Received: 9 September 2024 | Revised: 26 February 2025 | Accepted: 11 March 2025 | Published online: 27 March 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2025, Vol. 3(3) 305–315
DOI: 10.47852/bonviewAIA52024202

BON VIEW PUBLISHING

A Cryptocurrency Price Forecasting Model by Integrating Empirical Mode Decomposition and LSTM Neural Networks

Xiaowei Wang[1,*], **Ioana Cretu**[1] and **Hongying Meng**[1]

[1]*Department of Electronic and Electrical Engineering, Brunel University of London, UK*

Abstract: Cryptocurrencies, such as Bitcoin and Ethereum, are digital assets that use cryptographic techniques to enable secure and decentralized transactions over the internet. Cryptocurrency prices exhibit highly nonlinear and non-stationary behavior, influenced by a wide range of financial and nonfinancial factors, including market liquidity, regulatory developments, technological advancements, security incidents, and geopolitical events. The unpredictable nature of these price fluctuations underscores the need for robust predictive models to aid investors in making informed financial decisions. In this paper, we propose EMD-LSTM, a novel hybrid model that integrates empirical mode decomposition (EMD) and long short-term memory (LSTM) networks to enhance the accuracy of cryptocurrency price forecasting. EMD is utilized to decompose raw price signals into intrinsic mode functions (IMFs), which help in handling non-stationarity and extracting meaningful patterns. LSTM, with its capability to capture long-term dependencies, is then applied to the decomposed signals to learn relevant temporal features from high-frequency historical data. Our experimental results demonstrate that the EMD-LSTM model significantly outperforms traditional forecasting methods, achieving superior RMSE and MAE scores. These findings highlight the potential of EMD-LSTM as an effective tool for traders, investors, and researchers seeking reliable cryptocurrency price predictions in volatile market conditions.

Keywords: cryptocurrency price prediction, empirical mode decomposition, long short memory model, non-stationary time series, hybrid deep learning model

1. Introduction

Cryptocurrency is a virtual or digital currency [1]. Cryptocurrencies, such as Bitcoin and Ethereum, are decentralized digital assets that rely on cryptographic techniques to facilitate secure transactions. Unlike traditional fiat currencies, cryptocurrencies operate without centralized control, making their prices highly volatile and difficult to predict. The increasing adoption of cryptocurrencies in global finance has heightened the need for accurate forecasting models to support informed investment and trading decisions.

However, cryptocurrency price forecasting presents unique challenges due to high volatility, regulatory uncertainties, and the absence of intrinsic value. Traditional financial models struggle with these characteristics, underscoring the need for robust predictive models that can adapt to the dynamic and nonlinear nature of cryptocurrency markets. While machine learning and deep learning techniques have been increasingly applied to financial time series forecasting, cryptocurrency markets demand approaches that can effectively handle non-stationarity, noise, and multiscale price movements.

To address these challenges, we propose empirical mode decomposition-long short-term memory (EMD-LSTM), a hybrid model that combines EMD [2] with LSTM [3] networks. EMD decomposes the original price series into multiple intrinsic mode functions (IMFs), each representing different frequency components of the signal. This process helps mitigate non-stationarity, isolate meaningful trends, and reduce noise, making the data more suitable for deep learning models. LSTM, in turn, processes these structured IMF components, capturing long-term dependencies while avoiding overfitting to short-term fluctuations. This two-stage approach enhances predictive accuracy by extracting both high-frequency and long-term patterns from cryptocurrency price movements.

The proposed model was evaluated using historical price data from major cryptocurrencies, considering various forecasting horizons, from short-term to long-term predictions. Additionally, we compared the performance of EMD-LSTM against several benchmark models, including traditional time series methods, machine learning models, and hybrid deep learning models, to demonstrate its superiority.

The contributions of this paper are as follows:

1) A novel EMD-LSTM forecasting framework, where EMD decomposes price signals into structured components, enabling LSTM to effectively learn temporal dependencies and improve predictive performance.
2) A thorough empirical evaluation of the proposed model against benchmark models, demonstrating its effectiveness in cryptocurrency price prediction.
3) Insights and recommendations for investors and stakeholders based on the forecasting accuracy and practical implications of the model.

*Corresponding author: Xiaowei Wang, Department of Electronic and Electrical Engineering, Brunel University of London, UK. Email: xiaowei.wang@brunel.ac.uk

The remainder of this paper is structured as follows: Section 2 provides a comprehensive review of the related literature on cryptocurrency price forecasting. Section 3 details the methodology, including the integration of EMD and LSTM. Section 4 presents the experimental results and analysis of the proposed model. Finally, Section 5 offers concluding remarks and outlines potential future work.

2. Literature Review

Cryptocurrency price forecasting has attracted significant attention from both researchers and practitioners, driven by the increasing interest and investment in digital assets. Numerous studies have investigated various approaches for predicting cryptocurrency prices, spanning from traditional statistical models to advanced machine learning algorithms.

Numerous studies have indicated that incorporating multiple models in financial time series forecasting tends to yield better results than relying on a single model [4]. These models, known as ensemble models, consist of parallel base models that collectively generate an optimal predictive model. Based on the research conducted, ensemble models can be broadly categorized into two types: traditional ensembles and decomposition ensembles.

Ensemble models have proven more effective than individual models for forecasting financial time series. Traditional ensembles consist of multiple parallel base models that are combined to create an optimal predictive model. These base models work independently, and their predictions are aggregated either through meta-learning techniques or by using simple averages. Researchers have focused on enhancing the diversity of the base models to mitigate their respective weaknesses, leading to a more accurate collective prediction [5–10].

Recently, there has been growing interest in decomposition ensembles [11]. EMD, introduced by Huang, decomposes a time series into IMFs that capture different frequency components. These ensembles decompose the original time series using signal decomposition methods such as EMD, ensemble empirical mode decomposition (EEMD) [12], multivariate ensemble empirical mode decomposition (MEMD) [13], and complete ensemble empirical mode decomposition (CEEMD) [14]. Each decomposed spectrum is then fed into an independent deep learning model for training. The predictions obtained from each spectrum are combined to generate the final forecast for the response variable.

State-of-the-art decomposition ensembles include CEEMD-CNN-LSTM [15], MEMD-LSTM [16], and EEMD-Cluster-SVR-PSO-LSTM [17]. These models have shown superior performance in financial time series forecasting by leveraging the advantages of signal decomposition and deep learning techniques.

Compared to traditional statistical models like ARIMA [18], many researchers are increasingly interested in applying LSTM networks and their variants to predict cryptocurrency prices [19, 20]. Cerda and Reutter [21], for instance, introduced a multilayer LSTM deep learning model for Bitcoin price prediction. Their model's accuracy was enhanced by incorporating sentiment data from Twitter alongside Bitcoin prices.

A notable aspect of Cerda and Reutter's study is the Bitcoin dataset, which stands out compared to other cryptocurrency price prediction studies. This uniqueness arises from the use of high-frequency historical data, collected at minute intervals, whereas most studies rely on daily data. The use of minute-level data is particularly significant due to the nature of cryptocurrency markets, which operate 24/7, allowing trades to occur at any time. As a result, analyzing and predicting cryptocurrency prices using high-frequency data provides a more accurate representation of real-time market dynamics. This insight motivated us to use minute-level data in our work as well.

