IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 September 2023, accepted 15 October 2023, date of publication 23 October 2023, date of current version 3 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3327060

RESEARCH ARTICLE

Aspect-Based Sentiment Analysis of Twitter Influencers to Predict the Trend of Cryptocurrencies Based on Hybrid Deep Transfer Learning Models

KIA JAHANBIN AND MOHAMMAD ALI ZARE CHAHOOKI
Department of Computer Engineering, Yazd University, Yazd 8915818411, Iran
Corresponding author: Mohammad Ali Zare Chahooki (chahooki@yazd.ac.ir)

ABSTRACT With the expansion of social networks, sentiment analysis has become one of the hot topics in machine learning. However, in traditional sentiment analysis, the text is considered of a general nature and ignores the different aspects that may exist in the text. This paper presents a hybrid model of transfer deep learning methods for the aspect-oriented sentiment analysis of influencers' tweets to predict the trend of cryptocurrencies. In the first model, different aspects of tweets are extracted using the Concept Latent Dirichlet Allocation (Concept-LDA). Then, by using the pre-trained RoBERTa network and combining it with the Bidirectional Gated Recurrent Unit (BiGRU) deep learning network and attention layer, sentiments of different aspects of tweets are determined. In the following, the price trend of seven cryptocurrencies, Bitcoin, Ethereum, Binance, Ripple, Dogecoin, Cardano, and Solana, is determined using the historical price and the polarity of tweets with BiGRU combined deep neural network and the attention layer. Also, we used the gridsearch method to select dropout hyper-parameters, learning rate, and the number of GRU units, and the Akaike Information Criterion (AIC) criterion confirmed the results of this proposed combination. The results show that the proposed model in the aspect-based sentiment analysis section has been able to achieve 5.94% accuracy and 9.9% improvement in the f1-score on the SemEval 2015 dataset and 2.61% improvement on the SemEval 2016 dataset in f1-score compared to the state-of-arts. Also, the results of predicting the price trend of cryptocurrencies show that the proposed model has correctly recognized the price trend in the next five days in 77% of cases according to the ROC-AUC criterion.

INDEX TERMS Aspect based sentiment analysis, prediction trend price, cryptocurrencies, pre-trained networks, hybrid deep learning models.

I. INTRODUCTION

The sentiments of the market or investors are their general behavior and attitude towards financial markets such as cryptocurrencies. Analyzing market sentiment is a process during which the positive, negative, and neutral sentiments in a text or sentence are identified. By examining these sentiments, the market's future is predicted. Sentiment analysis helps traders to make decisions about market price trends. Suppose most people's feelings about the current state of the cryptocurrency market are positive. In that case, we can expect an increase in the price of that cryptocurrency or vice versa [1]. So, Sentiment analysis has been widely used in predicting cryptocurrency prices, particularly in the context of social media platforms such as Twitter. One approach is to combine historical price data with sentiment scores. Twitter, in particular, has become a significant source of news and discussion about cryptocurrencies, making it a valuable resource for sentiment analysis [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

The usability of different deep learning models for sentiment analysis in cryptocurrency trend prediction has been explored. However, not all sentiment analysis methods are equally accurate for this prediction task, and further research is needed to improve the accuracy of sentiment analysis in the cryptocurrency market.

From a statistical point of view, extensive research has been done on the effect of long-range dependence and heavy tail characteristics for remaining functional stock prediction [3]. Long-range dependence and heavy tail characteristics are used for remaining useful life (RUL) prediction in rolling bearing degradation. These terms refer to the statistical properties of the degradation process of rolling bearings, which can also be applied to other prediction models, such as cryptocurrency price prediction [4].

Long-range dependence refers to the persistence of the degradation process over time, where the system's current state depends on its past states. In other words, the degradation process exhibits a memory effect, and the system's future state can be predicted based on its past states. This property is essential in RUL prediction because it allows for developing models that can capture the system's long-term behavior. Heavy tail characteristics refer to the distribution of the degradation increments, which can be modeled using a heavy-tailed distribution. A heavy-tailed distribution has a higher probability of extreme events, such as sudden failures or rapid degradation, than a normal distribution. This property is important in RUL prediction because it allows for the development of models that can capture the rare but significant events that can affect the remaining useful life of the system [5].

Han et al. [6] explore the properties of the Ether market, the cryptocurrency associated with the Ethereum blockchain. The paper aims to provide insights into Ether investment by analyzing the Ether market's long-range dependence, multifractality, and asymmetry using detrended fluctuation analysis and asymmetric multifractal detrended fluctuation analysis. The article also studies the causality between returns and Ether's volume to find how investors' activity influences returns based on a nonparametric causality-in-quantiles test.

Arratia and López-Barrantes [7] examined the relationship between Bitcoin price and search interest in 2014. This paper uses linear and non-linear dependence tests to question the predictive ability of Google Bitcoin Trends for Bitcoin price behavior. This paper also examines the performance of ARIMA and neural network models augmented with this social sentiment index. Analyzes and models are based on standard statistical properties with financial returns created for Bitcoin, Ethereum, Ripple, and Litecoin.

This research is based on a wide range of state of the art. The related works to this study can be divided into two categories: Sentiment analysis and aspect extraction with deep transfer learning and cryptocurrency trend prediction using sentiment analysis.

Aspect Based Sentiment Analysis (ABSA) with transfer learning refers to a natural language processing technique involving pre-trained language models and transfer learning strategies to improve the accuracy and efficiency of aspect-based sentiment analysis tasks. Aspect-based sentiment analysis identifies the sentiment polarity (e.g., positive, negative, or neutral) of specific aspects or features of a given entity, such as a product or service mentioned in a text. Transfer learning involves leveraging knowledge gained from one task to improve performance on another related task. The use of transfer learning in ABSA can help address issues related to data scarcity and domain adaptation [8].

Aspect term extraction (ATE) can be divided into two categories: opinion target extraction (OE) and aspect category detection (ACD). For example, in the sentence "The car was so fuel-efficient and spacious that it was perfect for long road trips," the OE task extracts aspect terms ("car," "fuel-efficient," and "spacious"), and the ACD task removes aspect categories ("fuel economy" and "interior"). Most articles like this research focus on OE [9].

Poria et al. [9] introduce a deep learning approach for aspect extraction in opinion mining, a subtask of sentiment analysis. The proposed method combines a 7-layer deep convolutional neural network with linguistic patterns to tag each word as an aspect or a non-aspect word in opinionated sentences. This is the first deep learning approach to aspect extraction, and it represents a promising direction for improving opinion mining techniques.

Trang et al. [10] introduce a method for aspect extraction in sentiment analysis, utilizing a deep learning model that combines BiGRU and Conditional Random Field (CRF). Our model is trained on annotated data to identify and categorize feature sets within comments. Employing a BiGRU neural network with word embeddings obtained from training GloVe on the SemEval 2014 dataset achieves robust performance.

However, labels are scarce in many domains, such as cryptocurrencies. Therefore, pre-trained models such as the BERT family use large trains on corpora in such domains. Formally, aspects of transfer aspects with a pre-trained model are as follows [8]:

1) Let S be the source domain dataset, T be the target domain dataset, and Ds and Dt represent the corresponding feature-label pairs, where $Ds = \{(x\_s1, y\_s1), (x\_s2, y\_s2), \ldots, (x\_sN, y\_sN)\}$, $Dt = \{(x\_t1, y\_t1), (x\_t2, y\_t2), \ldots, (x\_tM, y\_tM)\}$, where x represents the input text and y denotes the sentiment label. A pre-trained model, denoted as M, is initially trained on Ds in a supervised manner, capturing sentiment understanding from the source domain. Transfer learning is then performed by adapting M to the target domain data Dt, where fine-tuning is applied to update the model parameters using a smaller labeled dataset. This process allows M to adapt to the nuances and specificities of the target domain, improving its performance in ABSA on the target domain.

Bensoltane and Zaki [11] propose a transfer learning approach that utilizes pre-trained language models for two

ABSA tasks in Arabic: aspect term extraction (ATE) and aspect category detection (ACD). The proposed models are based on AraBERT, the Arabic version of the BERT model. Various implementations of BERT, including fine-tuning and feature-based methods, are compared. The key findings of the study indicate that fine-tuning is more suitable for low-resource scenarios and that customizing downstream layers improves the performance of the default fine-tuned BERT model.

Tao and Fang [12] addressed the shortcomings of existing ABSA methods. These methods typically treat the problem as single-label classification and require substantial labeled data for training. The proposed approach addresses the limitations by introducing multi-label classification capabilities to ABSA methods. A novel sentiment analysis method called Aspect Enhanced Sentiment Analysis (AESA) is proposed, which considers entity aspects when classifying text into sentiment classes. Two state-of-the-art transfer learning models are extended as analytical frameworks for the multi-label ABSA and AESA tasks.

Xu et al. [13] propose an aspect-level sentiment classification model based on Attention-Bidirectional Long Short-Term Memory (Attention-BiLSTM) model and transfer learning. The authors propose three models, including Pre-training (PRET), Multitask learning (MTL), and Pre-training and multitask learning (PRET+MTL), which transfer the knowledge obtained from document-level training of sentiment classification to aspect-level sentiment classification. The performance of the proposed models is evaluated on four datasets, and the experiments demonstrate that the proposed methods improve the training of neural network models for aspect-level sentiment classification, which suffers from small dataset sizes. The proposed models represent a promising approach for improving aspect-level sentiment classification in natural language processing.

However, public domains are considered in most articles [14], [15], and less attention is paid to the specific domains of cryptocurrencies or drugs.

