IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 September 2024, accepted 19 November 2024, date of publication 26 November 2024, date of current version 9 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3506658

RESEARCH ARTICLE

Permuted Temporal Kolmogorov-Arnold Networks for Stock Price Forecasting Using Generative Aspect-Based Sentiment Analysis

AGUS TRI HARYONO [1,2], **RIYANARTO SARNO** [1], **(Senior Member, IEEE), RATIH NUR ESTI ANGGRAINI** [1], **(Member, IEEE), AND KELLY ROSSA SUNGKONO** [1], **(Member, IEEE)**

[1]Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia
[2]Department of Informatics, Universitas Mulawarman, Samarinda 75119, Indonesia

Corresponding author: Riyanarto Sarno (riyanarto@its.ac.id)

This work was supported by Institut Teknologi Sepuluh Nopember (ITS) under Penelitian Keilmuan, Penelitian Dana Departemen Program, Penelitian Flagship, Penelitian Kolaborasi Pusat and under the Project Scheme of the Publication Writing and Intellectual Property Rights (IPR) Incentive (Penulisan Publikasi dan Hak Kekayaan Intelektual/PPHKI) Program 2024.

ABSTRACT Stock prices are experiencing fluctuation daily. While stock price predictions typically rely on historical transaction data, other factors, such as news sentiment, also play an indirect role in influencing these changes. News sentiment, typically expressed as a qualitative sentiment label, cannot be directly used as an indicator in stock price analysis, as it differs from the quantitative stock transaction. The vital issue in stock price forecasting using news data is the quantification process from sentiment label to score. This research proposed generative Aspect-Based Sentiment Analysis (ABSA) to produce an aspect-sentiment quadruplet: aspect category, aspect term, sentiment polarity, and opinion term. The aspect-sentiment quadruplet produces daily ABSA sentiment scores using the Loughran and McDonald Sentiment Lexicon (LMSL) to handle out-of-vocabulary. The stock transaction history and daily ABSA sentiment score are used to unify stock price forecasting using permuted Temporal Kolmogorov-Arnold Network (pTKAN), which is a rearrangement of the position dimension for isolating the sequence of time series. The Movement-Weighted Regression Error (MWRE) evaluation method is proposed to measure the performance of unified stock price forecasting with representation movement direction error and regression error. The experimental results show that the daily ABSA sentiment score positively influences the performance of unified stock price forecasting using the Temporal Kolmogorov-Arnold Network (TKAN). The pTKAN architecture best performed in 25 of 27 stock issuers among 19 architectures, which includes traditional-, machine learning-, and deep learning-based architectures tested on the stock transaction dataset.

INDEX TERMS Movement-weighted regression error, paraphrase generation, permuted Kolmogorov-Arnold network, aspect-sentiment quadruplet, quantifying news sentiment, stock price forecasting.

I. INTRODUCTION

Research in sentiment analysis is rapidly evolving, enabling stock price analysis to move beyond the constraints of traditional fundamental and technical indicators. While sentiment analysis produces qualitative data, this cannot be directly integrated with the quantitative data from technical indicators for stock price prediction. Therefore, a process is required to convert qualitative data into a quantitative format [1]. Quantitative sentiment has been previously employed in research to indicate whether market conditions are positive or negative [2]. Utilizing the average sentiment value from training data has been a method used to forecast stock prices [3] to create models for predicting stock trends [4]. The sentiment of financial news, covering stock, corporate, economic news, and the market, is seen to reflect the stock

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

market [5]. Analyzing sentiment based on various aspects offers a more comprehensive representation of stock market conditions and helps identify significant factors that affect stock price forecasting.

In previous research, aspect-based sentiment analysis (ABSA) was separated into four processes: aspect categorization, sentiment classification, aspect term extraction, and opinion term extraction [6], [7]. The disadvantage of a separate process is that the result prediction has a separate context. Generative text-to-text models, such as machine translation [8] or paraphrase generation [9], can integrate separate processes in ABSA to maintain the source context in the prediction target. The paraphrase generation uses a sequence-to-sequence deep learning model for various tasks, such as question answering, information retrieval, or information extraction [9]. In previous research, Long Short-Term Memory (LSTM) networks with latent Bag of Word (BoW) [10] or pre-trained Transformer model [11] are used to learn source and target text. The generative ABSA uses raw text as the source text and aspect-sentiment quadruplet, which contains aspect category, sentiment polarity, aspect term, and opinion term, as the target text. The aspect-sentiment quadruplet is transformed into the natural text to fulfill the requirement of the paraphrase generation model [12]. The context from the source text is learned by contrastive learning, which employs an aspect-sentiment quadruplet label within the target text. Contrastive learning in paraphrase generation is used to distinguish the latent space of source text based on aspect-sentiment quadruplets. This research proposes paraphrase generation with contrastive learning with various pre-trained models to produce aspect-sentiment quadruplet, which has context [13]. The predicted aspect-sentiment quadruplet as qualitative data is processed into quantitative data using quantifying.

The sentiment analysis is already used in stock price forecasting with various methods, such as using sentiment embedding [14], confidence score [15], [16], sentiment factor [17], [18], and sentiment score [19], [20]. The sentiment embedding method combines word vector features with quantitative features, such as historical price, moving average, and relative strength index. The confidence score method uses the last logits from the output model to predict sentiment polarity. The confidence score represents the weight of knowledge from the sentiment analysis model. The sentiment factor method uses the amount of each label sentiment to produce the quantitative sentiment. The sentiment score method uses sentiment lexicon, such as the Valence Aware Dictionary and Sentiment Reasoner (VADER), TextBlob, and Loughran and McDonald, to produce quantitative sentiment from important terms. This research focuses on the sentiment score method to quantify sentiment labels from aspect and opinion terms. The problem with using a sentiment lexicon is that it is prone to out of vocabulary (OOV) and impacts important terms that do not have a sentiment score. This research proposes adding a list of words in the sentiment lexicon to tackle the limitation of OOV using semantic similarity. The sentiment score is created daily on each stock issuer. The sentiment score combined with daily stock transaction history for stock price forecasting is one step ahead.

Various tasks in stock analysis, such as price forecasting [21], volatile forecasting [22], profit forecasting [23], price movement prediction [24], trend prediction [25], and signal prediction [26], are already done. This research focuses on stock price forecasting with price movement evaluation. Stock price forecasting uses various deep learning architectures to achieve the best performance for each stock issuer. The deep learning architectures already used are Convolutional Neural Netword (CNN) [27], LSTM [28], Gated Recurrent Unit (GRU) [29], CNN-LSTM [30], Graph Convolutional Network (GCN) [31], and Transformer [32]. The previous research concluded that GRU architecture performs best on 315 of 727 Indonesian stock issuers [20]. The other research comparing Transformer architecture produces Transformer Encoder Gated Recurrent Unit (TEGRU) as the best performance [33]. Kolmogorov-Arnold Network (KAN) [34], as an emerging neural architecture to solve various problems in numerous domains, has been extended in the time-series forecasting domain. Existing works utilizing KAN spans from the crypto-currency market [35] to satellite traffic [36], and stock market [37]. The problem with the previous research is that the model must be created on each stock issuer, which implies that high computation is required. This research proposes a unified model for all stock issuers to reduce computation resources and time. This research uses a dataset with a panel data format with cross-sectional (stock issuer) and time series (transaction date) as data indexes. This research compares the traditional regression panel data, machine learning, and deep learning architecture to find the best architecture for a unified stock price forecasting model. The modified Temporal Kolmogorov-Arnold Network (TKAN) is proposed to produce unified stock price forecasting models. The input of TKAN is modified with the rearrangement of the position dimension of sequence length or lookback with feature dimension. The permuted dimension in TKAN helps learn temporal dependencies in multivariate data.