In summary, integrating EMD with LSTM has proven to be a promising approach for addressing the unique challenges of financial data, particularly in the context of cryptocurrencies. While previous studies have demonstrated the potential of this combination, there remains considerable scope for further research and improvement. This paper seeks to advance the field by developing an enhanced cryptocurrency price forecasting model that integrates EMD and LSTM and by evaluating its performance against existing models.

3. Research Methodology

This section presents the methodology employed in developing and evaluating the proposed cryptocurrency price forecasting model that integrates EMD and LSTM networks. We first start by preprocessing the signals (Figure 1), then we experiment with two different methodologies for the final price prediction. The process of data selection, preprocessing, and the two models developed for cryptocurrency price prediction are presented in detail in the sections below.

3.1. Empirical mode decomposition

The next step in the proposed methodology consists of the decomposition of the original cryptocurrency price series into IMFs using the EMD technique. EMD is a data-driven signal processing technique that separates a time series into different frequency components. Each IMF represents a specific scale or frequency present in the data, capturing different time scales of the cryptocurrency price dynamics. The decomposition process can be mathematically expressed as Equation (1).

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t) \tag{1}$$

Where:

1) $x(t)$ is the input signal being decomposed.
2) N is the number of modes extracted.
3) $c_i(t)$ represents the ith mode, which is an oscillatory component.
4) $r_N(t)$ is the final residue, which captures the trend or residual behavior.

The EMD algorithm iteratively applies two steps until a stopping criterion is met:

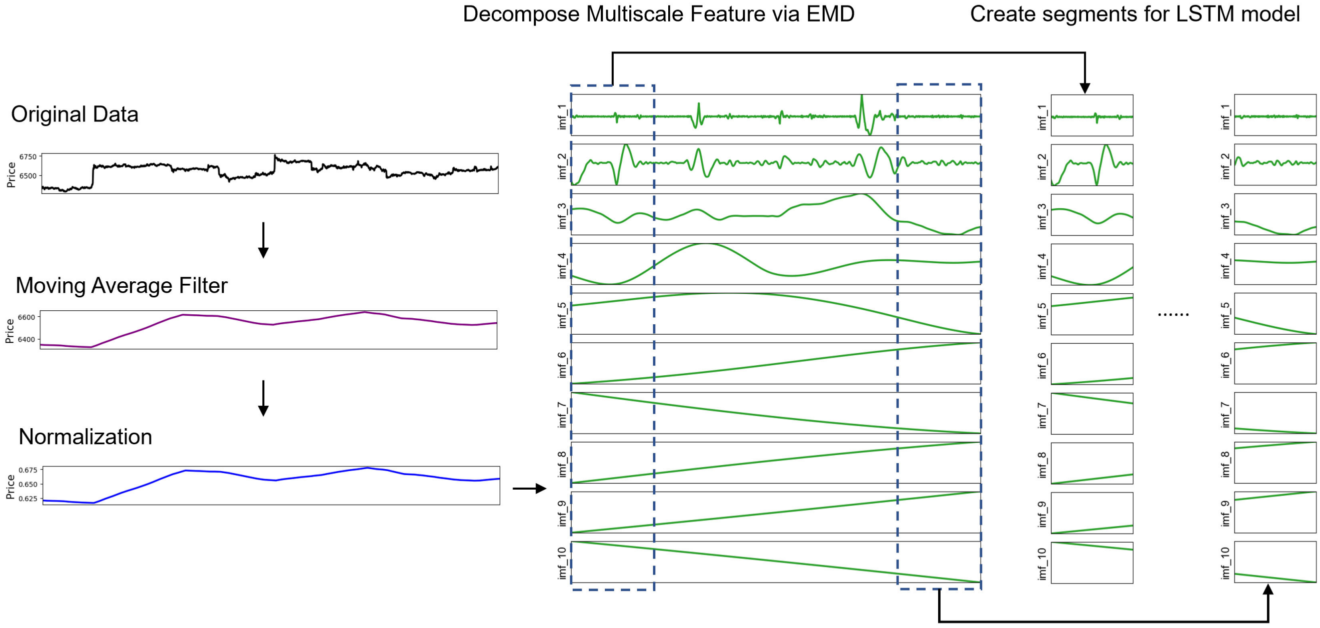
1) Find local maxima and minima called "extrema" in the signal.
2) Interpolate between adjacent extrema to obtain the upper and lower envelopes, which form the so-called IMFs.

The IMFs are obtained by repeating these steps until certain convergence criteria are satisfied. The final residue, $r_N(t)$, is obtained by subtracting all the extracted modes from the original signal.

3.2. Preprocessing

As shown in Figure 1, the preprocessing pipeline consists of three key steps: noise filtering using a moving average filter, normalization, and signal decomposition via EMD. The IMFs obtained from EMD are segmented into 288-data-point sequences, corresponding to 24-hour windows with a 5-minute resolution. These preprocessed segments serve as input for the forecasting model, ensuring the extraction of meaningful patterns while reducing noise interference.

Figure 1
Overview of the data pre-processing pipeline. The raw cryptocurrency price signal is smoothed with a moving average filter, normalized, and decomposed into Intrinsic Mode Functions (IMFs) using EMD. The IMFs are segmented into 288-data-point sequences (5-minute intervals), forming multidimensional input segments for the forecasting model



3.2.1. Noise filtering

Cryptocurrency signals exhibit susceptibility to multiple noise sources caused by the inherent volatility of this data, presenting significant challenges in their analysis. To identify the underlying trends that truly define the cryptocurrency price data, we applied a moving average noise filtering technique. Moving averages smooth out short-term fluctuations and highlight the long-term patterns in the time series. This filter calculates the average of a predefined number of consecutive data points within a time series. The equation for calculating the moving average is as follows in Equation (2):

$$Moving\ Average(t,\ k) = \frac{1}{k}\sum_{i=t-k+1}^{t} x_i \tag{2}$$

Where:
1) t represents the current time index or position in the time series.
2) k is the number of data points considered for averaging (also known as the window size).
3) x_i denotes the ith data point in the time series.

The moving average is computed by taking the sum of the k data points preceding the current time index and dividing it by k. This process slides along the time series, calculating the average for each window of k data points. In our work, we filter the data by using the moving average filter with a window size of 288 data points.

3.2.2. Normalization

Once the data are filtered, we apply normalization. Normalization helps the model converge faster and prevents large value differences from dominating the training process. We employed a min-max scaler to normalize the dataset, rescaling the values within the range of [0, 1].

