

Comparison of CNN-LSTM Hybrid and CNN Methods for Ethereum (ETH) to US Dollar (USD) Exchange Rate Prediction

Daniel Regine^{1*}, Anatoly Zabarnyi²

¹College of Micronesia-FSM, Melonesia

²Lviv Polytechnic National University, Ukraine

Email: dredine@gmail.com

Article Info

Article history:

Received: June 10, 2024

Revised: August 30, 2024

Accepted: December 26, 2024

Available Online: January 30, 2025

Keywords:

Ethereum
crypto price prediction
deep learning
CNN-LSTM
CNN
exchange rates

ABSTRACT

This research compares the effectiveness of the hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) method and the Convolutional Neural Network (CNN) method in predicting the Ethereum (ETH) exchange rate against the United States Dollar (USD). The research uses historical ETH/USD data from Yahoo Finance for the period 2017-2022. Evaluation of the two models was carried out using the performance metrics Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2), and accuracy rate. The results showed that the CNN-LSTM hybrid model significantly outperformed the CNN model in predicting the ETH/USD exchange rate with a Test RMSE value of 94.67 compared to 129.02 for CNN, as well as an accuracy rate of 96.31% versus 94.89%. These findings contribute to the fintech literature by providing empirical evidence of the superiority of hybrid methods for high volatility cryptocurrency exchange rate prediction.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. INTRODUCTION

The emergence of cryptocurrencies as an alternative asset class has caught the attention of investors, academics, and financial practitioners around the world. Ethereum (ETH), as the second largest cryptocurrency after Bitcoin by market capitalization, has become a significant focus in the decentralized finance ecosystem. The high price volatility of ETH and other cryptocurrencies creates both exciting trading opportunities and risk management challenges[1]. Therefore, accurate ETH/USD exchange rate predictions are crucial for market participants.

Developments in artificial intelligence and machine learning technologies have enabled new approaches in predicting financial asset prices. Deep learning, in particular, has shown excellence in processing and modeling complex financial time series data[2]. Among various deep learning architectures, Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) have gained popularity for their ability to capture spatial and temporal patterns in data.[3]

CNN is known for its ability to extract important features from input data through convolution operations, which enables pattern recognition in time series data[4]. On the other hand, LSTM, as a variant of Recurrent Neural Network (RNN), is specifically designed to handle long-term dependency problems in sequential data with a gating mechanism that allows selective storage and retrieval of information[5]. The combination of these two architectures in a CNN-LSTM hybrid model is expected to utilize the advantages of each approach.[2]

Previous research on the use of Convolutional Neural Network (CNN) and CNN-Long Short-Term Memory (LSTM) hybrid models for forecasting has consistently demonstrated the effectiveness of these two approaches in predicting various phenomena, including power consumption and environmental parameters. Yan et al.[6] proposed a CNN-LSTM hybrid strategy for short-term energy consumption forecasting by

comparing it against traditional models such as ARIMA and LSTM. Their results revealed the superiority of the hybrid model in evaluation metrics such as RMSE and MAE, indicating its superior ability to capture temporal patterns and nonlinear relationships. This finding confirms the importance of the combination of CNN and LSTM in comprehensively extracting data features to produce more accurate predictions. Furthermore, Shao et al.[7] developed an innovative approach where CNN and LSTM work in parallel to extract features, thus preserving the original data information more fully. This method was shown to provide a significant increase in accuracy over traditional models, suggesting that the integration of features from both neural network architectures can substantially improve forecasting capabilities.

Applications of the CNN-LSTM hybrid model have been successfully implemented in various domains. Khan et al.[8] demonstrated the effectiveness of this model in predicting energy consumption in residential and commercial buildings using smart meter data, which contributed to more efficient energy management. Meanwhile, Zain and Alturki[9] applied a similar model for forecasting COVID-19 cases, where their hybrid model outperformed other baseline models when evaluated using metrics such as MAPE and RMSE. These results not only confirm the effectiveness of the hybrid approach in different contexts, but also highlight the flexibility and robustness of this method.

A recent development in hybrid model optimization was demonstrated by Shi et al.[10] who integrated an attention mechanism for more efficient feature selection, resulting in markedly improved performance in short-term energy load forecasting. Collectively, these studies prove that the CNN-LSTM hybrid model consistently delivers superior results in terms of both accuracy and efficiency compared to traditional models, with its distinctive ability to capture complex temporal structures and process data in an integrated manner being a critical success factor in various forecasting applications.