In a previous study, ABSA used a separate process to determine aspect-sentiment quadruplet, which can affect the sentiment quantification process for stock price forecasting. To tackle that limitation, in this research, we present ABSA on a financial news dataset using generative text-to-text to produce a daily sentiment score for each stock issuer. The quantification process uses semantic similarity to tackle OOV limitations on the sentiment lexicon. Permuted Temporal Kolmogorov-Arnold Network (pTKAN) architecture produces the unified stock forecasting model to learn cross-sectional and time-series on each stock issuer in the panel dataset. The evaluation uses error analysis, representing value movement error and regression error with adaptive weighted, called Movement-Weighted Regression

Error (MWRE), to evaluate price movement (classification) and stock price (regression) predictions at once.

The significant contributions of this paper are summarized as follows.

1) Proposes a generative Aspect-Based Sentiment Analysis (ABSA) to accurately extract aspect and opinion terms for the sentiment quantification process.
2) Proposes a sentiment quantification process with semantic similarity using a sentiment lexicon for handling out-of-vocabulary problems.
3) Proposes a Permuted Temporal Kolmogorov-Arnold Network (pTKAN) to learn temporal dependencies in multivariate data.
4) Provide an in-depth analysis of price movement accuracy and regression error using the Movement-Weighted Regression Error (MWRE) evaluation method.

The remaining part of this article is organized as follows: Section II examines related work, Section III offers the suggested architecture, Section IV conducts experiments and results, Section V discusses, and Section VI concludes.

II. RELATED WORKS

This section presents a brief overview of the literature on stock price forecasting. It reviews previous research on aspect-based sentiment analysis and some related works in paraphrase generation. The issue of stock forecasting architecture, which contains traditional, machine learning, and deep learning architecture, is implemented in stock market or time series data.

A. ASPECT-BASED SENTIMENT ANALYSIS

ABSA has already made a significant contribution. The contribution focuses on extraction terms on aspect or sentiment, besides aspect categorization and sentiment polarity. In previous research, aspect sentiment classification was based on an aspect category using important term extraction with semantic similarity and Bidirectional Long Short-Term Memory (BiLSTM). The extracted important term is compared with the aspect keyword to get a similarity score employing word embedding and cosine distance to get aspect category and sentiment polarity. The BiLSTM helps in sentiment polarity in the majority class, and the semantic similarity method helps in the minority class [5]. The Target Aspect Sentiment Detection (TASD) task in financial news has already been done using word embedding from BERT and BiLSTM. The extracted aspect term is produced by Part of Speech (POS) tagging and dependency parse. The aspect category and sentiment polarity using semantic similarity from aspect term and aspect keyword are represented by BERT embedding [7].

The aspect sentiment quad prediction (ASQP) task in ABSA is to predict an aspect-sentiment quadruplet consisting of aspect category, aspect term, sentiment polarity, and opinion term. The previous result used two sequence tasks to produce an aspect-sentiment quadruplet. The aspect category and sentiment polarity employment binary text classification task to determine the target exists. The aspect and opinion term using BIO tagging to determine the word in the source text is targeted tag [38]. The ASQP task evolving using a generative model with the target is an aspect-sentiment quadruplet in naturalness form.

B. PARAPHRASE GENERATION

The generative model in ABSA uses sequence-to-sequence (seq2seq) architecture, the same as in machine translation. The seq2seq architecture consists of an encoder and a decoder. In previous research, the seq2seq used a stacked LSTM and residual layers in the encoder and decoder for paraphrase generation. The encoder accepts the source text as a word sequence, and the decoder predicts the next word sequence to produce the target text. The residual layer helps with the vanishing gradient problem in long sequences. The stacked LSTM with residual layer has limitations while paraphrasing short text, less than four words, because due to limitations on words candidate to be replaced [9]. Previous research used syntactic in input and output to control target words. The encoder parts accept the source text and source syntactic parse. The decoder parts have input target syntactic parse and target text. The encoder-decoder utilizes BART to learn the relation between the input encoder and the decoder. Adding the syntactic to the input increases performance in paraphrase generation [11].

The paraphrase generation for various ABSA tasks has already been done in previous research. The aspect-sentiment quadruplet from ABSA was changed to a natural language form as the target text. The paraphrase generation for ABSA, called PARAPHRASE, employs a T5 encoder-decoder pre-trained model for aspect-sentiment triplet extraction (ASTE), TASD, and ASQP tasks. The semantic label used in natural language form gains better performance but has limitations at predicted opinion terms in span form [12]. The other previous research in paraphrase generation for ABSA, called Generative Supervised Contrastive Learning Naturalness (GEN-SCL-NAT), employs supervised contrastive learning to keep away vector representation based on the labels. The labels used are aspect category, sentiment polarity, and opinion term. The SCL loss in GEN-SCL-NAT positively influences performance, but in the problem implicit aspect and implicit opinion, the PARAPHRASE achieved better performance than GEN-SCL-NAT [13].

C. STOCK MARKET ANALYSIS

Stock price forecasting has already been done in previous research using various architectures and features. The fundamental indicator [2], technical indicator, news sentiment [20], images [39], and stock transaction history [5]. Stock transactions are always used in stock price forecasting because they autocorrelate close prices. Additional features such as fundamental and technical indicators, news sentiment, and images do not always give better performance than stock

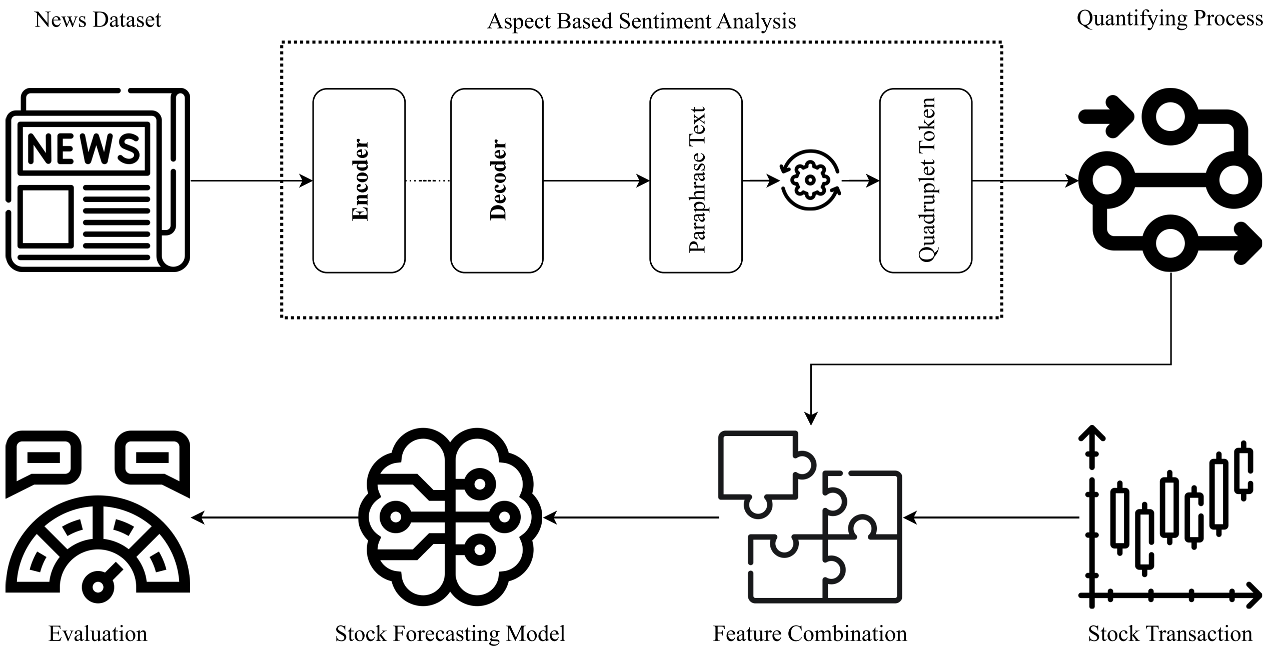


FIGURE 1. Proposed framework for unified stock price forecasting using generative aspect-based sentiment analysis.

