Forecasting Bitcoin Prices using Sentiment Analysis of News Headlines with Parallel GRU-CNN

Palveenraj Uma Kandan (U23103092), Madyan Omar (U23200049)

***Abstract*—Bitcoin is a well-known cryptocurrency, that revolutionized the financial landscape due to its decentralized and secure nature. However, its prices tend to fluctuate, hence the investors are unable to predict the future price of Bitcoin. This leads to poor strategic planning for Bitcoin investment. One of the factors that influence Bitcoin market price is the investors' sentiment towards cryptocurrency news headlines. In this paper, we proposed a deep learning model termed Parallel Gated Recurrent Unit-Convolutional Neural Network (Parallel GRU-CNN) for predicting future Bitcoin prices based on historical price and sentiment values of cryptocurrency news headlines. The proposed model leverages both GRU and 1D-convolutional layers to extract global and local features from the input data simultaneously. Compared to existing methods, Parallel GRU-CNN demonstrated lower mean absolute percentage error (MAPE) and root mean squared error (RMSE). The proposed model attained 0.02660 and 0.02291 for MAPE and RMSE respectively. Exclusion of sentiment values in model implementation reduced the accuracy of predicted Bitcoin prices, implying the significant influence of news headlines on Bitcoin price. Overall, the proposed model demonstrates that integration of sentiment values of news headlines in deep learning-based Bitcoin price predictive models could offer valuable insights for informed Bitcoin investment strategies.**

***Index Terms*—Bitcoin, sentiment analysis, cryptocurrency, Parallel GRU-CNN**

I. Introduction

BTC or more commonly known as Bitcoin is the most popular digital currency. It was introduced back in 2008 in [1] and signed by the pseudonym Satoshi Nakamoto. It is a decentralized currency and was introduced to address the known issues that fiat currency and traditional financial institutions presented. It has gained massive traction due to several reasons such as: The transparent nature of the Bitcoin system because transactions are recorded within Bitcoin's public ledger called the blockchain, providing investors with greater financial sovereignty and control over their finances by eliminating dependency on third party institutions such as banks, the secure nature of Bitcoin since transactions are secured by cryptography [2], [3], [4].

Despite Bitcoin being widely popular as an investment option, investors frequently encounter losses due to its inherent volatility and the unpredictability of factors influencing its valuation. For instance, regulatory announcements by major economies, trading volume, shifts in the market and investor sentiment influence the instability of Bitcoin price [5], [6]. Therefore, the prediction of Bitcoin prices holds significant importance as it enables investors to anticipate and mitigate potential financial losses, safeguarding their investments and enhancing overall financial security.

Previous works have performed sentiment analysis on social media posts, mainly on the tweets from twitter, using different deep learning models. Some of the widely used deep learning models are Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). [7], [8], [9], [10]. Both the models are designed to capture long-term dependencies in sequential data, such as time series data, except the difference is that the architecture of GRU is less complex compared to LSTM.

While existing studies have conducted sentiment analysis on the social media posts, fewer studies have been done using news from a dedicated cryptocurrency news source [5], [11]. Therefore, our research focuses on examining the impact of sentiment in news headlines on forecasting Bitcoin closing prices. To do this, we propose a Parallel Gated Recurrent Unit-Convolutional Neural Network (GRU-CNN) where the inputs, Bitcoin closing price and the sentiment values of news headlines are passed to GRU and 1D-convolutional layers simultaneously. GRU layer is used to extract global patterns, whereas 1D-convolutional layer is used to extract local patterns. Thus, enhancing the accuracy of Bitcoin closing price prediction.The followings are the key contributions of our work:

- In contrast to prior studies focusing on sentiment analysis of social media posts, we chose to leverage cryptocurrency news from a dedicated platform to explore potential influences on Bitcoin closing price forecasting.
- We introduced a Parallel GRU-CNN architecture, in which both Bitcoin closing price data and sentiment scores from news headlines are fed into GRU and 1D-convolutional layers concurrently.
- We opted Optuna hyperparameter tuning algorithm to find the best combination of parameters that optimizes our proposed model.

