Deep Learning-Based Forecasting of Cryptocurrency Prices : Short to Long Horizon

Dr Krishna.J[1], Kavya.J[2], Harsha Vardhan Reddy.R[3], Soma Sekhar.U[4], Manikanta.G[5], Venkata Thriveni.K[6]
[1]*Associate Professor, Dept of AIML, Annamacharya University, Rajampet, Andhra Prdasesh, India*
[2,3,4,5,6] *Student, Dept of CSE(AI),Annamacharya Institute Of Technology & Sciences, Rajampet, India*

Abstract— **The unstable cryptocurrency markets present new opportunities and threats. The possibility of exposure to the cryptocurrency assets is quite high as the exchange rates can change on a daily basis. The project estimates the price of cryptocurrency by the use of the potent machine learning tools. The Neutral Networks were the most efficient in the best forecast and best validation as compared to other two seven models having minor errors. The neural networks used in the prediction of the tendencies in the future have been long short term memory (LMST). Multiple associations of financial data can be assessed also, regarding LSTM model. Over fifty cryptocurrencies were subjected to the Exploratory Data Analysis (EDA), which first took the historical data collection, followed by the feature engineering, integrative binning, data preparation, and standardization. The most successful ones were identified based on the price movement, the market size and volumes. The Python written LSTM-based model has been applied in studying the intricate trends and associations in 90 days of price movement statistics. The performance measures were RMSE and MAE, which were used to monitor the performance of the model. These findings support the Adaptive Market Hypothesis (AMH) that states that adaptive changes in the behavior of investors and markets influence the dynamic efficiency of cryptocurrency market. The study highlights the potential of machine learning models in financial economics and how they will aid with risk management techniques and investment decision-making.**

Keywords— LSTM neural networks, machine learning, financial economics, model prediction, and currency forecasting

I. INTRODUCTION

The secretive Satoshi Nakamoto introduced bitcoin, a new financial innovation [19], [5]. When Bitcoin was released in 2009, it heralded the commencement of decentralized digital currency that allowed for safe peer-to-peer transactions [19]. After Bitcoin was discovered, a number of other cryptocurrencies, such as Ethereum, Ripple, and Litecoin, came into being. These are now widely used in many areas of the economy, including as banking, game creation, and supply chains [6]. The speed of growth and volatility of this decentralized market necessitate the use of complex predictive models in order to predict price trends and finely tune the risks [1]. Machine learning and especially LSTM networks are applied in the prediction of sequential data because the normal linear models cannot capture the complex nonlinear dynamics of financial time series [14].Research has demonstrated that LSTM models are superior to ARIMA and other conventional models in forecasting cryptocurrency and stock prices [16]. The recents also emphasize the success of their reliance on data preprocessing and model settings [10]. In this research, seven models are used to estimate the most effective model in terms of having the minimum error and the highest precision, and the LSTM Neural Network is chosen due to a 90-day cryptocurrency price prediction [3]. The research aims to strengthen forecasting accuracy through advanced feature engineering, normalization, and validation, contributing valuable insights into machine learning applications in dynamic financial markets [21].

II. LITERATURE REVIEW

An LSTM-based model is utilized by Suedumrong et al., 2021 [1] to forecast Bitcoin price changes. Using historical pricing data, a 72.4% accuracy rate was reached. Performance was poor when the market volatility was high. Zhang et al., 2022. [2] introduced a transformer-based long-term dependency learning paradigm report on a 78.6 percent accuracy rate on data

about bitcoin prices. It requires high processing power and the benefit over LSTM is not significant.
Cao et al., 2022. [3] developed a multivariate forecasting GRU model with attention addition. The accuracy of 91.3% on crypto datasets was achieved. The model had a problem of high memory consumption.
GWO in Ahmed et al. (2023) [4] proposed an LSTM with short and medium horizons prediction accuracy of 89.7%. had long training time and slow convergence.
A hybrid CNN-GRU model [5] proposed by Li et al. (2024) was proposed as a short-term forecasting model. The accuracy was 92.4%. At a higher cost, the long-horizon prediction led to low performance.
A comparison of ensemble models and DL models was conducted by Bouteska et al. (2024) [6]. GRU and LightGBM achieved up to 90.1 percent accuracy. None of the models proved to be the most suitable ones in all horizons.
Farooq et al. [7] established a Temporal Fusion model of multi-horizon prediction, which used Transformers. 85.6% long-term and 93.2% shortterm accuracy were achieved. limited by the complexity and high cost of training.