3.3. Long short-term memory network

An LSTM network was developed for our prediction models. The network consists of one LSTM layer, followed by a dropout layer and a dense layer. The core of the model consists of the LSTM layer, which has shown great success in various time series forecasting tasks. LSTM is a type of recurrent neural network specifically engineered to capture temporal dependencies and long-term patterns in sequential data. As shown in Figure 2, the LSTM model processes sequential data using gating mechanisms that regulate memory updates. The cell state enables the retention of long-term dependencies, which is particularly useful for time series forecasting. The basic steps of LSTM are as follows:

First, as defined in Equation (3), the input gate i_t falters and extracts new information from the input x_t at the current state (time t) and creates a candidate value $\widetilde{c}_t$ for updating the state, as shown in Equation (4).

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\widetilde{c}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \tag{4}$$

Next, the forget gate f_t filters and keeps the historical information that can indicate the long-term trends and discards the non-critical information, as defined in Equation (5).

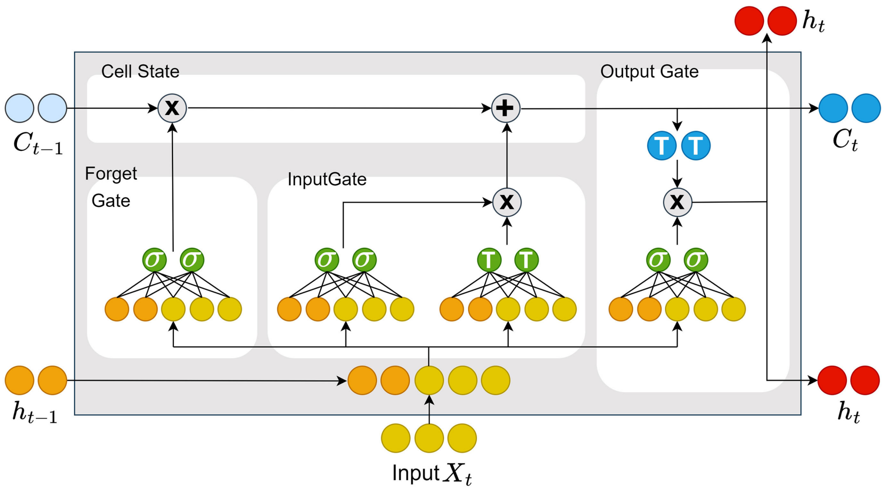
$$f_t = \sigma\left(W_f \times [h_{t-1}, x_t] + b_f\right) \tag{5}$$

By removing part of the information from the old cell and adding the filtered candidate value, as shown in Equation (6), the old cell state c_{t-1} is updated to the new cell state c_t.

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \widetilde{c}_t \tag{6}$$

Finally, as shown in Equation (7), the output gate o_t filters the updated state c_t, and the final output is calculated based on the updated state and the output gate state.

Figure 2
Structure of an LSTM cell. The cell state preserves long-term dependencies, while the hidden state captures short-term patterns, enabling effective sequential modeling for cryptocurrency price forecasting



$$o_t = \sigma\left(W_f \times [h_{t-1}, x_t] + b_o\right) \tag{7}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{8}$$

Here, h_t is the hidden layer state, namely the activation of the memory cell, as defined in Equation (8). W_c, W_f, and W_o represent the appropriate weight matrices, b_i, b_c, b_f, and b_o denote the corresponding bias vector, σ and tanh are the sigmoid functions and hyperbolic tangent functions, respectively.

The above-mentioned network is used to perform the final prediction, as shown in Figures 3 and 4. We feed the IMFs as input sequences to the LSTM model to capture the multiscale dynamics and temporal dependencies within the cryptocurrency price data. By combining the strengths of EMD and LSTM, we aim to enhance forecasting accuracy and capture the complex patterns present in the cryptocurrency market. After the decomposition of EMD, we can obtain different components of the original price series. Predicting each component can improve the prediction accuracy. Compared to the other EMD-LSTM models, this model does not train individual IMF signals with separate LSTM models individually, rather it converts each input sequence of the original input window into a concatenated sequence of multiple input sequences of IMF window inputs as shown in the bottom right part of the Figure 3.

3.4. EMD-LSTM model

We adopt a divide-and-conquer approach by leveraging EMD and LSTM networks to improve cryptocurrency price forecasting. EMD decomposes raw price signals into multiple IMFs, each capturing different frequency components, while LSTM models process these IMFs to learn meaningful temporal dependencies.

By decomposing the price series into IMFs, EMD isolates meaningful trends from noise, allowing LSTM to focus on

Figure 3
Overview of the Parallel EMD-LSTM model. The process consists of training (purple) and testing (green) phases. EMD decomposes the price series into Intrinsic Mode Functions (IMFs), each representing a distinct frequency. Separate LSTM models are trained on each IMF to learn multiscale patterns. During testing, each LSTM predicts its respective IMF, and the final forecast is reconstructed by summing all predictions, improving accuracy by reducing noise and capturing multiscale dependencies

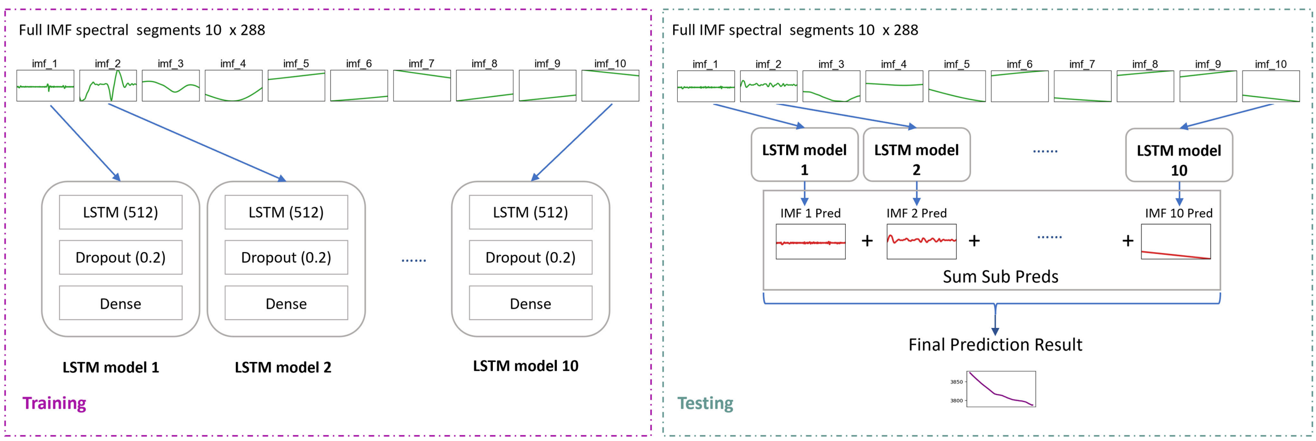
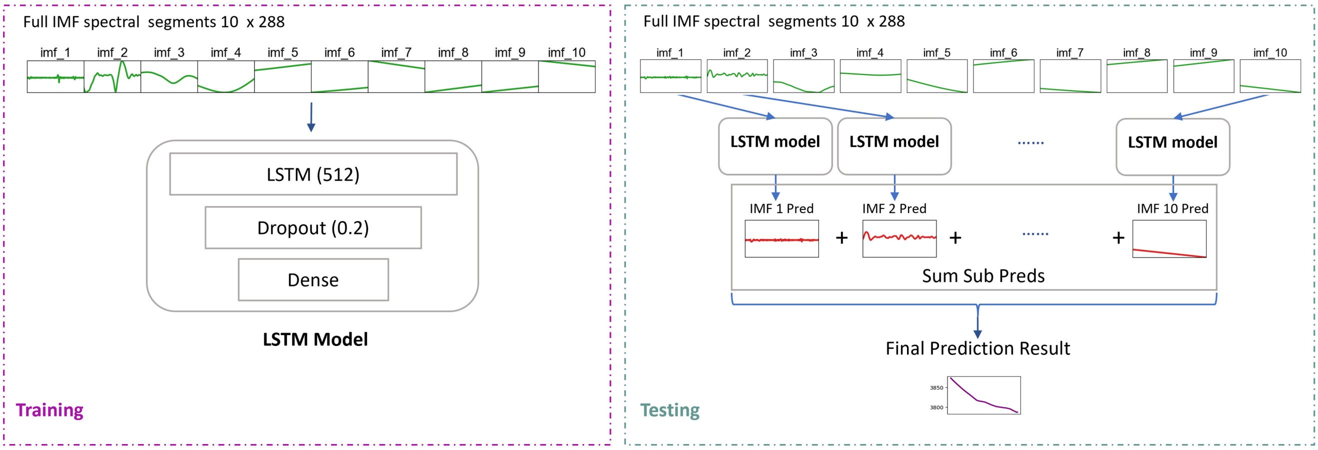