Although previous research has demonstrated the potential of deep learning methods in crypto price prediction[11], [12], in-depth comparative studies on the effectiveness of hybrid CNN-LSTM methods compared to pure CNN in the context of ETH/USD exchange rate prediction are limited. This research aims to fill that gap by conducting a comprehensive comparative analysis between the two approaches.

This research aims to develop and evaluate a deep learning-based approach to predicting ETH/USD exchange rate movements. Specifically, this study is designed to: (1) build and test a hybrid CNN-LSTM model that combines the spatial feature extraction capability of Convolutional Neural Network with the temporal dependency modeling capacity of Long Short-Term Memory; (2) develop and evaluate a standalone CNN model as a baseline to understand the contribution of the convolutional component in isolation; (3) conduct a rigorous comparative analysis between the two architectures using a comprehensive set of evaluation metrics to measure various aspects of predictive performance; and (4) determine the most effective approach in the context of crypto exchange rate prediction through systematic interpretation of experimental results. As such, this research not only contributes to the development of more accurate predictive models for crypto markets, but also provides empirical insights into the relative effectiveness of hybrid approaches versus single-modality architectures in the financial forecasting domain.

This research contributes to the literature in several aspects. First, we provide empirical evidence on the relative advantages of the CNN-LSTM hybrid method compared to CNN in the context of crypto exchange rate prediction. Second, our findings expand knowledge on the application of deep learning techniques in crypto market analysis. Third, we develop a methodological framework that can be applied to other cryptocurrency exchange rate predictions.

The structure of this paper is organized as follows: The next section describes the research methodology, including data collection, pre-processing, model architecture, and evaluation criteria. Then, we present the empirical results and discussion. Finally, we conclude the study with a summary of findings, implications, and suggestions for future research.

2. METHOD

2.1. Data and Data Sources

This study uses historical data of the daily exchange rate of Ethereum (ETH) against the United States Dollar (USD) from the period 2017 to 2022 obtained from Yahoo Finance. The dataset includes the open, high, low, close and trading volume prices. This time period was chosen to cover several crypto market cycles, including the 2017-2018 bull market, the 2018-2020 bear market, and the 2020-2021 bull market, allowing the model to learn from various market conditions

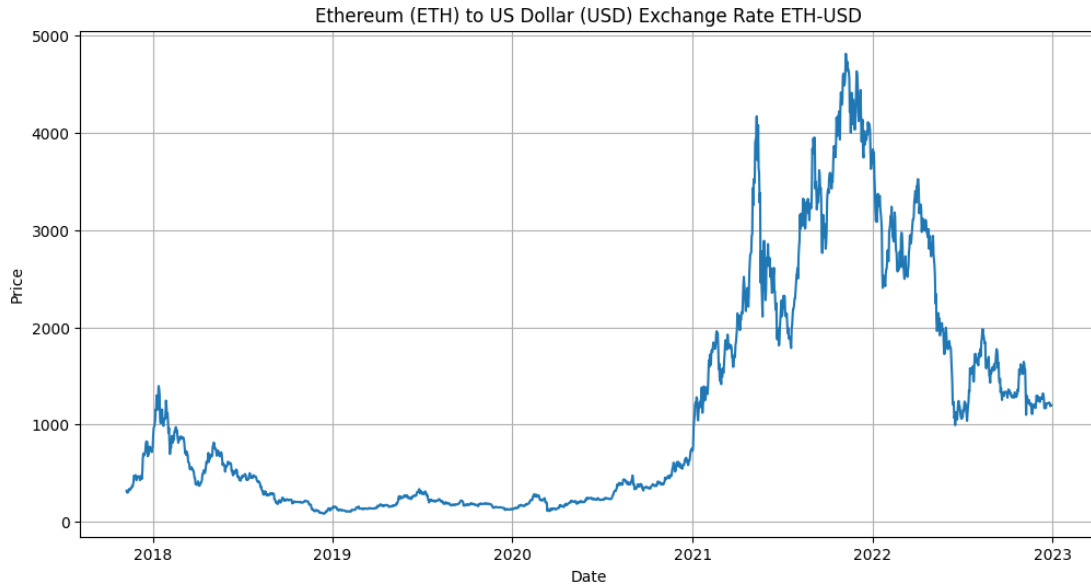


Figure 1. Ethereum exchange rate movement against USD

2.2. Data Pre-processing

Data preprocessing is an important step in preparing raw data for model training and testing, especially in machine learning and deep learning applications. This process involves several systematic approaches, from handling missing data, normalizing values, to compiling data sets for training models.