Contribution of the Proposed work

The available literature anticipates the price of a single cryptocurrency and is primarily oriented towards short-term predictions. Most of them use complicated models, which require a longer time and are more costly and cannot be applied effectively when it comes to long-term prediction.

The project employs a basic LSTM model to forecast price changes in 90 days of most cryptocurrencies. Through the appropriate data analysis and processing features, the model is highly accurate and less complex, thus applicable to the real world. The available literature anticipates the price of a single cryptocurrency and is primarily oriented towards short-term predictions. Most of them use complicated models, which require a longer time and are more costly and cannot be applied effectively when it comes to long-term prediction.

The project employs a basic LSTM model to forecast price changes in 90 days of most cryptocurrencies. Through the appropriate data analysis and processing features, the model is highly accurate and less complex, thus applicable to the real world.

III. METHODOLOGY

In order to make accurate predictions, a deep seated approach that has been confirmed to be correct was employed to analyze the possible market trends and patterns of cryptocurrencies[8]. This study process is organized in the form of figure 1[3].

Data Collection

The procedure begins by the systematic collection of historical bitcoin exchange data of reliable fiscal providers [17]. This data is based on the understanding of the decentralized character of cryptocurrencies, as conceived in the seminal research on Bitcoin and blockchain technology, which defines how subjective digital assets act and the way they are organized [19]. The studies about market interconnection and cryptocurrency volatility indicate the importance of obtaining reliable and constant data to model it [1]. In the case of widely used cryptocurrencies, such as Bitcoin, Ethereum, Ripple, and Bitcoin Cash, there are open, close, high, low, volume, and market capitalization data points in the dataset [6]. Recent studies exploring the relationship between prices and trading volumes of cryptocurrencies indicate the necessity to collect and identify nonlinear relationships in the rapidly dynamical markets [2].

Data Acquisition

They were first gathered using a Python-based selection of CoinMarketCap of the top 50 cryptocurrencies in Yahoo Finance [23] https://finance.yahoo.com/markets/crypto/all/)

Data Preprocessing

The raw data must undergo considerable preprocessing before it can be transformed into deep learning models [10].
Financial time-series data often contain gaps in numbers and values because of an exchange failure or difficulties in retrieving the data [17]. The studies on predictive modeling highlight the importance of interpolation or value propagation methods in a systematic fashion to fill such discrepancies [8]. As per the ideas presented in conventional statistical normalization methods, normalization will be conducted to put all the numerical values within relative scales once the data has undergone cleansing [12]. In order to identify temporal dependencies, the

data should be turned into sequence windows [16]. Such a sliding window adjustment guarantees the viability of the past data streams to forecast the future prices, which is in line with the time-series forecasting and feature engineering literature [21].

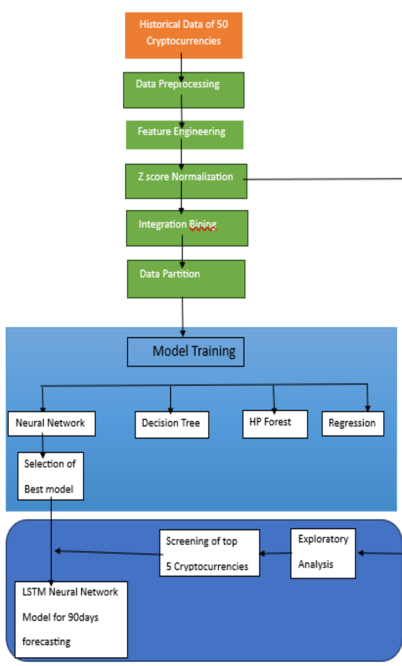
Feature Engineering

Technical indicators, volatility indices, and moving averages [8] were added to the content. To identify trends, Simple Moving Averages (SMA) and Exponential Moving Averages (EMA) were computed [21].