Figure 4
Overview of the Single EMD-LSTM model. Unlike the Parallel EMD-LSTM, this model uses a single LSTM network for all IMF components. EMD decomposes the price series into IMFs, which are fed into the LSTM sequentially to capture cross-frequency dependencies. During testing, IMF segments are processed through the same LSTM, and sub-predictions are summed to reconstruct the final forecast, maintaining EMD's benefits while reducing computational complexity



learning structured patterns rather than raw volatile price fluctuations. This decomposition improves feature extraction and helps LSTM capture both short-term fluctuations and long-term dependencies more effectively.

Instead of directly predicting the price, we forecast data for each IMF channel separately, leveraging EMD's ability to break down the original price series into different frequency components. By training LSTM models on individual IMFs, we allow each LSTM model to specialize in learning distinct price movement patterns at different temporal scales—high-frequency fluctuations, medium-term trends, and long-term trends. This divide-and-conquer approach reduces noise interference and enhances predictive accuracy. Since EMD is reversible, we reconstruct the final price prediction by summing all IMF forecasts, thereby preserving both short-term volatility and long-term market trends. During the training phase, we experimented with two distinct methods, leading to the creation of two separate models.

3.4.1. Parallel EMD-LSTM model

The Parallel EMD-LSTM model trains separate LSTM networks for each IMF component, rather than using the raw price series. The number of LSTM models equals the number of IMF channels, ensuring that each network specializes in learning distinct frequency-based patterns in price movements. During training, each IMF segment is fed into its corresponding LSTM model, which operates independently of the others. The final price prediction is reconstructed by aggregating the outputs of all LSTM models. This architecture allows each LSTM to focus on a specific frequency range, improving predictive performance in volatile market conditions. The entire process is illustrated in Figure 4.

3.4.2. Single EMD-LSTM model

The Single EMD-LSTM model differs from the Parallel approach by utilizing a single LSTM network instead of multiple independent models. Rather than processing each IMF separately, this model incorporates all IMF components as input features, enabling LSTM to capture cross-frequency dependencies within a unified architecture.

For example, if BTC price data is decomposed into 10 IMF channels, the LSTM model receives a sequence of 10 IMF segments for each time window. This approach preserves the advantages of EMD-based decomposition while significantly reducing computational cost, as only one LSTM network is required. However, since all IMF components are processed together, the Single EMD-LSTM model relies on LSTM's ability to internally capture frequency-specific patterns, rather than learning them independently, as in the Parallel EMD-LSTM approach. The entire process is illustrated in Figure 4.

While the Parallel EMD-LSTM model provides more specialized learning per frequency component, it comes at the cost of higher computational overhead due to multiple LSTM networks. In contrast, the Single EMD-LSTM model offers a computationally efficient alternative, processing all IMFs within a single LSTM while still leveraging EMD's decomposition benefits.

3.5. Parameter settings

When setting up an experiment for EMD-LSTM, meticulous attention must be given to defining and configuring the experiment parameters. These parameters play a critical role in determining the training process of the model, data preprocessing steps, and the overall setup of the experiment. In this section, some essential considerations when establishing the experiment parameters for EMD-LSTM are discussed.

Table 1 shows the parameters used for the training of the proposed EMD-LSTM model. It is well known that deeper architectures can capture more complex temporal patterns, but may require more computational resources. We found that 2 layers did provide slightly better results for BTC 2018 dataset, while it provided worse results for the rest of the 8 datasets. Therefore, to address the generalization, we chose 1 layer LSTM with 1 dropout and 1 dense layer. The number of units affects the model's capacity to learn complex patterns, but larger numbers can increase the risk of overfitting. Hence, 512 units were chosen with the best results after the experiments with a series of settings from 64 to 1024 units shown in Figure 5. A 0.2 dropout was also implemented, as it helps prevent overfitting and improves the generalization capacity of our model. The chosen loss function was the mean absolute error (MAE), which was selected to suit our regression problem. We employed the Adam optimizer with

Table 1
Experiment parameter settings

Parameter name	Value
Dataset cryptocurrency	BTC, ETH, and XRP
Dataset time range	H2 of 2018, 2019 and 2020
Dropout rate	0.2
Use Early Stop	Yes with Keras
Use ReduceLROnPlatean	Yes with Keras
Input sequence time steps	288
Initial learning rate	0.0005
LSTM layer activation function	Tanh
Loss function for LSTM layer	Mean Square Error
Moving average window length	288
No. of LSTM units	512
No. of neurons of the Dense layer	1
Optimizer for LSTM layer	Adam
Structure of the LSTM layer	1 LSTM layer, 1 dropout layer, and 1 dense layer
Scaling method	Min-Max scaling
Type of gradient descent	Batch gradient descent
Validation dataset	20% of the training dataset

an initial learning rate of 0.0005 and incorporated early stopping, along with the use of the adaptive learning rate technique ReduceLROnPlateau in TensorFlow. The experiments were conducted using the Python Keras and TensorFlow libraries as well as MATLAB, leveraging the computing power of a Windows station equipped with 32GB RAM, an RTX 3090 GPU, and an Intel i7 9700k CPU.

3.6. Evaluation metrics

To evaluate the performance of our proposed model, we employ various evaluation metrics commonly used in time series forecasting. These metrics include root mean squared error (RMSE) in Equation (9) and mean absolute error (MAE) in Equation (10). We compare the forecasting results of our two integrated EMD-LSTM models against a single LSTM benchmark model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{T} (\hat{y}_t - y_t)^2} \quad (9)$$

$$MAE = \sum_{t=1}^{T} |\hat{y}_t - y_t| \quad (10)$$

4. Results and Discussions

The performance of the cryptocurrency price forecasting models was analyzed using two methodologies:

1) Parallel EMD-LSTM, where each IMF is processed by an independent LSTM model.
2) Single EMD-LSTM, where all IMFs are processed within a single LSTM model, expanding the training set.