transaction history. The technical indicator combined with stock transactions produces the best performance on 5 of 8 stock issuers. The news sentiment with stock transaction features produces the best performance on 3 of 8 stock issuers [20]. The images contain a stock transaction chart combined with a stock transaction time series, gaining better performance than stock transaction time series only in one stock issuer [39].

The various architecture used in stock market analysis from traditional methods, such as fixed effect model (FEM) [40], the machine learning method, such as Random Forest [41], [42], the deep learning method, such as MLP, LSTM, GRU, and Transformer [33], [43]. The ensemble method is also used to achieve better performance in stock price forecasting [30]. The traditional method, Common Effect Model (CEM), Random Effect Model (REM), and FEM exploit the dataset panel to analyze the effect of interest rate on stock return in spatial and temporal [40]. The machine learning employed to panel data improves the performance of baseline linear regression in time series forecasting [42]. The MLP is used in panel data to forecast realized volatility and has increased forecasting performance compared to the traditional methods, heterogeneous autoregressive and ordinary least squares model [43].

D. TRADITIONAL METHOD

The traditional method is already used in panel data to analyze the influences of two features or forecasting tasks. The Common Effect Model (CEM) learns various stock issuers as cross-sectional and the timestep as time series with intercept is constant. The Fixed Effect Model (FEM) assumes each cross-sectional unit has unique characteristics and has not changed over time. The unique characteristic of each stock issuer produces intercept to represent fixed effect. The Random Effect Model (REM) has a constant intercept on each cross-sectional unit but has a random component for each stock issuer to represent the random effect. Respectively the CEM, FEM, and REM used the Eq. 1, Eq. 2, and Eq. 3.

$$y_{it} = \alpha + \beta X_{it} + \epsilon_{it} \tag{1}$$

$$y_{it} = \alpha_i + \beta X_{it} + \epsilon_{it} \tag{2}$$

$$y_{it} = \alpha + \beta X_{it} + u_i + \epsilon_{it} \tag{3}$$

where y_{it} is target feature in cross sectional unit i and timestep t. α is intercept. β is parameter for independent features X. ϵ is error. The α in Eq. 1 for CEM assumes it does not have different characteristics between cross-sectional and time series. The α_i in Eq. 2 for FEM assumes it has different characteristics in between cross-sectional or time series. The α in Eq. 3 for REM assumes it does not have different characteristics in between cross-sectional and time series but has unobserved random effect represented using u_i for cross-sectional or time series.

E. MACHINE LEARNING METHOD

The machine learning method, such as Support Vector Machine (SVM) and bagging variant, is implemented for time series analysis [44]. The SVM combined with the Genetic Algorithm (GA) successfully got better performance than the deep learning method in long-term stock price forecasting. The combined GA and rolling window forward technique optimizes the SVM hyperparameter to avoid overfitting the training data and generalize the model in

TABLE 1. Statistic of stock issuer.

Ticker	Sector	Indices	VIF	Mean	Deviation
ACES	Non-primary goods	LQ45-ESG	210.36	711.34	460.88
ADRO	Energy	LQ45-ESG	276.94	904.49	624.17
AKRA	Energy	LQ45-ESG	215.79	560.72	396.94
ANTM	Raw goods	LQ45-ESG	136.19	1,200.04	640.44
ASII	Industry	LQ45-ESG	435.82	3,006.34	2,151.38
BBCA	Finance	LQ45-ESG	1,140.02	2,906.23	2,694.75
BBNI	Finance	LQ45-ESG	522.39	1,933.40	1,358.01
BBRI	Finance	LQ45-ESG	885.04	1,598.70	1,463.78
BMRI	Finance	LQ45-ESG	1,122.96	1,762.70	1,458.30
BRPT	Raw goods	LQ45-ESG	48.66	302.42	382.99
CTRA	Property	ESG	232.68	609.85	429.70
HMSP	Primary goods	ESG	178.60	997.20	915.67
ICBP	Primary goods	LQ45-ESG	446.86	4,101.88	3,751.83
INCO	Raw goods	LQ45-ESG	180.82	3,456.32	1,508.60
INTP	Raw goods	LQ45-ESG	287.35	8,379.02	5,537.39
ITMG	Energy	LQ45-ESG	311.01	9,167.48	6,604.60
JSMR	Infrastructure	ESG	263.66	3,942.97	1,493.35
KLBF	Health	LQ45-ESG	409.18	894.08	608.44
MDKA	Raw goods	LQ45-ESG	116.88	1,693.40	1,426.82
MEDC	Energy	LQ45-ESG	92.28	493.04	264.69
PGAS	Energy	LQ45-ESG	136.01	1,725.90	930.72
PTBA	Energy	LQ45-ESG	166.93	873.83	653.03
SMGR	Raw goods	LQ45-ESG	283.97	7,223.34	3,212.00
TLKM	Infrastructure	LQ45-ESG	341.20	2,037.57	1,073.79
TPIA	Raw goods	ESG	32.07	1,021.49	1,071.92
UNTR	Industry	LQ45-ESG	298.30	9,772.53	7,675.86
UNVR	Primary goods	LQ45-ESG	389.30	3,839.03	2,706.59

testing data [45]. The XGBoost method predicts the direction of stock movement and then generates the trading signal. The stock movement direction was extracted using previous and current prices. The XGBoost, using the technical indicator, predicts the direction of movement. The trading signal is generated by the rule-based prediction of movement direction, and the first label in the series of up and down labels is converted to a buy or sell trade signal [46].

F. DEEP LEARNING METHOD

The deep learning methods, CNN, LSTM, and GRU single layer or combined of CNN, LSTM, and GRU, are already used for stock price forecasting [33]. The stacked method consists of decomposition, CNN, and LSTM gain outperformed the LSTM single layer and CNN-LSTM without decomposition on stock price forecasting. The time series feature decomposes into several intrinsic mode functions (IMF) and one residual. The CNN learns the patterns from decomposed features and passes the information to LSTM to learn the sequential patterns [47]. The cutting-edge Transformer Encoder in the Transformer approach can pull information from multivariate characteristics and feed it to the GRU layer, which learns sequential patterns. The Transformer Encoder has multi-head attention and is optimized for learning information from several features, technical indicators, news sentiment, and stock transactions. Adding GRU as a decoder allows the Transformer Encoder to learn sequential patterns and outperform other Transformer versions [20]. The other state-of-the-art architecture, TKAN, beat LSTM and GRU in the crypto-currency market in short-term or long-term forecasting. The TKAN with cell KAN imitated the LSTM architecture to learn sequence pattern [35].

