exchange Binance, covering 289 cryptocurrencies from June 2021 to June 2023, as detailed in Table 1. This dataset, transcending specific national boundaries, enhances the universality of the devised models. Mirroring findings from [5], these models navigate the volatility of financial instruments, maintaining a balance of stability and profitability, evident in Fig. 18, 19. Contrary to [6, 7], the models bypass the need for intricately structured neural networks, still achieving an acceptable 69 % accuracy, as depicted in Fig. 15, 16. Despite declining asset prices in the study's latter half, asset values generally exhibited an upward trajectory. Although both models presented analogous accuracies, the RNN-based GRU model outshined in terms of Sortino and Sharpe coefficients, suggesting superior forecasting stability. Such observations underscore the immense potential of GRU-based RNNs in crafting cryptocurrency exchange price-prediction models. A distinct scientific novelty is discerned in the algorithms crafted for feature formation (Fig. 8) and subsequent model training (Fig. 9, 10).

However, this research has limitations. Model results showcase sensitivity to particular parameters, potentially limiting the applicability and universality across diverse exchanges. The volatile nature of cryptocurrency exchanges might hamper result reproducibility. Additionally, textual data from various informational sources needed to be factored in, given the constraints in query volume and processing intricacies.

Employing RNN-based models in predicting cryptocurrency exchange prices is not without its setbacks. Despite the refinements in LSTM and GRU models, the quest for optimal gradients still poses complexities during the training phase. Such models are computationally demanding, inflating time and resource expenditures. Additionally, RNNs might succumb to overfitting in the turbulent realms of cryptocurrency markets. Their predictions might overlook crucial external determinants like political shifts or technological innovations. And while RNNs exhibit prowess with time-series data, their efficacy might wane for non-stationary datasets like cryptocurrency valuations. The absence of theoretical assurances regarding RNN forecast accuracy accentuates the prudence required in their application.

Prospective studies could incorporate external determinants, computational optimizations, algorithmic enhancements to curtail overfitting, and strategies to mitigate the vanishing gradients.

Anticipated mathematical hurdles might encompass the integration of external factors into predictive models. Methodological obstacles could revolve around pinpointing and verifying optimal model parameters over time. Experimentation challenges may stem from cryptocurrency market volatility, expansive data requisites for training, and adaptability to ever-evolving market conditions. Practical impediments orbit computational restrictions, focusing on streamlining computational efficiency and runtimes.

7. Conclusions

1. Three groups of features for the price forecasting model on cryptocurrency exchanges have been formed. The first group, which presents cryptocurrency clusters by daily trading volume, allows identifying the main trading patterns and trends for predicting future price fluctuations. The second group shows cryptocurrency clusters by price volatility, essential for forecasting and minimizing risks, especially in the highly volatile cryptocurrency market. The third group is based on signals from technical indicators, helping to identify key trends and determine the best time to enter and exit the market. These three features provide a holistic and comprehensive analysis of the cryptocurrency market, allowing the developed model to predict price movement direction on cryptocurrency exchanges effectively.

2. A target variable for the arithmetic mean model has been developed. Choosing the correct target variable is critical because it determines what specific task the model will solve. The dynamic component in the number of days allowed us to optimize the variable window to 5 days after training.

3. Based on the formed groups of features, predictive models based on recurrent neural networks were built and tested. These were two modifications: with long short-term memory (LSTM) and with a gated-recurrent unit (GRU). The accuracy for the GRU model was 69.47 %, and for the LSTM, 69.04 %. Sortino and Sharpe ratios were 3.13 and 2.45 for the GRU and 2.83 and 2.14 for the LSTM. The obtained ratios indicate the high efficiency of both models.

Conflict of interest

The authors declare that they have no conflicts of interest concerning this study, including financial, personal, authorship, or other, that could affect the research and its results presented in this paper.

Funding

The study was conducted without financial support.

Data accessibility

The dataset, code, and models are available here [12].