This section presents the experimental results of the proposed approaches. These results are then analyzed in the context of cryptocurrency price prediction using the developed LSTM

Figure 5
Comparison of RMSE and MAE for different LSTM unit configurations using BTC 2018 data. The figure shows how varying LSTM units affects model performance, with lower RMSE and MAE indicating better accuracy. The results help identify the optimal LSTM architecture for effective cryptocurrency price forecasting

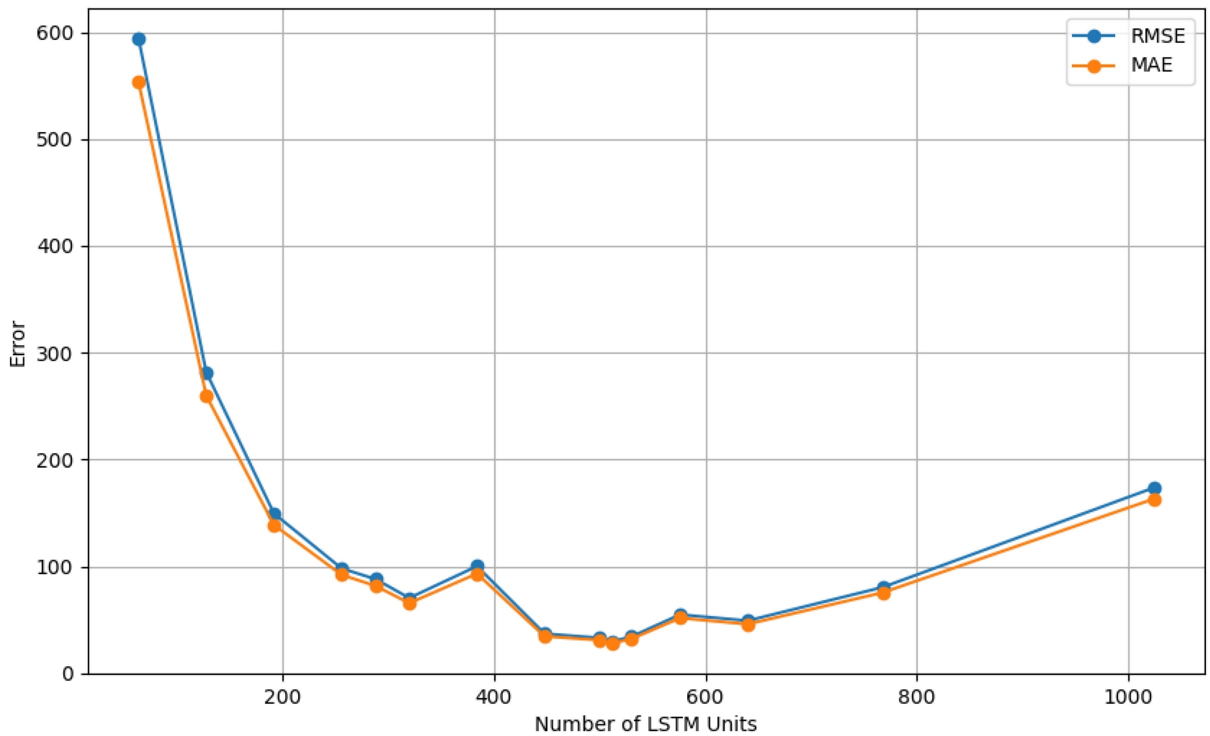
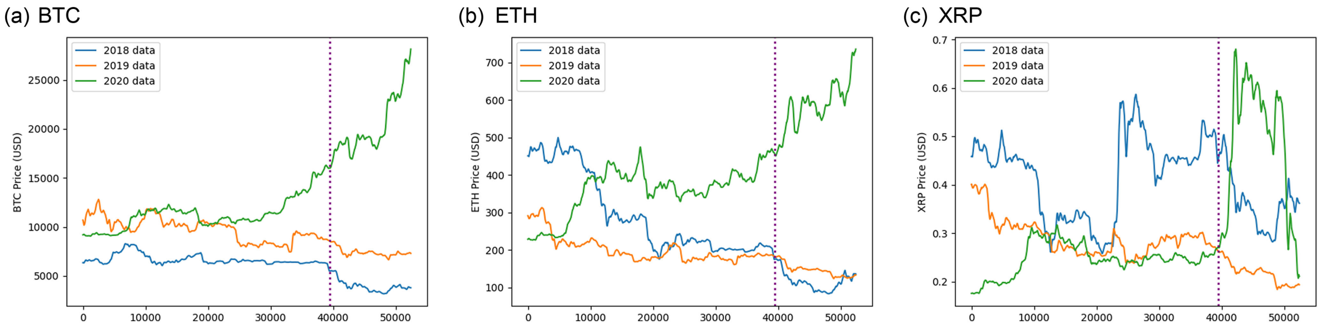


Figure 6
Historical closing prices of three cryptocurrencies in H2 2018, 2019, and 2020 (5-minute intervals). Each chart shows price fluctuations, with a purple line marking the split between the training set (39,456 points, 137 days) and the testing set (9,791 points, 35 days) used for model evaluation



model. Finally, we conclude this section with a comparison between our EMD-LSTM models and benchmark models, emphasizing the influence of integrating EMD-LSTM.

4.1. Data description and selection

The dataset consists of historical cryptocurrency price data collected from Bitstamp, a European-based exchange, covering the period January 1, 2018–December 30, 2020 (Figure 6). The selected cryptocurrencies—Bitcoin (BTC), Ethereum (ETH), and Ripple (XRP)—have a 5-minute price interval throughout the dataset. The dataset is partitioned into training and testing sets, with the split clearly indicated by a vertical line. The augmented Dickey–Fuller (ADF) test [22] was applied to assess the stationarity of each price series. If p-value is below a certain significance level (e.g., 0.05), it can reject the null hypothesis of non-stationarity and conclude that the time series is stationary. These datasets' statistic descriptions are listed in Table 2. As can be seen, none of their p-values is below 0.05, and therefore, none of them is stationary.

A suitable time interval is then chosen (daily or hourly) depending on the desired level of detail for forecasting. The dataset should cover a sufficiently long period to capture different market conditions and price trends. For our analysis, we focus on the closing prices from July 1st to December 30th for 2018, 2019, and 2020, resulting in 52,416 samples for each dataset. This data comprise three cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), and Ripple (XRP). We then divide each price series into an 80% training set and a 20% test set. The training set is used to train and optimize the ensemble EMD-LSTM model, while the test set assesses the model's prediction accuracy.

4.2. Segment construction for LSTM input

Since shorter sequences capture short-term dependencies, this study focuses on 5-minute segmentation to analyze short-term cryptocurrency price fluctuations. These intervals represent 288 data points across the entire IMF components, which results in multidimensional segments for the same time step. Following the segmentation step, we obtain 288 data point segments across all IMF components that are going to be used as input sequences in the LSTM model.