Certain technical indicators were calculated with the help of the pandas ta library which included Bollinger Bands, Relative Strength Index (RSI) among others [8].

Z-score Normalization

Z-score normalization was used to attain equal scaling of features used in our study, and this was done to the processed data obtained through feature engineering [12]. This was necessary in order to standardize the numerical properties [12]. Then, the StandardScaler from the scikit-learn package was used to find and pull out the numerical characteristics so they could be scaled [25]. To preserve the dataset's original structure, the normalized features were then computed and applied to the non-numeric data once more [10].



Modeling And Forecasting Experimental Flowchart

Exploratory Data Analysis (EDA)

The top achievers in the bitcoin market were determined utilizing a detailed data research approach [8]. The follow-up of cryptocurrency closing prices showed how things were changing [17]. The average close price and the quantity of the trades of the cryptocurrencies during a given period was compared using a bar graph [2]. Correlation between the attributes that were selected on the cryptocurrencies was also drawn up to assess the relationship amongst the attributes [11]. It consisted of a time series price analysis study in logarithmic form of the most popular cryptocurrencies [16]. The log transformation was employed [21] to stop the changes and find long-term trends.

Log (Close Price) = log Price

Model Training

The machine learning models were trained using SAS Enterprise Miner Client 15.2. The dataset was separated into three subsets: test (20%), validation (20%), and training (60%) [23]. Continuous variables were categorized using Integrative Binning [23]. Binning was employed in order to deal with the non-linear interactions and enhance the model performance [8]. There were four trained and evaluated models, namely Decision Trees, High Performance Forests, Regression Models, and Neural Networks [18]. The neural networks could capture complex nonlinear patterns with the help of multi-layered architectures [14]. In order to achieve greater accuracy in predictions with the help of an ensemble learning, HP Forest relied on a high-performance random forest model [18]. The performance of each of the models was compared using the normalized data [12].

LSTM Model for forecasting future trends

The Long Short-Term Memory (LSTM) Neural Network model was employed to forecast the future price of the top 5 cryptocurrencies because it signaled the best performance during the training of the model [16]. The memory cell structure of LSTMs also stores information at long time steps and thus, they do not experience the vanishing gradient issue of simple RNNs [14]. The input gate, output gate and forget gate are the three crucial components that make this feasible. What data must be erased from the cell state is determined by the forget gate. It generates a value between 0 and 1 and the activation function is a sigmoid whereby 0 implies discarding all and 1 implies retaining all [16].

$$g_t = \sigma\,(X_g\,[\,k_t{-}1,\, y_t\,]{+}d_g\,)$$

where, dg is the bias term, y_t is the fresh input, Xg is the weight matrix, g_t is the forget gate's output at time t, and σ is the sigmoid activation function.

The new data is characterized as being stored in the cell's input gate. It is represented by a layer of sigmoid and a layer of tanh which decides what values to update and a vector of possible candidate values respectively is generated [16].

$$j_t = \sigma\,(X_j\,[k_t{-}1,\, y_t] + d_j)$$

$$P_t = \tanh\,(X_p\,[k_t{-}1,\, y_t] + d_P)$$

Finally, the output will be generated at the output gate that will refresh the updated cell state[16].

$$P_t = g_t \times P_t{-}1 + j_t \times P_t$$

$$q_t = \sigma\,(X_q\,[k_t{-}1\,, y_t] + d_q)$$

$k_t = q_t \times \tanh\,(P_t)$

Model Validation

A range of statistical evaluation criteria, such as Mean Absolute Error (MAE), Root Mean, Final Prediction Error, and Squared Error (RMSE), were utilized to adequately gauge the model's performance.