The rest of the paper is outlined as follows: The Related Works section discusses various deep learning and machine learning models utilized for predicting Bitcoin closing prices, both with and without incorporating sentiment analysis. Data preparation, sentiment analysis, proposed model and Optuna hyperparameter tuning algorithm are explained in detail in Methodology section. The Results and Discussion section elaborates on the performance of the proposed model, assessing its effectiveness through different evaluation metrics. The Conclusion section concludes the paper.

II. Related Works

A. Cryptocurrency Price Prediction

When working with time series data such as Bitcoin prices, stock prices and temperature changes, various approaches have been explored. One of the approaches employs Recurrent Neural Network (RNN) which was introduced by Pant et al. [12]. The issue with RNN is that they suffer from vanishing gradient problem, which makes it difficult for RNN to learn long term dependencies from the data. Although this could be problematic for several types of time series data, the authors have shown that it performs well in Bitcoin price prediction.

There exists several solutions to the vanishing gradient problem such as LSTM and GRU. It is widely used for time series data and forecasting and has been used by several researchers [2], [4], [13], [10] for Bitcoin price prediction. Since both GRU and LSTM were effective at preserving long-term dependencies, the authors achieved improved results.

Periketi et al. [10] aimed to enhance Bitcoin price forecasting accuracy by combining historical and real-time data using two deep learning models: Stacked LSTM and Prophet. The dataset from Tiingo was chosen for its performance and support for extensive filters. By integrating these models and data sources, the study provides valuable insights for traders, investors, and researchers, enabling informed decision-making and a better understanding of Bitcoin market behavior. Kazeminia et al. [14] utilized Bitcoin historical data and introduced a novel hybrid 2D-CNN-LSTM model with OPTUNA hyperparameter tuning for predicting the next day's closing price. The model was trained on data obtained through automated web scraping and achieved high accuracy, surpassing the performance of CNN, LSTM, and GRU models. This hybrid approach combines the strengths of convolutional neural networks (CNN) and long short-term memory (LSTM) networks, making it robust for real-time forecasting and supporting informed investment decisions in the cryptocurrency market.

Moreover, statistical models such as Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) have been utilized for Bitcoin price prediction. In a paper published by Glenski et al. [8], they trained a variety of ARIMA models and evaluated the models. Despite achieving satisfactory outcomes, they discovered that LSTM outperformed and generated superior results.

Various machine learning algorithms like Linear Regression models and Support Vector Regression (SVR) models have been employed for prediction purpose. Alahmari et al. [15] utilized SVR models with Linear, Polynomial, and Radial Basis Function (RBF) kernels to forecast the prices of major cryptocurrencies (Bitcoin, XRP, and Ethereum). They employed a bivariate time series approach, with cryptocurrency daily closed prices as the dependent variable and the Morgan Stanley Capital International (MSCI) World Index and daily closed prices as predictor variables. The findings highlighted that the RBF kernel outperformed other kernels. Ali et al. [16] developed a method to select appropriate data chunks for training a prediction model, specifically focusing on bitcoin price prediction using a simple linear regression algorithm. They identified relevant features from the dataset with strong correlations to bitcoin prices and ensured proper data selection to improve prediction accuracy.

B. Sentiment Analysis of Social Media Posts

It is quite common to use Twitter news sentiment to predict bitcoin prices. This is because the age distribution of bitcoin investors and twitter users typically range between ages 25-33 [17]. One of the widely used models for sentiment analysis is Valence Aware Dictionary and Sentiment Reasoner (VADER) due to it's impressive performance in obtaining sentiment from the social media domain. Nouira et al. [5] used the VADER model and trained it with data sourced from Twitter and Google News. They devised an equation that considers factors such as a Twitter user's follower count, likes, and retweet count. This equation aims to incorporate the influence wielded by widely followed creators.

Pant et al. [12] used a machine learning approach for performing sentiment analysis on tweets obtained from twitter to predict the final Bitcoin prices. Other researchers such as Yao et al. [18] used combined methodologies to extract features from news articles using commonly used text feature extraction algorithms such as N-Gram, TD-IDF, Doc2vec methods and SentiGraph which takes advantages of sentiment analysis and transforms a news article into a graph structure.

Aslam et al. [19] used TextBlob and Text2Emotion for sentiment annotation and emotion detection respectively. Sentiment and Emotion are extracted from tweets that have been preprocessed after being retrieved from Twitter. NLP techniques such as: Stemming and Lemmatization is used in order to preprocess the tweets.

