Comparative Analysis of LSTM, GRU and Transformer Deep Learning Models for Cryptocurrency ZEC Price Prediction Performance

Jiakun Lian

Northwestern Polytechnical University, 710129 Dongxiang Road, Chang'an District, Xi'an, Shaanxi, China

lianjiakun@mail.nwpu.edu.cn

Abstract. This paper delves into the intriguing realm of cryptocurrency price prediction, with a specific focus on Zcash (ZEC), employing a cutting-edge deep learning approach. The study introduces two crucial features, "close_off_high" and "volatility", then systematically analyzes the correlations between these variables and the price of ZEC. By investigating the predictive accuracy of three prominent neural network architectures-Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and the Transformer model-the study discerns that LSTM and GRU models outperform the others in forecasting ZEC's price movements. Furthermore, the paper scrutinizes the influence of different activation functions on model performance, shedding light on the effectiveness of the linear activation function in this context. The research also addresses common challenges in predictive modeling, such as overfitting and multicollinearity. Moreover, it candidly acknowledges the limitations associated with solely focusing on a single cryptocurrency, recognizing that broader research efforts and interdisciplinary collaboration are required for a more comprehensive understanding of the ever-evolving cryptocurrency landscape. As the cryptocurrency market continues to evolve rapidly, this study provides invaluable insights for investors, offering a rational perspective on cryptocurrency investment. It underscores the importance of utilizing appropriate models and embracing interdisciplinary cooperation to navigate the complex and dynamic world of cryptocurrency. By bridging the gap between the cutting-edge world of deep learning and the financial market, this research paves the way for enhanced future investigations and more informed investment decisions.

Keywords: ZEC Price Forecast, Deep Learning Models, Neural Network.

1 Introduction

In the realm of electronic payments, two prevalent methods stand out: credit cards and cash. Credit card transactions typically involve a complex web of intermediaries, including banks and credit card companies. This method, while convenient, often lacks direct communication between buyers and sellers, relying on various institutions

K. Elbagory et al. (eds.), *Proceedings of the 9th International Conference on Financial Innovation and Economic Development (ICFIED 2024)*, Advances in Economics, Business and Management Research 281,
https://doi.org/10.2991/978-94-6463-408-2_45

to facilitate transactions. While credit card usage can shield personal information from the other party, it inevitably results in intermediaries collecting customer data. For instance, platforms like PayPal act as intermediaries between buyers and sellers, enhancing security and privacy but adding an extra layer of complexity [1]. In stark contrast, cash payments offer a higher level of anonymity. Since credit cards are issued in an individual's name, banks can meticulously track spending. However, when cash is the medium of exchange, banks are not involved, and the other party does not require knowledge of the payer's identity. Furthermore, cash transactions allow for offline exchanges, eliminating the need for immediate third-party approval [1].

Despite these advantages, cash payments confront a crucial challenge: preventing double spending. This issue was ingeniously tackled by David Chaum, in collaboration with cryptographers Amos Fiat and Moni Naor [1], through the concept of offline electronic cash. Their solution not only paved the way for cryptocurrencies to function as a form of cash payment but also ensured the security of electronic transactions. Cryptocurrency, in contrast to traditional financial systems, operates as an electronic payment system based on cryptographic proof rather than trust. Its emergence has effectively addressed inherent issues in conventional financial payment systems, which rely on third-party financial institutions as trusted intermediaries, resulting in increased transaction costs and risks [2].

However, cryptocurrencies, initially designed to facilitate direct transactions using blockchain technology, seem to exhibit fewer characteristics of traditional currencies. Using Bitcoin as an example, it falls short in fulfilling the fundamental functions of a currency, such as serving as a medium of exchange, a store of value, and a unit of account. Its low consumer transaction volume, high price volatility, lack of correlation with traditional currencies and gold, inconvenient price representation, and security risks make it resemble more of a speculative investment than a true currency [3]. This speculative nature of cryptocurrencies can give rise to a variety of socio-economic issues. Firstly, the extreme price volatility of cryptocurrencies creates financial instability for investors and holders, as rapid price fluctuations make it challenging to predict market trends, exposing them to significant risks. Secondly, cryptocurrency speculation may lead to a speculative bubble, attracting large capital inflows driven by the pursuit of quick profits rather than long-term investment goals. When the bubble eventually bursts, investors may suffer substantial losses, potentially causing adverse effects on the overall economy. Moreover, the significant price fluctuations in the cryptocurrency market may have repercussions on the financial system. Additionally, it may foster a certain degree of libertarianism [4, 5].