These approaches produce quantitative predictions regarding the quality and applicability of the model's forecasts [17]. The result of these tests illustrated how these tests indicated the complex time interrelation that existed in the cryptocurrency market data[16].

Feature Engineering

There are important observations that have been made by computing the technical indicators, volatility and moving averages[8]. The indicators of volatility such as ATR helped give a better picture of the market whereas the SMA and EMA helped in locating trends[1]. The Bollinger Bands, SMA 20, and EMA 20 plots are shown in Figure 2[8]. The SMA 20 shows the average asset price over 20 periods, where prices above it indicate an uptrend. The EMA 20 gives greater weight to recent prices, highlighting short-term momentum[21]. Bollinger Bands, placed two standard deviations from the SMA, indicate overbought conditions when prices reach the upper band and oversold conditions when they touch the lower band[8].

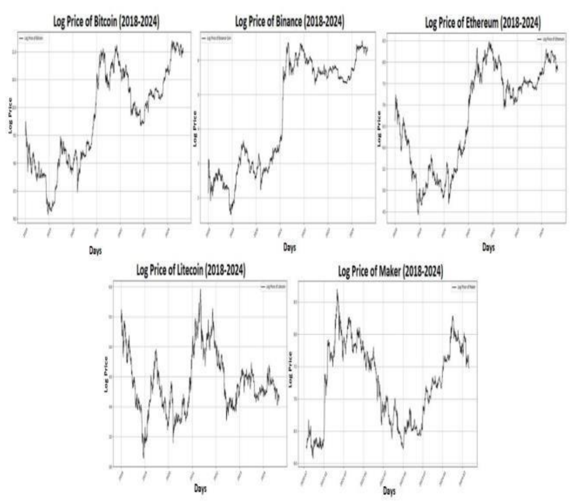
Z-score Normalization

The z-score standardization was able to standardize the numerical characteristics of the dataset. This action served to reduce the influences of all the feature scales and ranges and the data were made ready to perform further statistical processing and machine learning applications[12]. The visual analysis of the post-normalized data shows that the transformation was successful since all the numerical variables revealed a similar scale[10]. The data did not have any outliers that could distort the results of the analysis. To enhance model convergence and forecast accuracy, the standardized dataset was preserved and employed in succeeding modeling [25].

Looking at Exploratory Data

The most successful of the 50 cryptocurrencies analyzed were Bitcoin, Ethereum, Maker, Binance Coin, and Litecoin due of their continuously high trading volume and unique average price [6]. The correlation matrix showed that the closing prices had varying levels of linkage, with some having large positive associations [11]. To take an example, the correlation to Bitcoin and Ethereum was also strong which means that they affected the whole market[1]. Important trends and development patterns within the cryptocurrencies were determined by plotting the logarithmic-transformed time series graphs of each of them [16]. To accurately depict periods of rapid growth, adjustments and mergers, log plots are normally used to depict financial variables [21].

The logarithmic graphs are shown in fig 2.