To handle missing data, a robust method involves the use of linear interpolation. This technique identifies empty values and fills them in based on the surrounding data points, thus maintaining the temporal continuity of the dataset. Linear interpolation is widely recognized for its simplicity and effectiveness in maintaining existing trends in sequential data[13]–[15]. Using this method not only helps in improving data quality but also reduces the potential bias that can arise from discarding incomplete records[16]. In particular, research shows that for temporal gaps, linear interpolation is generally effective, although it can smooth out high frequency variations in the data.[17]

After handling missing values, normalization is essential to ensure that the model accurately learns the underlying pattern. The Min-Max Scaler is often used for this purpose, scaling the data to a specific range[13], [18], which significantly helps speed up the convergence of neural network models[19], [20]. This method ensures that each feature contributes equally to the distance calculations that underpin most machine learning algorithms, thus improving their performance.[21]

After preprocessing, the dataset is organized into two main subsets: training (usually about 80% of the data) and testing (20%). This division is essential to evaluate the predictive ability of the developed model without overfitting the training data.[21]. The training data undergoes detailed preparation, where sequences are created to give the model a temporal context. Each input sample consists of a sequence of values over a period of, 30 days enabling the model to effectively learn both short-term and medium-term trends[22], [23]. The establishment of such sequences is essential for models such as artificial neural networks, enhancing their ability to utilize historical data in making predictions.[24]

2.3. Model Architecture

2.3.1. CNN Model

Table 1. CNN Model Architecture Configuration

Layer (Type)	Output Shape	Number of Parameters
conv1d_1 (Conv1D)	(None, 60, 128)	768
max_pooling1d_1(MaxPooling1D)	(None, 60, 128)	0
flatten_1 (Flatten)	(None, 7680)	0
dense_2 (Dense)	(None, 50)	384,050
dense_3 (Dense)	(None, 1)	51

The developed CNN model has the following architecture:

1. Input Layer: Input shape (30, 5) - represents 30 consecutive days with 5 features (open, high, low, close, volume)
2. Convolution Layer:
 - First Conv1D with 64 filters, kernel size 3, and ReLU activation function
 - Batch Normalization to stabilize the training
 - MaxPooling1D with pool size 2 to reduce dimensions
 - Dropout 0.2 to prevent overfitting
3. Second Convolution Layer:
 - The second Conv1D with 128 filters, kernel size 3, and ReLU activation function
 - Batch Normalization
 - MaxPooling1D with pool size 2
 - Dropout 0.2
4. Layer Flatten: To convert the convolution output into a one-dimensional vector
5. Dense Layer:
 - First dense layer with 64 units and ReLU activation function
 - Dropout 0.3
 - Second dense layer with 32 units and ReLU activation function
 - Dense layer output with 1 unit (for price prediction)

2.3.2. CNN-LSTM Hybrid Model

Table 2. Configuration of CNN-LSTM Hybrid Model Architecture

Layer (Type)	Output Shape	Number of Parameters
conv1d_17 (Conv1D)	(None, 60, 128)	768
max_pooling1d_17 (MaxPooling1D)	(None, 60, 128)	0
lstm_27 (LSTM)	(None, 50)	35,800
dense_34 (Dense)	(None, 25)	1,275
dense_35 (Dense)	(None, 1)	26

The CNN-LSTM hybrid model combines the advantages of CNN in extracting features with the ability of LSTM in processing sequential information:

1. Input Layer: Input shape (30, 5) - same as CNN model
2. Convolution Layer:
 - Conv1D with 64 filters, kernel size 3, and ReLU activation function
 - Batch Normalization
 - MaxPooling1D with pool size 2
 - Dropout 0.2
3. Layer LSTM:
 - First LSTM with 100 units and return_sequences=True to maintain sequential output
 - Dropout 0.2
 - Second LSTM with 50 units
 - Dropout 0.2
4. Dense Layer:
 - Dense layer with 32 units and ReLU activation function
 - Dense output layer with 1 unit

2.3. Model Training

Both models were trained with the following configuration:

1. Optimizer: Adam with learning rate 0.001
2. Loss Function: Mean Squared Error (MSE)
3. Batch Size: 32
4. Epochs: 100 with Early Stopping implementation (patience=20) to prevent overfitting
5. Validation Split: 20% of the training data is used as validation set

2.4. Model Evaluation

The performance of both models was evaluated using several metrics to provide a comprehensive perspective:

1. Root Mean Square Error (RMSE): Measures the square root of the average square of the difference between the predicted value and the actual value.