III. Methodology

This section describes the dataset, data preparation, sentiment analysis, model design and optimization, and model evaluation. An overview of the methodology framework is illustrated in Fig. 1

A. Dataset

This study utilizes two different datasets, Cryptocurrency News Headlines and Bitcoin Historical Price. The Cryptocurrency News Headlines dataset includes news headlines of multiple cryptocurrencies from diverse news sources, ranging from January 01, 2022 to December 19, 2023. For every date, there were multiple news headlines, each accompanied by its sentiment score, resulting in a total of 31,037 data points. This dataset was obtained from Kaggle.

The Bitcoin Historical Price dataset comprises daily historical price data, spanning from March 8, 2019 to January 21, 2024, structured in a time-series format and including information on opening, closing, high and low prices as well as trading volume. It includes a total of 1,710 data points. This dataset was downloaded from Nasdaq.

TABLE I
DETAILS OF BITCOIN HISTORICAL PRICE AND CRYPTOCURRENCY NEWS HEADLINES DATASET BEFORE AND AFTER PREPROCESSING

	Cryptocurrency News Headlines Dataset		Bitcoin Historical Price Dataset	
	Number of Features	Number of Data Points	Number of Features	Number of Data Points
Before Processing	7	31037	5	1710
After Processing	4	5393	2	702

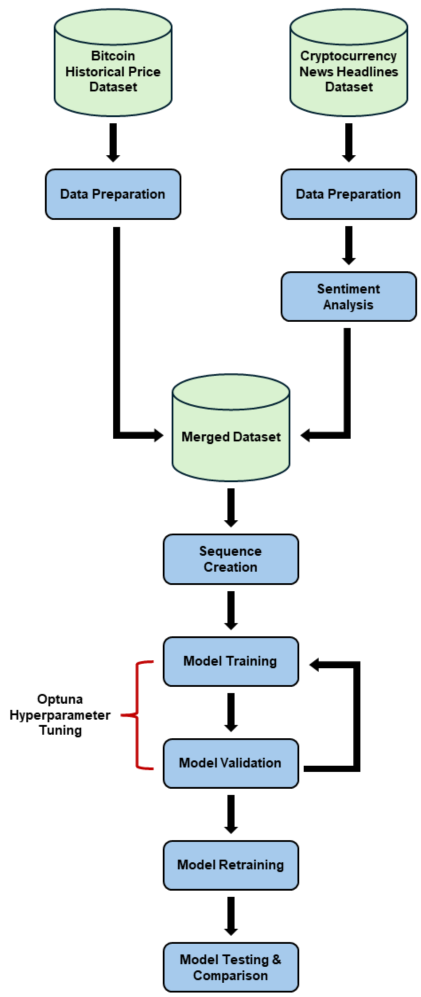

Fig. 1. Workflow of the methodological framework

B. Data Preparation

In Cryptocurrency News Headlines dataset, CoinTelegraph had the highest number of news headlines among all the news sources (Fig. 2). Therefore, we selected CoinTelegraph as our only news source to avoid redundancy, resulting from similar news headlines from other news sources. As the dataset consists of various cryptocurrencies, we excluded all the cryptocurrencies except Bitcoin as it is the primary cryptocurrency of our study. As a result, the dataset was reduced to 5,393 data points, covering from January 13, 2022 to December 19, 2023. From this range of dates, 11 dates had no news articles published. Additionally, we removed the sentiment score feature from the dataset as we planned to perform our own sentiment analysis.

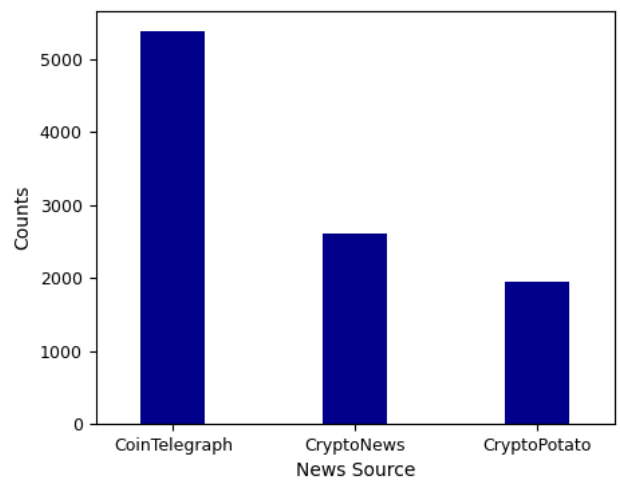

Fig. 2. Number of Bitcoin-related news articles in three different news sources comprising the dataset.