In order to address these issues, researchers have increasingly focused on cryptocurrency price prediction. In the early exploration phase (2009-2013), as Bitcoin emerged in 2009, initial research primarily concentrated on the technical and security aspects of Bitcoin. There was relatively less research on price prediction because the market was not mature enough, and there was insufficient data for in-depth analysis [2]. Entering the phase of technical analysis (2014-2017), as the price of Bitcoin steadily increased, researchers began to employ technical analysis methods such as chart patterns and trendlines to attempt price trend predictions. During this period, research heavily relied on historical price data and market indicators [6].

With the development of machine learning and artificial intelligence technologies, the era of machine learning (2018-present) was ushered in. Researchers started using machine learning algorithms, such as neural networks and decision trees, to analyze more data and variables to enhance prediction accuracy [7-9]. Recent research trends have focused on combining social media and news sentiment analysis with price prediction, entering the phase of social media sentiment analysis (2019-present). Researchers have recognized the influence of social media and news events on market sentiment and have begun using sentiment analysis to better understand market dynamics [10].

Entering the phase of blockchain data-based research (2020-present), as blockchain technology matures, researchers have started exploring how to use blockchain data for price prediction. This data includes transaction volume, transaction history, and market depth, among other information. Over time, cryptocurrency price prediction research has evolved, providing investors with more valuable tools and insights. Zcash's utilization of the zk-SNARK method significantly enhances the privacy of cryptocurrency transactions. The author has a keen interest in such technology. Therefore, the author chooses this as the research background and plan to employ three distinct models, namely LSTM, GRU, and Transformer, for cryptocurrency price prediction. The author will compare their prediction errors and performance differences. This research aims to provide investors with a more rational perspective on cryptocurrencies while deepening my own understanding of deep learning.

2 Methodology

2.1 Data Source

The cryptocurrency market data used in this study was sourced from tokeninsight.com. The dataset covers the period from October 29, 2016, to October 10, 2023. The data provides information on Zcash (ZEC) cryptocurrency, including historical price and volume data. The author processed the original data through feature engineering as follows. Firstly, as the Table 1 shows that the 'Date' column was converted into a datetime format to better represent time series information. Secondly, missing values in the 'Volume' column were replaced with zeros to ensure data integrity and availability. Subsequently, the 'Volume' data was converted into an integer data type for numerical calculations and analysis. Next, new features were created, including 'close_off_high' (the difference between the closing price and the highest price on the same day) and 'volatility' (the ratio of 'close_off_high' to the closing price of the previous day). The introduction of these new features helps capture a more comprehensive view of changes and fluctuations in cryptocurrency price data, providing additional insights into market trends and price volatility. The author then selected a subset of columns from the processed data, including date, price, trading volume, 'close_off_high,' and 'volatility.' These columns are essential features required for building deep learning models for cryptocurrency price prediction. Lastly, the dataset was sorted by date to facilitate further analysis and modeling. Through this series of preprocessing steps, the data was prepared for use in deep learning models for the prediction of cryptocurrency prices (Table 1).

Table 1. Table about data processing.

Preprocessing Step	Description
Data Conversion	'Date' column converted into datetime
Feature Engineering	'close_off_high' and 'volatility'
Feature Selection	'Date,"Price,"TradingVolume,"close_off_high,'and 'volatility'
Data Sorting	Dataset sorted by date

2.2 Methods Introduction

In this study, the authors adopted a systematic approach to employ different deep learning models for cryptocurrency price prediction. Let's break down the details into manageable paragraphs for better comprehension.