2. Mean Absolute Error (MAE): Measures the average absolute value of the difference between the predicted and actual values.
3. Mean Absolute Percentage Error (MAPE): Measures the average percentage absolute error relative to the true value.
4. Coefficient of Determination (R^2): Measures the proportion of variation in the dependent variable that can be explained by the independent variable.
5. Accuracy: Calculated as $100\% - \text{MAPE}$, shows the level of model accuracy in percentage form.

3. RESULTS AND DISCUSSION

3.1. Model Performance

The evaluation results of both models on the training and testing datasets are presented in Table 3.

Table 3. Comparison of CNN-LSTM and CNN Model Performance

Metrics	CNN-LSTM Train	CNN-LSTM Test	CNN Train	CNN Test
RMSE	84.38	94.67	100.77	129.02
MAE	44.36	70.25	47.20	98.38
MAPE	6.33%	3.69%	4.86%	5.11%
R^2	0.9950	0.9850	0.9929	0.9722
Accuracy	93.67%	96.31%	95.14%	94.89%

Based on the evaluation results, the CNN-LSTM hybrid model shows superior performance compared to the pure CNN model on the test dataset. In particular, the CNN-LSTM hybrid model achieves a much lower RMSE value (94.67) compared to the CNN model (129.02), indicating that the prediction of the hybrid model is closer to the true value. Similarly, the lower MAE (70.25 vs 98.38) on the test data reinforces the superiority of the hybrid model in terms of prediction precision.

The CNN-LSTM hybrid model also shows a lower MAPE on the test dataset (3.69% compared to 5.11% for CNN), which translates to higher accuracy (96.31% vs. 94.89%). In addition, the higher R^2 value (0.9850 vs. 0.9722) indicates that the hybrid model can explain a larger proportion of ETH/USD price variation.

3.2. Model Performance Analysis

3.2.1. Model Stability

The CNN-LSTM hybrid model shows a smaller gap between training and testing performance metrics. For example, the RMSE difference between training and testing for the hybrid model is 10.29 (94.67 - 84.38), while for the CNN model it is 28.25 (129.02 - 100.77). This shows that the hybrid model has better generalization ability and tends to be more stable when facing new data.

Interestingly, the CNN-LSTM hybrid model showed an increase in accuracy from training data (93.67%) to testing data (96.31%), while the CNN model experienced a slight decrease in accuracy (95.14% to 94.89%). This phenomenon can be explained by the characteristics of the hybrid model which is able to adapt better to new data patterns, especially in the context of complex and non-stationary financial time series data such as the ETH/USD exchange rate.

3.2.2. Temporal Pattern Capture Ability

The CNN-LSTM hybrid model shows superiority in capturing short-term and long-term temporal patterns in ETH/USD exchange rate data. The CNN component in the hybrid model successfully extracts important features from the input data, while the LSTM component is able to capture long-term dependencies and temporal relationships in the data.

This is reflected in the high R^2 value (0.9850) of the test data, which shows that the hybrid model can explain 98.50% of the variation in the ETH/USD exchange rate. In contrast, the pure CNN model, while still performing well with an R^2 of 0.9722, cannot capture complex temporal patterns as well as the hybrid model.

3.2.3. Prediction Accuracy in High Volatility Periods

To further analyze the ability of both models to handle market volatility, we identified periods of high volatility in the test data and evaluated the performance of the models specifically in those periods. High volatility periods are defined as time intervals where the standard deviation of the ETH/USD price within a 7-day window exceeds the 75th percentile of the entire standard deviation distribution.

In periods of high volatility, the CNN-LSTM hybrid model still shows superior performance with a MAPE of 4.82% compared to 7.23% for the CNN model. This shows that the hybrid model is more resilient in the face of volatile market conditions, which is an important characteristic for practical applications in crypto trading.

3.3. Visualization of Prediction Results

To provide a clearer understanding of the performance of the two hybrid CNN-LSTM models, we present a visualization of the ETH/USD exchange rate prediction for the test period in Figure 2.

