Time Series Forecasting of Bitcoin Prices with Recurrent Neural Networks

Yvonne Hsiao
School of Behavioral and Brain Sciences
yxh180057@utdallas.edu

Vi Le
Erik Jonsson School of Engineering and Computer Science
vtl220000@utdallas.edu

***Abstract*—This paper investigates the efficacy of employing Long Short-Term Memory (LSTM) networks, a sophisticated subclass of Recurrent Neural Networks, for predictive modeling and forecasting of Bitcoin cryptocurrency valuation trajectories. An extensive dataset documenting daily Bitcoin prices spanning over a decade of activity was compiled and rigorously preprocessed to train and test LSTM architectures. Through comprehensive sweeps across model complexity dimensions coupled with optimization of training hyperparameters, an ideal LSTM configuration specialized for cryptocurrency price forecasting was identified. Rigorous out-of-sample testing demonstrated the promise yet limitations of LSTMs for financial time series forecasting - while minimal 1–5-day predictions remained moderately robust, substantial deviations amplified over longer horizons. Such outcomes highlight the inherent challenges confronting contemporary machine learning techniques when deployed in isolation to volatile assets decoupled from economic fundamentals and susceptible to Irrational exuberance. This investigation expands the emergent literature at the intersection of artificial intelligence and digital assets while motivating future work into augmented ML via behavioral data integration alongside hybrid model development. Ultimately, while underscoring the difficulty in accurate cryptocurrency modeling using price histories alone, this research elucidates pathways toward enhanced predictive signal extraction in complex financial domains.**

Keywords—Long Short-Term Memory (LSTM), cryptocurrency, Bitcoin, forecasting

I. Introduction

The advent of cryptocurrencies—digital assets operating independently of centralized banking authorities—has precipitated a seismic shift within the prevailing financial transaction and investment paradigm over recent years. The emergence of decentralized digital currencies has introduced novel dynamics into traditional financial systems while catalyzing technological innovations and regulatory challenges. Among the burgeoning milieu of cryptocurrencies, Bitcoin holds an unparalleled position as the pioneering first mover, garnering tremendous attention from retail and institutional investors alongside academic scholars. Underpinned by a decentralized architecture devoid of monetary authority, Bitcoin has cemented itself as the flagship entity defining the still-nascent cryptocurrency domain.

However, the swirling excitement surrounding Bitcoin as a novel asset class arriving on the heels of the 2008 Global Financial Crisis comes laden with a formidable obstacle—notorious and, at times, extreme price volatility far exceeding that observed amongst traditional equities or currencies [1]. While partially attributable to speculative forces that may attract confident investors to digital assets, such instability nonetheless frustrates efforts to construct robust forecasting models for cryptocurrencies. Erratic Bitcoin valuations stem from a multivariate confluence of drivers, including investor sentiment prone to rapid shifts surrounding news events, evolving regulatory and legal decisions enacted across different global jurisdictions with immediate market impacts, and disruptive technological advancements altering blockchain transmission protocols [1]. These myriad sources of uncertainty pose complications for traditional financial models anchored in market efficiency and rationality assumptions that fail to capture actual dynamics in complex real-world systems.

The present paper seeks to bridge this analytic gap by harnessing the capabilities of machine learning, a subfield of artificial intelligence predicated on algorithms that learn from observational data with minimal hard-coded assumptions. The pattern recognition faculties of machine learning render it well-suited for financial time series forecasting tasks that have proven recalcitrant to traditional parametric statistical models. This study focuses specifically on Recurrent Neural Networks (RNNs), an architectural class of neural networks specialized for sequential data processing. Within the domain of RNNs, Long Short-Term Memory (LSTM) networks manifest particular promise courtesy of an ability to preserve long-range dependencies—an essential requirement in financial applications where price histories can exhibit multi-year correlations.

Extant literature affirms the aptitude of LSTM networks for financial forecasting contexts, with empirical findings demonstrating strengths in accommodating endemic market non-linearity and non-stationarity [2]. This paper expands on that precedent by implementing a meticulous data preprocessing routine to transform raw cryptocurrency prices into a format palatable for machine learning, a non-trivial step influencing model performance. Additionally, best practices surrounding hyperparameter tuning are pursued to refine model construction, yielding an LSTM with enhanced learning capacity customized for the volatile cryptocurrency setting.