The authors began by implementing a Long Short-Term Memory (LSTM) model, a recurrent neural network (RNN) suitable for time series data. For this LSTM model, the following parameters were set. Firstly, the input data had a shape of (1515), which is ideal for processing temporal information in time series data. Secondly, 20 neurons were used in the model, representing a balanced choice between complexity and simplicity to avoid both underfitting and overfitting. A linear activation function was chosen, as a nonlinear activation function like sigmoid might negatively affect predictive accuracy in this context. A dropout rate of 0.25 was set to prevent overfitting. The Mean Absolute Error (MAE) was selected as the loss function, suitable for regression problems. The model underwent training for 50 epochs with a batch size of 1.

The authors then explored another deep learning model, specifically a Gated Recurrent Unit (GRU) model. GRU is similar to LSTM and designed for processing time series data but offers advantages like a simpler internal structure with reduced computational complexity and shorter information propagation paths. The model's parameters were set similarly to the LSTM model, including input data shape, the number of neurons, activation function, dropout rate, loss function, and optimizer. The model underwent training for 50 epochs with a batch size of 1.

In the next phase, the authors introduced a Transformer model for cryptocurrency price prediction. The Transformer model departs from traditional RNNs like LSTM and GRU by employing self-attention mechanisms and feedforward neural networks, eliminating the need for explicit time steps. This architecture enhances parallelism and the ability to process entire sequences simultaneously. The model was constructed with multiple self-attention layers (MultiHeadAttention) and feedforward neural network layers. Key parameters were specified, such as the number of attention heads, embedding dimension, dimension of the feedforward network layers, and the number of layers. Similar to the previous models, MAE was chosen as the loss function, and Adam as the optimizer. The model underwent 50 epochs of training, and the author recorded its performance.

These different models, LSTM, GRU, and Transformer, offer distinct advantages and approaches to cryptocurrency price prediction, allowing for a comprehensive evaluation of their performance.

3 Results and Discussion

3.1 Correlation Analysis

In this comprehensive study, the author embarked on a data exploration journey, enriching the dataset with two pivotal features: "close_off_high," which serves as a measure of price trends, and "volatility," offering insights into price fluctuations. This strategic data augmentation was followed by a meticulous correlation analysis, leading to a series of intriguing findings and thought-provoking conclusions.

Table 2 reveals that the correlation coefficient between ZEC_Price and ZEC_Volume is approximately zero. This finding suggests a relatively weak linear relationship between these two variables. In simpler terms, it indicates that there is minimal linear interdependence between cryptocurrency price and trading volume. Moving forward, we encountered a more compelling discovery: the correlation coefficient between ZEC_Price and Market_Cap stands at 0.397950. This figure suggests a noteworthy positive correlation between price and market capitalization. This positive association indicates that as the price of the cryptocurrency rises, so does its market capitalization. This correlation is a testament to the inherent connection between the two aspects of the cryptocurrency ecosystem.

The exploration didn't stop there. It was observed that ZEC_Price and ZEC_close_off_high exhibited a correlation coefficient of 0.459493. This value signifies a positive trend between the cryptocurrency's price and the extent to which the closing price deviates from the highest price attained during the trading period. In essence, this points to a pattern where, when the cryptocurrency's price is relatively closer to the highest point of the trading period, it exhibits a particular positive relationship.

Another remarkable discovery was the correlation coefficient between ZEC_Price and ZEC_volatility, which was measured at 0.443765. This correlation highlights a positive connection between the cryptocurrency's price and its inherent price volatility. In simpler terms, as the cryptocurrency's price exhibits higher levels of volatility, it is also likely to experience price increases. This underscores the dynamic nature of cryptocurrency markets.

Table 2. Correlation results

Variable Pair	Correlation	Coefficient Interpretation
ZEC_Price vs. ZEC_Volume	Close to 0	Weak linear relationship
ZEC_Price vs. Market_Cap	0.397950	positive correlation
ZEC_Price vs. ZEC_close_off_high	0.459493	Positive trend
ZEC_Price vs. ZEC_volatility	0.443765	Positive connection

3.2 Discussion about Models

The study went a step further by implementing predictive models, specifically using the LSTM architecture, with an exploration of different activation functions. Two activation functions were examined: the linear activation function and the sigmoid activation function.