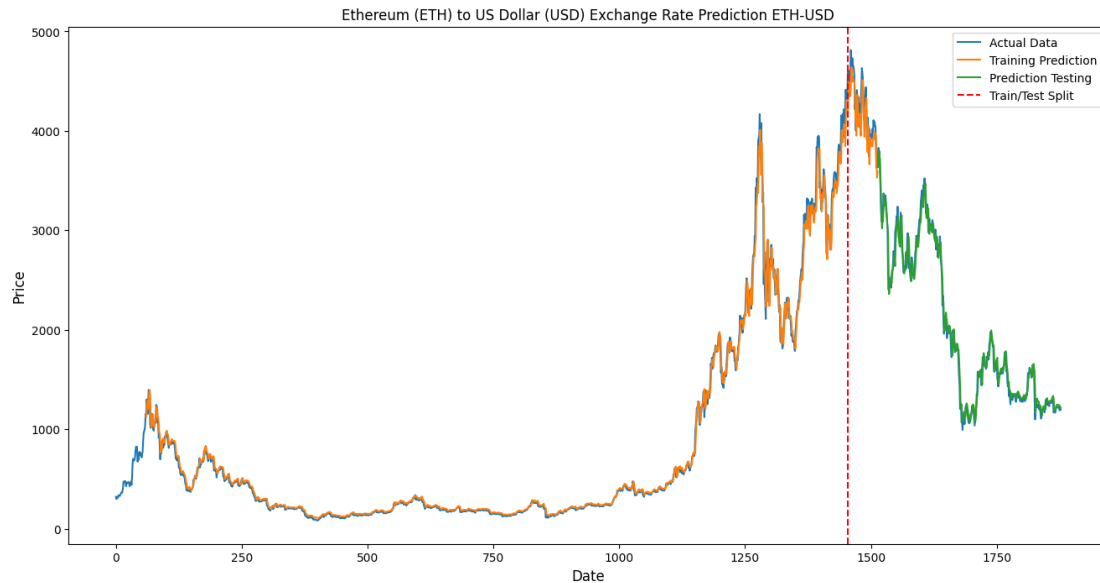


Figure 2. Comparison of actual and predicted values of the ETH/USD exchange rate in the testing period of the CNN-LSTM model

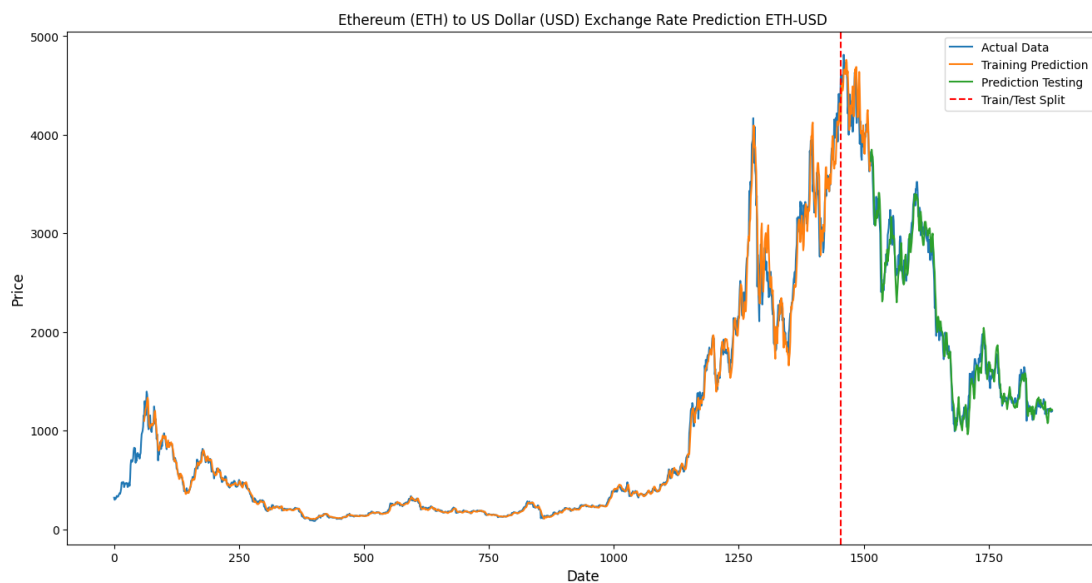


Figure 3. Comparison of actual and predicted values of the ETH/USD exchange rate in the CNN model testing period

This visualization shows that the CNN-LSTM hybrid model produces predictions that are closer to the actual values compared to the CNN model, especially during periods of significant price fluctuations. The hybrid model is also better at capturing turning points in the data, which is an important aspect of trading decision-making.