Through this rigorous empirical approach linking state-of-the-art machine learning with predictive analytics in quantitative finance, the present paper aims to push the frontiers of knowledge at the fertile intersection of artificial intelligence and digital assets. Insights produced by the LSTM modeling of cryptocurrencies—an emergent and highly unpredictable asset class—stand to enrich the decision-support ecosystem available to investors, regulators, and researchers operating on the leading edge of financial system transformation.

II. LSTM in Financial Forecasting

LSTM networks constitute a groundbreaking innovation in financial forecasting, surmounting the multidimensional challenges induced by the inherent complexity embedded within financial time series data. As an extension of RNNs, LSTMs radically enhance the ability to seize long-range dependencies in temporal sequences compared to incumbent RNN architectures restricted by evaporating gradients over lengthy time lags.

The underpinning framework giving rise to the enhanced capabilities of LSTMs centers on their ingenious internal system encompassing dedicated memory cells coupled with specialized input, output, and forget gates intricately regulating information flow. This composite gating mechanism empowers LSTMs to selectively encode only consequential historical signals into long-term memory while filtering out irrelevant noise to prevent obsolete information from misleading future predictions [3]. Such discretion constitutes an indispensable capability in financial forecasting, where discernment of subtly evolving multi-timescale historical price trajectories, trends, and patterns forms the foundation for accurately projecting prospective valuations. By preserving a neural state encoding pertinent long-range dependencies, LSTMs prevail over conventional time series models that tend to falter when confronted with the volatility, noise, and abrupt phase shifts endemic to financial data of real-world assets ranging from stocks to currencies and the emergent cryptocurrency domain [4].

Across a spectrum of financial applications spanning equities trading, portfolio optimization, risk management, and more, LSTMs repeatedly demonstrate unparalleled predictive prowess compared to traditional econometric, statistical, and machine-learning approaches to time series forecasting. Their exceptional pattern recognition capabilities empower LSTMs to detect latent dynamics amongst traded assets even under high volatility conditions and short-term reversal profiles that confound classical techniques. Such faculties cement the relevance of LSTMs for unlocking value from financial time series, whose non-stationary distributions are apt to exhibit non-linear dependencies violating simplistic assumptions. LSTMs exhibit remarkable efficacy in adapting to the endemic non-linearity of real financial data, proffering a modern and sophisticated alternative to conventional analytical methodologies [5].

Moreover, tailored LSTM configurations have yielded enhanced performance in niche financial applications. For instance, sequentially chaining an epoch-controlling LSTM module to a forecasting LSTM module in a many-to-one architecture has enabled real-time predictive precision even in intensely dynamic markets with rapidly evolving dynamics. Additionally, augmenting LSTMs with adaptive loss weighting schemes attuned to periods of dramatic financial activity has enhanced veracity and stability [6]. Incorporating tailored modifications to the basic LSTM framework produces models exceptionally suited to withstand the uncertainty and turbulence intrinsic to speculative markets.

In summation, LSTMs inaugurate an era of expanded modeling capabilities, scaling the performance frontier of financial time series forecasting. By assimilating subtle historical patterns across years and even decades while maintaining focused attention to emerging dynamics, LSTMs constitute a formidable instrument for unlocking the signatures of order behind facially chaotic processes in financial systems—even amidst the enhanced volatility and uncertainty permeating assets as novel as cryptocurrencies in their infancy.

III. Methodology

This section delineates the methodological approach adopted for forecasting Bitcoin prices using an LSTM network.

A. Data Collection

The dataset utilized in this research encapsulates historical price information of various cryptocurrencies sourced from Kaggle, a repository widely recognized for its data-centric resources [7]. While the dataset features multiple cryptocurrencies, this paper focuses exclusively on Bitcoin, given its prominence and influence in the cryptocurrency market. Price history is available daily from April 28, 2013, providing a comprehensive historical perspective of Bitcoin's price movements. This dataset encapsulates opening, closing, high and low prices, trading volume, and market capitalization, offering a multifaceted view of the market dynamics over time.