4.3. Results using Parallel EMD-LSTM model

The Parallel EMD-LSTM model demonstrated strong predictive performance across BTC, ETH, and XRP datasets over three years. However, accuracy declined for XRP in 2020. By dividing the RMSE values by the mean value of each dataset, most of the ratios are below 1%, especially for the 2018 and 2019 datasets. It is similar to MAE results. But when using the 2020 XRP dataset, the ratio is 2.63%. On average across the 9 datasets, the ratio of RMSE/Mean is 0.43%. On average across the 9 datasets, the ratio of MAE/Mean is 0.37%.

4.4. Results using Single EMD-LSTM model

This model consistently delivered strong performance across all BTC, ETH, and XRP datasets over three years. When the RMSE values are divided by the mean value of each dataset, the majority of the resulting ratios are below 0.1%, especially for the 2018 and 2019 datasets, which aligns with the MAE results. In the case of the 2020 dataset, the ratios consistently approach or exceed 0.1%. On average, in the nine datasets, the RMSE/Mean ratio stands at 0.06%, while the MAE/Mean ratio is averaged at 0.05%.

4.5. Comparison with benchmark models

The EMD-LSTM model was evaluated against a standard single-layer LSTM to assess its comparative performance. As shown in Tables 3 and 4, the EMD-LSTM consistently outperformed the

Table 2
The statistical description of all 9 cryptocurrency close price datasets

Ccy	BTC			ETH			XRP		
Year	2018	2019	2020	2018	2019	2020	2018	2019	2020
Mean	5992.26	9182.82	13558.26	250.11	187.28	414.89	0.41	0.27	0.31
Std	1277.93	1448.84	4355.12	119.62	38.49	115.75	0.08	0.05	0.13
Min	3178.46	6645.57	9054.20	82.77	123.30	226.07	0.26	0.18	0.18
Max	8284.72	12803.23	28135.61	499.77	312.59	735.90	0.59	0.40	0.68
P-Value	0.81	0.39	0.99	0.56	0.23	0.91	0.05	0.24	0.16

Table 3
Prediction results evaluation using RMSE for all BTC, ETH, and XRP prices in H2 of 2018, 2019, and 2020

Currency	Year	LSTM	Parallel EMD LSTM model	Single EMD LSTM model
BTC	2018	29.73	5.05	**1.217**
	2019	2.64	9.97	**2.418**
	2020	70.68	17.96	**13.81**
ETH	2018	0.336	0.25	**0.177**
	2019	0.16	0.05	**0.033**
	2020	0.77	1.81	**0.382**
XRP	2018	0.00022	0.00077	**0.00017**
	2019	**0.00012**	0.00022	0.00013
	2020	0.00402	0.00815	**0.00033**

Table 4
Prediction results evaluation using MAE for all BTC, ETH, and XRP prices in H2 of 2018, 2019, and 2020

Currency	Year	LSTM	Parallel EMD LSTM model	Single EMD LSTM model
BTC	2018	27.9	4.41	**1.018**
	2019	2.1	9.1	**1.858**
	2020	60.63	16.06	**12.864**
ETH	2018	0.319	0.24	**0.173**
	2019	0.13	0.04	**0.026**
	2020	0.68	1.78	**0.353**
XRP	2018	0.00016	0.00067	**0.00013**
	2019	**0.00009**	0.0002	0.00012
	2020	0.00334	0.00705	**0.00027**

benchmark model across multiple evaluation metrics, demonstrating its ability to capture the unique characteristics and volatility of cryptocurrency prices. One exception was the XRP 2019 dataset, where performance did not follow this trend. These results highlight that the integration of EMD with LSTM significantly enhances forecasting capabilities by improving feature extraction and noise reduction.

Figure 7 presents a direct comparison between the standard LSTM model and the proposed EMD-LSTM model using BTC 2018 data. The EMD-LSTM achieved an RMSE of 1.217, significantly lower than the benchmark model's RMSE of 10.87 as reported in Cerda and Reutter [21], highlighting a substantial improvement in forecasting accuracy.

As shown in Figure 7(a), the standard LSTM model struggles to capture the price trend accurately, particularly during volatile market conditions. Predictions exhibit noticeable deviations from actual price movements, resulting in higher forecasting errors. In contrast, the EMD-LSTM model (Figure 7(b)) demonstrates improved alignment with actual prices, particularly in the test prediction phase.

The zoomed-in section (Figure 7(c)) highlights this improvement: the EMD-LSTM model closely tracks actual price movements, whereas the standard LSTM exhibits greater divergence. This improvement is attributed to EMD's ability to decompose the price series into structured components, enabling the LSTM network to focus on meaningful price trends while mitigating the impact of market noise.

The integration of EMD and LSTM provides several key advantages in cryptocurrency price forecasting:

1) Feature decomposition: EMD decomposes the cryptocurrency price series into IMFs, isolating different frequency components and their respective dynamics.
2) Noise reduction: By breaking down the raw price signal, EMD filters out noise, allowing LSTM to process structured and meaningful input rather than noisy raw data.
3) Capturing temporal dependencies: LSTM models long-term patterns and dependencies within the decomposed IMFs, effectively learning from both short-term fluctuations and long-term trends.

These factors contribute to EMD-LSTM's superior predictive accuracy, as shown in the experimental results. By decomposing price series into IMFs, the model effectively adapts to different frequency components, capturing short-term volatility, and long-term market trends separately. This multiscale learning approach enhances forecasting performance, making EMD-LSTM particularly well-suited for volatile financial time series like cryptocurrency markets.