On the other hand, in Bitcoin Historical Price dataset, only the dates and closing prices were selected as it represents the final price at the end of the trading day. It had four missing dates as well as their corresponding closing prices, which were found manually from Yahoo Finance.The dataset was selected between January 13, 2022 to December 19, 2023 to match with the range of dates in Cryptocurrency News Headlines dataset, thereby reducing to 702 data points. At last, we normalized the closing prices feature in the range of [0,1] using min-max normalization. Table I contains the details of the datasets.

C. Sentiment Analysis

Sentiment Analysis is a Natural Language Processing (NLP) approach, used to assess emotional sentiment expressed in a text. In our case, we conducted sentiment analysis to evaluate the emotional tone of news headlines, assigning a sentiment score to each news headline. This was performed to observe whether sentiment analysis of news headlines had any influence on the Bitcoin closing price. We used Natural Language Toolkit (NLTK), a popular library that includes Valence Aware Dictionary for sEntiment Reasoning (VADER), which is a sentiment analyzer. VADER utilizes a lexicon comprising thousands of words, each assigned with a sentiment score for positivity, negativity and neutrality, indicating the sentiment strength of each word. The cumulative sentiment score of each news headline is computed based on this lexicon.

Several text preprocessing techniques were carried out on the news headlines prior to conducting sentiment analysis. This was performed to remove any irrelevant information, making the news headlines suitable for analysis. Firstly, we converted the news headlines to lowercase and eliminated any special characters from them. Following that, each news headline underwent word tokenization where sentences were split into list of words or tokens. After that, a process known as lemmatization was done to reduce words to their base form. Lastly, we filtered out stop words from the word list of each news headline as they were considered semantically insignificant.

The preprocessed news headlines were passed to VADER, to obtain cumulative sentiment score, ranging from [-1,1]. A sentiment score of less 0 indicates negative sentiment, whereas greater than 0 indicates positive sentiment. 0 represents neutral sentiment. In the Data Preparation subsection, it was noted that there were 11 missing dates, and we have assigned a value of 0 to each of them, indicating no significant impact on Bitcoin closing price fluctuation. As stated in the Dataset subsection, given the numerous news headlines available for each date, we computed the average sentiment score for each respective date. Subsequently, we merged the sentiment score feature with the Bitcoin historical price dataset, which would be passed to our proposed model. Lastly, we normalized the merged data in the range of [0,1] using min-max normalization.

D. Sequence Creation for Time-Series Data

We implemented a sliding window approach to create input-output pairs for a time-series prediction task. Sentiment score and closing prices features are the input while the following day price is the output. We defined a time step of value one, indicating that the window shifts by one unit with each iteration. This allows us to predict daily closing prices.

E. Parallel GRU-CNN Model Architecture

The proposed model consist of an input layer, a GRU layer, a 1D-convolutional layer, three dense layers, a concatenate layer, a dropout layer, a flatten layer and an output layer. The model architecture is depicted in (Fig. 3). The input layer is shaped to receive two input features which are the normalized closing price and sentiment score. After that, the input features are passed to the GRU and 1D-convolutional layers at the same time. This allows for simultaneous extraction of both local and global features, enhancing model capability. Afterward, the features extracted from the GRU and 1D-convolutional layers are individually forwarded to two separate dense layers. Both the dense layers have the same number of units, allowing them to subsequently be concatenated by the concatenated layer. To prevent overfitting, a dropout layer is incorporated into the model. The concatenated features were flattened via a Flatten layer, after which they are directed to a third dense layer with a different number of units. The output layer predicts the Bitcoin closing price of the following day. Additionally, the biases within the GRU, 1D-convolutional and the three dense layers were set to an initial value of 0.1, while the filter matrices are initialized using the Glorot normal initializer. The GRU layer and the three dense layers use hyperbolic tangent (tanh), whereas the 1D-convolutional layer uses Rectified Linear Unit (ReLU) activation function. The proposed model employs the Adam optimizer along with a specified learning rate and utilizes Mean Absolute Error (MAE) as the loss function.