Sweidan et al. [16] proposed a hybrid ontology-XLNet sentiment analysis classification approach for sentence-level aspects. The objective of the proposed method is to discover user social data and provide in-depth inference about sentiment based on the context. The paper explores using a lexicalized ontology to enhance aspect-based sentiment analysis by capturing indirect relationships in user social data. In this paper, To evaluate the performance of the proposed approach, several experiments are conducted on seven real-world datasets related to drug-related social data, specifically focusing on Adverse Drug Reactions (ADRs) discovery.

However, Tao and X. Fang [12] have no research data set on cryptocurrencies. Therefore, research in the field of ASBA related to cryptocurrencies is considered a research challenge because aspects of the content of public tweets are very different from those of crypto-related tweets.

Another aspect of this research is related to predicting the trend of cryptocurrencies with sentiment analysis. Several articles [17], [18], [19] have investigated predicting the trend of cryptocurrencies in the short term (a few hours to a day) and in a long time (more than a day). Formally speaking, the problem of forecasting cryptocurrencies or any other stock with sentiment analysis is formulated as follows.

1) Let X be the set of features extracted from sentiment analysis of textual data related to cryptocurrencies, and let Y be the set of corresponding cryptocurrency price movements. The problem of predicting cryptocurrencies based on sentiment analysis can be formulated as finding a function $f : X \rightarrow Y$ that maps the sentiment features to the price movements, such that *f(X)* minimizes the prediction error.

Very few researchers have used transfer learning to analyze the sentiment of tweets to predict the trend or price of cryptocurrencies.

Davchev et al. [20] proposed a methodology for predicting the price of Bitcoin by combining Twitter data and historical prices. They used RoBERTa pre-trained model for sentiment analysis and time-series forecasting to predict the next day's price. The study showed that including sentiment extracted from financial micro-blogs as input improved the accuracy of the predictions compared to using only historical price data. The authors also demonstrated that transfer-learning methodologies improved the efficiency of sentiment extraction in financial micro-blogs compared to standard sentiment extraction methods.

Xin and Peng [21] proposed a prediction scheme for chaotic time series using a combination of autoencoders and convolutional neural networks (AE-CNN) and transfer learning. They developed a method named AE-CNN-TL that showed improved performance compared to other methods such as AE-CNN, ARMA, and LSTM. The study demonstrated the effectiveness of AE-CNN-TL in capturing the intrinsic certainty of chaotic time series and improving prediction performance in the medium-to-long term.

As mentioned, in most articles that have used social media to predict the trend or price of a cryptocurrency, basic deep learning methods or ready-made tools such as Vader [22] have been used to determine the polarity of text sentiments.

Huang et al. [23] proposed a method for predicting cryptocurrency price movements by analyzing sentiment in Chinese social media posts. They developed a pipeline to capture Weibo posts, created a crypto-specific sentiment dictionary, and used a long short-term memory (LSTM) based recurrent neural network and historical price data to predict future price trends. The study showed that their approach outperformed the state-of-the-art auto-regressive-based model in precision and recall, demonstrating the effectiveness of analyzing sentiment in Chinese social media for cryptocurrency price prediction.

Abraham et al. [24] proposed a method for predicting changes in Bitcoin and Ethereum prices using Twitter and

Google Trends data. They found that tweet volume, rather than sentiment (with Vader), indicates price direction. Using a linear model that takes tweets and Google Trends data as input, they accurately predicted the direction of price changes. The study demonstrated the potential of using social media and search engine data for making informed purchase and selling decisions related to cryptocurrencies.

Valencia et al. [25] proposed using machine learning tools and social media data to predict the price movement of Bitcoin, Ethereum, Ripple, and Litecoin cryptocurrencies. They compared the performance of neural networks (NN), support vector machines (SVM), and random forest (RF) using input features from Twitter and market data. The study showed that machine learning and sentiment analysis (with Vader) could be used to predict cryptocurrency markets and that NN outperformed the other models. Additionally, they found that Twitter data alone could be used to predict specific cryptocurrencies.

Parekh et al. [26] proposed a hybrid and robust DL-Gues framework for cryptocurrency price prediction. The framework considers interdependency among cryptocurrencies and market sentiments. The authors tested the framework on Dash and Bitcoin Cash using price history and tweets of the respective cryptocurrencies and Bitcoin and Litecoin. Various loss functions were used for validation. The results suggest the framework's usability for predicting the prices of other cryptocurrencies.

In these articles, such as Abraham et al. [24] and Valencia et al. [25], the VADER tool is used to analyze the sentiments of social media texts. VADER tool is a lexicon-based tool developed for public texts in social networks. This tool takes a text and declares its general polarity. In addition to the fact that VADER is not produced for specialized texts, it does not care about the aspect of the text and assigns a general score that indicates the polarity of the text. However, some tweets, especially in cryptocurrencies, have several aspects. For example, a tweet like "Bitcoin is up, but Ethereum and Ripple are down" has three aspects of Bitcoin, Ethereum, and Ripple, which is positive for Bitcoin and negative for the other two cryptocurrencies. Now, if the polarity of the sentiments of this tweet is considered positive or negative with methods like VADER, it can cause the cryptocurrency price trend to be mispredicted. Therefore, this research has tried to extract different aspects of tweets and determine their emotions' polarity.

This paper will use sentiment analysis of influencer tweets to predict the price trend of seven popular cryptocurrencies: Bitcoin, Ethereum, Binance, Ripple, Dogecoin, Cardano, and Solana. Cryptocurrencies such as Bitcoin, Ethereum, and Binance are frequently mentioned in tweets daily. However, cryptocurrencies such as Dai, Torn, Polygon, Avalanche, Polkadat, Wrapped BTC, Electroneum, and Monero are less consistently covered by specific crypto influencers. Figure 1 shows the difference in the volume of tweets published for five cryptocurrencies: Bitcoin, Ethereum, Ripple, Polkadate, and Monero.

Figure 1 shows that tweets about Bitcoin or Ethereum are much higher than those of cryptocurrencies such as Monero or Polkadot. Therefore, fewer influencers comment on cryptocurrencies like Monero or Polkadot. One of the reasons could be the higher price volatility of these cryptocurrencies compared to major cryptocurrencies such as Bitcoin or Ethereum. While in terms of trading volume, according to Figure 2, except Bitcoin and Ethereum, the other three cryptocurrencies have almost the same trading volume.

We collected tweets related to the top 52 influencers in the field of cryptocurrencies, which were extracted and cleaned with the Twitter API. Cleaning operations include removing stopwords, Lemmatization, equalization, hyperlinks, and redundant characters such as @ and #. After the cleaning operation, a tweet impact coefficient column is added to each row of the tweet's dataset. This column shows the importance of each tweet. In this research, a separate impact factor is assigned to each tweet because some tweets of users with fewer followers may have more feedback than tweets of other users with more followers.

After assigning an influence coefficient to each tweet, the tweets are injected into the hybrid deep neural network of RoBERTa [27] and Bidirectional Gated Recurrent Unit (BiGRU) (HDRB) for aspect extraction and sentiment analysis. In this research, the Concept Latent Dirichlet Allocation (Concept-LDA) [28] model has been used to extract aspects. Then, the sentiments of each aspect are determined by the HDRB model. In the DHBR model, the attention layer is used for more accuracy on the important aspects of the text. The reason for using RoBERTa is that it is trained on a much larger corpus than BERT. Also, in the RoBERTa training process, Larger mini-batch training is used, which makes the training process more efficient.

Another reason for using the RoBERTa pre-trained neural network to analyze the sentiments of cryptocurrency tweets is to use the transfer learning technique in sentiment analysis. Because labeled datasets related to cryptocurrencies are scarce. Therefore, it is necessary to use cross-domain data for deep neural network training. RoBERTa is a model introduced in 2019 by the AI development group at Facebook and trained on a large corpus of Internet texts (about 160 GB). Therefore, using the knowledge available in RoBERTa can help prepare the HDRB model for sentiment analysis of cryptocurrencies.

After extracting the aspect and analyzing the sentiments of the tweets, the sentiment score associated with each cryptocurrency is multiplied by the tweet's influence factor and normalized between 0 and 1. Then, the maximum normalized score for negative and positive polarities is considered the polarity of emotions associated with the day "i" for cryptocurrency "j". For example, if there are three positive polarities, each with a normalized score of 0.2, and 2 negative polarities, each with a normalized score of 0.4 for

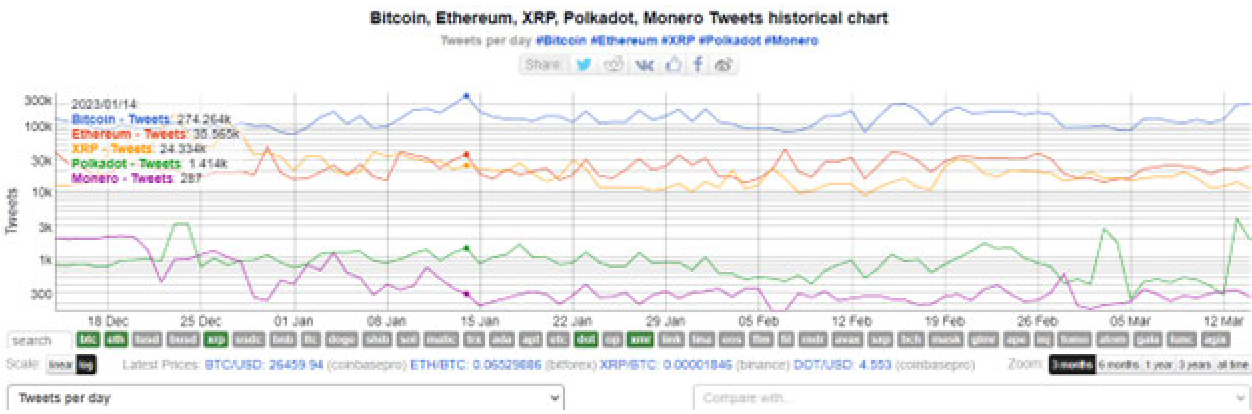


FIGURE 1. **The volume of tweets for Bitcoin, Ethereum, Ripple, and Dai (source bitinfocharts.com).**