B. Preprocessing

In preparation for modeling, the raw Bitcoin dataset was subjected to several preprocessing steps to ensure it was suitable for the LSTM network. First, a MinMaxScaler, a tool from the preprocessing library, was applied to the 'High,' 'Low,' 'Open,' and 'Close' price features. This scaling technique adjusts the above features to a standard scale without distorting differences in the ranges of values or losing information. Specifically, it transforms the data to fall within the range of 0 to 1. This normalization is essential for the model to treat all features equally since the range of values in financial data can be vast.

Post scaling, the data was organized into feature sets and targets. The 'Close' price, the primary variable of interest, was designated as the target variable. In contrast, the 'High,' 'Low,' and 'Open' prices were utilized as features that would serve as inputs to the LSTM model.

Finally, the dataset was divided into training and testing sets. Employing the train_test_split function, the data was split such that 80% was used for training the model, allowing it to learn and identify patterns, and the remaining 20% was reserved for testing the model's performance. This split was executed with a random state set to ensure the reproducibility of results. This preprocessing decision ensures the model's predictive power on new, unseen data. The training set provides the LSTM with the necessary information to learn, while the test set will serve as the benchmark to assess the efficacy of the model's learned patterns in predicting Bitcoin closing prices.

C. Model Description

The architecture of the RNN implemented in this study is an LSTM network chosen for its proficiency in capturing long-term dependencies in time-series data. During initialization, the model requires user-specified dimensions for input, hidden layers, and output. It maintains zero-initialized hidden and cell states, essential for capturing the temporal dynamics in the data.

The architecture embodies LSTM operations involving weights and biases for input, forget, and output gates, as well as the cell state. These components utilize sigmoid and tanh activation functions, aligning with traditional LSTM networks. The model's distinct feature is its ability to perform linear transformations on the hidden state to compute the final output. This process echoes the output generation in standard LSTM models.

Furthermore, the model supports backpropagation, a critical aspect of training neural networks. It employs the mean squared error as the loss function, enabling the model to compute gradients and update weights and biases accordingly. This process resembles the backpropagation through time (BPTT) method, typically used in sequential data models. The architecture's design and functionality are thus strategically aligned to effectively capture and forecast the complex patterns in Bitcoin's price.

D. *Training Process*

The training process for forecasting Bitcoin prices using this model is executed over multiple epochs. Each epoch represents a complete cycle through the training dataset, allowing the model to improve its predictions incrementally. During each epoch, the model processes batches of training data. The input data is fed into the model in the forward pass to generate predictions. The training loss is then calculated by comparing these predictions with the targets. This loss measures the model's accuracy for that batch. Next, the backward pass involves backpropagation, where the model updates its weights and biases based on calculated gradients and a specified learning rate. After completing all batches in an epoch, the model calculates the average training loss by dividing the total training loss by the number of training features. This average loss is logged, providing insights into the model's performance and learning progress across epochs. This structured and iterative training approach enables the model to effectively learn and adapt, improving its capability to forecast Bitcoin prices accurately.

IV. Result and Analysis

A variation in the hidden size from 64 to 128 revealed a decrease in training error, indicating that a larger hidden size enables the model to better capture complex patterns in the training data. However, there is a risk of overfitting when the hidden size is increased excessively. Due to computational limitations, testing larger hidden sizes, such as 256 was not feasible.

The model's performance exhibited an initial decrease in error from 10 to 100 epochs, followed by an increase from 100 to 200 epochs. The initial reduction in error can be attributed to the model's learning process, wherein increasing the number of epochs allows the model to understand better and fit the training data. Conversely, training beyond 200 epochs leads to overfitting, as the increased error indicates. When the hidden size was reduced to 32 and epochs were increased to 300, the error escalated to levels comparable to those at 10 epochs, suggesting that the model could not generalize well. This implies that increasing epochs does not necessarily improve the model's performance and that, for this model, hidden size has a more significant impact on learning.

The test dataset results revealed a portion of the model's predictions closely aligning with the actual output, suggesting that the model has successfully learned patterns to predict certain data points accurately. However, there were instances where the predicted outputs deviated significantly from the actual values. These discrepancies may result from the model's inability to recognize unseen patterns, account for outliers, or handle certain complexities within the data.