4.6. Ablation study: Evaluating the impact of EMD and LSTM variants

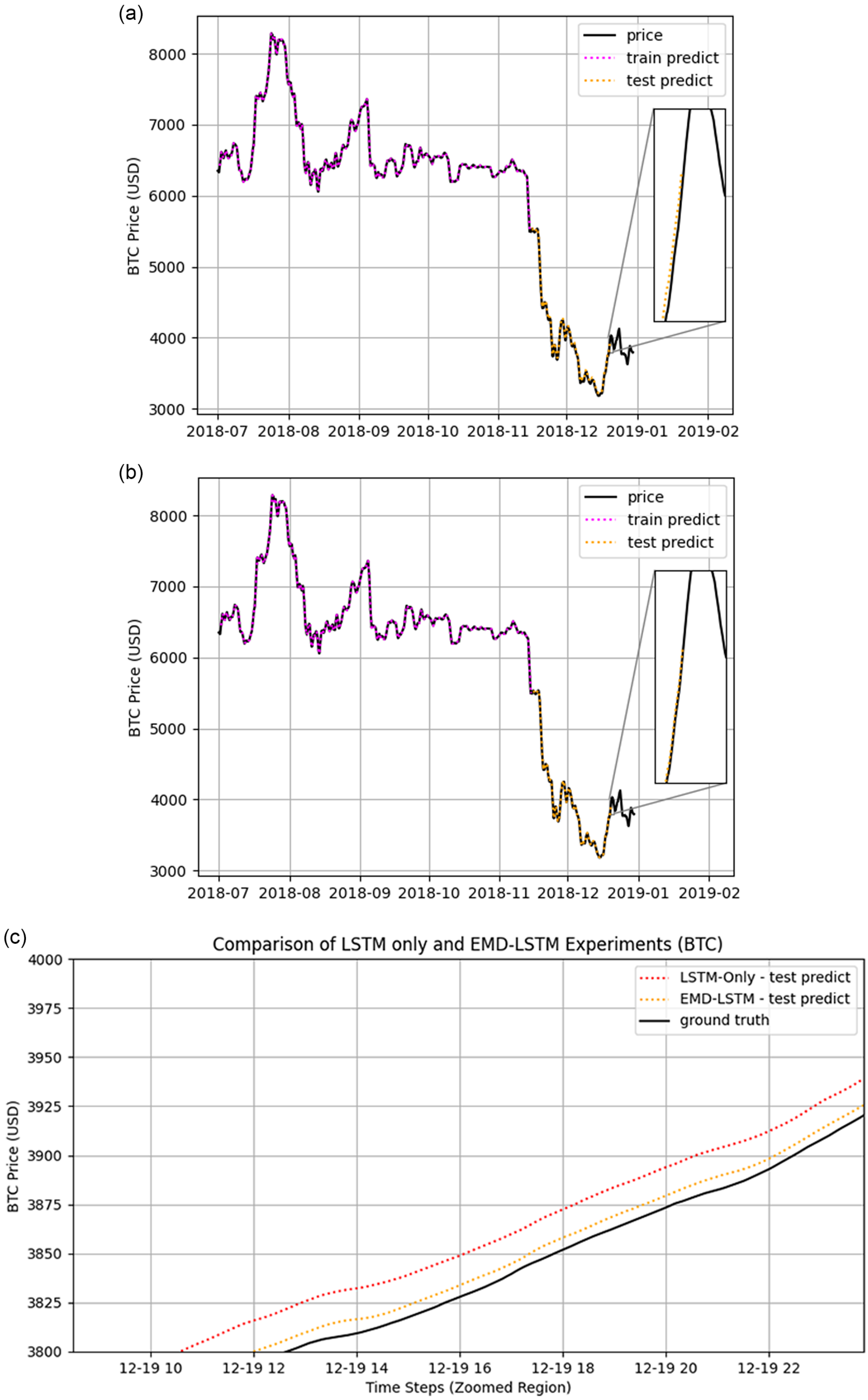
This section serves as an ablation study to assess the contribution of EMD and different LSTM configurations in improving cryptocurrency price forecasting. We compare three model variants:

1) Baseline LSTM (No EMD): Standard LSTM trained on raw price data.
2) Parallel EMD-LSTM: Each IMF is processed by an independent LSTM model.
3) Single EMD-LSTM: All IMFs are input into a single LSTM model.

Tables 3 and 4 present the RMSE and MAE evaluation results for these models across BTC, ETH, and XRP datasets (2018–2020). The results confirm that EMD enhances forecasting accuracy, with the Single EMD-LSTM model consistently delivering the best performance across all datasets.

In Table 2, the 2018 dataset exhibits a substantially lower mean value (5992.26) compared to 2019 (9182.82). Surprisingly, despite this discrepancy, the RMSE value for BTC price prediction in 2018 exceeds that of 2019 for the LSTM model. In particular, when employing both Parallel EMD-LSTM model and Single EMD-LSTM Model, the RMSE values for the predictions in each year consistently align with the ranking of the mean of the dataset for that year, as detailed in Table 3. Consequently, the Single EMD-LSTM Model consistently delivers significantly improved RMSE values (1.217, 2.418, 13.81) and MAE values (1.018, 1.858, 12.864) throughout the three years for the prediction of BTC prices. For ETH, the disparities in RMSE and MAE values among the models are relatively minor and the selection of the model does not appear to have a substantial impact on the accuracy of predictions. The values are consistently low, signifying strong predictive accuracy. Similarly to the BTC price prediction results, both EMD-LSTM models significantly outperform the LSTM, especially with Single EMD-LSTM Model RMSE values of 0.177, 0.033, 0.382 compared to 0.336, 0.16, 0.77, respectively. Moreover, the Single EMD-LSTM Model also outperforms the Parallel EMD-LSTM model consistently. Regarding XRP, both EMD-LSTM models delivered superior performance when applied to the 2018 and 2020 datasets. However, in the case of the 2019 data, they both performed less effectively than the LSTM model. Yearly trend analysis reveals fluctuations in RMSE and MAE values across different years. For example, in 2020, errors were more pronounced for most currencies and models compared to the years 2018 and 2019. This could be attributed to heightened volatility or other factors influencing the cryptocurrency market in 2020.

Figure 7
Comparison of BTC 2018 prediction results (5-minute intervals) (a) Prediction results using the standard LSTM model. (b) Prediction results using the proposed EMD-LSTM model. (c) The zoomed parts in (a) and (b) were put together into single plot to compare the difference more clearly. The EMD-LSTM model aligns more closely with actual price trends than the standalone LSTM, effectively capturing both short-term fluctuations and long-term trends by reducing noise and learning meaningful temporal patterns



Based on these results, it appears that incorporating EMD into the LSTM model can lead to improved predictive accuracy for cryptocurrencies for most of the datasets compared to using the pure LSTM model, 84.1% improvement for BTC, 54.3% improvement for ETH, and 85.5% improvement for XRP on average across 3 years. When considering a 3-year window, EMD-LSTM achieved an average RMSE of 5.815 for BTC, 0.197 for ETH, and 0.00021 for XRP, while the Parallel EMD-LSTM model resulted in an average RMSE of 10.99 for BTC, 0.70 for ETH, and 0.003 for XRP. Notably, for pricing data exclusively, decomposing the original data using EMD yields more detailed information for the model input. Furthermore, the way of storing these details in the same place will improve the prediction better than storing the details in different places.

4.7. Robustness and generalization

To assess the robustness and generalization of the proposed model, additional experiments were conducted using different cryptocurrencies, various periods, and distinct market conditions. The model's performance was evaluated consistently across different datasets to ensure its reliability and applicability in real-world scenarios. Therefore, as demonstrated by our experimental results presented above, the proposed EMD-LSTM model exhibited robustness and generalization capabilities. It consistently produced accurate and reliable predictions across different cryptocurrencies and periods, indicating its potential for practical applications in the cryptocurrency market. The model's ability to generalize well across different datasets enhanced its reliability and usefulness for various cryptocurrency assets.

5. Limitations and Future Directions

While the experimental results showcased the effectiveness of the proposed EMD-LSTM model, certain limitations should be acknowledged. The model's reliance on historical price data may restrict its ability to capture unforeseen market events or sudden changes. Additionally, extreme price fluctuations and external factors not considered in the dataset may impact the model's performance.