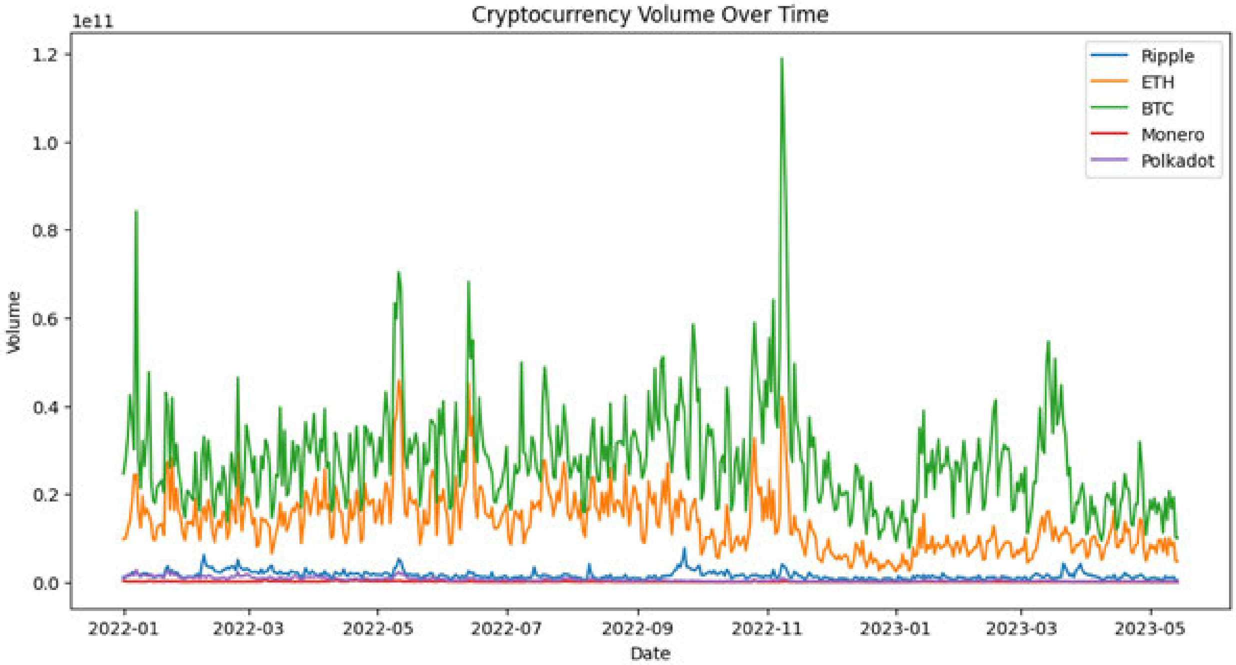


FIGURE 2. **Compare four cryptocurrencies per volume.**

cryptocurrency "j" on day "I", the sentiment polarity of this cryptocurrency on a day "I" is considered negative.

After determining the sentiment polarity of each cryptocurrency in the time interval [I, I+T], the price profile of that cryptocurrency in the same time interval is extracted by Yahoo Finance API. Then, in a column titled Change price, the price trend of each cryptocurrency is recorded according to whether the Colse price is rising or falling on the current day compared to the previous day. For example, if the Colse price of the cryptocurrency has increased on the i-th day compared to the i-1th day, a positive value is entered in the Change price column. Finally, by using the combination of BiGRU deep neural network and Attention layer (called BGA model), the price trend of each cryptocurrency is predicted in the coming days. Also, in this research, the number of days is considered a super parameter to predict the price trend of cryptocurrencies. The value of this hyperparameter is calculated based on the highest accuracy obtained in 3 to 10 days by running the BGA model. The significant contributions of this research are:

1. We present the HDRB model for aspect extraction and sentiment analysis related to cryptocurrencies, a combination of the Concept-LDA model for aspect extraction and pre-trained RoBERTa neural network, BiGRU deep neural network, and attention layer for sentiment analysis.
2. We are using transfer learning to analyze the sentiments of cryptocurrencies and train the HDRB deep learning model.

TABLE 1. List of abbreviations.

Abbreviations	Main word
Concept-LDA	Concept Latent Dirichlet Allocation
BiGRU	Bidirectional Gated Recurrent Unit
AIC	Akaike Information Criterion
RUL	Remaining Useful Life
HDRB	Hybrid Deep neural network of RoBERTa and Bidirectional Gated Recurrent Unit
BGA	BiGRU deep neural network and Attention layer
ABSA	Aspect Based Sentiment Analysis
ATE	Aspect Term Extraction
OE	Opinion target Extraction
ACD	Aspect Category Detection
CRF	Conditional Random Field
ADRs	Adverse Drug Reactions
AE-CNN	Autoencoders and Convolutional Neural Networks
LSTM	Long Short-Term Memory
IC	Importance Coefficient
TP	True Positive
FP	False Positive
GCN	Graph Convolutional Network
TD	Targeted sentiment Detection
FC	Fully Connected

3. We are determining the price trend of seven different cryptocurrencies using influencers' tweets.
4. The best time frame for predicting the price trend is determined using the accuracy criterion.

Table 1 shows a list of abbreviations in the order of being mentioned in the paper.

Continuing this article and the second section, we review state-of-arts. The DHRB and BGT models are introduced in detail in the third section. In the fourth section, the experimental results obtained from implementing the DHRB and BGT models for the analysis of aspect-oriented sentiments and for predicting the price trend of cryptocurrencies are presented.

II. HDRB+BGA MODEL

This research is divided into sentiment analysis, aspect extraction, and cryptocurrency trend forecasting. I use the HDRB model for sentiment analysis, aspect extraction, and BGA to predict cryptocurrency trends using sentiment analysis and historical prices. Tweets related to cryptocurrency are extracted first, and then the aspect and sentiment of the tweets are extracted from pre-processing with the HDRB model. Then, the sentiments are classified based on each cryptocurrency's name, and the cryptocurrency's trend is predicted with the BGA model using historical price. Figure 3 shows this process.

A. SENTIMENT ANALYSIS AND ASPECT EXTRACTION WITH THE HDRB MODEL

HDRB model is a hybrid approach that uses a pre-trained RoBERTa model, attention layer, and BiGRU deep neural network. According to Figure 3, first, tweets related to 52 influencers specializing in cryptocurrency are extracted. Several websites have been used to find influencers. After extracting the influencer IDs, the ID of the users who subscribed to all these sites was used as a seed to extract tweets. The tweets were extracted using the API of the apify.com website, and the importance coefficient was assigned to the extracted tweets after cleaning. This coefficient indicates the level of acceptance of the contents of this tweet by other users and is calculated through Eq (1).

$$Importance\ coefficient\ (IC) = retweet + (2 * favorite) + (0.5 * reply) \tag{1}$$

Eq (1) assigns an importance coefficient or IC to each tweet. This coefficient helps that the sentiment polarity of more important tweets significantly impacts selecting the sentiment polarity related to the i-th cryptocurrency on day "j". The reason for choosing the coefficients in Eq (1) is that in most of the content related to social networks, the number of likes is more than the number of comments or reposts.

The components in Eq (1) represent different interactions that a tweet can receive:

- Retweet: Indicates the number of times other users have retweeted the tweet. Retweets signify that the tweet's content has resonated with a larger audience, suggesting its broader impact and relevance.
- Favorite: Represents the number of times the tweet has been marked as a favorite by other users. Favoriting a tweet is a common way for users to show their approval and interest in its content.
- Reply: Reflects the number of replies or comments the tweet has received from other users. Replies demonstrate engagement and can often involve discussions and conversations, making them a valuable indicator of the tweet's influence.

The reason for selecting these specific parameters and coefficients in Eq (1) is rooted in the observation that, in the context of social networks, specific interactions tend to hold more weight in gauging a tweet's impact:

1. Retweets and Favorites: In our analysis of social media behavior, we have observed that users often retweet and favorite tweets to share them with their followers and to express their agreement or appreciation. As a result, we assign a higher weight to the number of favorites (2) than retweets (1), given the more comprehensive visibility and approval associated with favorites.
2. Replies: Replies or comments represent engagement and discussions around a tweet's content. While replies are essential in assessing the depth of engagement, they typically exhibit a different dynamic than retweets and favorites. To account for this, we assign a lower weight

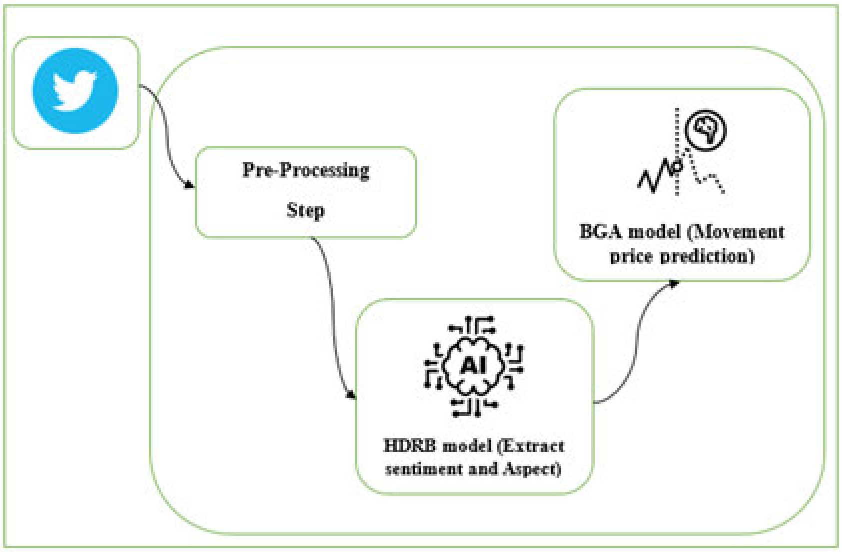


FIGURE 3. The general framework of the proposed model.