When employing the linear activation function, the model exhibited an intriguing pattern in Figure 1. As the number of training iterations, or epochs, increased, the Mean Absolute Error (MAE) gradually decreased. Notably, there was a significant drop in MAE during the initial 10 epochs, followed by a gradual reduction that eventually stabilized at an impressive 0.0126 according to the Figure 2. This implies that the linear activation function enabled the model to rapidly grasp the underlying data patterns. With an extended training duration of 100 epochs, this function continued to outperform its counterpart.

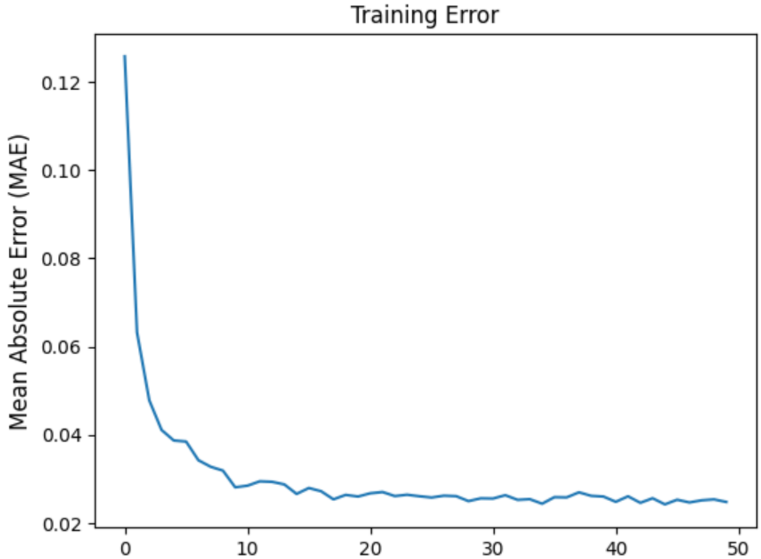

Fig. 1. The training error of LSTM (linear) (Photo/Picture credit: Original).

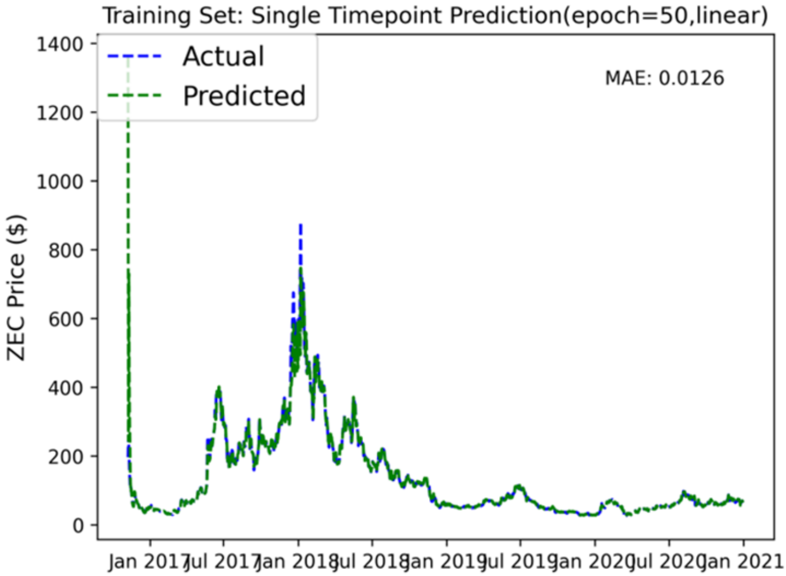

Fig. 2. LSTM (linear) (Photo/Picture credit: Original).

In contrast, the sigmoid activation function exhibited a similar trend in the early stages of training, with a rapid reduction in MAE. However, as the training process continued, the rate of decrease in MAE slowed down, and the final MAE settled at a higher value of 0.0874 as the Figure 3 shows.

