Crypto-sentiment Detection in Malay Text Using Language Models with an Attention Mechanism

Nur Azmina Mohamad Zamani [1] (ID) **, Norhaslinda Kamaruddin** [2]* (ID)

[1] *College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM), Selangor, Malaysia*
[1]azmina@uitm.edu.my, namz.ina@gmail.com

[2] *Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Universiti Teknologi MARA (UiTM), Selangor, Malaysia*
[2]norhaslinda@tmsk.uitm.edu.my

Abstract

Background: Due to the increased interest in cryptocurrencies, opinions on cryptocurrency-related topics are shared on news and social media. The enormous amount of sentiment data that is frequently released makes data processing and analytics on such important issues more challenging. In addition, the present sentiment models in the cryptocurrency domain are primarily focused on English with minimal work on Malay language, further complicating problems.

Objective: The performance of the sentiment regression model to forecast sentiment scores for Malay news and tweets is examined in this study.

Methods: Malay news headlines and tweets on Bitcoin and Ethereum are used as the input. A hybrid Generalized Autoregressive Pretraining for Language Understanding (XLNet) language model in combination with Bidirectional-Gated Recurrent Unit (Bi-GRU) deep learning model is applied in the proposed sentiment regression implementation. The effectiveness of the proposed sentiment regression model is also investigated using the multi-head self-attention mechanism. Then, a comparison analysis using Bidirectional Encoder Representations from Transformers (BERT) is carried out.

Results: The experimental results demonstrate that the number of attention heads is vital in improving the XLNet-GRU sentiment model performance. There are slight improvements of 0.03 in the adjusted R^2 values with an average MAE of 0.163 (Malay news) and 0.174 (Malay tweets). In addition, an average RMSE of 0.267 and 0.255 were obtained respectively for Malay news and tweets, which show that the proposed XLNet-GRU sentiment model outperforms the BERT sentiment model with lesser prediction errors.

Conclusion: The proposed model contributes to predicting sentiment on cryptocurrency. Moreover, this study also introduced two carefully curated Malay corpora, CryptoSentiNews-Malay and CryptoSentiTweets-Malay, which are extracted from news and tweets, respectively. Further works to enhance Malay news and tweets corpora on cryptocurrency-related issues will be expended with implementing the proposed XLNet Bi-GRU deep learning model for greater financial insight.

Keywords: Cryptocurrency, Deep learning model, Malay text, Sentiment analysis, Sentiment regression model

Article history: Received 28 May 2023, first decision 2 August 2023, accepted 1 September 2023, available online 28 October 2023

I. INTRODUCTION

Currency is an essential component of financial transactions. According to history, money has evolved from its most primitive forms—the barter trading system—to the current money system, known as fiat money. Trade is now more convenient than before thanks to this developed currency system, which can be utilized both physically and digitally via online transfer [1]. Over the years, currency has evolved to a higher level via the introduction of a cryptocurrency known as Bitcoin, a digital currency introduced by Nakamoto [2] in 2008. Malaysia started to adopt a cryptocurrency system in 2019, following the policies and regulations of the Securities Commission Malaysia (SC) and Bank Negara Malaysia (BNM) [3], [4]. To support the rising public interest among Malaysian merchants in the cryptocurrency industry [4], the Malaysian government has approved the use of a digital exchange platform named "Luno" to enable Malaysians to trade any of the nine cryptocurrency coins that are now supported, including Bitcoin, Ethereum, Ripple, Litecoin, Cardano, Solana, Chainlink, Uniswap, and Bitcoin Cash [5].

Over the years, cryptocurrency market is highly influenced by cryptocurrency enthusiasts' comments and social media presence. In which, these enthusiasts make remarks, and the sentiment of the remark will sway the

---

* Corresponding author

cryptocurrency's market behaviour, such as to sell, buy or hold the cryptocurrency. However, works focusing on impact of sentiment on cryptocurrency price prediction are limited. To complicate matters, most recent research works [6]–[9] concentrate on sentiment analysis of English text for classification [8], [11], [12] and regression [13], [14] tasks. Regression is not popular due to subjectivity of sentiment polarity determination. In addition, it can be observed that resources for the Malay language are few [10], and there are essentially no sentiment-labeled Malay corpora in the cryptocurrency domain, making it challenging to create effective Malay sentiment models.

This study aims to investigate the impact of sentiment on Malay cryptocurrency text using regression model on two newly curated Malay sentiment corpora of Bitcoin and Ethereum news headlines and tweets. First, a hybrid XLNet language model with GRU deep learning model is proposed to model the sentiment regression that can precisely distinguish sentiment levels based on a numeric scale [10], [15], [16]. In addition, the performance of the proposed sentiment regression model is then tested using a multi-head self-attention mechanism [17]. Subsequently, the performance of the proposed sentiment model is compared against a benchmark, the Bidirectional Encoder Representations from Transformers (BERT) [18]. BERT is a language model commonly applied in several current research studies on cryptocurrency [8], [19], [20] and is also available for Malay text.

This paper is organized as follows—Section II reviews sentiment analysis methods used in the existing cryptocurrency-related research studies. The proposed XLNet-GRU sentiment regression model methodology is presented in Section III. The construction of the Malay corpus on the news and tweets are also outlined. Then, experimental results are discussed in Section IV. Finally, Section V concludes and provides potential future research directions.

II. LITERATURE REVIEW

Positive and negative news have been observed to influence cryptocurrency values and movements, particularly during a bubble period [21]. Cryptocurrency prices typically decline when news or social media posts have a negative tone and vice versa [22]. One of the earliest works is presented in [23], which used IBM Watson and Stanford CoreNLP to analyze the sentiment of Reddit and Twitter users. It is reported that sentiment captured from both social media sources contributed to cryptocurrency price change in the market. Therefore, it is worthwhile to investigate how sentiment affects the price of cryptocurrencies.

The existing sentiment analysis methods can be predominantly categorized into two categories, namely the lexicon-based approach and the machine learning approach, which also incorporates deep learning. Prior research studies mainly focused on the lexicon-based approach for sentiment analysis on cryptocurrency-related words or sentences. Examples of lexicon-based approaches are the application of existing financial dictionaries [24], [25] or sentiment-based dictionaries such as the Affective Norms for English Words (ANEW) [26] in order to determine the sentiment polarity of each word or sentence [27]–[30].

In more recent work, Valence Aware Dictionary and Sentiment Reasoner (VADER) [31] is applied to label the sentiment texts in English based on four types of scoring (i.e., positive, negative, neutral, and compound score). The VADER positive and negative scores were used in [32]–[34] for online forums and tweets, while the VADER compound score for tweets was implemented in [35], [36]. The experimental results demonstrate that the sentiment features are statistically significant for the task and can be used as the sentiment features for cryptocurrency price forecasting. Another comparable tool for sentiment analysis is TextBlob [37], [38], which determines the polarity and subjectivity of a sentence. The range of polarity values is from -1 to +1, with -1 denoting a negative sentiment and +1 denoting a positive sentiment. Subjectivity also ranges between 0 and 1, with 0 denoting very objective and 1 denoting very subjective. A higher subjectivity suggests that the text contains more subjective information, such as personal opinion, than factual information [37]. Unfortunately, despite being well-developed in English, these technologies are not readily available for Malay text.