TABLE 2. Concept-LDA in extacrt Aspect.

1	Let K be the number of aspects.
2	For each document $d \in D$:
3	Choose the concept distribution $\theta\_d \sim Dirichlet(\alpha)$
4	For each word position n in document d:
5	Choose a concept $c\_n \sim Multinomial(\theta\_d)$
6	Choose a topic $z\_n \sim Multinomial(\beta\_c\_n)$
7	Choose a word $w\_n \sim Multinomial(\varphi\_z\_n)$

(0.5) to replies, acknowledging their relevance while avoiding potential skewing of the importance measure.

Next, the tweets are injected into the Concept-LDA model to extract the aspect. Concept-LDA is a probabilistic topic modeling technique extending LDA by incorporating concepts. In Concept-LDA, a concept represents a higher-level abstraction that captures the relationship between multiple topics. It introduces a concept matrix that relates topics to concepts, allowing for the simultaneous modeling of both topics and concepts. Mathematically, Concept-LDA can be represented as:

1) Given a corpus of documents D, where each document d is represented as a sequence of words w, Concept-LDA assigns topics z and concepts c to each word. The generative process of Concept-LDA can be expressed as follows pseudocode in Table 2.

Here, α and β are hyperparameters, and φ represents the word distribution for each topic. By modeling topics and concepts simultaneously, Concept-LDA enables the discovery of underlying concepts that connect related topics and provides a richer representation of the corpus. In this study, we calculated φ using GloVe because Concept-LDA can benefit from the semantic relationships and contextual information captured in these embeddings by utilizing pre-trained word embeddings. This can potentially improve the modeling accuracy and the representation of topics and concepts in Concept-LDA.

After extracting the aspects related to each tweet, the data is injected into the RoBERTa pre-trained model for sentiment analysis. Formally, Given an input sequence of tokens $X = [x_1, x_2, \ldots, x_n]$, RoBERTa aims to learn contextualized representations for each token. It applies a series of transformer layers to capture the dependencies and contextual information within the sequence. Each transformer layer shows as Eq (2):

$$H^{(l)} = TransformerLayer(H^{(l-1)}) \tag{2}$$

where $H^{(0)} = Embeddings(X)$ represents the initial token embeddings, and l denotes the layer index. The Transformer layer comprises two sub-modules: a multi-head self-attention mechanism and a position-wise feed-forward neural network. These sub-modules can be shown as Eq (3).

$$\begin{aligned} H\_att^{(l)} &= MultiHeadAttention(H^{(l-1)}), \\ H\_ffn^{(l)} &= PositionwiseFeedForward(H\_att^{(l)}). \end{aligned} \tag{3}$$

The MultiHeadAttention module computes a weighted combination of different self-attention heads, capturing various dependencies among the tokens. The Positionwise-FeedForward module independently applies a non-linear transformation to each position, allowing for more expressive representations.

The process is repeated for L layers, where L is the total number of transformer layers in the RoBERTa model. The final hidden state representation H_final is then used for downstream tasks such as sentiment analysis, text classification, named entity recognition, or question-answering.

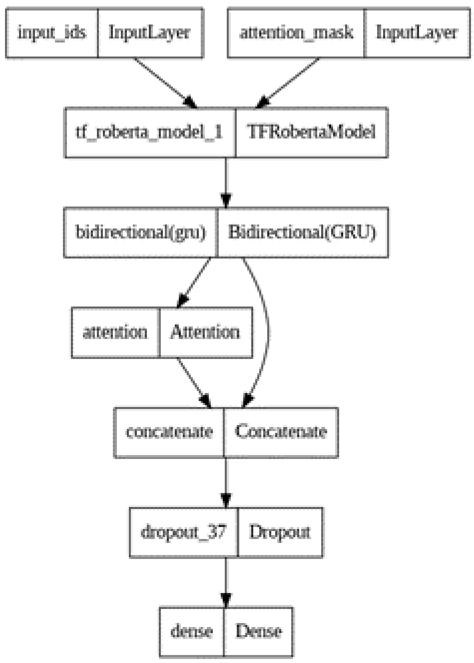


FIGURE 4. HDRB model.

Then, the extracted features are injected into BiGRU deep neural network. The reason for using BiGRU is the power of this method in processing sequences. Combining RoBERTa with BiGRU offers several advantages in natural language processing tasks. The integration of these models enhances the representation of learning capabilities and captures both local and global contexts within the input sequence. Mathematically, the advantage of combining RoBERTa with BiGRU can be summarized as follows:

- Given an input sequence of tokens $X = [x_1, x_2, \ldots, x_n]$, RoBERTa captures the contextualized representations of each token as $H\_roberta = RoBERTa(X)$. These representations provide rich semantic and syntactic information. The BiGRU model, on the other hand, introduces recurrent bidirectional connections to capture the sequential dependencies within the input sequence. It shows as Eq (4).

$$H\_bi\_gru = Bi - GRU(H\_roberta). \tag{4}$$

- Where H_bi_gru represents the hidden output states of the BiGRU model, by combining the RoBERTa representations with the BiGRU hidden states, we obtain a fused representation that benefits from both the contextualized information learned by RoBERTa and the sequential modeling of BiGRU. This fusion can be represented as Eq (5).

$$H\_fused = Concatenate(H\_roberta, H\_bi\_gru). \tag{5}$$

- The fused representation H_fused captures the local and global contextual information within the input sequence, allowing for a more comprehensive understanding of the text.

The features extracted by BiGRU are given to the attention layer. The attention layer can extract the essential features of the text and ignore other features. Then, the results are concatenated by the layer, and projection is performed. The dropout layer prevents the model from overfitting. Finally, the features are added to the Dense layer after the extrapolation layer, and the aspect, sentiment of the aspects, and the overall polarity of the tweet are calculated. This model is shown in Figure 4.

B. BGA MODEL

BiGRU combined model and attention layer have been used to predict the trend of cryptocurrencies. One advantage of using BiGRUs is that they are designed to capture long-term dependencies in sequential data. When predicting movement prices, it is essential to consider the history of the prices over time, as well as any external factors that may have influenced them. BiGRUs can model these long-term dependencies, allowing them to make more accurate predictions. Another advantage of BiGRUs is their ability to process input sequences in both forward and backward directions. In movement price prediction, the movement of prices can depend on both past and future events, so this bidirectional processing is beneficial. BiGRUs are also able to handle variable-length input sequences. In movement price prediction, the length of the historical price data used as input may vary depending on the time horizon of the forecast. BiGRUs can handle this variability, making accurate predictions regardless of the input sequence length.

Let's define the mathematical representation of the BGA model to predict the movement of cryptocurrency prices using historical price data and sentiment tweets. Given historical price data $P = [p_1, p_2, \ldots, p_n]$ and sentiment tweets $S = [s_1, s_2, \ldots, s_n]$, where p_i represents the price at time step i, and s_i represents the sentiment score of the corresponding tweet at time step i, we aim to predict the movement of cryptocurrency prices as either positive or negative. We first pass the historical price data and sentiment scores through a BiGRU layer to capture sequential dependencies and contextual information. The BiGRU layer can be represented as Eq (6):

$$H\_bi\_gru = BiGRU([P, S]) \tag{6}$$

Then, we apply an attention mechanism to focus on relevant information within the BiGRU hidden states. The attention layer computes a context vector based on the weighted combination of the BiGRU hidden conditions. The attention mechanism can be represented as Eq (7).

$$\begin{aligned} A &= Softmax(W\_att * \tanh(U\_att * H\_bi\_gru)) \\ C &= A * H\_bi\_gru \end{aligned} \tag{7}$$

Here, W_att and U_att are weight matrices used in the attention mechanism, *tanh* represents the hyperbolic tangent

TABLE 3. Pseudocode HDRB+BGA.

1. Extract tweets about cryptocurrency for 50 influencers
- For each influencer:
- Use API to extract tweets containing cryptocurrency-related keywords
2. Preprocess tweets
- Tokenize tweets
- Remove stop words
- Perform stemming/lemmatization
3. Calculate Importance Coefficient (IC) per tweet
- Use a formula to calculate IC based on factors such as retweets, likes, and replies
4. Extract aspect tweets with Concept-LDA
- Use a supervised learning algorithm such as Concept-LDA to extract aspect-based tweets
5. Determine sentiment for each tweet and aspect
- Use a sentiment analysis algorithm to determine the sentiment of each tweet and aspect
6. Determine overall sentiment with RoBERTa+BiGRU and attention mechanism (HDRB model)
- Use RoBERTa+BiGRU with an attention mechanism to determine the overall sentiment of each tweet and aspect
7. Calculate and merge sentiment for i'th days to j'th crypto
- Calculate the sentiment for each crypto by merging the sentiment of tweets for i'th days to j'th crypto
8. Extract historical price between [i..i+t] days
- Use a cryptocurrency price API to extract historical price data for the specified time range
9. Use BiGRU and attention mechanism to predict price movement
- Use a BiGRU with an attention mechanism to predict the price movement (positive or negative) for each crypto
- If the number of sentiment values for j'th crypto is not equal to the historical price data, use tweets volume to predict price movement.

function, and *Softmax* performs the softmax operation to obtain attention weights. The context vector C obtained from the attention layer is then passed through a dense layer to predict cryptocurrency price movement. The dense layer can be represented as Eq (8):

$$O = Dense(C) \tag{8}$$

To obtain a probabilistic output, we apply a sigmoid activation function to the output of the dense layer. This activation function maps the output to the range [0, 1], representing the probability of the price movement being positive or negative. The final prediction for the movement of cryptocurrency prices can be defined as $Prediction = Activation(O)$.