III. PROPOSED METHOD

This section elucidates the dataset of news titles and stock transactions, process labeling aspect-sentiment quadruplet, generative ABSA, sentiment quantifying process, the architecture of the forecasting model, evaluations, and experiment setup. The dataset explains the statistics and the collection method. The labeling aspect-sentiment quadruplet ABSA uses different models from various tasks. The architecture of forecasting models contains a modified transformer and TKAN architecture. Shown in Fig. 1, the proposed framework uses generative ABSA to process news datasets and produces aspect-sentiment quadruplet in qualitative form. The aspect-sentiment quadruplet is used in the quantifying process to change qualitative into quantitative form, which is the daily sentiment score. The daily sentiment score combined with stock transactions on each stock issuer in panel data form. The proposed forecasting method, pTKAN, is used to learn panel data and is compared with previous traditional, machine learning and deep learning methods. The best forecasting method is selected by employing six evaluation methods.

A. DATASET

This sub-section describes the dataset collection method, statistics, and preparation. The dataset news title was prepared to fulfill the requirement of the generative ABSA model. The stock transactions on each stock issuer are prepared for unified stock price forecasting models, which contain splitting training and testing datasets, scaling normalization, and creating different inputs for traditional, machine learning, and deep learning models.

This research uses the stock issuer from LQ45 and the Environmental, Social, and Governance (ESG) indices in the Indonesia Stock Exchange (IDX). The LQ45 indices are a stock issuer group with a big capitalization in the market in 12 months at the latest. The ESG indices are stock issuer groups that are aware of the environment and social and good corporate governance, which can be affected by market sentiment. The stock issuers in LQ45 and ESG indices are representations of the stock market in Indonesia, which has 83 companies, but this research only uses 27 companies. The criterion for a selected company, a stock transaction, has ranged from 1 January 2017 to 31 May 2024. The stock issuer must have a multicollinearity of more than five, and news on each stock issuer must have more than 400 news titles. Seven years of daily stock transaction data ensure that the selected companies manifest diverse movement patterns within the stock market and provide a sufficient dataset suitable for deep learning methodologies, which necessitate extensive datasets. A high multicollinearity value among companies is crucial in panel data analysis, as it captures the interdependencies of stock prices across different entities. The requirement for a minimum number of news titles guarantees that each company is significantly influenced by news sentiment.

1) DATASET OF STOCK PRICES

The yfinance library in Python is utilized to gather stock transaction history. Daily transaction data is obtained, which includes details such as transaction date, open price (O), lowest price (L), highest price (H), close price (C), transaction volume (V), dividend (D), and stock split (S). The research utilizes stock transaction information, which consists of OLHCV and transaction date.

The multicollinearity score on each stock issuer is calculated using the Variance Inflation Factor (VIF) method. As shown in Table. 1, the VIF score between stock issuers is above 5, which is the threshold for low multicollinearity [20], which indicates close stock price between stock issuers is interrelated. The three highest VIF scores from the finance sector are BBCA, BMRI, and BBRI. The three lowest VIF scores are TPIA (raw goods), BRPT (raw goods), and MEDC (energy). The interesting aspect of the stock transaction dataset is that the standard deviation from TPIA and BRPT is larger than the mean, which means that two stock issuers have high volatility.

2) DATASET OF NEWS HEADLINES

The news portal kontan.co.id in the Indonesian language is used as a source of the news dataset with news published dates between 1 January 2017 and 31 May 2024. The news dataset contains the title, description, published date, and category gathered using scraping technique under sections investment, industry, and finance. This research uses the news section in the news portal as an aspect category. The rule of collecting the dataset is the stock issuer ticker is shown explicitly in the news title to distinguish news on each stock issuer.

As shown in Fig. 2, BBRI has the most significant news records containing the investment and finance categories. The lowest record is ADRO, which contains investment and industry categories. The finance category only appears on stock issuers in finance sectors. This research uses the news published date for integration with the stock transaction dataset as a key value, the news category as an aspect category, the stock ticker for separate news on each stock issuer, and the news title for generative ABSA as the source text. In this research, one news title can have more than one stock issuer, so one source text can have some target text.

B. LABELING ASPECT-SENTIMENT QUADRUPLET

Generative ABSA requires an aspect-sentiment quadruplet to create target text. The aspect-sentiment quadruplet comprises an aspect category, sentiment polarity, aspect term, and opinion term. The news dataset already has an aspect category from the news category. The sentiment polarity, aspect term, and opinion term are produced using the framework shown in Fig. 3. The news dataset in Indonesian is translated into English using a translation application because the trained model siebert/sentiment-roberta-large-english is only

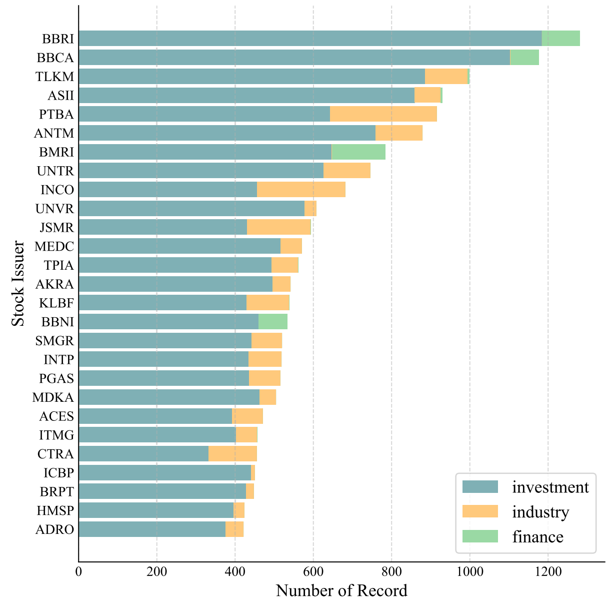


FIGURE 2. Number of record News titles on each stock issuer based on aspect category.

available in English. The trained model siebert/sentiment-roberta-large-english is a binary sentiment classification with an accuracy of 93.2% [48] and is used to determine sentiment polarity, negative and positive. The news dataset contains news published date, aspect category, sentiment polarity, stock ticker, and news title.

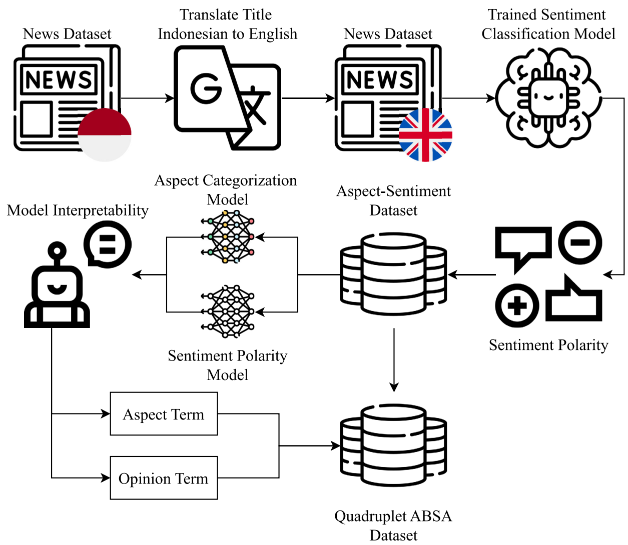


FIGURE 3. Framework of labeling aspect-sentiment quadruplet.