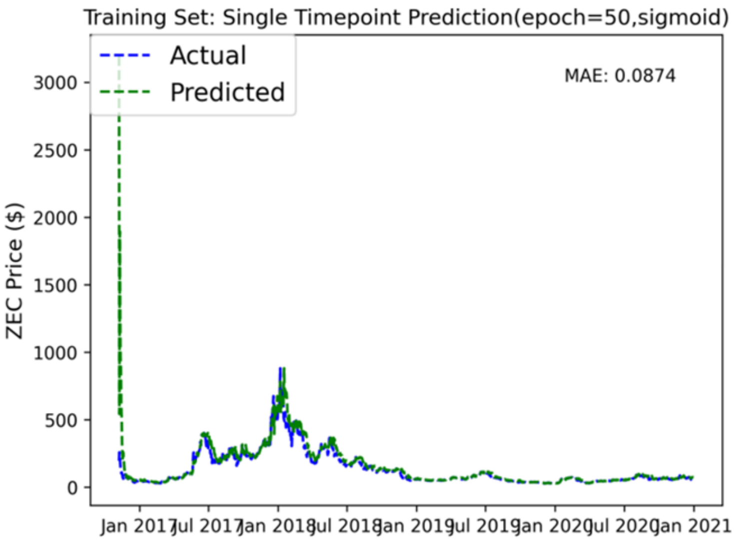

Fig. 3. LSTM (sigmoid) (Photo/Picture credit: Original).

This discrepancy indicates that the sigmoid activation function was less effective in modeling and predicting the data, especially when the training duration was extended to 100 epochs. It suggests that the sigmoid activation function might not be suitable for this specific time series data prediction task or may require further training iterations to capture the data's complexity fully. Comparing Model Performances: The study did not conclude with activation functions. It expanded into a comprehensive comparison of three different models: LSTM, GRU, and Transformer, each with 50 training iterations.

The LSTM model exhibited an MAE of 0.0126, showcasing its robust performance in fitting and predicting time series data. Its ability to accurately capture the underlying data patterns made it a standout choice. The GRU model, on the other hand, achieved an even lower MAE of 0.0096, indicating superior performance with the same number of training iterations. This suggests that the GRU architecture excelled in capturing intricate time series data patterns efficiently. In contrast, the Transformer model showed a higher MAE of 0.0269, indicating a less effective fit to the data. This implies that the Transformer architecture struggled to adequately capture the data's intricate patterns, in order to make it easy to understand, the author use the Table 3 to illustrate it (Table 3).

Table 3. MAE of the 3 models.

Model	MAE	Remarks
LSTM (linear)	0.0126	Superior performance, effective pattern recognition.
LSTM (sigmoid)	0.0874	The worst performance around the 4
GRU	0.0096	Excellent performance, efficient pattern capture.
Transformer	0.0269	Lower performance, struggled with pattern recognition

3.3 Limitations and Prospect

One of the primary limitations of this study is the potential for overfitting. Overfitting occurs when a model learns the training data too well, capturing noise and random fluctuations rather than genuine underlying patterns. To address this issue, it is crucial to explore regularization techniques, such as L1 and L2 regularization, dropout layers, and early stopping. Implementing these strategies can help prevent overfitting and improve the model's generalizability. Another challenge that warrants attention is the presence of multicollinearity within the dataset. Multicollinearity occurs when independent variables in a regression model are highly correlated with each other. This can lead to unstable coefficient estimates and difficulties in interpreting the model's predictive power. Future research should delve into feature selection methods, such as Principal Component Analysis (PCA) or Recursive Feature Elimination (RFE), to mitigate multicollinearity and enhance the model's robustness. Moreover, the current study primarily focused on a single cryptocurrency, Zcash (ZEC). While the findings provide valuable insights into ZEC price prediction, the cryptocurrency market is highly diverse and dynamic. Future research should aim to expand the scope to include multiple cryptocurrencies, thus providing a more comprehensive analysis of the broader market trends. This broader perspective could unveil unique dynamics, correlations, and trends that are specific to different cryptocurrencies.