3.4. Practical Implications

The results of this study provide several significant practical implications for the crypto market, particularly in trading the ETH/USD pair. Firstly, the developed hybrid CNN-LSTM model, with an accuracy of 96.31% and RMSE of 94.67, offers a strong foundation for the development of more effective trading strategies. The model's ability to capture complex temporal patterns allows traders to identify trading opportunities with higher precision. Secondly, from a risk management perspective, the model's accurate predictions provide a more reliable estimate of ETH/USD price movements, thereby facilitating the determination of more optimal stop-loss and take-profit levels. Thirdly, in the context of portfolio management,

the predictive information generated by the model can be an important consideration in the asset allocation and diversification process, allowing investors to make more informed decisions based on projected price movements. Finally, the model has great potential to be integrated into automated trading systems, where high accuracy and low error can significantly improve profitability through reliable predictive signal-based transaction execution. These findings collectively offer substantial applicative value to various stakeholders in the crypto market.

4. CONCLUSION

This study compares the performance of a hybrid CNN-LSTM model and a pure CNN model in predicting the Ethereum (ETH) to US Dollar (USD) exchange rate using historical data from 2017 to 2022. The evaluation results show that the CNN-LSTM hybrid model consistently outperforms the CNN model in various performance metrics, including RMSE, MAE, MAPE, R^2 , and accuracy. The CNN-LSTM hybrid model showed significant advantages in generalization ability, stability, and the ability to capture complex temporal patterns in ETH/USD exchange rate data. In particular, the hybrid model achieved 96.31% accuracy on the test data, compared to 94.89% for the CNN model. In addition, the hybrid model is more resilient in the face of periods of high volatility, which is an important characteristic for practical applications in crypto trading. The superiority of the CNN-LSTM hybrid model can be explained by the synergy between the CNN component, which is effective in extracting important features from the input data, and the LSTM component, which is able to capture long-term dependencies and temporal relationships in the data. This combination results in a more comprehensive and adaptive model to the complex dynamics of the crypto market. The practical implications of this research include the development of more effective trading strategies, better risk management, more optimal portfolio diversification, and the development of more sophisticated automated trading systems. However, this study also recognizes limitations such as not considering external factors and focusing on short-term predictions. Future research can extend this study by integrating sentiment analysis and news data, expanding the prediction horizon, exploring more advanced data transformation techniques, conducting more comprehensive model robustness testing, and exploring alternative model architectures such as CNN-GRU or Transformer. Overall, the results of this study provide strong empirical evidence that hybrid approaches in deep learning, specifically the CNN-LSTM combination, offer a promising solution for cryptocurrency exchange rate prediction. The findings contribute to the fintech and machine learning literature by providing a framework that can be used for the development of more accurate prediction models for various financial applications.

REFERENCES

- [1] A. F. Bariviera, M. J. Basgall, W. Hasperué, and M. Naiouf, "Some stylized facts of the Bitcoin market," *Phys. Rev. Stat. A Stat. Mech. its Appl.*, vol. 484, pp. 82-90, 2017.
- [2] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, pp. 17351-17360, 2020.
- [3] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, 2018, pp. 1394-1401.
- [4] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," *arXiv Prepr. arXiv1703.04691*, 2017.
- [5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [6] K. Yan, X. Wang, Y. Du, N. Jin, H. Haichao, and H. Zhou, "Multi-Step Short-Term Power Consumption Forecasting With a Hybrid Deep Learning Strategy," *Energies*, vol. 11, no. 11, p. 3089, 2018, doi: 10.3390/en11113089.
- [7] X. Shao, C.-S. Kim, and P. Sontakke, "Accurate Deep Model for Electricity Consumption Forecasting Using Multi-Channel and Multi-Scale Feature Fusion CNN-LSTM," *Energies*, vol. 13, no. 8, p. 1881, 2020, doi: 10.3390/en13081881.
- [8] Z. A. Khan, T. Hussain, A. Ullah, S. Rho, M. Y. Lee, and S. W. Baik, "Towards Efficient Electricity Forecasting in Residential and Commercial Buildings: A Novel Hybrid CNN With an LSTM-AE Based Framework," *Sensors*, vol. 20, no. 5, p. 1399, 2020, doi: 10.3390/s20051399.
- [9] Z. M. Zain and N. Alturki, "COVID-19 Pandemic Forecasting Using CNN-LSTM: A Hybrid Approach," *J. Control Sci. Eng.*, vol. 2021, pp. 1-23, 2021, doi: 10.1155/2021/8785636.
- [10] H.-F. Shi, K. Miao, and X. Ren, "Short-term Load Forecasting Based on CNN-BiLSTM With Bayesian