Graphs of specific cryptocurrencies logarithmic time series

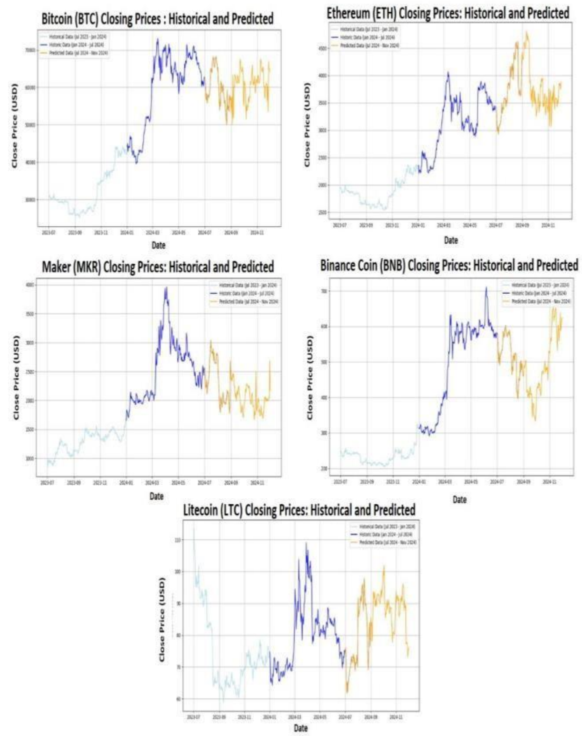
Model Comparision

Comparisons of models showed that both models possess certain performance peculiarities[18]. Neural networks work well with intricate patterns and non-linear interactions [14]. HP Forest worked well and gave better forecast accuracy because it used an ensemble technique. Decision trees provided decision directions and rendered them comprehensible[18]. It can be seen in Table 1 that the lowest squared error was Neural Networks followed by Decision Trees[17].

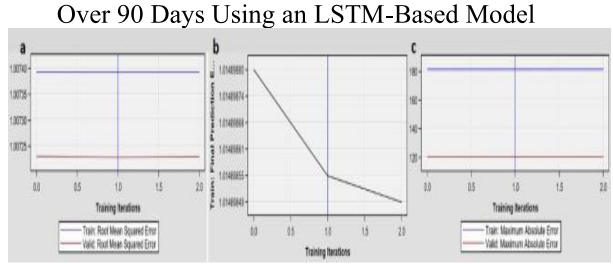
90-day LSTM model forecasting with a valid average squared error of applied models

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Selection Criteria:Valid Average Squared Error
Y	Neural	Neural	Neural Network	Close	1.01450912
	Tree	Tree	Decision Tree	Close	1.014666461
	Reg	Reg	Logistic Regression	Close	1.016109443
	Reg2	Reg2	Stepwise Regression	Close	1.016111252
	Reg3	Reg3	Forward Regression	Close	1.016111252
	Reg4	Reg4	Backward Regression	Close	1.016111252
	HPDM Forest	HPDM Forest	HP Forest	Close	1.01674716

The sequential patterns and oscillations of bitcoin price data have been successfully recognized by the LSTM model [16]. The model can reasonably predict market turning points and general trends over 90 days[17]. The findings indicate the technique effectively predicts complicated market activity and provides a helpful tool for predicting future Bitcoin values. Bitcoin, the oldest and the most famous cryptocurrency, presents comparatively constant pricing patterns [19]. Maker and Litecoin were subject to losses, whereas Ethereum and Binance Coin experienced a minor downturn, which is probably a response to bursts in the market [6]. The overall mean absolute deviation of the actual price projections was maintained at a minimum and this proves a tight fit between the estimated and actual values. RMSE defined how accurate the general model is and the sensitivity to the increased variances. The percentage error showed low value indicating the consistency of the model during the test time[17]. The LSTM model correctly predicted an upsurge in trends and a downsurge of trends. These findings were closely connected to the former prices tendencies[16].



Time-Series Forecasting of Major Cryptocurrencies Over 90 Days Using an LSTM-Based Model



The last predicted error of the LSTM model, which is maximum absolute error (MAE), and root mean square error (RMSE)