Enhanced Regularization Techniques is the first. Future research should delve into more advanced regularization techniques to combat overfitting effectively. This could include Bayesian regularization, which allows for the incorporation of prior knowledge, or the exploration of newer regularization methods developed in the field of deep learning. And Multimodal Data Integration also matters**.** Integrating multiple sources of data, such as social media sentiment, trading volumes, and macroeconomic indicators, can significantly enhance the accuracy of cryptocurrency price predictions. Researchers should explore the integration of diverse data sources and consider advanced data fusion methods. The next is Interdisciplinary Collaboration. Collaboration between machine learning experts, economists, and cryptocurrency specialists can bring a holistic perspective to cryptocurrency price prediction. Such interdisciplinary teams can leverage domain-specific knowledge to enhance models and provide more accurate insights. Market Sentiment Analysis is also a promising way**.** Sentiment analysis of social media and news sources can play a vital role in predicting cryptocurrency price movements. Future research should explore advanced natural language processing techniques and sentiment analysis tools to incorporate sentiment data into predictive models. The next is Blockchain Data Analysis**.** Blockchain technology offers a wealth of data that can be harnessed for predictive modeling. Researchers can explore the integration of on-chain data, such as transaction volumes and network activity, to enhance predictive models.

In conclusion, while this study has provided valuable insights into cryptocurrency price prediction, it is essential to acknowledge its limitations and look ahead to future research opportunities. The dynamic and rapidly evolving nature of the cryptocurrency market offers a wealth of possibilities for further investigation. By addressing the identified limitations and embracing interdisciplinary collaboration, the field of cryptocurrency price prediction can continue to evolve and provide increasingly accurate and valuable insights for investors and stakeholders.

4 Conclusion

In a nutshell, the study sheds light on the remarkable performance of the RNN architecture, particularly LSTM and GRU, in the realm of time series data prediction. This was in stark contrast to the Transformer model with self-attention, which exhibited inferior predictive capabilities. The relatively minor difference in prediction accuracy among these models suggests that machine learning models can indeed serve as valuable theoretical references for cryptocurrency investments. These models have the potential to provide insights into the complex dynamics of cryptocurrency markets and guide investment decisions. As the cryptocurrency landscape continues to evolve, the application of machine learning in this domain holds promise for more accurate and informed investment strategies. This study underscores the importance of choosing the right tools and models to navigate the intricate world of cryptocurrency investments.

References

1. Narayanan, A., Bonneau, J., Felten, E., Miller, A., Goldfeder, S.: Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press, (2016).
2. Yermack, D.: Is Bitcoin a real currency? An economic appraisal. Handbook of digital currency, 31-43 (2015).
3. Golumbia, D.: The politics of Bitcoin: Software as right-wing extremism. Minnesota, (2016).
4. Wu, C. H., Lu, C. C., Ma, Y. F., Lu, R. S.: A new forecasting framework for bitcoin price with LSTM. In 2018 IEEE International Conference on Data Mining Workshops,168-175 (2018).
5. Chen, Z., Li, C., Sun, W.: Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics 365, 112395 (2020).
6. Patel, M. M., Tanwar, S., Gupta, R., Kumar, N.: A deep learning-based cryptocurrency price prediction scheme for financial institutions. Journal of information security and applications 55, 102583 (2020).
7. Khedr, A. M., Arif, I., El-Bannany, M., Alhashmi, S. M., Sreedharan, M.: Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. Intelligent Systems in Accounting, Finance and Management 28(1), 3-34 (2021).
8. Chen, Z., Li, C., Sun, W.: Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 365 (2020).
9. Khurshid, A. R.: Cryptocurrency price prediction using sentiment analysis. In Proceedings of conference. Washington, DC, USA, (2017).
10. Morgia, L. M., et al.: The Doge of Wall Street: Analysis and Detection of Pump and Dump Cryptocurrency Manipulations. ACM Transactions on Internet Technology (2023).