TABLE II
PARAMETERS OF THE PARALLEL GRU-CNN

Input Layer	
Shape	(1,2)
GRU Layer	
GRU Units	80
Activation Function	tanh
Conv1D Layer	
Filters	
Kernel Size	1
Activation Function	relu
Dense Layers (1 & 2)	
Dense Units	160
Activation Function	tanh
Concatenate Layer	
Concatenation Axis	2
Dropout Layer	
Dropout Rate	0.2
Dense Layer 3	
Dense Units	112
Activation Function	tanh
Output Layer	
Dense Units	1
Other Parameters	
Learning Rate	0.0001
Epochs	95
Batch Size	64

F. Optuna Hyperparameter Tuning Algorithm

The data is split into 70% training, 15% validation and 15% testing sets. The training and validation sets were used for hy-

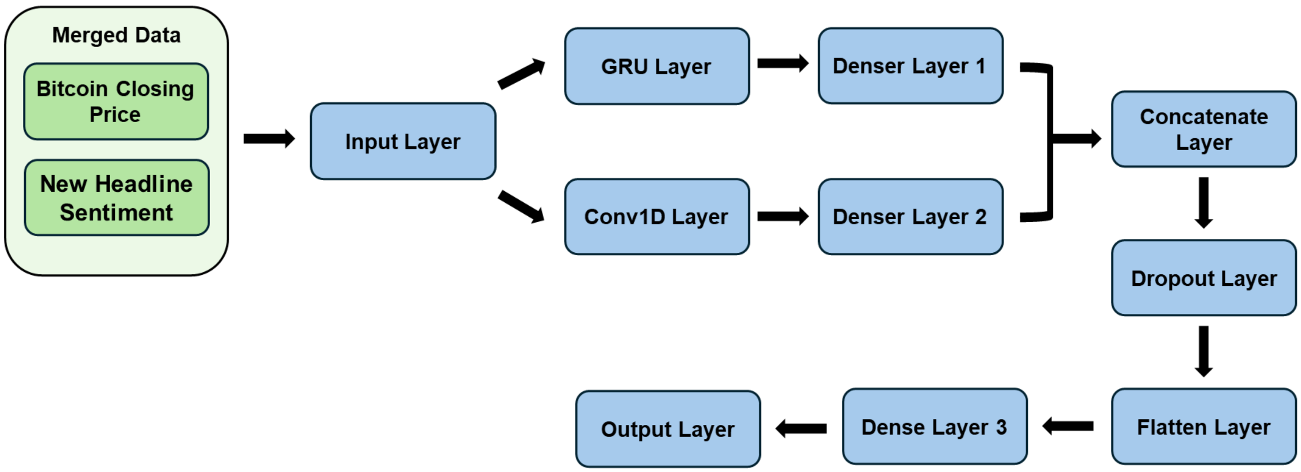

Fig. 3. Parallel GRU-CNN model architecture for predicting the future Bitcoin prices by using Bitcoin historical price and sentiment values of the news headlines.

perparameter tuning. We used Optuna hyperparameter tuning algorithm to identify the best combination of parameters that optimizes the performance of the proposed model based on MAE loss function. The optimization was carried out for 200 trials. We fine tuned several parameters, including number of units within the GRU layer, number of filters within the 1D-convolutional layer, number of units within the dense layers, dropout rate, learning rate, number of epochs and batch size. Table Table II outlines the ideal values of the parameters used in the proposed model.

G. Software Requirements

In this research, we primarily used Jupyter Notebook as our interactive computational environment, alongside with the Python programming language. We utilized Pandas and Numpy libraries for data manipulation and data preprocessing. Matplotlib library was employed for visualizing the data. Furthermore, the proposed Parallel GRU-CNN model was constructed using the TensorFlow library.