The use of machine learning methods for sentiment analysis is also feasible. However, such an approach requires sentiment-labeled corpora and word embeddings to be available. For example, Lamon et al. [39] implemented traditional machine learning algorithms using the bag-of-words (BoW) representation on news headlines and tweets labelled with binary labels (1 or 0). It was reported that logistic regression achieved greater accuracy than Support Vector Machine (SVM) and Naïve Bayes. Another study by [40] used multi-class classification to identify a range of sentiment levels from highly negative to very positive by implementing Recurrent Neural Network (RNN) - Long Short-Term Memory (LSTM) deep learning approaches. However, the findings did not achieve promising results as most news, and tweets were labeled neutral, making it less impactful on cryptocurrency prices. Therefore, developing accurate sentiment models requires the correct language resources and tools to ensure the effect of sentiment features is feasible for input to predict cryptocurrency price.

The dynamics of cryptocurrency and their market behavior heavily depend on sentiment. It influences the behaviors of cryptocurrency traders in the market [27], [35], [41], [42]. Thus, this study attempts to further elaborate the sentiment analysis method by examining the topic of cryptocurrency in social media (Twitter) and news platforms. The input selected is the Malay web news and tweets on Bitcoin and Ethereum because there is a glaring need for Malay language resources.

III. METHODS

The research framework consists of data collection, preprocessing, and experiments on XLNet language models in developing the sentiment regression model. Fig. 1 presents the overview of the proposed solution. The details of each module are presented in the following subsections.
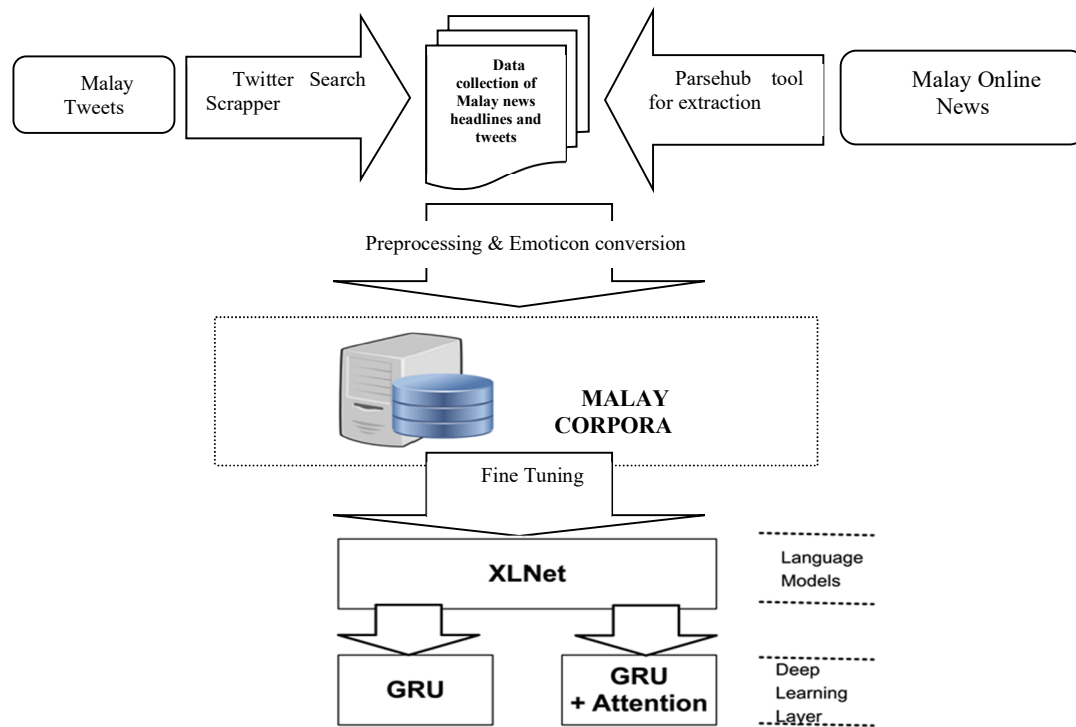

Fig. 1 Research framework

A. Data collection and Pre-processing

Malay news headlines and tweets are collected for data acquisition. Malay news headlines were extracted using the Parsehub tool from various websites, including Intraday.my, Utusan Malaysia, Berita Harian, Harian Metro, Astro Awani, and Malaysiakini. Malay tweets were extracted using the TwitterSearchScrapper Python package. The keywords "bitcoin," "btc," "ethereum," and "eth" are used to collect information about Bitcoin and Ethereum. The data was collected for a year, from January 1 to December 31, 2021. The reason for this data scrapping duration was due to the effect from COVID-19 pandemic season that happened since end of 2019 and at its peak throughout the year 2021, thus, the data are more adequate to explore the economic impact through the public's sentiment, particularly towards cryptocurrency-related matter [43], [44]. During this period, social media and online news are optimally used due to Movement Control Order (MCO) that restrict people from going out. Most activities were performed online, and people are using online transaction very actively during this period. This also aligns with the data collected in [45] which stated that the year 2020 to 2021 was being crucially focused by media worldwide. Examples of the raw data that was extracted for Malay news headlines and tweets, respectively, are shown in Figs. 2(a) and 2(b).

Date/Time	Source	Title
03/01/2021 16:48	Harian Metro	Sindiket Macau Scam beli hartanah RM336 juta hasil penipuan
04/01/2021 11:00	Intraday.my	Ethereum Tembus $1,000, Bitcoin Capai Tahap Tertinggi $34,700!
04/01/2021 15:30	Intraday.my	Litecoin Atasi XRP, #3 Terbesar Selepas Bitcoin Dan Ethereum
04/01/2021 17:00	Intraday.my	Sehari Je Ethereum Catat Rekod Tertinggi Lebih $1,100!
04/01/2021 17:30	Intraday.my	Puluhan Altcoins Catat Lonjakan Tertinggi! Adakah Alt-Season Telah Kembali?
07/01/2021 10:40	Intraday.my	Permodalan Pasaran Kripto Kini Dah Cecah $1 Trilion!
07/01/2021 11:20	Intraday.my	1,620 Ladang Perlombongan Kripto Di Iran Ditutup Serta-Merta
07/01/2021 13:09	Berita Harian	Nilai pasaran kripto cecah AS$1 trilion, Bitcoin catat rekod tertinggi
07/01/2021 17:01	Intraday.my	Ethereum Laju Berdesssuup, Dijangka Cecah $10,000 Tahun Ini!
08/01/2021 12:00	Intraday.my	BTC Susut Selepas Cecah $40,000, Tiba-tiba Coinbase Down Pula!