By combining the BiGRU layer and attention mechanism, the BGA model leverages historical price data and sentiment tweets to predict the movement of cryptocurrency prices as either positive or negative. The pseudocode of Table 3 shows the stages of this research, including the extraction and analysis of sentiments of tweets and the prediction of price trends.

TABLE 4. Evaluation metrics.

$F1-score = 2 * (precision * recall) / (precision + recall)$
$Accuarcy = number\ of\ accurattely\ classified\ examples / Total\ examples$
$ROC-AUC = \int TP(FP)\, dFP$
$Akaike\ Information\ Criterion\ (AIC) = 2 * k - 2 * \ln(L)$

TABLE 5. SemEval dataset aspect-based task.

Dataset	Train			Validation			Test		
	Pos	Neg	Neu	Pos	Neg	Neu	Pos	Neg	Neu
Sem15	808	29	228	147	5	44	340	28	195
Sem16	706	454	406	191	9	60	474	29	127

The pseudocode provides a model for predicting cryptocurrency price movements based on historical price data and sentiment analysis of tweets. The pseudocode starts by extracting tweets related to cryptocurrencies from 52 influencers. These tweets are then preprocessed to prepare them for analysis. Importance coefficients (IC) are calculated for each tweet to determine their significance.

Next, Concept-LDA is applied to extract aspect tweets, focusing on specific topics or aspects related to cryptocurrencies. Sentiment analysis is performed on each tweet and aspect using the RoBERTa+BiGRU model with an attention mechanism (HDRB model). This step enables the model to capture the sentiment associated with individual tweets and aspects.

The overall sentiment for each cryptocurrency is then calculated by aggregating the sentiment scores of associated tweets and aspects. Historical price data is extracted within a specified time frame to provide context for the prediction task.

Finally, BiGRU with an attention mechanism is employed to predict each cryptocurrency's price movement (positive or negative). The sentiment information and tweet volume are input to the model if required. The predicted price movements are outputted for each cryptocurrency.

III. EXPERIMENTAL RESULT

This research uses Accuracy, F1-Score, and Roc-AUC criteria to evaluate HDRB and BGA models. Table 4 shows the relationships between these criteria.

True Positive (TP) equals the number of cases the model correctly predicted positive, and False Positive (FP) equals the number of cases the model incorrectly predicted positive.

The Akaike Information Criterion (AIC) is a statistical metric used for model selection, particularly in comparing the goodness of fit of different models. It balances the trade-off between the complexity of a model and its fit to the data. In AIC, k is the number of parameters in the model (model complexity), and $\ln(L)$ is the natural logarithm of the likelihood function of the data given the model [29].

TABLE 6. Configuration setup HDRB.

Tokenizer and pre-trained model	Twitter-roberta-base
Dropout	One layer, 0.5 rate
BiGRU	128 units
Attention	Self-Attention
Activations	Relu, Softmax
Optimizers	Adam(learning_rate=1e-5)
Loss	Categorical cross-entropy
Epochs	50
EarlyStopping	monitor='val_loss', patience=5

TABLE 7. Configuration setup BGA.

Dropout	One layer, 0.2 rate
BiGRU	64 units
Attention	Self-Attention
Activations	Relu, Sigmoid
Pooling	GlobalAveragePooling1D
Optimizers	Adam(learning_rate=1e-5)
Loss	Binary cross-entropy
Epochs	100
EarlyStopping	monitor='val_loss', patience=5

A. DATASETS

To evaluate the model like Huang et al. [30], SemEval datasets, which include public tweets in different areas such as restaurants and laptops, have been used. The datasets evaluated in this research are SemEval-2015 Task 12 and SemEval-2016 Task 5. These datasets are introduced in Table 5.

Tweets were collected between January 01, 2020, and June 20, 2023. Due to the drastic price changes and political and social events from 2020 until now, tweets related to 2023 have been considered to evaluate the price trend of cryptocurrencies. Of the 29,860 tweets after cleaning and removing public tweets (unrelated to cryptocurrencies), 16,512 tweets remained. Most tweets are related to three famous cryptocurrencies: Bitcoin, Ethereum, and Binance. To predict the price trend, the historical price of six cryptocurrencies, Bitcoin, Ethereum, Binance, Ripple, Dogecoin, Cardano, and Solana, has been extracted by YahooFiance API in the period from January 1, 2023, to June 12, 2023. Historical prices include open, low, high, close, and volume prices. Then, the three variables of the compound value of tweet polarity, tweet polarity, and daily price trend have been added to these prices. The two compound columns of the polarity of the tweets are the polarity of the tweet output of the HDRB model and the column of the daily price trend, which includes positive or negative values, shows the price trend of the cryptocurrency compared to the previous day.

B. CONFIG MODELS

In the HDRB model, the config described in Table 6 is used.

TABLE 8. Configuration setup grid-search.

Tokenizer and pre-trained model	Twitter-roberta-base
Dropout	$[0.2, 0.5, 1.0]$
BiGRU units	$[64, 128, 256]$
Learning rate	$[1e-5, 1e-4, 1e-3]$

In the HDRB model, the pre-trained model and RoBERTa tokenizer (Twitter version) are used. Also, BiGRU and Attention Layer are used in this model to interpret the texts in two directions and focus on the essential parts of the text. In addition, the EarlyStopping technique with five attempts has been used to prevent overfit and save the best model. The settings in Table 7 have also been used for the BGA model.

C. RESULTS

In this section, the efficiency of the HDRB model is examined in the first part, and the efficiency of the BGA model is studied in the third part.

1) FIND HYPER-PARAMTERS

To find the best values of hyper-parameters of the HDRB model, we have used grid-search to check the three parameters: Learning rate, dropout, and the number of BiGRU deep neural network neurons. The tested values of these parameters are listed in Table 8.

After running the HDRB model in 50 epochs, it was determined that the best results related to Learning rate= 1e-5, dropout = 0.5, and Bigruunit = 64. These values are also shown in table 6. In addition, the values related to the BGA model in Table 7 were also confirmed with a separate test.

2) HDRB MODEL RESULTS

By executing the HDRB model on the Sem2015 and Sem2016 datasets, the results of Table 9 have been obtained. These results are promising compared to similar studies [30]. We have reached the HDRB model with the following studies:

- CPA-SA [30]: The model employs two different contextual position weight functions to adjust the importance of contextual words based on their positions relative to the aspect words in sentences. This approach helps mitigate the impact of word imbalances on sentiment judgments caused by differing word counts on either side of the aspect words. By utilizing bidirectional GRU layers at both the single-sentence and multiple-sentence levels, the model captures the contextual influence of each sentence in the document on the sentiment polarity of individual sentences. Additionally, the paper addresses the class imbalance issue in sentiment analysis by analyzing the distribution properties of challenging samples and introducing a novel loss function.
- TAS-BERT [31]: The method utilizes a pre-trained language model to capture the dependency between targets, aspects, and sentiment for prediction.

TABLE 9. Evaluation of HDRB.

Models	Res2015		Res2016	
	ACC	F1	ACC	F1
CPA-SA	79.57	60.26	**88.80**	71.47
TAS-BERT	-	70.42	-	70.66
BERT-pair-NLI-B	-	70.78	-	80.25
ASGCN	79.34	60.78	88.69	66.64
THA + STN		71.46	-	73.61
XL-NET+ BiGRU	76.25	71.28	80.32	78.25
HDRB	**87.61**	**87.96**	85.14	**82.86**

TABLE 10. Evaluation of HDRB with AIC.

Models	AIC	K	Ln(L)
Model 1(Base BERT)	180,671,682.29	123,336,067	238.30
Model 2(Xl-NET+BiGRU)	190,671,695.86		225.15
HDRB	114,816,923.54	121,405,120	211.85

- BERT-pair-NLI-B [31]: This model utilizes BERT for aspect sentiment detection. We used the public code of this model to assess its performance on aspect sentiment using the Res15 and Res16 datasets.
- ASGCN [30]: ASGCN leverages a Graph Convolutional Network (GCN) to capture semantic information and establish connections between distant elements (Aspect-level sentiment paper).
- THA+STN [31]: THA + STN is a neural model for targeted sentiment detection (TD). This model incorporates bi-linear attention and fully connected (FC) layers.
- XL-NET+ BiGRU: This paper combines Xl-NET pre-trained and BiGRU deep neural networks for aspect sentiment analysis.

Table 9 shows that the HDRB model has a stable performance on both Sem2015 and Sem2016 datasets, and the rates of both F1-score and accuracy are very close. While in different models, these two criteria have a relatively large distance from each other, which shows that these models have weaknesses in precision or recall criteria. In addition, the performance of the HDRB model on both datasets is close to each other and at an optimal level. In contrast, in other models, their performance is favorable in one dataset and decreased in the different datasets. Figure 5 shows the ROC-AUC criterion to understand the HDRB model's efficiency better.