H. Model Retraining and Evaluation

We combined the training set and validation set into a final training set for model testing. The final training set constitutes 85% of the dataset. After training, the model was evaluated with 15% test set. Then, the model was compared with existing models such as LSTM [10] and 2D-CNN-LSTM [14]. In addition, we also assessed the proposed model when only Bitcoin historical price is used as an input (Parallel GRU-CNN without sentiment). This allowed us to understand the effect of news headline sentiment on the accuracy of predicted Bitcoin closing prices. We evaluated the performance of our proposed model as well as the existing models using Root mean squared error (RMSE) and Mean Absolute Percentage Error (MAPE) as shown below. The symbols y_i and $\hat{y}_i$ denote actual and predicted Bitcoin closing prices respectively, whereas n denotes number of data points.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{2}$$

IV. Results and Discussion

The proposed Parallel GRU-CNN achieved a MAPE of 0.02660 and RMSE value of 0.02291. We compared our proposed model with LSTM [10] and 2D-CNN-LSTM [14] (Table III). The proposed model surpassed both LSTM and 2D-CNN-LSTM, demonstrating a 15.9% enhancement over LSTM and a 19.3% improvement over 2D-CNN-LSTM in reducing MAPE. Additionally, the model showed a 13.8% improvement over LSTM and a 6.7% improvement over 2D-CNN-LSTM in decreasing RMSE. The proposed model's superior performance is due to extraction of local and global features in which 2D-CNN-LSTM and LSTM did not employ.

We also evaluated the efficacy of our proposed model in absence of sentiment analysis, aiming to determine the impact of sentiment on bitcoin price fluctuations. Without sentiment values, the MAPE and RMSE of the proposed model increased by 1.8% and 3% respectively. We also plotted a line graph describing the actual and predicted Bitcoin prices with and without sentiment (Fig. 4 and Fig. 5). It was evident that the predicted values were closer to the actual values when integrating sentiment values to the model. This shows that news headline sentiments have an impact on Bitcoin price prediction.

Our project's drawback lies in the limited quantity of data points used for training. The large number of data points

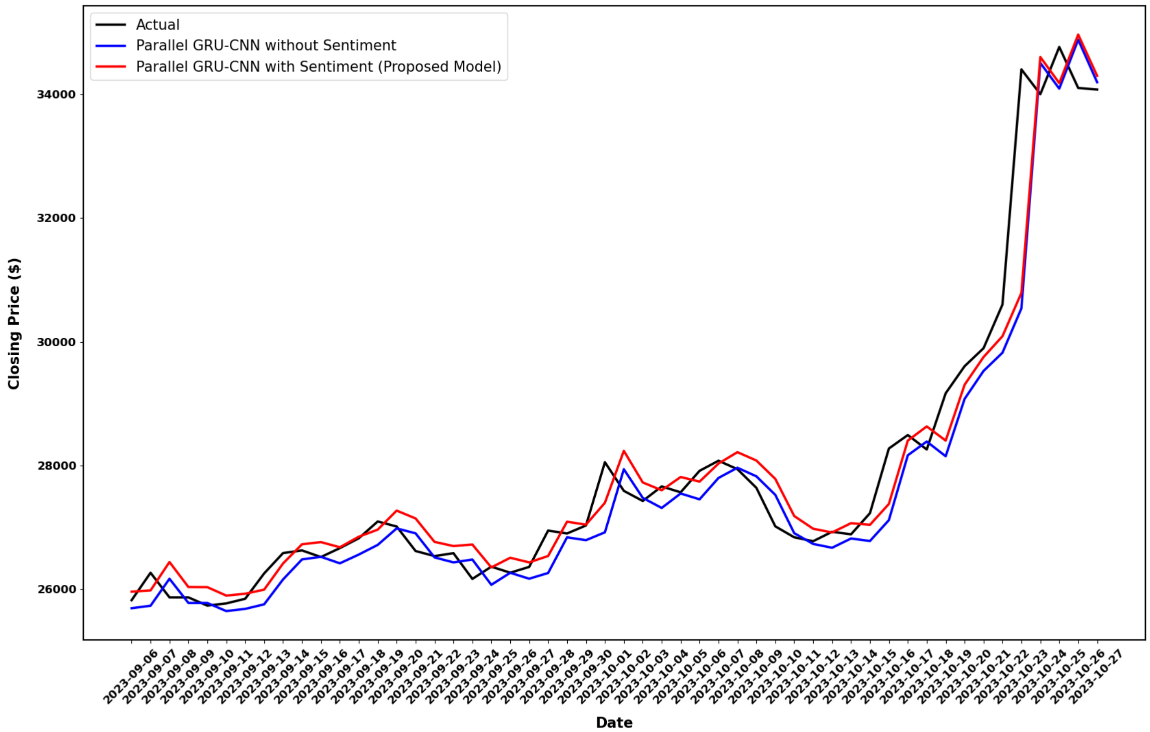

Fig. 4. Actual and predicted Bitcoin price with and without Sentiment, ranging from 2023-09-06 to 2023-10-27