(a)

created	tweet	name
01/04/2021 00:00	Mar 31, 2021 23:59:00 UTC | 58,771.40$ | 50,120.90◆◆ | 42,640.50◆◆ | #Bitcoin #BTC $BTC #crypto #financial #news	NowBitcoinPrice
01/04/2021 00:00	$BTC: $58918.833 (0.02%)$ETH: $1918.362 (3.97%)$BNB: $302.139 (-2.55%)$USDT: $1.0 (-0.09%)$ADA: $1.194 (-1.75%)$DC	BastionExchange
01/04/2021 00:01	#Bitcoin hari ini...Rp 850,770,00024h : ◆◆◆◆ 0.78%7d : ◆◆◆◆ 8.43%#btc	BotFinansial
01/04/2021 00:01	DASH/BTC yeni sinyal:buy @ fiyat:0.003785 #DASHBTC #CryptoCat #KriptoKedisi #KriptoSinyalleri #TeknikAnaliz #Trend	cryptocatvisor
01/04/2021 00:02	Strong #Downtrends in 30/60/120 mins (Binance $BTC pairs):$CHZ $KSM $REN $RVN $SUSHI $UNI #crypto #coin #RSIhttps:/	yekturus
01/04/2021 00:03	@1MillionTokens @300Token Amazing. @Martins44160536 @PhanNhuNguyen @xyfhy @Bendafit #bitcoin $BTC $ETH $DO	SubamosN
01/04/2021 00:03	◆◆◆◆◆◆◆$2,500,000 #bitcoin LONGED @$59,169.7334 [01/04/21 00:03:39]◆◆◆BitMEX | $XBTUSD◆◆◆◆S◆	WhaleTrades
01/04/2021 00:05	$Bidao $Bid $10 #EOY$ETH $TRON $ONT $LINK DGBBTC $BNB $HUSD $TELLOR $NEM $DASH $IOST $BCD $QTUM $KMD	DDivs3
01/04/2021 00:05	@MEkinGungor CRYPTOCAP:BTC.D/(100-CRYPTOCAP:BTC.D)	TraderMikabot
01/04/2021 00:05	BTC.D ◆◆◆◆◆◆◆	Hesam55168820
01/04/2021 00:09	$LINK $DOT $KYL $ADA $HTR $DOGE $ROOM $UMB $ORAI $AVAX $LTC $BTC $ETH	CryptoG0ld

(b)

Fig. 2. Examples of the extracted data. (a) From Malay News headlines. (b) From Malay Tweets.

Before beginning the annotation task, the raw Malay text underwent preprocessing, which involves cleaning the texts. All special characters, URLs, and hashtags were removed during the preprocessing. Next, the emoticons in the messages were catered using a Python tool called Emot. Emot was created initially to handle English text. However, Emot dictionary is modified by converting the English text to Malay text and translates an emoticon into text that expresses emotion.

B. Malay cryptocurrency news and tweets sentiment corpora

After the texts are cleaned, all non-Malay texts for news and tweets were manually removed. Then, all cleaned instances were selected for the news corpus, which is identified as CryptoSentiNews-Malay. On the other hand, for the tweets corpus designated as CryptoSentiTweets-Malay, only 6,000 documents from each Bitcoin and Ethereum dataset were randomly selected. Table 1 shows the total size of each Malay sentiment corpus.

TABLE 1
NUMBER OF INSTANCES IN EACH MALAY SENTIMENT CORPUS

Corpus	BTC	ETH
CryptoSentiNews-Malay	1,520	1,452
CryptoSentiTweets-Malay	6,000	6,000

Each text unit, such as a news item or a tweet, was manually annotated. The sentiment annotation scoring was conducted by three trained annotators who were native Malay speakers fluent in Malay. The highly negative to the very positive sentiment is reflected in the sentiment scores, which range from -1 to +1. In order to establish an inter-annotator agreement using Krippendorff's alpha measurement [46] of more than 60%, all annotators were given a codebook for the annotation task and underwent several rounds of training. Due to the challenging work of sentiment scoring, where it is very challenging to detect the degree of positivity or negativity in words, more than 60% agreement is deemed appropriate [47]. The final sentiment score for each news headline and tweet was determined by averaging the sentiment scores across the three annotators. Fig. 3 illustrates the flowchart of the annotation process.
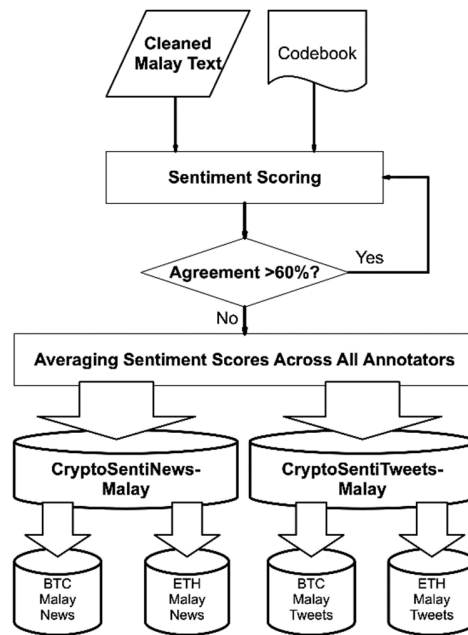
Fig. 3 Annotation process

C. Fine-tuning process

First, to create the vector representations, each sentence from the CryptoSentiNews-Malay and CryptoSentiTweet-Malay corpora was tokenized using the XLNetTokenizer for Malay text (also known as "xlnet-base-bahasa-cased"), which has 12 layers of transformer blocks, 768 hidden layers (dimensions), and 12 self-attention heads [48]. Then, each dataset, namely, the BTC Malay news, ETH Malay news, BTC Malay tweets, and ETH Malay tweets, was fine-tuned before being fed into the sentiment regression models. The purpose of fine-tuning is to replace the current values of the pre-trained XLNet language model derived from Hugging Face (https://huggingface.co/) with the sentiment weight that was learned from the annotated corpora. The step for the fine-tuning process is shown in Algorithm 1 and was applied to all four datasets.

Once the datasets are cleaned, the regression task will be performed. This research explores the performance of the CryptoSentiNews-Malay and CryptoSentiTweets-Malay corpora using Bidirectional-Gated Recurrent Unit (Bi-GRU) deep learning. This approach is adopted based on the comparable result of Cantonese tweet rumour detection [49]. On top of that, the different multi-head self-attention mechanism is also applied to the Bi-GRU model to see if performance can be enhanced.

D. XLNet-GRU sentiment model

The pre-trained XLNet model that has been fine-tuned is then added to the Bi-GRU layer. The sentiment model is anticipated to give each sentence's sentiment score a continuous value between -1 and +1 for regression analysis. For each dataset, four unique sentiment models were created. A 5-fold validation utilizing an 80% training - 20% testing ratio is used to get a generalization result. The Bi-GRU model with 16 hidden states, one output layer and a dropout value of 0.5 Tanh activation function was adopted. The Grid Search of 5-fold cross validation was used to discover the optimized hyperparameter values. The hyperparameter settings for both Malay news and tweet sentiment models are outlined in Table 2. The XLNet-GRU sentiment model architecture is presented in Fig. 4(a) and Fig. 4(b) illustrating XLNet-GRU with multi-head self-attention mechanism model architecture.