The pair aspect category-sentiment polarity with news titles was trained using various pre-trained models such as indobert-large-p2 [49], mxbai-embed-large-v1 [50], multilingual-e5-large [51], distiluse-base-multilingual-cased-v2 [52], and bge-base-en-v1.5 [53] to develop aspect categorization and sentiment classification models. The indobert-large-p2 model is specifically designed for Indonesians, while the other models are multilingual. The performance of the classification models is determined based on metrics such as precision, recall, F1-score, and accuracy. The top-performing model for aspect categorization and sentiment classification is utilized for extracting aspect terms and opinion terms using the deep learning model interpretability provided by the captum library [54].

The captum library computes input contribution scores by evaluating the probability of each output token in relation to the input tokens. Integrated gradients within captum calculate these contribution scores by aggregating the gradients from each layer, starting from the output and moving towards the input. More precisely, Integrated gradients are defined as the path integral of the gradients along a path between the source and the target [55]. The contribution score from integrated gradients is used to select the aspect and opinion terms with the highest absolute value.

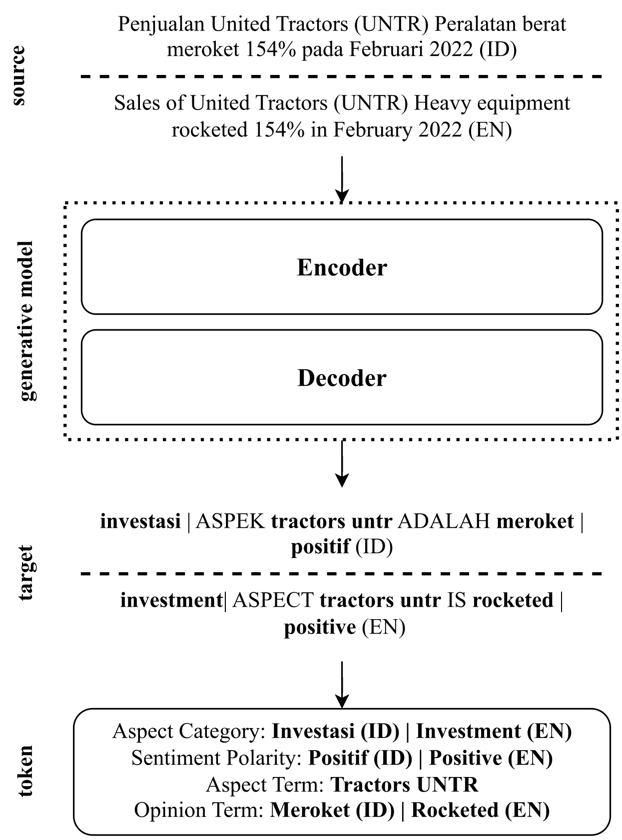


FIGURE 4. Generative aspect-based sentiment analysis framework.

C. PARAPHRASE GENERATION AND QUANTIFYING PROCESS

Generative ABSA uses a paraphrase generation framework to process sources and target text. The target text contains aspect category, sentiment polarity, aspect term, opinion term, and separator token. As shown in Fig. 4, the target text is parsed using a separator token to produce an aspect-sentiment quadruplet ABSA. This research uses a paraphrase generation framework from previous research, which is Generative Supervised Contrastive Learning Naturalness (GEN-SCL-NAT) [13]. This research modified the separator token and sequence-to-sequence (seq2seq) pre-trained model. We added a new seq2seq pre-trained model, FlanT5-Small, FlanT5-Base, and Bart-Base, then compared it with T5-Small and T5-Base [8], [56], [57]. The SCL loss in the training process uses aspect category, sentiment polarity, and opinion term as anchors to distinguish word vector representation on each label. The SCL loss achieved better performance than the cross-entropy loss on the paraphrase generation task to generate an aspect-sentiment quadruplet.

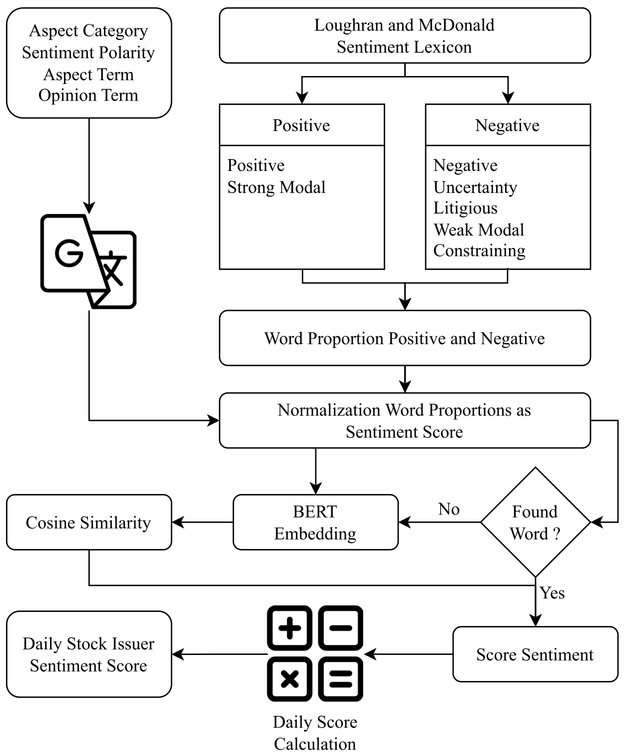


FIGURE 5. Quantifying sentiment label to daily stock issuer sentiment score.

As shown in Fig. 5, the aspect-sentiment quadruplet from Generative ABSA is used to produce the daily stock issuer sentiment score. The aspect and opinion terms were translated from Indonesian to English to have the same language as the Loughran and McDonald Sentiment Lexicon (LMSL). The LMSL had seven sentiment polarities: positive, strong modal, negative, uncertainty, litigious, weak modal, and constraining. This research grouped seven categories into positive and negative, the same as sentiment polarity from the aspect-sentiment quadruplet. The negative polarity contains negative, uncertainty, litigious, weak modal, and constraining categories, and the rest of the categories are positive polarity. The sentiment score in LMSL uses word proportions because of the representation of important words at frequency in

TABLE 2. Stock transaction panel data with a daily sentiment score.

Ticker	Date	Open	Low	High	Close	Volume	Positive	Negative
ACES	2017-01-02	711.806	711.806	711.806	711.806	0.000	0.000	-0.012
ACES	2017-01-03	711.806	711.806	677.708	690.495	10,432,600.000	0.000	-0.004
ACES	2017-01-04	690.495	690.495	664.921	681.970	6,997,800.000	0.001	0.000
ACES	2017-01-05	686.232	690.495	669.183	677.708	2,313,800.000	0.005	0.000
ACES	2017-01-06	681.970	681.970	669.183	677.708	2,208,600.000	0.003	-0.002
...	...	...	...	...	...	...	...	...
BBRI	2020-05-15	1,814.921	1,830.172	1,685.284	1,708.161	654,135,428.000	0.000	-0.007
BBRI	2020-05-18	1,715.786	1,731.038	1,647.155	1,654.781	501,017,513.000	0.002	-0.002
BBRI	2020-05-19	1,715.786	1,845.424	1,715.786	1,807.295	584,451,204.000	0.004	0.000
BBRI	2020-05-20	1,799.669	1,906.430	1,769.167	1,891.178	433,842,983.000	0.002	-0.009
BBRI	2020-05-26	1,944.558	1,944.558	1,883.552	1,921.681	444,514,707.000	0.002	0.000
...	...	...	...	...	...	...	...	...
UNVR	2024-05-27	3,090.000	3,270.000	3,090.000	3,260.000	54,566,200.000	0.008	-0.009
UNVR	2024-05-28	3,290.000	3,390.000	3,160.000	3,280.000	29,382,200.000	0.004	0.000
UNVR	2024-05-29	3,280.000	3,280.000	2,970.000	3,010.000	28,321,900.000	0.000	0.000
UNVR	2024-05-30	3,000.000	3,190.000	2,970.000	3,110.000	32,045,000.000	0.005	-0.001
UNVR	2024-05-31	3,160.000	3,230.000	3,120.000	3,120.000	93,435,800.000	0.001	0.000

each category. Scaling normalization from zero to one is implemented in word proportions.