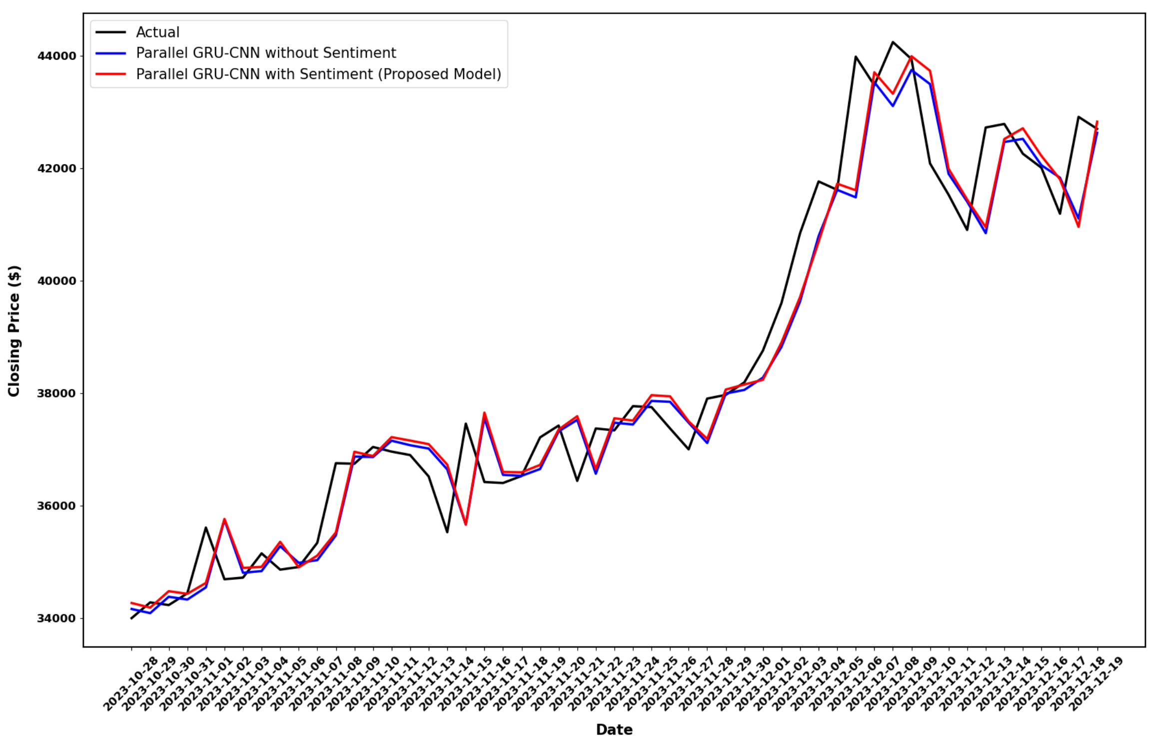

Fig. 5. Actual and predicted Bitcoin price with and without Sentiment, ranging from 2023-10-28 to 2023-12-19

TABLE III
ROOT MEAN SQUARE ERROR (RMSE) AND MEAN ABSOLUTE PERCENTAGE ERROR (MAPE) OF THE PROPOSED MODEL AND EXISTING MODELS

Model	MAPE	RMSE
LSTM	0.03163	0.02659
2D-LSTM-CNN	0.03295	0.02456
Parallel GRU-CNN without Sentiment	0.02709	0.02362
Parallel GRU-CNN with Sentiment (Proposed Model)	0.02660	0.02291

contributes to building a more robust and stable model that can generalize well to unseen time-series data, leading to improved predictive performance.