References

1. About Crypto wallets. Available at: https://www.liga.net/crypto/ua/wallets
2. Cryptocurrencies statistics. Available at: https://coinmarketcap.com
3. High Data Growth and Modern Applications Drive New Storage Requirements in Digitally Transformed Enterprises. Available at: https://www.delltechnologies.com/asset/en-us/products/storage/industry-market/h19267-wp-idc-storage-reqs-digital-enterprise.pdf
4. Chen, C., Zhao, L., Bian, J., Xing, C., Liu, T.-Y. (2019). Investment Behaviors Can Tell What Inside. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. doi: https://doi.org/10.1145/3292500.3330663
5. Marques, N. C., Gomes, C. (2010). Maximus-AI: Using Elman Neural Networks for Implementing a SLMR Trading Strategy. Lecture Notes in Computer Science, 579–584. doi: https://doi.org/10.1007/978-3-642-15280-1_55
6. Lee, R. S. T. (2020). Chaotic Type-2 Transient-Fuzzy Deep Neuro-Oscillatory Network (CT2TFDNN) for Worldwide Financial Prediction. IEEE Transactions on Fuzzy Systems, 28 (4), 731–745. doi: https://doi.org/10.1109/tfuzz.2019.2914642
7. Tran, D. T., Iosifidis, A., Kanniainen, J., Gabbouj, M. (2019). Temporal Attention-Augmented Bilinear Network for Financial Time-Series Data Analysis. IEEE Transactions on Neural Networks and Learning Systems, 30 (5), 1407–1418. doi: https://doi.org/10.1109/tnnls.2018.2869225
8. Bao, W., Yue, J., Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLOS ONE, 12 (7), e0180944. doi: https://doi.org/10.1371/journal.pone.0180944
9. Achkasova, S. (2020). Implementation the fuzzy modeling technology by means of fuzzyTECH into the process of management the riskiness of business entities activity. Eastern-European Journal of Enterprise Technologies, 5 (3 (107)), 39–54. doi: https://doi.org/10.15587/1729-4061.2020.209836
10. Gers, F. A., Schmidhuber, J., Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. Neural Computation, 12 (10), 2451–2471. doi: https://doi.org/10.1162/089976600300015015
11. Binance API. Available at: https://www.binance.com/en/binance-api
12. Development of recurrent neural networks for price forecasting at cryptocurrency exchanges. doi: https://doi.org/10.5281/zenodo.8193302
13. Resta, M., Pagnottoni, P., De Giuli, M. E. (2020). Technical Analysis on the Bitcoin Market: Trading Opportunities or Investors' Pitfall? Risks, 8 (2), 44. doi: https://doi.org/10.3390/risks8020044
14. Deng, Y., Bao, F., Kong, Y., Ren, Z., Dai, Q. (2017). Deep Direct Reinforcement Learning for Financial Signal Representation and Trading. IEEE Transactions on Neural Networks and Learning Systems, 28 (3), 653–664. doi: https://doi.org/10.1109/tnnls.2016.2522401
15. Wang, J., Wang, J., Fang, W., Niu, H. (2016). Financial Time Series Prediction Using Elman Recurrent Random Neural Networks. Computational Intelligence and Neuroscience, 2016, 1–14. doi: https://doi.org/10.1155/2016/4742515
16. Sharpe, W. F. (1966). Mutual Fund Performance. The Journal of Business, 39 (S1), 119. doi: https://doi.org/10.1086/294846
17. Sortino, F. A., Price, L. N. (1994). Performance Measurement in a Downside Risk Framework. The Journal of Investing, 3 (3), 59–64. doi: https://doi.org/10.3905/joi.3.3.59
18. Lloyd, S. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory, 28 (2), 129–137. doi: https://doi.org/10.1109/tit.1982.1056489
19. Li, Z., Yang, D., Zhao, L., Bian, J., Qin, T., Liu, T.-Y. (2019). Individualized Indicator for All. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. doi: https://doi.org/10.1145/3292500.3330833
20. Neely, C. J., Rapach, D. E., Tu, J., Zhou, G. (2014). Forecasting the Equity Risk Premium: The Role of Technical Indicators. Management Science, 60 (7), 1772–1791. doi: https://doi.org/10.1287/mnsc.2013.1838