The aspect and opinion term matchmaking with a word list in the LMSL is conducted to get a sentiment score. If the term is OOV, then the term and lexicon in the LMSL are converted to word vector using the Bidirectional Encoder Representation of Transformer (BERT) pre-trained model [58]. Word vectors from terms and lexicons are used to calculate word distance using cosine similarity. The lexicon with the lowest distance from the term is used to get the sentiment score. After getting the sentiment score on each news title, calculate the daily stock issuer sentiment score based on aspect category and sentiment polarity. The negative sentiment has a negative value, and the positive sentiment has a positive value. The daily stock issuer sentiment score using the mean sentiment score on each day trading, if not found, filled with zero value. The daily stock issuer sentiment score can be combined with stock transaction history. The illustration of the final dataset in the form of panel data is shown in Table. 2.

D. FORECASTING MODELS

This sub-section presents modified Transformer architecture, which are Transformer GRU (FormerGRU), Transformer LSTM (FormerLSTM), Transformer Encoder GRU (TEGRU), and Transformer Encoder LSTM (TELSTM). Modified Transformer was chosen because TEGRU outperformed other Transformer variants, such as vanilla Transformer, Informer, FEDformer, and Transformer Encoder in stock price forecasting [20]. This research proposes permuted Temporal Kolmogorov-Arnold Networks (pTKAN) architecture to develop a unified stock forecasting model since TKAN performed better than GRU and LSTM in crypto-currency market forecasting [35].

1) MODIFIED TRANSFORMER

The encoder-decoder architecture is used in Transformer architecture with positional encoding and multi-head activation functions. Two encoder layers are stacked with each encoder. The consecutive layers are multi-head attention, normalization, feed-forward networks (FFN), and normalization. Two decoder blocks are stacked with each decoder's consecutive layers: multi-head attention, normalization, multi-head attention, normalization, feed-forward networks (FFN), and normalization. The encoder accepts a sequence of multivariate stock transactions, produces the encoder output matrix, and fed to the decoder. The decoder block generates the decoder output using the time sequence of multivariate stock transactions and the output encoder block.

Vanilla Transformer uses a linear layer to predict stock price one day ahead using a decoder output projection. The FormerGRU uses a GRU layer and a dropout layer for decoder output projection, followed by a linear layer for a predicted stock price. FormerLSTM utilizes decoder output to predict stock price one step ahead using consecutive layers: LSTM layer, dropout layer, and linear layer.

TEGRU only uses the Transformer encoder, which produces the encoder output matrix and then processes it into the GRU, dropout, and linear layers. TELSTM uses an LSTM layer, dropout layer, and linear layer to process the encoder output matrix from the Transformer encoder to predict stock prices one step ahead. Modified Transformer architecture is conducted to find the effect of the encoder and decoder section, which is not built natively for sequence tasks because it needs positional encoding for the representation of sequence, combined with LSTM and GRU, which are native to sequence tasks.

2) TEMPORAL KOLMOGOROV-ARNOLD NETWORK

Temporal Kolmogorov-Arnold Networks (TKAN) introduced Recurrent Kolmogorov-Arnold Networks (RKAN) to handle time dependency with KAN ability. The recurrent memory in hidden state RKAN is defined in Eq. 4.

$$h_{l,n}(t) = W_{hh}h_{l,n}(t-1) + W_{hz}x_{l,n}(t) \tag{4}$$

where $hl, i(t)$ is hidden state in node n in l-th layer, t is time-step, x is input, and W is weight vector. Inspired by LSTM, TKAN uses hidden state RKAN to build neural network architecture. TKAN consists of an input gate, a forget gate, and an output gate.

Eq. 5 to Eq. 11 show architecture TKAN, with i as input gate, f as forget gate, o as output gate. The i, f, and $\widetilde{c}$ calculate using weight vector W, input x, output previous state h_{t-1}, and bias b. The o calculate using KAN layer which contain RKAN. The f and i is used to produce current state c. Output current state h produce by o and c. Final predict y using weight vector W and b.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{5}$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{6}$$

$$o_t = \sigma(\mathrm{KAN}(x, t)) \tag{7}$$

$$\widetilde{c}_t = \sigma(W_{x\widetilde{c}}x_t + W_{h\widetilde{c}}h_{t-1} + b_{\widetilde{c}}) \tag{8}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t \tag{9}$$

$$h_t = o_t \odot \tanh(c_t) \tag{10}$$

$$y_t = W_{hy}h_t + b_y \tag{11}$$

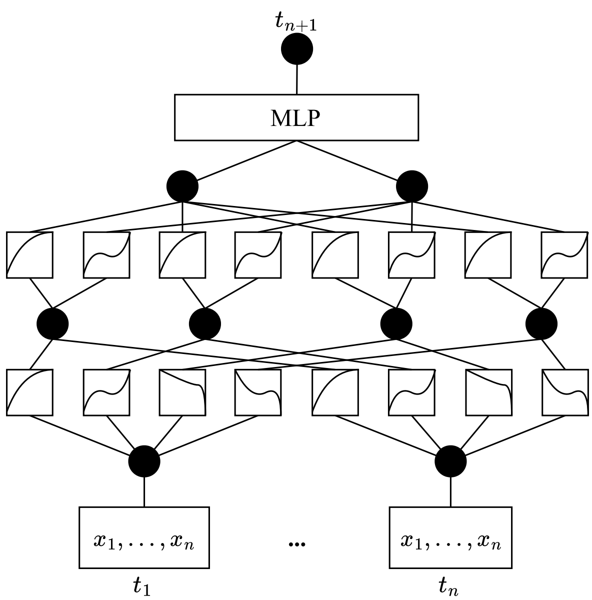


FIGURE 6. **Vanilla temporal kolmogorov-arnold networks architecture.**

3) PERMUTED TEMPORAL KOLMOGOROV-ARNOLD NETWORK

The original TKAN uses 3-dimensional input: batch size, lookback, and features. As shown in Fig. 6, the original TKAN process batch size and features on each time step to learn interaction between features over time. The original TKAN has a problem with the unified stock forecasting model using panel data because the feature on each stock issuer is continuous in batch. The features on each stock issuer with low correlation can happen in batch, which may