V. CONCLUSION

In conclusion, this paper developed Parallel GRU-CNN to predict future Bitcoin prices based on historical price and sentiment values of news headlines. Cryptocurrency related news articles play a significant role in modulating the sentiment of Bitcoin investors, thus affecting the market price. Through incorporation of sentiment values, the proposed model predicted the future Bitcoin prices with high accuracy. In addition, the proposed model was designed to extract both local and global features, hence leading to better performance compared to other state-of-the-art methods. We expect to observe an improvement in the performance by increasing the size of the dataset. Nevertheless, the proposed Parallel GRU-CNN serve as a promising tool for future Bitcoin price prediction and help investors make informed decisions about buying or selling Bitcoin.

REFERENCES

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[2] M. H. B. M. Sabri, A. Muneer, and S. M. Taib, "Cryptocurrency price prediction using long short-term memory and twitter sentiment analysis," in *2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA*, 2022, pp. 1–6.

[3] M. Rizwan, S. Narejo, and M. Javed, "Bitcoin price prediction using deep learning algorithm," in *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, 2019, pp. 1–7.

[4] N. Latif, J. D. Selvam, M. Kapse, V. Sharma, and V. Mahajan, "Comparative performance of lstm and arima for the short-term prediction of bitcoin prices," *Australasian Accounting, Business and Finance Journal*, vol. 17, no. 1, pp. 256–276, 2023.

[5] A. Y. Nouira, M. Bouchakwa, and Y. Jamoussi, "Bitcoin price prediction considering sentiment analysis on twitter and google news," in *Proceedings of the 27th International Database Engineered Applications Symposium*, 2023, pp. 71–78.

[6] Y. Sovbetov, "Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero," *Journal of Economics and Financial Analysis*, vol. 2, no. 2, pp. 1–27, 2018.

[7] K. P. Sharma, S. K. Singh, A. Choudhary, and H. Goel, "Price prediction of bitcoin using social media activities and past trends," in *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2023, pp. 516–521.

[8] M. Glenski, T. Weninger, and S. Volkova, "Improved forecasting of cryptocurrency price using social signals," *arXiv preprint arXiv:1907.00558*, 2019.

[9] P. Bansal and S. Jain, "Cryptocurrency price prediction using twitter and news articles analysis," in *Proceedings of the 2022 Fourteenth International Conference on Contemporary Computing*, 2022, pp. 225–233.

[10] S. Periketi, R. Kulkarni, and U. N. Dulhare, "Lstm and prophet model fusion framework for bitcoin candlestick visualization and price prediction forecasting," in *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2023, pp. 1190–1194.

[11] C. Lamon, E. Nielsen, and E. Redondo, "Cryptocurrency price prediction using news and social media sentiment," *SMU Data Sci.Rev*, vol. 1, no. 3, pp. 1–22, 2017.

[12] D. R. Pant, P. Neupane, A. Poudel, A. K. Pokhrel, and B. K. Lama, "Recurrent neural network based bitcoin price prediction by twitter sentiment analysis," in *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, 2018, pp. 128–132.

[13] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting cryptocurrency prices using lstm, gru, and bi-directional lstm: A deep learning approach," *Fractal and Fractional*, vol. 7, no. 2, p. 203, 2023.

[14] S. Kazeminia, H. Sajedi, and M. Arjmand, "Real-time bitcoin price prediction using hybrid 2d-cnn lstm model," in *2023 9th International Conference on Web Research (ICWR)*, 2023, pp. 173–178.

[15] S. Alahmari, "Predicting the price of cryptocurrency using support vector regression methods," *J.Mech.Continua Math.Sci*, vol. 15, no. 4, pp. 313–322, 2020.

[16] M. Ali and S. Shatabda, "A data selection methodology to train linear regression model to predict bitcoin price," in *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*. IEEE, 2020, pp. 330–335.

[17] J. Bohr and M. Bashir, "Who uses bitcoin? an exploration of the bitcoin community," in *2014 Twelfth Annual International Conference on Privacy, Security and Trust*. IEEE, 2014, pp. 94–101.

[18] W. Yao, K. Xu, and Q. Li, "Exploring the influence of news articles on bitcoin price with machine learning," in *2019 IEEE Symposium on computers and communications (ISCC)*. IEEE, 2019, pp. 1–6.

[19] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble lstm-gru model," *IEEE Access*, vol. 10, pp. 39 313–39 324, 2022.