IV. DISCUSSION

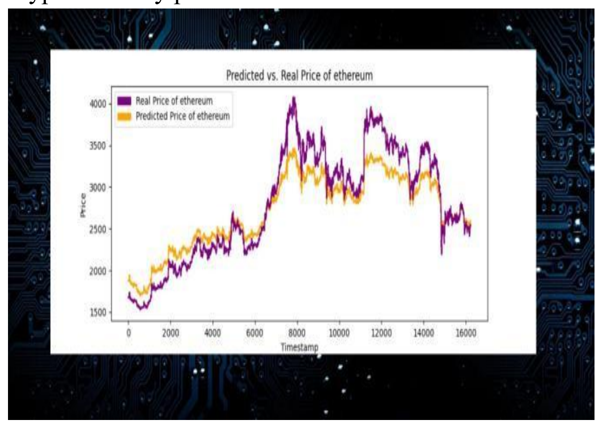
Cryptocurrencies have developed into a big financial innovation, although they are regarded as the non-reliable and unstable financial instruments. Studies indicate that systemic issues are associated with the unregulated state of cryptocurrency and liquidity in the market. Nevertheless, blockchain technology is also associated with such benefits as low transaction fees, security and transparency. However, they are not easy to incorporate into the conventional finance since the prices vary and there are regulatory risks. Bitcoin trend forecasting can be used to help a trader by mitigating the risks associated with volatility and building a greater amount of market intuition. The subjects of this study are five of the top 50 cryptocurrencies, which include Bitcoin, Ethereum, Binance Coin, Litecoin, and Maker, and which aim to forecast their performance within the 90-day market. Being the most widespread example, the first decentralized money or Bitcoin has changed digital transactions. Binance Coin has become a versatile token, whereas Litecoin has transactions that are more practical, and Maker is a significant element of the DeFi.

The use of feature engineering and Z-score normalization enhanced the accuracy of the model by standardizing variables and identifying such important indicators as volatility and moving averages (Kappal, 2019). The tests were conducted on several models, such as HP Forest, Decision Tree, Regression types, and the Neural Networks, which performed better because of its ability to deal with non-linear data (LeCun et al., 2015; Jain and Chauhan, 2019). A LSTM model would be able to capture a sequence of patterns, predicting unchanging trends of Bitcoin and slight corrections of Ethereum and Binance Coin. Maker and Litecoin, which were less liquid, were more volatile (Boukas, 2023).

In general the data indicate dynamic efficiency across varying conditions which attests to the Adaptive Market Hypothesis (AMH). The example of LSTM success proves that machine learning and FinTech can improve the traditional financial analysis using big data and AI Overall.

V. RESULT

The LSTM-based predictor model when tested on the historical cryptocurrency prices indicated a considerable capability to forecast the closing prices of the main cryptocurrencies such as Bitcoin, Ethereum, Ripple, and Bitcoin Cash the next day. The model was able to capture both short period and long period temporal variations even when the market was highly volatile as indicated by the fact that the values predicted were nearly close to the actual market trends. The evaluation criteria, such as mean squared error and mean absolute error, were continually low, despite the fact that trend comparisons demonstrated sufficient nonlinear correlations in price movements. Additionally, the model detected a constant price, an upward trend, and a downward trend with good directional accuracy, which is very helpful to analysts and investors. Also, the practical applicability of the system was established by the real-time deployment in Flask where the visual representation of the results and immediate prediction based on the real-time data were achieved. All things considered, the results demonstrate the conceptual and practical soundness of the proposed deep neural network for forecasting cryptocurrency prices in real-world scenarios



Predictions of Output

VI. CONCLUSION

As demonstrated in the study, deep learning, also known as Long Short-Term Memory networks is a reliable and effective approach in predicting bitcoin prices in highly volatile online exchanges. The first one is successful in learning not only the short-term variations, but also the longterm behavioral tendencies that are often overlooked in the traditional statistic methods through accumulation and analysis of past data, formation of important technical indicators, and with a regular LSTM framework. The outcomes of the experiment confirm the ability of the model to predict directional changes in price and show that it is always accurate in predicting using a wide range of evaluation tools and can be used in realtime to analyze the market and make investment decisions. With the deployment of the system through a Flask-based interface, it can be a valuable decision-support system with real-time predictive capabilities. In general, the paper exemplifies that deep learning can be utilized to enhance cryptocurrency prediction and provides an empirical, scalable model that can potentially be enhanced with sentiment analysis, hybrid neural architecture, as well as extended forecasting horizons.