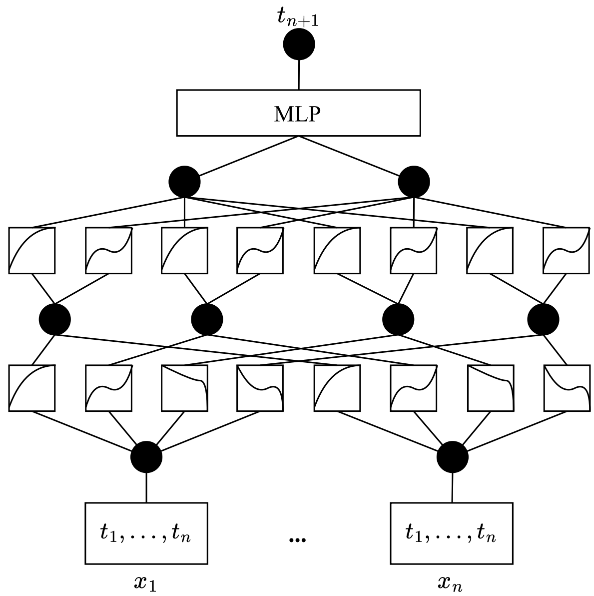


FIGURE 7. **Permuted temporal kolmogorov-arnold networks architecture.**

lead to overfitting on inter-feature correlations. Our proposed architecture permuted 3-dimensional input from a batch size, lookback, and features to a batch size, features, and lookback. As shown in Fig. 7, the permuted TKAN (pTKAN) processed isolated lookback on each feature. The isolated lookback is treated as univariate to learn temporal dependencies on itself. The problem of overfitting in batch in the original TKAN can be solved with an isolated lookback. The TKAN and pTKAN have the same architecture, except for the input, two stacked TKAN layers with 128 units, and one MLP layer for the final output.

This research uses TKAN and pTKAN with positional encoding and attention mechanisms from the Transformer Encoder to reveal the effect of positional encoding and attention mechanisms. The architecture of the TKAN attention mechanism (TKANatt), shown in Fig 8, and pTKAN attention mechanism (pTKANatt), shown in Fig. 9, is the same except for input. The input passes into positional encoding to get positional information and then processes with a single TKAN layer. The output from the TKAN layer is processed with the original Transformer Encoder to get multi-head attention information and then passes to a single TKAN layer before producing the final output with the MLP layer.

E. EVALUATION METRICS

The measured performance of each model is conducted on each process. The model used in the labeling aspect-sentiment quadruplet process was evaluated using F1-score, precision, accuracy, and recall. The best performance model is selected to produce aspect and opinion terms. The paraphrase generation process is evaluated using F1-score,

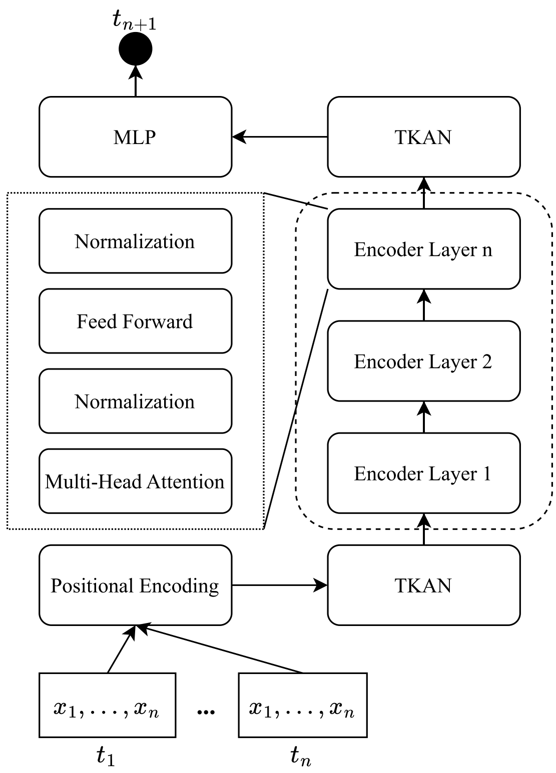


FIGURE 8. **Temporal Kolmogorov-Arnold with Encoder Blocks of Transformer Architecture.**

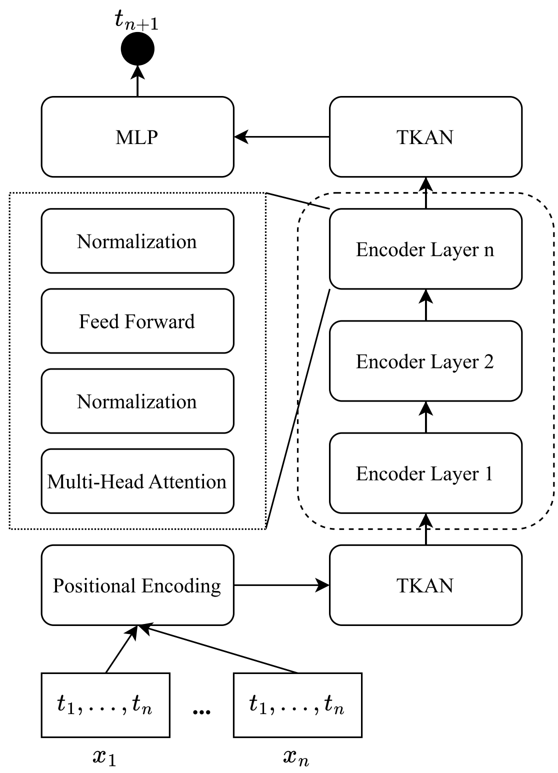


FIGURE 9. **Permuted Temporal Kolmogorov-Arnold with Encoder Blocks of Transformer Architecture.**

precision, accuracy, and recall to measure the predicted aspect category, predicted sentiment polarity, predicted pair of aspect categories, and predicted sentiment polarity. Predicted aspect-sentiment quadruplets from paraphrase generation are evaluated with generative evaluation methods, which are Bilingual Evaluation Understudy (BLEU), Word Error Rate (WER), and Recall-Oriented Understudy for Gisting Evaluation (ROUGE). The BLEU method is used to evaluate the quality of the predicted aspect-sentiment quadruplet in the form of natural language with reference text without concern for grammatical correctness. Two types of ROUGE methods are used to evaluate generated and reference text, which are ROUGE-1, which uses unigram, and ROUGE-L, which uses longest common subsequences (LCS). The WER method evaluates substitutions, insertions, and deletions in the generated text. The WER method is sensitive to changes in the aspect-sentiment quadruplet in the reference text that are not in the generated text or vice versa.

Inspired from previous research [20], evaluate regression model representing value movement accuracy and error regression. Our research proposes a general evaluation regression model, which combines the F1-score from the confusion matrix and Symmetric Mean Absolute Percentage Error (sMAPE). We added α variables for weighting the movement accuracy or regression error. A higher α places more emphasis on movement accuracy, while α lower places more emphasis on regression error. The unified stock price forecasting architecture approach was evaluated using the coefficient of determination (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), sMAPE, F1-score, and Movement-Weighted Regression Error (MWRE). The MWRE using Eq. 12, the sMAPE using Eq. 13, and the F1-score using Eq. 14.

$$\text{MWRE} = (1-\alpha)\cdot \text{sMAPE} + \alpha \cdot (1-\text{F1}) \tag{12}$$

$$\text{sMAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{2\cdot|y_t-\hat{y}_t|}{|y_t|+|\hat{y}_t|} \tag{13}$$

$$\text{F1-score} = 2\cdot\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}} \tag{14}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{15}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{16}$$

Up, down, and sideways movement labels are determined using Eq. 17 and Eq. 18. For actual movement direction M, a price change is labeled as sideways when the previous price y_{t-1} matches the current price y_t. A decrease in price from y_{t-1} to y_t results in a down label, while an increase corresponds to an up label. Similarly, for predicted movement direction $\hat{\text{M}}$, labels are assigned based on comparisons between the previous actual price y_{t-1} and the predicted current price $\hat{y}_t$: equal prices indicate sideways, a decrease indicates down, and an increase indicates up. These labels are then used to calculate True Positives (TP), False Positives (FP), and False Negatives (FN) using Eq. 14 to Eq. 16.