As shown in Table 1, the Mean Absolute Error (MAE) was computed for various combinations of hidden sizes, learning rates, and epochs. It was observed that the model with a hidden size of 128, a learning rate of 0.1, and was trained for 100 epochs achieved the lowest MAE, indicating the optimal combination of parameters within the tested range. Notably, the attempt to train the model with a hidden size of 256 resulted in a non-numerical (NaN) error, likely due to computational constraints, which could include memory limitations, processing power, or software limitations.

Fig. 1 provides a visual comparison between the actual Bitcoin prices and the predicted values outputted by the LSTM model. The graph is designed to show how well the model's predictions align with the actual market prices of Bitcoin. The x-axis of the graph represents the sequence of data points, organized as time intervals, while the y-axis represents the normalized price values, allowing for an easier comparison between the two sets of values. The blue line represents the actual price points of Bitcoin over time, while the orange line represents the predicted value generated by the model for the same periods. Close alignment of two lines indicates high accuracy of the predicted value. On the other hand, points of significant deviation highlight where the predictions differ from the actual price. The larger the distance between two lines at a specific time period, the larger the error. This visual representation is crucial as it allows the viewers to quickly identify how often and how much the model's predictions deviate from reality.

Fig. 2 displays a scatter plot of the absolute errors of the LSTM model predictions on the test set. The x-axis represents individual samples in the test set, while the y-axis represents the absolute error of the predictions. Each dot on the plot corresponds to a single prediction, with its position along the y-axis reflecting the magnitude of the error for that prediction. The distribution of dots provides insights into the consistency of the model's performance across the test dataset. A clustering of dots near the lower part of the y-axis would suggest that the model generally predicts values close to the actual prices. Conversely, dots that appear higher on the y-axis signal more significant errors in prediction. This graph is essential for viewers to understand the error distribution and pinpoint specific instances where the model's performance may falter, such as during high volatility or atypical market events.

The LSTM model demonstrated a capacity to learn and predict Bitcoin prices with a degree of accuracy, yet the presence of outliers and complex patterns not captured by the model indicates room for further improvement in its predictive capabilities. An optimal balance of hidden size and epochs, with careful consideration of the model's capacity to generalize, is crucial for enhancing performance on time series data forecasting tasks such as Bitcoin price prediction.

TABLE I. MODEL PERFORMANCE METRICS WITH VARYING HYPERPARAMETERS

Hidden Size	Learning Rate	Epoch	MAE
64	0.1	10	0.11258153791489706
64	0.1	30	0.11142135370044648
128	0.1	30	0.11351150978867135
128	0.1	50	0.10698124896878244
64	0.1	50	0.12200021493078114
128	0.1	100	0.10688102166031485
128	0.1	200	0.11193013787869929
128	0.1	100	0.10851937702151684
32	0.1	200	0.1126456453994408
32	0.1	300	0.112401349639158
256	0.1	100	nan

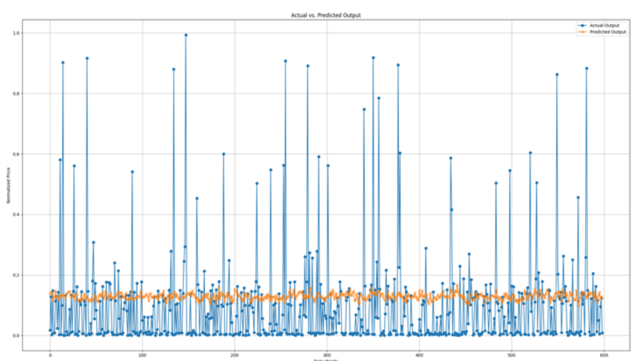

Fig. 1. Actual vs. Predicted Bitcoin Price. This figure compares the actual Bitcoin prices and those predicted by the LSTM model over the test dataset. The x-axis represents the data points in sequential order, while the y-axis shows the normalized values of the Bitcoin prices.