In Figure 5, class "0" equals positive polarity, class "1" equals neutral polarity, and class "2" equals negative polarity. In the Res2015 and 2016 datasets, the neutral class has the least number of tweets. Figure (5) shows that the model has recognized all three polarities of positive, negative, and neutral for different aspects of the text. Also, the benchmark rate of ROC-AUC for the Res2015 dataset is equal to 79.50, and for Res2016 is equal to 82.93.

Also, to prove the effectiveness of the HDRB model, we compared three HDRB models, Xl-NET+BiGRU and Base BERT, using the AIC criterion on the Res2016 dataset. The results are shown in Table 10.

TABLE 11. Full result evaluation of HDRB.

Models	Datasets					
HDRB	Res2015			Res2016		
	Recall	Precision	F1-score	Recall	Precision	F1-score
	87.30	**88.63**	**87.96**	82.21	**83.53**	82.86
CNN_stacked_BiLSTM_Attn	81	81	81	82	82	82
CNN_stacked_BiLSTM_Multi_Attn	81	81	81	**83**	83	**83**

TABLE 12. HDRB execution time.

Execute time (ms)	Res2015	Epochs	Res2016	Epochs
	6.25e5	31	7.41e5	39

Upon examining these AIC values in Table 10, several observations can be made:

- Comparison within Models: The HDRB Model has significantly lower AIC values than Models 1 and 2. This value indicates that the former set of models better fits the observed data, considering both the goodness of fit and model complexity.
- Model Complexity: The HDRB Model has lower AIC values despite the introduction of model complexity. This value suggests that the increased complexity is warranted by the improved fit to the data. However, the increase in complexity should still be justified based on domain knowledge and theoretical understanding.
- Outliers and Residuals: It's essential to investigate whether the models with lower AIC values effectively capture patterns in the data, including potential outliers. Models with lower AIC values might better accommodate deviations from the expected behavior.
- Negative Log-Likelihood (ln(L)): The negative log-likelihood values provide insights into how well the models predict the observed outcomes. Lower values of ln(L) indicate better model predictions.

In conclusion, the AIC values suggest that the HDRB model offers a better balance between fit to the data and model complexity than Models 1 and 2. Then, models with lower AIC values are preferred due to their improved balance between fit and complexity. We compare the HDRB model regarding recall, precision, and f1-score criteria with the results of Trueman et al. [32] The authors proposed the hybrid model, a convolutional stacked BiLSTM with a multiplicative/single attention mechanism. The convolutional layers capture local patterns and features from the input data, while the bidirectional LSTM allows the model to consider past and future contexts. The multiplicative attention mechanism helps the model focus on relevant aspects and sentiments in the input. The results are shown in Table 11.

The results of Table 11 show that the HDRB model has an acceptable performance on both datasets, and the model's efficiency has not decreased with the reduction of the data volume in the Res2015 dataset. In addition, the HDRB model

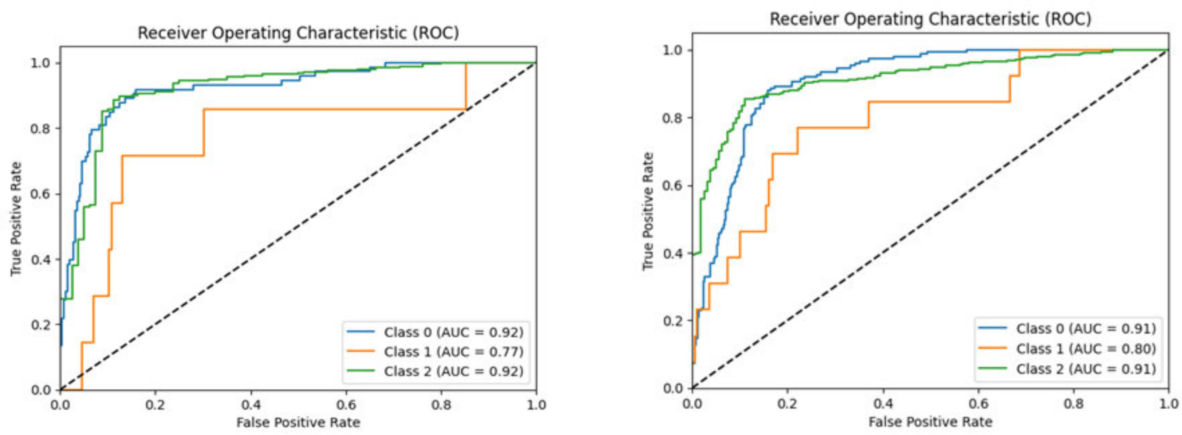


FIGURE 5. Efficiency of HDRB model with ROC-AUC criterion on Res 2015, 2016 datasets.

is not sensitive to the tweet context. This model can be adapted by changing the context of the tweet (from general to specialized, such as cryptocurrency) and can be used as a backbone in the sentiment analysis of all types of texts, as shown in the next section. Table 12 also shows the duration of the model execution to converge.

As shown in Table 12, The HDRB model has converged in 31 epochs (Average) with a duration of 625,000ms on the res2015 dataset and in 39 epochs with 741000ms on the res2016 dataset. The duration and the low number of epochs of the HDRB model mean that the model is well fine-tuned; one reason is the choices of appropriate hyper-parameters and another reason for the existence of EarlyStopping to prevent overfitting of the model and to save on the consumption of hardware resources.

3) BGA MODEL RESULTS

In this research, the price trends of Bitcoin, Ethereum, Binance, Ripple, Dogecoin, Cardano, and Solana have been predicted using sentiment analysis and historical prices in five-day intervals. Figure 6 shows the relationship between the price and sentiments of Bitcoin and Ethereum between January 1, 2023, and June 12, 2023.

As seen in Figure 6, there is a significant relationship between the price of Bitcoin and Ethereum and the sentiments of tweets. We used the correlation criterion to confirm the relationship between sentiment analysis and the price with Eq (9). In state-of-art [33], [34], [35], the relationship between the closing price of cryptocurrencies and the level of sentiment has been proven using the Pearson correlation.

$$r = \left(\Sigma((X_i - \hat{X}) * (Y_i - \hat{Y}))\right) / \left(sqrt(\Sigma(X_i - \hat{X})^\wedge 2) * sqrt(\Sigma(Y_i - \hat{Y})^\wedge 2)\right) \quad (9)$$

X_i and Y_i indicate the numerical value of the i-th row of the two fields whose correlation is compared. Also, $\hat{X}$ and $\hat{Y}$ shows the average values of two variables, X and Y. Based on this, the sentiment polarity and compound columns correlate with the price trend column of cryptocurrencies with a correlation of nearly 70%. The lowest correlation is related to the Volume, Low, and High columns. Also, in this research, based on the study of Burnie et al. [36], Spearman's rank correlation has been used to prove the correlation between sentiment and close price. The reason for using the Spearman coefficient is that the two variables of the sentiment score and the closed price do not always follow a linear relationship and an utterly normal distribution, so using the Spearman coefficient in these situations is recommended. Using Spearman's coefficient, like Pearson's, shows that emotion ratings and close price have a significant relationship. After the Spearman test, it can be seen that the correlation coefficient is 0.642 and p=0.032. Therefore, because P<0.05, the null hypothesis is rejected, so it can be concluded that there is a statistically significant relationship between the sentiment test and Close Price.

TABLE 13. Select best days to predict trend.

Days	3	4	5	6	7	8	9	10
Acc	85.61	85.72	86.15	84.25	80.6	71.5	66.32	54.21

Therefore, in this research, the BGA model, compound, sentiment, and close and open columns are used to predict the price trend of cryptocurrencies. To obtain the best time frame for predicting the price trend of cryptocurrencies, the BGA model was examined in time frames of 3-10 days in 100 epochs with the accuracy criterion for Bitcoin cryptocurrency, and these results are shown in Table 13.

As shown in Table 13, the best time frame for predicting the price trend based on sentiment analysis of tweets is 3 to 5 days. The accuracy criterion is chosen because we know that the cryptocurrency price trend is positive or negative on the I+k days that k∈[3 − 10] compared to the 1st day. Figure (7) shows the loss and accuracy rate of the BGA model for predicting the price trend of Bitcoin, Ethereum, Binance, Ripple, Dogecoin, Cardano, and Solana.

Figure 7 shows that the BGA model has achieved acceptable accuracy and loss in the test and train data.

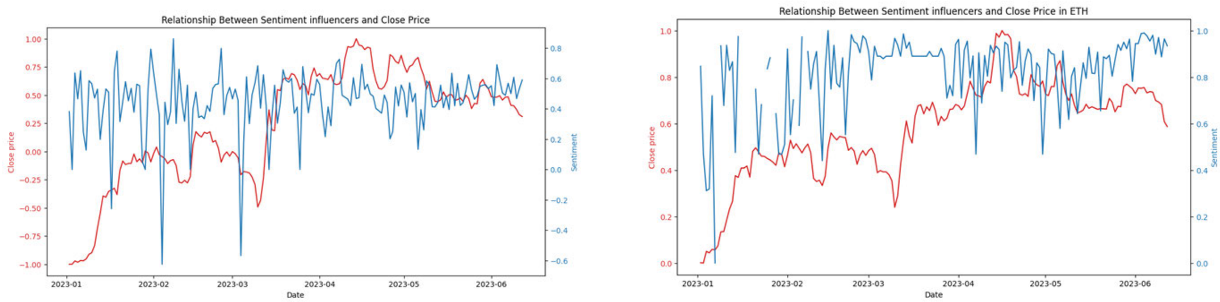


FIGURE 6. The relationship between sentiment analysis and the price of Bitcoin and Ethereum.