TABLE 2
HYPERPARAMETER SETTINGS FOR MALAY NEWS AND TWEET MODELS

	Malay News	Malay Tweets
Batch size	8	16
AdamW optimizer learning rate	0.001	0.001
Number of epochs	20	20

Algorithm 1

Fine-tuning process

get corpus
for each text and sentiment score
   set list of text and sentiment score as values
end for
get maximum sentence length
set test_size as 0.1
set random_state as 56
set tokenizer as XLNetTokenizer.from_pretrained('xlnet-base-bahasa-cased')
get input_ids, attention_mask and label tokens
for each input_ids, attention_mask and label tokens
   set tokens as TensorDataset
   if corpus is news headlines then
     set batch_size as 8
   else if corpus is tweets then
     set batch_size as 16
   end if
   set TensorDataset as DataLoader
end for

set optimizer as AdamW
set learning_rate as 2e-5
if corpus is news headlines then
   set epoch as 50
else if corpus is tweets then
   set epoch as 30

set model as XLNetForSequenceClassification.from_pretrained("xlnet-base-bahasa-cased", num_labels=1)

function train()
   get model
   for each DataLoader
     compute model training
     get trained_model
   end for
   return trained_model

function evaluate()
   get trained_model
   for each DataLoader
     get Dataloader items
     compute metrics for evaluation
   end for
   return metrics

E. XLNet-GRU sentiment model with an attention mechanism

In order to observe any potential performance improvement of the sentiment model, the self-attention mechanism [17] was introduced to the GRU deep learning layer model architecture. The multi-head self-attention mechanism aims to learn the semantic information of some sentences by encapsulating the significant words connected to sentiment or specific features of the issue [50]. Additionally, it is advantageous to prevent the model from overfitting [51].
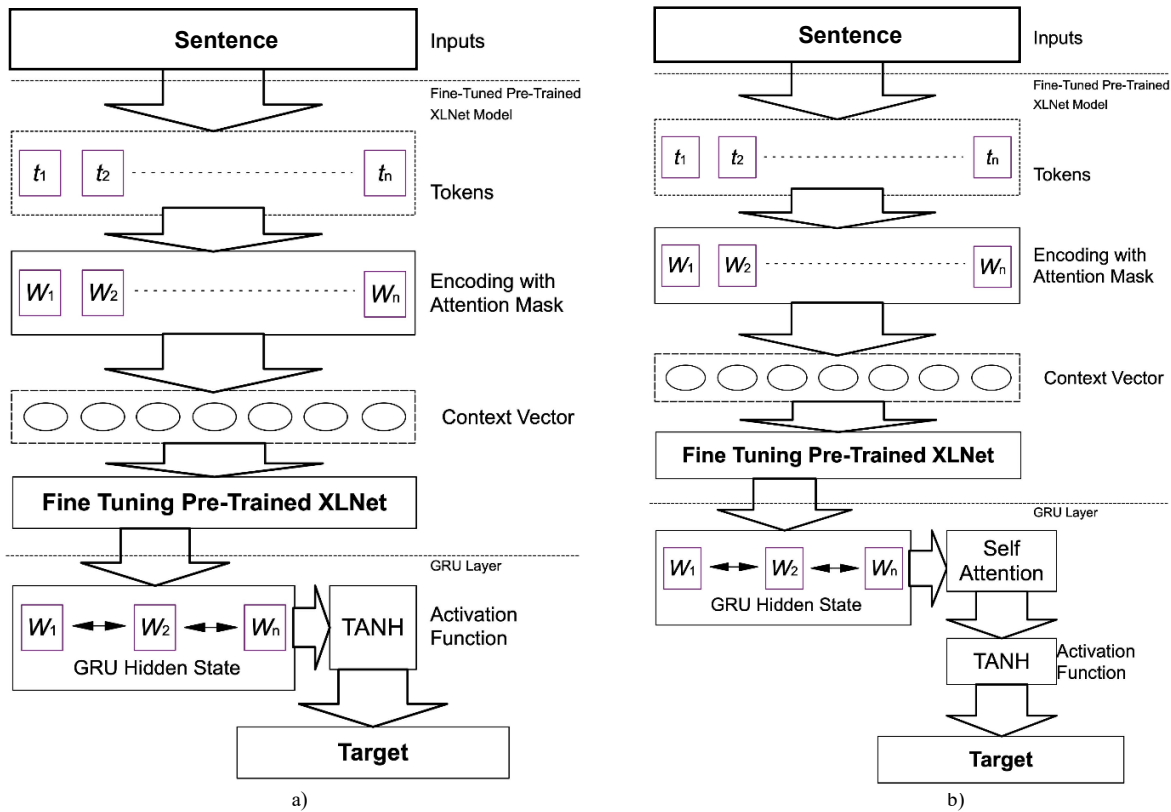
Fig. 4 XLNet-GRU sentiment model (a) XLNet-GRU sentiment model architecture and (b) XLNet-GRU sentiment regression model with attention mechanism

This study uses four different numbers of attention heads – single-head, 2-head, 4-head, and 8-head attention. The value of attention heads was selected based on the findings by [50], which found that if there are more than eight attention heads, some will learn the same attention weights, which will cause noise in the sentiment regression.

The multi-head self-attention mechanism will produce a particular weight for each word in a sentence, increasing proportionately to how strongly the important word is being focused. Fig. 5 and Fig. 6 illustrate the sentiment words learnt using the generated attention weights from the XLNet-GRU sentiment model implementation. The words with a darker color draw more attention than a lighter color, which translates to a more significant effect on the sentence.
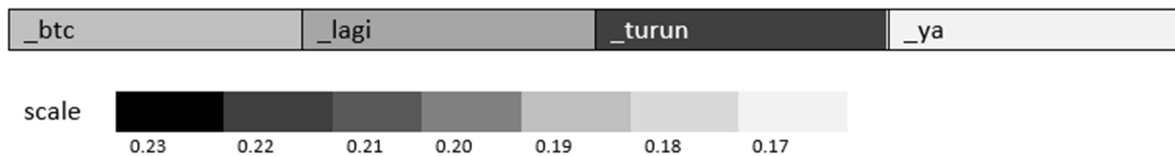

Fig. 5 Visualization of attention weights on a negative sentiment word

It can be observed from Fig. 5 that the Malay word "turun" (which is translated as "drop" in English) in the sentence has the highest attention weight. Such a situation indicates that the attention mechanism allows the model to focus more on words related to a sentiment that is important to predict the overall negative sentiment. Similar patterns are observed with positive sentiment in a sentence shown in Fig. 6. According to Fig. 6, the word "naik" in Malay, which means "up" in English, has more significance than the other words in the sentence. These two examples show two simple situations in which the attention mechanism can be helpful in sentiment regression challenges. Even though attention mechanisms have been quite successful recently, our study's findings indicate that adding an attention layer only sometimes ensures a noticeable performance gain.
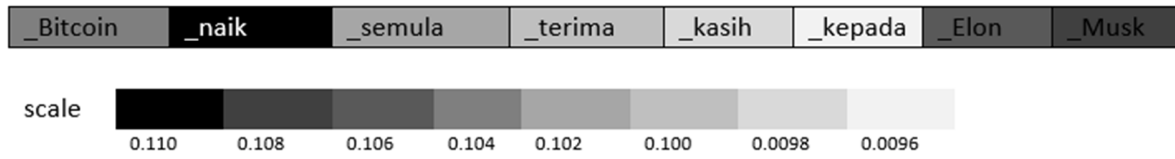
_Bitcoin	_naik	_semula	_terima	_kasih	_kepada	_Elon	_Musk

scale

0.110	0.108	0.106	0.104	0.102	0.100	0.0098	0.0096

Fig. 6 Visualization of attention weights on a positive sentiment word

F. XLNet-GRU sentiment model vs BERT-GRU sentiment model

BERT is a benchmark for comparative analysis because it has demonstrated strong performance in sentiment analysis in the cryptocurrency domain [8]and its pre-trained language model ('bert-base-bahasa-cased') can be employed on Malay text. This work is motivated by a study by [52] that used the identical pipeline and hyperparameter configurations to compare the XLNet-GRU sentiment model with the BERT-GRU sentiment model. This study tested Malay-BERT on tweets for sentiment classification utilizing hierarchical attention mechanisms. Consequently, six performance results will be presented in the next section for the following experiments:

1) *XLNet sentiment model (XLNet language model only without any deep learning model)*
2) *XLNet-GRU sentiment model (XLNet language model with the application of Bi-GRU deep learning model)*
3) *XLNet-GRU+ sentiment model (XLNet-GRU with multi-head self-attention mechanism)*
4) *BERT sentiment model (BERT language model only without any deep learning model)*
5) *BERT-GRU sentiment model (BERT language model with the application of Bi-GRU deep learning model)*
6) *BERT-GRU+ sentiment model (BERT -GRU with multi-head self-attention mechanism)*

G. Evaluation metrics

Three metrics are used to evaluate the proposed sentiment regression model: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and adjusted R^2. The RMSE metric is good for getting unbiased forecasts but in some cases RMSE is bound to penalize significant errors. The MAE metric is used to protect outliers and is more robust than RMSE, making it a more appropriate metric for error assessment [53]. In addition, the adjusted R^2 was also selected as the performance metric because it is more accurate and informative [54]. Adjusted R^2 is different from R^2 that it tests many independent variables against the model, which is not being done in R^2. Therefore, adjusted R^2 measurement is favoured by most investors because it provides a more accurate assessment of the correlation between one variable and another [55].

IV. RESULTS

The performance of the XLNet sentiment model, followed by the XLNet-GRU sentiment model, and the XLNet-GRU sentiment model across different attention heads for the multi-head self-attention mechanism is examined. Finally, the results of three experiments using the BERT sentiment model, the BERT-GRU sentiment model, and the BERT-GRU sentiment model with the best-performing number of attention heads produced for XLNet are presented. The experiments are performed independently for news and tweets.

A. Malay News

The results of the XLNet sentiment model for Bitcoin and Ethereum Malay news are shown in Table 3, with and without the use of the Bi-GRU deep learning model. When utilizing the XLNet with Bi-GRU deep learning model, the adjusted R^2 shows a minor improvement of 0.025 (Bitcoin Malay news) and 0.008 (Ethereum Malay news) from Table 3. The BERT sentiment model yields result like those in Table 4.

When using the Bi-GRU to BERT language model, the adjusted R^2 increases by 0.051 for Bitcoin Malay news and 0.004 for Ethereum Malay news, as presented in Table 4. The results show that a deep learning model can improve performance. Additionally, it is noted that in both XLNet and BERT applications, the adjusted R^2 for Ethereum Malay news is higher than that for Bitcoin Malay news. This might be because there are more texts in Bitcoin Malay news than Ethereum Malay news, and more texts produce more noise.

For Bitcoin and Ethereum Malay news, the XLNet-GRU sentiment regression model results with and without an attention mechanism are shown in Table 5 and 6, respectively. Without an attention mechanism, the adjusted R^2 for Bitcoin Malay news was 0.628, whereas the adjusted R^2 for Ethereum Malay news was 0.653. By incorporating single-head attention to the model, it is observed that there is a modest improvement in performance, with adjusted R^2 values for Bitcoin and Ethereum Malay news being 0.608 and 0.631, respectively. Subsequently, after testing with different

number of attention heads, it is found that 2-head attention yields the best performance for Bitcoin Malay news (adjusted $R^2 = 0.628$) and 8-head attention gives a good result for Ethereum Malay news (adjusted $R^2 = 0.653$).

TABLE 3
XLNET SENTIMENT REGRESSION RESULTS FOR BITCOIN & ETHEREUM MALAY NEWS WITH AND WITHOUT BI-GRU DEEP LEARNING MODEL

	BITCOIN		ETHEREUM	
	XLNet	XLNet-GRU	XLNet	XLNet-GRU
RMSE	0.301	**0.292**	0.258	**0.255**
MAE	0.183	**0.179**	0.141	**0.144**
Adjusted R^2	0.583	**0.608**	0.623	**0.631**

TABLE 4
BERT SENTIMENT REGRESSION RESULTS FOR BITCOIN & ETHEREUM MALAY NEWS WITH AND WITHOUT BI-GRU DEEP LEARNING MODEL

	BITCOIN		ETHEREUM	
	BERT	BERT-GRU	BERT	BERT-GRU
RMSE	0.353	**0.337**	0.303	**0.302**
MAE	0.224	**0.205**	0.180	**0.179**
Adjusted R^2	0.427	**0.478**	0.480	**0.484**

TABLE 5
XLNET-GRU SENTIMENT REGRESSION RESULTS FOR BITCOIN MALAY NEWS WITH AND WITHOUT ATTENTION MECHANISM

	BITCOIN				
	No Attention	Single-head Attention	**2-head Attention**	4-head Attention	8-head Attention
RMSE	0.292	0.290	**0.285**	0.293	0.287
MAE	0.179	0.176	**0.186**	0.175	0.186
Adjusted R^2	0.608	0.615	**0.628**	0.605	0.621

TABLE 6
XLNET-GRU SENTIMENT REGRESSION RESULTS FOR ETHEREUM MALAY NEWS WITH AND WITHOUT ATTENTION MECHANISM

	ETHEREUM				
	No Attention	Single-head Attention	2-head Attention	4-head Attention	**8-head Attention**
RMSE	0.255	0.252	0.255	0.250	**0.248**
MAE	0.144	0.140	0.145	0.139	**0.139**
Adjusted R^2	0.631	0.641	0.632	0.647	**0.653**

B. Malay Tweets

The results of the XLNet sentiment model with and without using the Bi-GRU deep learning model for Bitcoin and Ethereum Malay tweets are shown in Table 7. It is observed in Table 7 that XLNet-GRU achieved an adjusted R^2 of 0.426, which is 0.014 higher than XLNet adjusted R^2 of 0.412 for Bitcoin Malay tweets, whereas, for Ethereum Malay tweets, an improvement of 0.019 was acquired (adjusted $R^2 = 0.381$). The same experimental setup was used with BERT; the findings are displayed in Table 8.

TABLE 7
XLNET SENTIMENT REGRESSION RESULTS FOR BITCOIN & ETHEREUM MALAY TWEETS WITH AND WITHOUT BI-GRU DEEP LEARNING MODEL

	BITCOIN		ETHEREUM	
	XLNet	XLNet-GRU	XLNet	XLNet-GRU
RMSE	0.292	**0.289**	0.239	**0.236**
MAE	0.191	**0.189**	0.161	**0.159**
Adjusted R^2	0.412	**0.426**	0.362	**0.381**

TABLE 8
BERT SENTIMENT REGRESSION RESULTS FOR BITCOIN MALAY TWEETS WITH AND WITHOUT BI-GRU DEEP LEARNING MODEL

	BITCOIN		ETHEREUM	
	BERT	BERT-GRU	BERT	BERT-GRU
RMSE	0.302	**0.291**	0.268	**0.257**
MAE	0.221	**0.214**	0.192	**0.189**
Adjusted R^2	0.371	**0.415**	0.196	**0.366**

Similar findings have been obtained, as seen in Table 8, where the Bi-GRU deep learning model has been demonstrated to deliver a modest improvement over the BERT sentiment model. An adjusted R^2 of 0.415 was attained for Bitcoin Malay tweets, while an adjusted R^2 of 0.366 was attained for Ethereum Malay tweets. The results for Malay tweets differ from those for Malay news. When employing XLNet and BERT, the adjusted R^2 for Bitcoin-related Malay tweets is superior to Ethereum-related Malay tweets. Since the number of documents is the same for Bitcoin and Ethereum Malay tweets, it may be assumed that Bitcoin tweets have more understandable Malay content than Ethereum tweets.