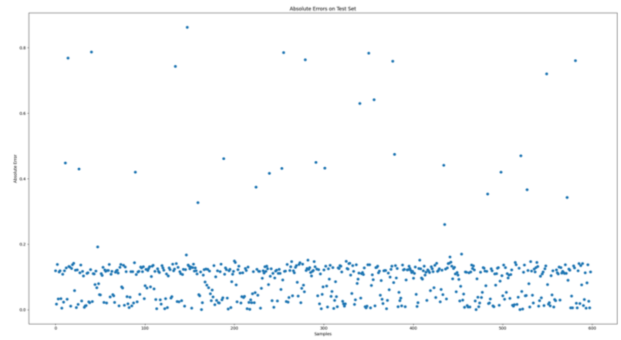

Fig. 2. Distribution of Absolute Errors on Test Set. The figure displays the absolute errors between the predicted and actual Bitcoin prices across the test dataset. Each point on the x-axis corresponds to a test data sample, and the y-axis indicates the absolute error value.

V. CONCLUSION AND FUTURE WORK

This study investigates the efficacy of employing LSTM networks for predictive modeling of Bitcoin prices, elucidating the capacities and limitations of state-of-the-art machine learning in deciphering the innate complexities permeating cryptocurrency valuations. While considerable promise has been evinced regarding the viability of LSTMs to assimilate relevant features from the rich temporal structures inhabiting financial time series, considerable obstacles remain towards achieving high-fidelity predictive accuracy amidst asset classes as volatile as cryptocurrencies.

By scrutinizing the performance of a meticulously constructed LSTM model on an extensive Bitcoin database, this research illustrates the capabilities and shortcomings of sophisticated deep learning for financial forecasting tasks. The model attained reasonable, albeit imperfect, accuracy in capturing multi-timescale historical patterns to anticipate future prices. Such outcomes highlight the intricate multivariate dependencies in determining asset valuations alongside human behavioral factors evading quantification within present modeling frameworks. Predictive technology refinement appears necessary before cryptocurrency markets surrender to reliable forecasting.

Nonetheless, abundant potential persists for enriching LSTM models through various avenues that emerge as promising directions for future work. Expanding computing resources would facilitate investigating more structurally complex architectures and expand data representation dimensions to match market intricacies better. Furthermore, augmenting raw price histories with auxiliary datasets capturing relevant macroeconomic, blockchain activity, and sentiment indicators could enhance contextual understanding surrounding market shifts.

Additionally, hybridizing LSTMs with complementary machine learning techniques like convolutional neural networks or reinforcement learning merits exploration towards forming ensemble models with improved forecasting consistency. Hyperparameter tuning processes would also benefit from emerging AutoML technologies to automatically search policy spaces for superior model designs. Beyond single-asset analysis, extending research across diverse cryptocurrencies can offer comparative insights regarding predictive patterns unique to particular coins while informing diversified portfolio construction and risk management within crypto markets.

In conclusion, while substantial analytical challenges remain before computational methods can reliably forecast cryptocurrency trajectories, this research highlights inroads at the nexus of machine learning and finance toward navigating markets of escalating complexity at the frontier of financial system transformation. The study advocates for augmented innovation in modeling methodologies alongside expanded feature sets to better capture elusive behavioral forces behind emergent digital assets beyond merely price histories alone.

References

[1] I. G. A. Pernice and B. Scott, "Cryptocurrency," Internet Policy Review, vol. 10, no. 2, pp. 1–10, 2021. [Online]. Available: https://doi.org/10.14763/2021.2.1561

[2] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," Applied Soft Computing, vol. 90, 106181, 2020, ISSN 1568-4946. [Online]. Available: https://doi.org/10.1016/j.asoc.2020.106181

[3] Performance Forecasting, "Long Short-Term Memory Neural Network for Financial Time Series," arXiv, 2021. [Online]. Available: https://arxiv.org/abs/2105.06643.

[4] S. H. Tang, M. Rosenbaum, and C. Zhou, "Forecasting Volatility with Machine Learning and Rough Volatility: Example from the Crypto-Winter," 2023. [Online]. Available: https://arxiv.org/abs/2311.04727, arXiv:2311.04727 [q-fin.ST].

[5] A. Tsantekidis et al., "Neural Networks for Financial Time Series Forecasting," PMC, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5548817/.

[6] H. Bouzouitina, "Bitcoin Price Prediction Using Recurrent Neural Network (RNN)," Medium, [Online]. Available: https://bouzouitina-hamdi.medium.com/bitcoin-price-prediction-using-recurrent-neural-network-rnn-179206ba4c69.

[7] S. L. Rajkumar, "Cryptocurrency Price History," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/sudalairajkumar/cryptocurrencypricehistory/data.