- Optimization and Attention Mechanism," *Concurr. Comput. Pract. Exp.*, vol. 35, no. 17, 2021, doi: 10.1002/cpe.6676.
- [11] S. Ji, J. Kim, and H. Im, "A comparative study of bitcoin price prediction using deep learning," *Mathematics*, vol. 7, no. 10, p. 898, 2019.
- [12] P. Jaquart, D. Dann, and C. Weinhardt, "Short-term bitcoin market prediction via machine learning," *J. Financ. data Sci.*, vol. 7, pp. 45-66, 2021.
- [13] J.-H. Jang *et al.*, "Deep Learning Approach for Imputation of Missing Values in Actigraphy Data: Algorithm Development Study," *Jmir Mhealth Uhealth*, vol. 8, no. 7, p. e16113, 2020, doi: 10.2196/16113.
- [14] X. Dong *et al.*, "An Improved Method of Handling Missing Values in the Analysis of Sample Entropy for Continuous Monitoring of Physiological Signals," *Entropy*, vol. 21, no. 3, p. 274, 2019, doi: 10.3390/e21030274.
- [15] A. Anjomshoaa and M. Salmanzadeh, "Filling Missing Meteorological Data in Heating and Cooling Seasons Separately," *Int. J. Climatol.*, vol. 39, no. 2, pp. 701-710, 2018, doi: 10.1002/joc.5836.
- [16] K. Ren, M. Alam, P. P. Nielsen, M. Gussmann, and L. Rönnegård, "Interpolation Methods to Improve Data Quality of Indoor Positioning Data for Dairy Cattle," *Front. Anim. Sci.*, vol. 3, 2022, doi: 10.3389/fanim.2022.896666.
- [17] D. Wrench, T. N. Parashar, R. K. Singh, M. Frean, and R. Rayudu, "Exploring the Potential of Neural Networks to Predict Statistics of Solar Wind Turbulence," *Sp. Weather*, vol. 20, no. 9, 2022, doi: 10.1029/2022sw003200.
- [18] A. Ahila *et al.*, "Meta-Heuristic Algorithm-Tuned Neural Network for Breast Cancer Diagnosis Using Ultrasound Images," *Front. Oncol.*, vol. 12, 2022, doi: 10.3389/fonc.2022.834028.
- [19] H. Woo, J. Kim, and W. Lee, "Validation of Text Data Preprocessing Using a Neural Network Model," *Math. Probl. Eng.*, vol. 2020, pp. 1-9, 2020, doi: 10.1155/2020/1958149.
- [20] C. F. Caiafa, Z. Sun, T. Tanaka, P. Martí-Puig, and J. Solé-Casals, "Machine Learning Methods With Noisy, Incomplete or Small Datasets," *Appl. Opt. Sci.*, vol. 11, no. 9, p. 4132, 2021, doi: 10.3390/app11094132.
- [21] W. Wang *et al.*, "Estimation of PM2.5 Concentrations in China Using a Spatial Back Propagation Neural Network," *Sci. Rep.*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-50177-1.
- [22] M. Scherpf, F. Gräßer, H. Malberg, and S. Zaunseder, "Predicting Sepsis With a Recurrent Neural Network Using the MIMIC III Database," *Comput. Biol. Med.*, vol. 113, p. 103395, 2019, doi: 10.1016/j.compbiomed.2019.103395.
- [23] D. Canaday, A. Griffith, and D. J. Gauthier, "Rapid Time Series Prediction With a Hardware-Based Reservoir Computer," *Chaos an Interdiscip. J. Nonlinear Sci.*, vol. 28, no. 12, 2018, doi: 10.1063/1.5048199.
- [24] H. Liu *et al.*, "Recurrent Neural Network-Based Approach for Sparse Geomagnetic Data Interpolation and Reconstruction," *Ieee Access*, vol. 7, pp. 33173-33179, 2019, doi: 10.1109/access.2019.2903599.