Future research directions could address these limitations and further enhance the proposed model. Incorporating additional features such as volume data, sentiment analysis, or macroeconomic factors into the model could improve its accuracy and robustness. Exploring alternative deep learning architectures, hybrid models, or ensemble techniques may also contribute to advancing cryptocurrency price forecasting capabilities.

6. Conclusion

In this work, we have proposed a novel cryptocurrency price forecasting model that integrates EMD and LSTM networks. The aim was to develop an accurate and reliable model capable of capturing the unique characteristics and volatility of cryptocurrency prices. Through a comprehensive evaluation of the model's performance, we have demonstrated its effectiveness and superiority over benchmark models.

The experimental results have shown that both proposed EMD-LSTM models outperform the traditional LSTM model commonly used in cryptocurrency price forecasting, especially with the Single EMD-LSTM Model. It achieves higher prediction accuracy, lower errors, and improved directional accuracy. The integration of EMD with LSTM allows the model to capture multiscale features and temporal dependencies within the data, leading to more accurate and reliable forecasts. The robustness and generalization of the proposed model have also been validated through additional experiments using different cryptocurrencies, periods, and market conditions. The consistent performance across diverse datasets indicates the model's reliability and applicability in real-world scenarios. However, it is important to acknowledge certain limitations of the proposed model. It relies on historical price data and may be influenced by factors not accounted for in the dataset. Additionally, sudden market changes, extreme price fluctuations, and unforeseen events can impact the model's performance. Future research directions could focus on addressing these limitations and further enhancing the proposed model. This may involve incorporating additional features such as volume data, sentiment analysis, or external factors that influence cryptocurrency prices. Exploring different deep learning architectures, hybrid models, or ensemble techniques could also contribute to improving the accuracy and robustness of cryptocurrency price forecasting.

In conclusion, the proposed EMD-LSTM models offer a powerful framework for cryptocurrency price forecasting. Its integration of EMD and LSTM allows for the capture of complex patterns, multiscale dynamics, and temporal dependencies within cryptocurrency price data. The model's superior performance and reliability make it a valuable tool for investors, traders, and policymakers seeking to make informed decisions in the dynamic cryptocurrency market.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available at https://www.cryptodatadownload.com/.

Author Contribution Statement

Xiaowei Wang: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Ioana Cretu:** Writing – review & editing, Visualization. **Hongying Meng:** Supervision, Project administration.

References

[1] Investopedia. (2024). *Cryptocurrency explained with pros and cons for investment*. Retrieved from: https://www.investopedia.com/terms/c/cryptocurrency.asp

[2] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., . . . , & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 454*(1971), 903–995. https://doi.org/10.1098/rspa.1998.0193

[3] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[4] Bishop, C. M. (2006). *Pattern recognition and machine learning*. UK: Springer New York.

[5] Chong, L. S., Lim, K. M., & Lee, C. P. (2020). Stock market prediction using ensemble of deep neural networks. In *IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology*, 1–5. https://doi.org/10.1109/IICAIET49801.2020.9257864

[6] Liu, J., & Huang, X. (2021). Forecasting crude oil price using event extraction. *IEEE Access*, *9*, 149067–149076. https://doi.org/10.1109/ACCESS.2021.3124802

[7] Lu, H., & Peng, X. (2021). Better effectiveness of multi-integrated neural networks: Take stock big data as an example. *Wireless Communications and Mobile Computing*, *2021*(1), 3938409. https://doi.org/10.1155/2021/3938409

[8] Patil, P., Wu, C. S. M., Potika, K., & Orang, M. (2020). Stock market prediction using ensemble of graph theory, machine learning and deep learning models. In *Proceedings of the 3rd International Conference on Software Engineering and Information Management*, 85–92. https://doi.org/10.1145/3378936.3378972

[9] Wang, J., & Wang, J. (2021). A new hybrid forecasting model based on SW-LSTM and wavelet packet decomposition: A case study of oil futures prices. *Computational Intelligence and Neuroscience*, *2021*(1), 7653091. https://doi.org/10.1155/2021/7653091

[10] Yang, S., Guo, H., & Li, J. (2022). CNN-GRUA-FC stock price forecast model based on multi-factor analysis. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *26*(4), 600–608. https://doi.org/10.20965/jaciii.2022.p0600

[11] Zhang, C., Sjarif, N. N. A., & Ibrahim, R. B. (2022). Decision fusion for stock market prediction: A systematic review. *IEEE Access*, *10*, 81364–81379. https://doi.org/10.1109/ACCESS.2022.3195942

[12] Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, *1*(01), 1–41. https://doi.org/10.1142/S1793536909000047

[13] Prasad, R., Ali, M., Kwan, P., & Khan, H. (2019). Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation. *Applied Energy*, *236*, 778–792. https://doi.org/10.1016/j.apenergy.2018.12.034

[14] Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 4144–4147. https://doi.org/10.1109/ICASSP.2011.5947265

[15] Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Stock price prediction using deep learning and frequency decomposition. *Expert Systems with Applications*, *169*, 114332. https://doi.org/10.1016/j.eswa.2020.114332

[16] Deng, C., Huang, Y., Hasan, N., & Bao, Y. (2022). Multi-step-ahead stock price index forecasting using long short-term memory model with multivariate empirical mode decomposition. *Information Sciences*, *607*, 297–321. https://doi.org/10.1016/j.ins.2022.05.088

[17] Guo, Y., Guo, J., Sun, B., Bai, J., & Chen, Y. (2022). A new decomposition ensemble model for stock price forecasting based on system clustering and particle swarm optimization. *Applied Soft Computing*, *130*, 109726. https://doi.org/10.1016/j.asoc.2022.109726

[18] Wang, Z., Yang, Z., Zheng, Z., & Zhu, Y. (2023). Bitcoin price forecasting based on ARIMA model and multifactorial linear regression. In *International Conference on Networking, Informatics and Computing*, 15–20. https://doi.org/10.1109/ICNETIC59568.2023.00009

[19] Livieris, I. E., Pintelas, E., Stavroyiannis, S., & Pintelas, P. (2020). Ensemble deep learning models for forecasting cryptocurrency time-series. *Algorithms*, *13*(5), 121. https://doi.org/10.3390/a13050121

[20] Vijaya Kumar, T., Santhi, S., Shanthi, K. G., & Gokila, M. (2023). Cryptocurrency price prediction using LSTM and recurrent neural networks. In *2nd International Conference on Applied Artificial Intelligence and Computing*, 1–5. https://doi.org/10.1109/ICAAIC56838.2023.10141048

[21] Cerda, G. C., & Reutter, J. L. (2019). Bitcoin price prediction through opinion mining. In *Companion Proceedings of the 2019 World Wide Web Conference*, 755–762. https://doi.org/10.1145/3308560.3316454

[22] Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366a), 427–431. https://doi.org/10.1080/01621459.1979.10482531

How to Cite: Wang, X., Cretu, I., & Meng, H. (2025). A Cryptocurrency Price Forecasting Model by Integrating Empirical Mode Decomposition and LSTM Neural Networks. *Artificial Intelligence and Applications*, *3*(3), 305–315. https://doi.org/10.47852/bonviewAIA52024202