Then, the XLNet-GRU sentiment regression model with and without an attention mechanism was tested for Malay tweets about Bitcoin and Ethereum. Results from Malay tweets are consistent with the observations from Malay news. The XLNet-GRU sentiment model with 2-head attention still yielded the best performance for Bitcoin Malay tweets with an adjusted R^2 of 0.457, showing a slight improvement of 0.038 from the no-attention head. The 8-head attention XLNet-GRU sentiment model performed the best while analyzing Ethereum Malay tweets, achieving an adjusted R^2 of 0.418 with an improvement of between 0.031 and 0.037 of the adjusted R^2 from the no-attention head XLNet-GRU sentiment model. The results are presented in Table 9 and Table 10.

TABLE 9
XLNET-GRU SENTIMENT REGRESSION RESULTS FOR BITCOIN MALAY TWEETS WITH AND WITHOUT ATTENTION MECHANISM

	BITCOIN				
	No Attention	Single-head Attention	**2-head Attention**	4-head Attention	8-head Attention
RMSE	0.289	0.282	**0.281**	0.287	0.283
MAE	0.189	0.208	**0.183**	0.207	0.205
Adjusted R^2	0.426	0.453	**0.457**	0.434	0.444

TABLE 10
XLNET-GRU SENTIMENT REGRESSION RESULTS FOR ETHEREUM MALAY TWEETS WITH AND WITHOUT ATTENTION MECHANISM

	ETHEREUM				
	No Attention	Single-head Attention	2-head Attention	4-head Attention	**8-head Attention**
RMSE	0.236	0.238	0.235	0.235	**0.229**
MAE	0.159	0.177	0.161	0.162	**0.165**
Adjusted R^2	0.381	0.366	0.385	0.384	**0.418**

C. Summary of results for XLNet and BERT sentiment models

Table 11 and Table 12 presents the performance results of XLNet without Bi-GRU, XLNet-GRU sentiment regression models without attention mechanism, XLNet-GRU with the best number of attention heads (XLNet-GRU+), and BERT-GRU with attention mechanism sentiment model (BERT-GRU+) for Bitcoin and Ethereum Malay news and tweets accordingly. The XLNet-GRU+ model with a multi-head attention mechanism outperformed the BERT sentiment model for both Bitcoin and Ethereum news and tweets. Model performance on news is consistently higher than tweets, which can be attributed to the tweet corpus naturally containing more noise. However, we observe a smaller performance gap between news and tweets for Bitcoin while this gap is noticeably wider for Ethereum.

TABLE 11
COMPARISON OF RESULTS WITH BERT-GRU+ SENTIMENT MODEL FOR BITCOIN

	BITCOIN							
	XLNet		XLNet-GRU		XLNet-GRU+		BERT-GRU+	
	Malay News	Malay Tweets	Malay News	Malay Tweets	Malay News	Malay Tweets	Malay News	Malay Tweets
RMSE	0.301	0.292	0.292	0.289	**0.285**	**0.281**	0.327	0.286
MAE	0.183	0.191	0.179	0.189	**0.186**	**0.183**	0.219	0.185
Adjusted R^2	0.583	0.412	0.608	0.426	**0.628**	**0.457**	0.513	0.436

TABLE 12
COMPARISON OF RESULTS WITH BERT-GRU+ SENTIMENT MODEL FOR ETHEREUM

	ETHEREUM							
	XLNet		XLNet-GRU		XLNet-GRU+		BERT-GRU+	
	Malay News	Malay Tweets	Malay News	Malay Tweets	Malay News	Malay Tweets	Malay News	Malay Tweets
RMSE	0.258	0.239	0.255	0.236	**0.248**	**0.229**	0.279	0.233
MAE	0.141	0.161	0.144	0.159	**0.139**	**0.165**	0.151	0.157
Adjusted R^2	0.623	0.362	0.631	0.381	**0.653**	**0.418**	0.575	0.393

V. Discussion

The results demonstrate that the Malay news and tweets for Bitcoin (2-head attention) and Ethereum (8-head attention) perform significantly better using the multi-head self-attention technique. A mean adjusted R^2 of 0.620 (no-attention) and mean adjusted R^2 of 0.641 (multi-head self-attention) were obtained by averaging across Bitcoin and Ethereum for Malay news. Moreover, a mean adjusted R^2 of 0.404 (no-attention) and mean adjusted R^2 of 0.438 (multi-head self-attention) were achieved for Bitcoin and Ethereum Malay tweets.

The results indicate that the multi-head self-attention mechanism can slightly improve the precision of the XLNet-GRU sentiment regression model, which is in line with the experiment done by [49]. The XLNet language model incorporating a deep learning model does produce significant results for sentiment analysis. To the best of our knowledge, there is no existing research study that reported on the sentiment regression evaluation on cryptocurrency-related matter in Malay text. Therefore, we compare our results with several related work that used similar dataset domain that addressed regression problem which is sentiment analysis on financial news headlines by [15], which reported their result in R^2 metric, and sentiment analysis on financial microblogging by [16] using MAE and RMSE as their evaluation metrics. However, these research studies used English text. The work by [15] implemented simple regression models such as linear regression, support vector regression, and XGBoost regression and obtained maximum R^2 result of 0.38 for English financial news headlines, while our XLNet-GRU+ model managed to achieve a higher adjusted R^2 of 0.628 (Bitcoin Malay news) and 0.653 (Ethereum Malay news). Adjusted R^2 is always less than or equals to the R^2, thus, this proves that albeit Malay text has poor language resources, it managed to perform well using our XLNet-GRU+ model.

Moreover, this also supported by comparing the MAE and RMSE achieved by [16] with our work. [16] applied basic machine learning regressors such as linear regression, decision tree, and neural network regression on English microblogging texts such as tweets. It was reported that the lowest MAE achieved was 0.143 using a neural network regression and lowest RMSE of 0.183. Even though our Malay tweets only achieved an average MAE result of 0.174 and an average RMSE result of 0.255, it is still considered good since there is more noise in tweet texts compared to news headlines, plus the low language resource of Malay text. Based on the findings, it provides a potential for further research on Malay language particularly in the financial domain in terms of predicting market prices.

The experiment provides new insight into the relationship between the language models with and without deep learning model integration, as well as the effect of using different numbers of attention heads. Although it can be seen that there is only a slight increase in the adjusted R^2 (precision), it is still considered a beneficial outcome as the model provides a more focused approach toward the important terms in sentiment texts. The study can experiment with two sources (i.e., online news and tweets). However, manual annotations are performed on a few documents due to the limited computational resources, time constraints, and inadequate qualified annotators.

VI. Conclusions

This study builds explicitly two curated Malay sentiment corpora (CryptoSentiNews-Malay and CryptoSentiTweet-Malay) for the cryptocurrency domain. This will motivate future research to evaluate the effect of sentiment features in cryptocurrency price prediction using various contextual word embeddings and deep learning methods.

Due to computational resource constraints, we could only present the best hyperparameter settings within the investigated value ranges in our studies. These settings are by no means all-inclusive. Additionally, there are mixtures of Indonesian and English words in some of the Malay sentences, and the lack of a cryptocurrency-related dictionary for Malay spelling correction when dealing with tweets during text pre-processing prevents the model from producing a significantly higher accuracy.