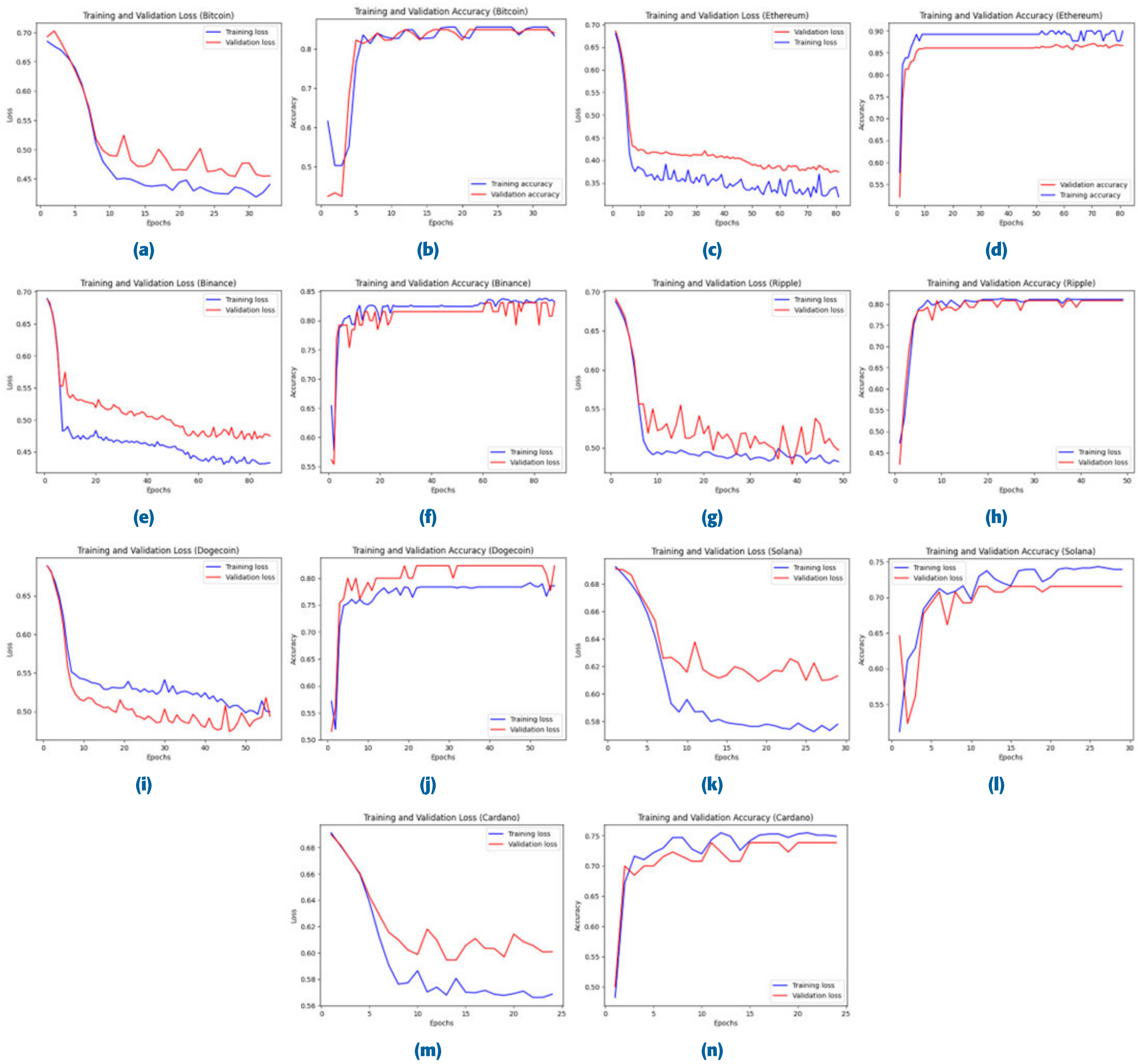


FIGURE 7. Loss rate and Accuracy in the BGA model.

The results obtained for predicting the price of the studied cryptocurrencies are encouraging, considering the fluctuating nature of the cryptocurrency market. However, as shown in the graphs in Figure 7, the BGA model has obtained higher accuracy in predicting Bitcoin and Ethereum due to the lower price fluctuations and the higher volume of novelties of these two cryptocurrencies compared to other cryptocurrencies. The lower price fluctuations and the higher volume of analyses make the analysis provided by experts regarding the future price trend of Bitcoin and Ethereum more reliable than other cryptocurrencies. To better understand the accuracy of the BGA model, the ROC-AUC diagram

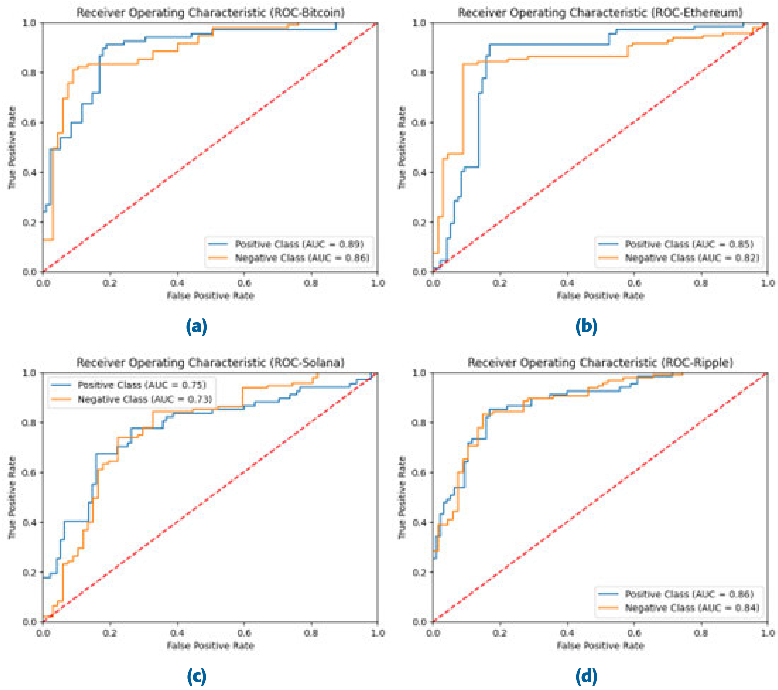


FIGURE 8. ROC-AUC diagram of the BGA model for Ethereum, Bitcoin, Solana, and Ripple cryptocurrencies.

TABLE 14. Evaluation The BGA model.

Crypto	Accuracy	F1-score	ROC-AUC (avg)
Bitcoin	87.25	86.14	85.25
Ethereum	86.41	82.81	83.61
Binance	82.46	81.24	81.38
Ripple	80.95	80.06	75.88
Dogecoin	78.35	74.82	72.49
Solana	73.07	70.23	70.59
Cardano	78.35	74.82	75.88

of this model is shown in Figure 8 on test data for four cryptocurrencies: Bitcoin, Ethereum, Solana, and Ripple.

Figure 8 shows that the positive and negative classes that indicate the upward or downward trend of the cryptocurrency have been identified by the BGA model with a suitable balance. Table 14 shows the BGA model's Accuracy, F1-score, and ROC-AUC values.

Table 14 shows that the BGA model has been able to predict the price trend of cryptocurrencies using historical price and sentiment information. Also, Table 14 shows a significant relationship between the sentiment analysis of experts in cryptocurrencies and their price trends.

IV. CONCLUSION

Sentiment analysis of influencer tweets is essential because these people's opinions can directly impact the cryptocurrency market trend. However, the rich tagged dataset is very rare in the cryptocurrency domain. Therefore, this study uses the RoBERTa pre-trained neural network to transfer learning and adapt the HDRB model to the sentiment analysis task based on cryptocurrencies. The HDRB model uses a combination of RoBERTa pre-trained network, BiGRU deep learning network, and attention layer for facet-based sentiment analysis of cryptocurrencies. Also, the HDRB model's hyperparameters, including dropout rate, learning rate, and the number of Gru-units, have been evaluated and confirmed by the gridsearch method. In addition, the results of the AIC criterion show that the HDRB model, compared to Xl-NET+BiGRU and Base-BERT, could better fit the observed data, considering both the goodness of fit and model complexity. The results show that the HDRB model has improved the accuracy and f1-score criteria on SemEval 2015 and 2016 datasets compared to previous research. After analyzing the sentiments of influencers' tweets, the polarity of sentiments along with the historical price is injected into the combined BGA model (a combination of attention layer and BiGRU) to predict the price trend of seven cryptocurrencies: Bitcoin, Ethereum, Binance, Ripple, Dogecoin, Cardano and Solana. The results of implementing the BGA model in 3 to 10 days show that the most accurate prediction of the price trend of cryptocurrencies is related to the period of 3 to 5 days. In general, the results of this research show that influencer comments on social networks such as Twitter can affect the price trend of cryptocurrencies. Also, these comments can help traders choose the right stocks and manage their investment portfolio. In future research, we will analyze the sentiments of influencers on other social platforms such as Instagram, YouTube, and Reddit and combine these sentiments with Google Trend and cryptocurrency blockchain information.

REFERENCES

[1] A. Keller and M. Scholz, "Trading on cryptocurrency markets: Analyzing the behavior of Bitcoin investors," in *Proc. Int. Conf. Inf. Syst.*, 2019.