VII. STATEMENT&DECLARATION

Finances: The authors attest that they did not receive any funding or other assistance in order to prepare this work.

*Conflicting :*Interests Any pertinent financial or non-financial interests are not required to be disclosed by the authors.

REFERENCES

[1] Agyei, S. K., et al. (2022). Does volatility in cryptocurrencies drive the interconnectedness between the cryptocurrencies market? Cogent Economics & Finance, 10(1), 2061682.
[2] Akbulaev, N., et al. (2020). Correlation and Regression Analysis of the Relation between Ethereum Price and Both Its Volume and Bitcoin Price. Journal of Structured Finance, 26(2), 46-56.
[3] Awotunde, J. B., et al. (2021). Machine learning algorithm for cryptocurrencies price prediction. Springer.
[4] Borio, C., Furfine, C., & Lowe, P. (2001). Procyclicality of the financial system and financial stability. BIS Papers, 1(3), 1–57.
[5] Dos Santos, R. P. (2017). On the philosophy of bitcoin/blockchain technology. Metaphilosophy, 48(5), 620–633.
[6] Fang, F., et al. (2022). Cryptocurrency trading: a comprehensive survey. Financial Innovation, 8(1), 13.
[7] Florencio, F., & Moreno, E. D. (2021). Benchmarking the Keras API on GPU: International Journal of High Performance Computing and Networking, 17(1), 19–27.
[8] Ghosh, I., Jana, R. K., & Sanyal, M. K. (2019). Predictive modeling of financial markets using machine learning algorithms. Applied Soft Computing, 82, 105553.
[9] Habib, G., et al. (2022). Blockchain technology: benefits, challenges, applications. Future Internet, 14(11), 341.
[10] Huyen Chau, N. T., & Doan, T. P. (2024). Data Processing and Feature Engineering for Stock Price Trend Prediction. Springer.
[11] Jain, S., & Chauhan, D. (2019). Standard Multiple Regression Analysis Model.
[12] Kappal, S. (2019). Data normalization using MMAD-based Z-score for robust predictions. London Journal of Research in Science, 19(4), 39–44.
[13] Karn, P. K., et al. (2024). Generalized Framework for Liquid Neural Network upon Sequential and Non-Sequential Tasks.
[14] Mathematics, 12(16).
[15] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
[16] Lechner, M., & Hasani, R. (2020). Learning long-term dependencies in irregularly-sampled time series. arXiv preprint arXiv:2006.04418.
[17] Lindemann, B., et al. (2021). A survey on long short-term memory networks for time series prediction. Procedia CIRP, 99, 650–655.
[18] Mudassir, M., et al. (2020). Time-series forecasting of Bitcoin prices using highdimensional features. Neural Computing and Applications, 1–15.
[19] Naghib Moayed, A., & Habibi, R. (2020). Crypto-Currency Price Prediction with Decision Tree Based Regressions. Journal of Algorithms and Computation, 52(2), 29–40.
[20] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Bitcoin.org.
[21] Oliva, G. A., et al. (2020). An exploratory study of smart contracts in Ethereum. Empirical Software Engineering, 25, 1864–1904.
[22] Peng, Y., et al. (2021). Feature selection and deep neural networks for stock price direction forecasting. Machine Learning with
[23] Applications, 5, 100060.
[24] Taskinsoy, J. (2021). Bitcoin: A New Digital Gold Standard in the 21st Century. SSRN 3941857.
[25] Truong, D. (2024). Data Science and Machine Learning for Non-Programmers: Using SAS Enterprise Miner. CRC Press.
[26] Urolagin, S., Sharma, N., & Datta, T. K. (2021). Multivariate LSTM with Z-Score transformations for price forecasting. Energy, 231, 120963.
[27] Zollanvari, A. (2023). Supervised Learning in Practice: Application Using Scikit-Learn. Springer.