More experiments can be conducted to test different embeddings and deep learning algorithms. Despite a year of data collection, the size of the curated Malay sentiment corpora, especially for news, still needs to be considered limited compared to sentiment corpora that are available in other resource-rich languages such as English. Nevertheless, the performance of the Malay corpus, despite its modest size and lack of language resources, is comparable to that of other similar experiments on sentiment regression.

Therefore, this study has successfully addressed the primary objective which is to construct a Malay sentiment regression model. Regression model is better than classification because it provides precise sentiment score as opposed to limited positive/negative sentiment polarity. Furthermore, this study has provided a new set of data that focuses on Malay cryptocurrency text that are gathered during the peak of the Covid19 pandemic that depict the volatile and chaotic cryptocurrency market.

For future work, we plan to continue to increase the size of the Malay news and tweet corpora and include more text sources relevant to cryptocurrency that can help advance research, particularly in the financial domain. Additionally, this model can be integrated as a part of the engine for cryptocurrency price prediction to obtain the

sentiment features from Malay news and tweets that can further aid the traders to practically predict cryptocurrency prices based on the sentiment features.

ORCID:
Nur Azmina Mohamad Zamani: https://orcid.org/0000-0001-7376-5379
Norhaslinda Kamaruddin: https://orcid.org/0000-0003-0827-2417

REFERENCES

[1] K. Danial, *Cryptocurrency investing for dummies*. John Wiley & Sons, 2019.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized Business Review*, p. 21260, 2008.

[3] K. Farhana, and S. Muthaiyah, "Behavioral Intention to Use Cryptocurrency as an Electronic Payment in Malaysia," *J. Syst. Manag. Sci.,* vol. 12, no. 4, pp. 219-231, 2022.

[4] S. Sukumaran, T. S. Bee, and S. Wasiuzzaman, "Cryptocurrency as an investment: The Malaysian context," *Risks*, vol. 10, no. 4, 86, 2022, doi: 10.3390/risks10040086.

[5] M. F. Yusof, L.A. Rasid, and R. Masri, "Implementation Of Zakat Payment Platform For Cryptocurrencies," *AZKA International Journal of Zakat and Social Finance,* vol. 2, no. 1, pp. 17-31, 2021, doi: 10.51377/azjaf.vol2no1.41.

[6] S.A. Farimani, M.V. Jahan, A.M. Fard, and S.R.K. Tabbakh, "Investigating the informativeness of technical indicators and news sentiment in financial market price prediction," *Knowledge-Based Systems*, vol. 247, 108742, 2022, doi: 10.1016/j.knosys.2022.108742.

[7] A.M. Balfagih, and V. Keselj, "Evaluating sentiment classifiers for Bitcoin tweets in price prediction task," in *IEEE International Conference on Big Data (Big Data)*, pp. 5499–5506, 2019, doi: 10.1109/BigData47090.2019.9006140.

[8] G.N.C. Cerda, "Bitcoin price prediction through stimulus analysis: On the footprints of Twitter's crypto influencers," Master's Thesis, Pontificia Universidad Católica de Chile, Santiago de Chile, 2021. [Online]. Available: https://repositorio.uc.cl/xmlui/bitstream/handle/11534/60881/TESIS_GCheuque_Firma%20Final.pdf?sequence=1

[9] E. Edgari, J. Thiojaya, and N.N. Qomariyah, "The impact of Twitter sentiment analysis on Bitcoin price during COVID-19 with XGBoost," in *5th International Conference on Computing and Informatics (ICCI)*, pp. 337–342, 2022, doi: 10.1109/ICCI54321.2022.9756123.

[10] N.A.M. Zamani, J.S.Y. Liew, and A.M. Yusof, "XLNET-GRU sentiment regression model for cryptocurrency news in English and Malay," in *Proceedings of the 4th Financial Narrative Processing Workshop @ LREC 2022*, pp. 36–42, 2022.

[11] F.H. Jahjah and M. Rajab, "Impact of Twitter Sentiment Related to Bitcoin on Stock Price Returns," *jcoeng*, vol. 26, no. 6, pp. 60–71, Jun. 2020, doi: 10.31026/j.eng.2020.06.05.

[12] S. Mohapatra, N. Ahmed, and P. Alencar, "KryptoOracle: A real-time cryptocurrency price prediction platform using Twitter sentiments," in *IEEE International Conference on Big Data (Big Data)*, pp. 5544–5551, 2019, doi: 10.1109/BigData47090.2019.9006554.

[13] U. Maqsood, F.Y. Khuhawar, S. Talpur, F.H. Jaskani, and A.A. Memon, "Twitter Mining based Forecasting of cryptocurrency using sentimental analysis of Tweets," in *Global Conference on Wireless and Optical Technologies (GCWOT)*, pp. 1-6, 2022, doi: 10.1109/GCWOT53057.2022.9772923.

[14] Z. Ye, Y. Wu, H. Chen, Y. Pan, and Q. Jiang, "A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin," *Mathematics*, vol. 10, no. 8, 1307, 2022, doi: 10.3390/math10081307.

[15] V. John and O. Vechtomova, "UW-FinSent at SemEval-2017 Task 5: Sentiment Analysis on Financial News Headlines using Training Dataset Augmentation," in *Proceedings of the 11th International Workshop on Semantic Evaluation*, pp. 872–876, 2017, doi: 10.18653/v1/S17-2149.

[16] S. Symeonidis, J. Kordonis, D. Effrosynidis, and A. Arampatzis, "Sentiment predictability in financial microblogging and news articles," in *Proceedings of the 11th International Workshop on Semantic Evaluation* pp. 861–865, 2017, doi: 10.18653/v1/S17-2147.

[17] A. Vaswani *et al.*, "Attention is All you Need," arXiv:1706.03762, 2017.

[18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 4171–4186, 2019.

[19] N. Passalis, L. Avramelou, S. Seficha, A. Tsantekidis, S. Doropoulos, G. Makris, and A. Tefas, "Multisource financial sentiment analysis for detecting Bitcoin price change indications using deep learning," *Neural Comput & Applic*, vol. 34, no. 22, pp. 19441-19452, 2022, doi: 10.1007/s00521-022-07509-6.

[20] M. Ortu, N. Uras, C. Conversano, S. Bartolucci, and G. Destefanis, "On technical trading and social media indicators for cryptocurrency price classification through deep learning," *Expert Systems with Applications*, vol. 198, 116804, 2022, doi: 10.1016/j.eswa.2022.116804.

[21] L. Rognone, S. Hyde, and S.S. Zhang, "News sentiment in the cryptocurrency market: An empirical comparison with Forex," *International Review of Financial Analysis*, vol. 69, 101462, 2020, doi: 10.1016/j.irfa.2020.101462.

[22] E. Stenqvist and J. Lönnö, "Predicting Bitcoin price fluctuation with Twitter sentiment analysis," Degree Project, KTH Royal Institute of Technology School of Computer Science and Communication, 2017.

[23] T.M. Dulău and M. Dulău, "Cryptocurrency-sentiment analysis in social media," *Acta Marisiensis. Seria Technologica*, vol. 16, no. 2, pp. 1-6, 2019, doi: 10.2478/amset-2019-0009.