[2] M. Frohmann, M. Karner, S. Khudoyan, R. Wagner, and M. Schedl, "Predicting the price of Bitcoin using sentiment-enriched time series forecasting," *Big Data Cognit. Comput.*, vol. 7, no. 3, p. 137, Jul. 2023.

[3] W. Ding, J. Li, W. Mao, Z. Meng, and Z. Shen, "Rolling bearing remaining useful life prediction based on dilated causal convolutional DenseNet and an exponential model," *Rel. Eng. Syst. Saf.*, vol. 232, Apr. 2023, Art. no. 109072.

[4] A. Tayade, S. Patil, V. Phalle, F. Kazi, and S. Powar, "Remaining Useful Life (RUL) prediction of bearing by using regression model and Principal Component Analysis (PCA) technique," *Vibroeng. Proc.*, vol. 23, pp. 30–36, Apr. 2019.

[5] W. Song, H. Liu, and E. Zio, "Long-range dependence and heavy tail characteristics for remaining useful life prediction in rolling bearing degradation," *Appl. Math. Model.*, vol. 102, pp. 268–284, Feb. 2022.

[6] Q. Han, J. Wu, and Z. Zheng, "Long-range dependence, multi-fractality and volume-return causality of ether market," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 30, no. 1, Jan. 2020, Art. no. 011101.

[7] A. Arratia and A. X. López-Barrantes, "Do Google Trends forecast Bitcoins? Stylized facts and statistical evidence," *J. Banking Financial Technol.*, vol. 5, pp. 45–57, Mar. 2021.

[8] A. Nazir, Y. Rao, L. Wu, and L. Sun, "Issues and challenges of aspect-based sentiment analysis: A comprehensive survey," *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, pp. 845–863, Apr. 2022.

[9] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowl.-Based Syst.*, vol. 108, pp. 42–49, Sep. 2016.

[10] T. U. Tran, H. T. T. Hoang, and H. X. Huynh, "Aspect extraction with bidirectional GRU and CRF," in *Proc. IEEE-RIVF Int. Conf. Comput. Commun. Technol. (RIVF)*, Mar. 2019, pp. 1–5.

[11] R. Bensoltane and T. Zaki, "Towards Arabic aspect-based sentiment analysis: A transfer learning-based approach," *Social Netw. Anal. Mining*, vol. 12, no. 1, pp. 1–16, Dec. 2022.

[12] J. Tao and X. Fang, "Toward multi-label sentiment analysis: A transfer learning based approach," *J. Big Data*, vol. 7, no. 1, pp. 1–26, Dec. 2020.

[13] G. Xu, Z. Zhang, T. Zhang, S. Yu, Y. Meng, and S. Chen, "Aspect-level sentiment classification based on attention-BiLSTM model and transfer learning," *Knowl.-Based Syst.*, vol. 245, Jun. 2022, Art. no. 108586.

[14] H. Li, B. X. B. Yu, G. Li, and H. Gao, "Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews," *Tourism Manage.*, vol. 96, Jun. 2023, Art. no. 104707.

[15] F. Tang, L. Fu, B. Yao, and W. Xu, "Aspect based fine-grained sentiment analysis for online reviews," *Inf. Sci.*, vol. 488, pp. 190–204, Jul. 2019.

[16] A. H. Sweidan, N. El-Bendary, and H. Al-Feel, "Sentence-level aspect-based sentiment analysis for classifying adverse drug reactions (ADRs) using hybrid ontology-XLNet transfer learning," *IEEE Access*, vol. 9, pp. 90828–90846, 2021.

[17] Ş. Ö. Birim and F. E. Sönmez, "Social sentiment analysis for prediction of cryptocurrency prices using neuro-fuzzy techniques," in *Proc. Int. Conf. Intell. Fuzzy Syst.* Cham, Switzerland: Springer, 2022, pp. 606–616.

[18] E. Loginova, W. K. Tsang, G. van Heijningen, L.-P. Kerkhove, and D. F. Benoit, "Forecasting directional Bitcoin price returns using aspect-based sentiment analysis on online text data," *Mach. Learn.*, pp. 1–24, Nov. 2021.

[19] S. Pathak and A. Kakkar, "Cryptocurrency price prediction based on historical data and social media sentiment analysis," in *Proc. 7th Innov. Comput. Sci. Eng. (ICICSE)*, 2020, pp. 47–55.

[20] J. Davchev, K. Mishev, I. Vodenska, L. Chitkushev, and D. Trajanov, "Bitcoin price prediction using transfer learning on financial micro-blogs," in *Proc. 16th Annu. Int. Conf. Comput. Sci. Educ. Comput. Sci.*, 2020, pp. 1–6.

[21] B. Xin and W. Peng, "Prediction for chaotic time series-based AE-CNN and transfer learning," *Complexity*, vol. 2020, Sep. 2020, Art. no. 2680480.

[22] C. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 8, no. 1, 2014, pp. 216–225.

[23] X. Huang, W. Zhang, X. Tang, M. Zhang, J. Surbiryala, V. Iosifidis, Z. Liu, and J. Zhang, "LSTM based sentiment analysis for cryptocurrency prediction," in *Proc. 26th Int. Conf. Database Syst. Adv. Appl. (DASFAA)*. Cham, Switzerland: Springer, Taipei, Taiwan, 2021, pp. 617–621.

[24] J. Abraham, D. Higdon, J. Nelson, and J. Ibarra, "Cryptocurrency price prediction using tweet volumes and sentiment analysis," *SMU Data Sci. Rev.*, vol. 1, no. 3, pp. 1–22, 2018.

[25] F. Valencia, A. Gómez-Espinosa, and B. Valdés-Aguirre, "Price movement prediction of cryptocurrencies using sentiment analysis and machine learning," *Entropy*, vol. 21, no. 6, p. 589, Jun. 2019.

[26] R. Parekh, N. P. Patel, N. Thakkar, R. Gupta, S. Tanwar, G. Sharma, I. E. Davidson, and R. Sharma, "DL-GuesS: Deep learning and sentiment analysis-based cryptocurrency price prediction," *IEEE Access*, vol. 10, pp. 35398–35409, 2022.

[27] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBERTa: A robustly optimized BERT pretraining approach," 2019, *arXiv:1907.11692*.

[28] E. Ekinci and S. İlhan Omurca, "Concept-LDA: Incorporating Babelfy into LDA for aspect extraction," *J. Inf. Sci.*, vol. 46, no. 3, pp. 406–418, Jun. 2020.

[29] Y. Fan, J. Chen, G. Shirkey, R. John, S. R. Wu, H. Park, and C. Shao, "Applications of structural equation modeling (SEM) in ecological studies: An updated review," *Ecol. Processes*, vol. 5, no. 1, pp. 1–12, Dec. 2016.

[30] B. Huang, R. Guo, Y. Zhu, Z. Fang, G. Zeng, J. Liu, Y. Wang, H. Fujita, and Z. Shi, "Aspect-level sentiment analysis with aspect-specific context position information," *Knowl.-Based Syst.*, vol. 243, May 2022, Art. no. 108473.

[31] H. Wan, Y. Yang, J. Du, Y. Liu, K. Qi, and J. Z. Pan, "Target-aspect-sentiment joint detection for aspect-based sentiment analysis," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 5, 2020, pp. 9122–9129.

[32] K. J. Ashok, T. E. Trueman, and E. Cambria, "A convolutional stacked bidirectional LSTM with a multiplicative attention mechanism for aspect category and sentiment detection," *Cognit. Comput.*, vol. 13, no. 6, pp. 1423–1432, Nov. 2021.

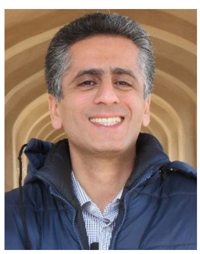
[33] T. R. Li, A. S. Chamrajnagar, X. R. Fong, N. R. Rizik, and F. Fu, "Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model," *Frontiers Phys.*, vol. 7, p. 98, Jul. 2019.

[34] A. Raheman, A. Kolonin, I. Fridkins, I. Ansari, and M. Vishwas, "Social media sentiment analysis for cryptocurrency market prediction," 2022, *arXiv:2204.10185*.

[35] M. Trigka, A. Kanavos, E. Dritsas, G. Vonitsanos, and P. Mylonas, "The predictive power of a Twitter user's profile on cryptocurrency popularity," *Big Data Cognit. Comput.*, vol. 6, no. 2, p. 59, May 2022.

[36] A. Burnie and E. Yilmaz, "Social media and Bitcoin metrics: Which words matter," *Roy. Soc. Open Sci.*, vol. 6, no. 10, Oct. 2019, Art. no. 191068.

KIA JAHANBIN received the associate degree from the Shahid Bahonar Faculty of Engineering and Technology, Shiraz, and the bachelor's degrees in software engineering and information technology engineering from Islamic Azad University, Kerman Branch, with a focus on computer networks. He is currently pursuing the Ph.D. degree in computer engineering, major in software. His research interests include sentiment analysis and opinion mining in social networks, machine learning, and deep learning. He is a continuous member of the Iran Computer Association and the Iran Informatics Association and a Volunteer Reviewer in IEEE Access.



MOHAMMAD ALI ZARE CHAHOOKI is currently an Associate Professor with the Department of Computer Engineering, Yazd University. Since 2015, he has been an Investor with Idekav, a platform for marketing in Telegram. Telegram is a messenger and social network with more than 550 million users. Idkav is now a platform for sending requests from employers to find target customers among 200 million users in Telegram groups. His research interests include machine learning in software engineering, image retrieval, and text mining. His academic page is: yazd.ac.ir/en/people/chahooki.