[24] T. Loughran and B. Mcdonald, "When is a liability not a liability? textual analysis, dictionaries, and 10-Ks," *The Journal of Finance*, vol. 66, no. 1, pp. 35-65, 2011, doi: 10.1111/j.1540-6261.2010.01625.x.

[25] T. Loughran and B. Mcdonald, "Measuring readability in financial disclosures," *The Journal of Finance*, vol. 69, no. 4, pp. 1643-1671, 2014, doi: 10.1111/jofi.12162.

[26] C. Y.H. Chen, R. Després, L. Guo, and T. Renault, "What makes cryptocurrencies special? Investor sentiment and return predictability during the bubble," *Comparative Political Economy: Monetary Policy eJournal*, pp. 1–36, 2019.

[27] C. Gurdgiev and D. O'Loughlin, "Herding and anchoring in cryptocurrency markets: Investor reaction to fear and uncertainty," *Journal of Behavioral and Experimental Finance*, vol. 25, 100271, 2020, doi: 10.1016/j.jbef.2020.100271.

[28] V. Karalevicius, N. Degrande, and J. De Weerdt, "Using sentiment analysis to predict interday Bitcoin price movements," *The Journal of Risk Finance*, vol. 19, no. 1, pp. 56-75, 2018, doi: 10.1108/JRF-06-2017-0092.

[29] F. Mai, Q. Bai, and J. Shan, "The impacts of social media on Bitcoin performance," in *International Conference on Information Systems*, pp. 1-16, 2015.

[30] F. Mai, Z. Shan, Q. Bai, X. Wang, and R.H.L. Chiang, "How does social media impact Bitcoin value? A test of the silent majority hypothesis," *Journal of Management Information Systems*, vol. 35, no. 1, pp. 19-52, 2018, doi: 10.1080/07421222.2018.1440774.

[31] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the 8th International Conference on Weblogs and Social Media,* pp. 10, 2015.

[32] Y. B. Kim *et al.*, "Predicting fluctuations in cryptocurrency transactions based on user comments and replies," *PLoS ONE*, vol. 11, no. 8, e0161197, 2016, doi: 10.1371/journal.pone.0161197.

[33] 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, 589, 2019, doi: 10.3390/e21060589.

[34] K. Wołk, "Advanced social media sentiment analysis for short-term cryptocurrency price prediction," *Expert Systems*, vol. 37, no. 2, pp. 1-16, 2019, doi: 10.1111/exsy.12493.

[35] O. Kraaijeveld and J. De Smedt, "The predictive power of public Twitter sentiment for forecasting cryptocurrency prices," *Journal of International Financial Markets, Institutions and Money*, vol. 65, 101188, 2020, doi: 10.1016/j.intfin.2020.101188.

[36] G. Serafini *et al.*, "Sentiment-driven price prediction of the Bitcoin based on statistical and deep learning approaches," in *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2020, doi: 10.1109/IJCNN48605.2020.9206704.

[37] S. Loria, "Textblob Documentation." Apr. 26, 2020. [Online]. Available: https://buildmedia.readthedocs.org/media/pdf/textblob/latest/textblob.pdf

[38] A. Jain, S. Tripathi, H. D. Dwivedi, and P. Saxena, "Forecasting price of cryptocurrencies using tweets sentiment analysis," in *2018 Eleventh International Conference on Contemporary Computing (IC3)*, pp. 1-7, 2018, doi: 10.1109/IC3.2018.8530659.

[39] C. Lamon, E. Nielsen, and E. Redondo, "Cryptocurrency price prediction using news and social media sentiment," *SMU Data Science Review*, pp. 1-22, 2017.

[40] A. Inamdar, A. Bhagtani, S. Bhatt, and P.M. Shetty, "Predicting cryptocurrency value using sentiment analysis," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, pp. 932–934, 2019, doi: 10.1109/ICCS45141.2019.9065838.

[41] M.M. Patel, S. Tanwar, R. Gupta, and N. Kumar, "A deep learning-based cryptocurrency price prediction scheme for financial institutions," *Journal of Information Security and Applications*, vol. 55, 102583, 2020, doi: 10.1016/j.jisa.2020.102583.

[42] Y. Wang and R. Chen, "Cryptocurrency price prediction based on multiple market sentiment," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, pp. 1092-1100, 2020.

[43] J. Luo, "Bitcoin price prediction in the time of COVID-19," in *Management Science Informatization and Economic Innovation Development Conference (MSIEID)*, pp. 243–247, 2020, doi: 10.1109/MSIEID52046.2020.00050.

[44] L. Barbaglia, L. Frattarolo, L. Onorante, F.M. Pericoli, M. Ratto, and L.T. Pezzoli, "Testing big data in a big crisis: Nowcasting under Covid-19," *International Journal of Forecasting*, S0169207022001431, 2022, doi: 10.1016/j.ijforecast.2022.10.005.

[45] W. Ahmad, B. Wang, P. Martin, M. Xu, and H. Xu, "Enhanced sentiment analysis regarding COVID-19 news from global channels," *J Comput Soc Sc*, vol. 6, no. 1, pp. 19-57, 2023, doi: 10.1007/s42001-022-00189-1.

[46] K. Krippendorff, *Content analysis: An introduction to its methodology*, Fourth. SAGE Publications, 2018.

[47] R. Artstein and M. Poesio, "Inter-Coder agreement for computational linguistics," *Computational Linguistics*, vol. 34, no. 4, pp. 555-596, 2008, doi: 10.1162/coli.07-034-R2.

[48] X. R. Gong, J.X. Jin, and T. Zhang, "Sentiment analysis using autoregressive language modeling and broad learning system," in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1130-1134, 2019, doi: 10.1109/BIBM47256.2019.8983025.

[49] X. Chen, L. Ke, Z. Lu, H. Su, and H. Wang, "A novel hybrid model for Cantonese rumor detection on Twitter," *Applied Sciences*, vol. 10, no. 20, 7093, 2020, doi: 10.3390/app10207093.

[50] X. Li, L. Ding, Y. Du, Y. Fan, and F. Shen, "Position-Enhanced Multi-Head Self-Attention Based Bidirectional Gated Recurrent Unit for Aspect-Level Sentiment Classification," *Front. Psychol.*, vol. 12, 799926, 2022, doi: 10.3389/fpsyg.2021.799926.

[51] X. Zhang, Z. Wu, K. Liu, Z. Zhao, J. Wang, and C. Wu, "Text sentiment classification based on BERT embedding and sliced multi-head self-attention Bi-GRU," *Sensors*, vol. 23, no. 3, 1481, 2023, doi: 10.3390/s23031481.

[52] Y.H. Lim and J.S.Y. Liew, "English-Malay word embeddings alignment for cross-lingual emotion classification with hierarchical attention network," in *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, pp. 113–124, 2022, doi: 10.18653/v1/2022.wassa-1.12.

[53] S. M. Robeson and C. J. Willmott, "Decomposition of the mean absolute error (MAE) into systematic and unsystematic components," *PLoS ONE*, vol. 18, no. 2, e0279774, 2023, doi: 10.1371/journal.pone.0279774.

[54] D. Chicco, M.J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, e623, 2021, doi: 10.7717/peerj-cs.623.

[55] J. Karch, "Improving on Adjusted R-Squared," PsyArXiv, preprint, Sep. 2019. doi: 10.31234/osf.io/v8dz5.

Publisher's Note: Publisher stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.