$$M = \begin{cases} \text{sideways}, & \text{if } y_{t-1} = y_t \\ \text{down}, & \text{if } y_{t-1} > y_t \\ \text{up}, & \text{if } y_{t-1} < y_t \end{cases} \tag{17}$$

$$\hat{M} = \begin{cases} \text{sideways}, & \text{if } y_{t-1} = \hat{y}_t \\ \text{down}, & \text{if } y_{t-1} > \hat{y}_t \\ \text{up}, & \text{if } y_{t-1} < \hat{y}_t \end{cases} \tag{18}$$

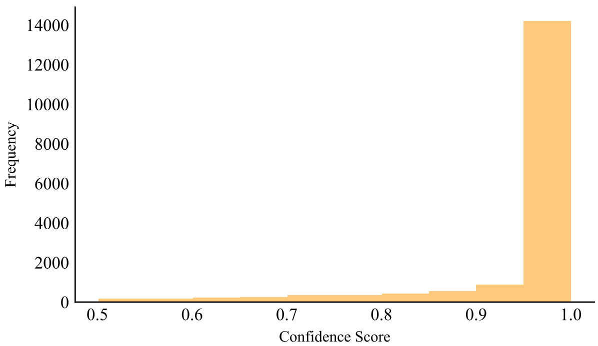


FIGURE 10. Confidence score on sentiment polarity labeling.

F. EXPERIMENT SETUP

In labeling the aspect-sentiment quadruplet, five pre-trained encoder models are employed to determine the optimal model for aspect categorization and sentiment polarity classification. The selection is made using 80% of the data for training and 20% for testing, with the dataset being split using the stratified method based on label aspect and sentiment. The pre-trained model is fine-tuned over three epochs, utilizing a batch size of eight and a learning rate of 0.00001. The best model was chosen based on the evaluation of F1-score, precision, accuracy, and recall, and it will be utilized to extract aspect and opinion terms across all datasets.

In paraphrase generation, the five pre-trained encoder-decoder models and GEN-SCL-NAT framework are employed to choose the highest performance model on generative ABSA. The dataset is split by published date from 1 January 2017 to 31 May 2023 as training data and from 1 June 2023 to 31 May 2024 as testing data. The dataset split by published date is conducted to restrict knowledge generative ABSA model on testing data, which is also used as testing data in stock price forecasting. The training GEN-SCL-NAT used 50 epochs, 16 batch sizes, three beams, and 50 epochs with five patience early stopping. The best generative ABSA model is evaluated by precision, recall, f1-score, accuracy, BLEU, ROUGE-1, ROUGE-L, and WER. The best generative ABSA model is used to predict aspect-sentiment quadruplet on all datasets and then passed to the sentiment quantification process.

TABLE 3. Statistic of news dataset.

Aspect	Positive	Negative
Investment	11,161	3,952
Industry	1,670	351
Finance	317	74
Total	13,148	4,377

TABLE 4. Comparison of aspect categorization model results.

Model	Pre	Rec	F1	Acc
mxbai-embed-large-v1	0.760	0.640	0.690	0.890
multilingual-e5-large	0.290	0.330	0.310	0.860
distiluse-base-multilingual-cased-v2	0.730	0.700	0.710	0.890
bge-base-en-v1.5	0.740	0.700	0.710	0.890
indobert-large-p2	**0.770**	**0.720**	**0.740**	**0.900**

TABLE 5. Comparison of sentiment classification model results.

Model	Pre	Rec	F1	Acc
mxbai-embed-large-v1	0.900	0.870	0.880	0.910
multilingual-e5-large	0.920	0.910	0.920	0.940
distiluse-base-multilingual-cased-v2	0.890	0.880	0.880	0.920
bge-base-en-v1.5	0.900	0.850	0.870	0.910
indobert-large-p2	**0.930**	**0.920**	**0.920**	**0.940**

In unified stock price forecasting, the dataset splitting uses transaction date from 1 January 2017 to 31 May 2023 as training data and from 1 June 2023 to 31 May 2024 as testing data. The dataset splitting rule on each stock issuer is equal to fulfill the requirement panel dataset, which has the same data amount for each stock issuer. In evaluating the unified stock price forecasting model, our proposed architecture, pTKAN, is compared with traditional, ML, and DL architecture. The traditional architecture is represented by the Common Effect Model (CEM), the Fixed Effect Model (FEM), and the Random Effect Model (REM), which are specializations of the panel data regression model. XGBoost, LightBGM, and CatBoost deputize the ML architecture. MLP, LSTM, GRU, CNN-LSTM, Transformer, and TKAN represent DL architecture. This research conducted modified Transformer and TKAN, which are TranformerGRU, TransformerLSTM, TEGRU, and TELSTM, TKANatt, and pTKANatt, to reveal positional encoder and attention mechanism effect in stock price forecasting. The traditional and ML architecture uses 1 of lookback to predict one-step price because both architectures only accept 2-dimensional data. The DL architecture uses 30 lookbacks, 32 batch sizes, 150 epochs with 50 patience early stopping, and a 0.001 learning rate. The best weight model based on minimum MSE loss is used to predict data testing. The dataset stock transaction combined with the daily sentiment score is compared with the dataset stock transaction only to analyze the effect of the sentiment score. The analysis effect of sentiment score is conducted on each stock issuer using the best-unified stock forecasting model. The best model was chosen by R^2, MAE, MSE, sMAPE, F1-score, and MWRE evaluation metrics.

TABLE 6. Comparison of generative ABSA in aspect categorization results.

Model	Pre	Rec	F1	Acc
Bart-Base	0.890	0.880	0.880	0.880
T5-Small	0.890	0.890	0.890	0.890
T5-Base	0.860	0.870	0.840	0.870
FlanT5-Small	0.870	0.890	0.880	0.890
FlanT5-Base	**0.900**	**0.890**	**0.900**	**0.890**
FlanT5-Base*	0.800	0.640	0.690	0.900

Best-performing model in generative ABSA
*Conventional classification using a generative model

TABLE 7. Comparison of generative ABSA in sentiment classification results.

Model	Pre	Rec	F1	Acc
Bart-Base	0.880	0.890	0.880	0.890
T5-Small	0.870	0.870	0.860	0.870
T5-Base	0.890	0.890	0.890	0.890
FlanT5-Small	0.890	0.890	0.880	0.890
FlanT5-Base	**0.890**	**0.900**	**0.890**	**0.900**
FlanT5-Base*	0.870	0.840	0.850	0.900

Best-performing model in generative ABSA
*Conventional classification using a generative model

IV. EXPERIMENT RESULT

This section describes the outcomes of each process: labeling aspect-sentiment quadruplet, generative aspect-based sentiment analysis, and stock price forecasting. The labeling aspect-sentiment quadruplet result contains statistics from the news dataset and a comparison model on aspect categorization and sentiment polarity classification. The comparison of the generative model is presented in the generative ABSA result. The stock price forecasting presents a comparison of 19 architectures on two dataset scenarios.

A. LABELING ASPECT-SENTIMENT QUADRUPLET

The sentiment polarity labeling using siebert/sentiment-roberta-large-english trained model. The confidence score distribution is shown in Fig. 10. The mean confidence score is 0.95, and the number of rows above mean confidence is 80.07%. The result indicates the trained model has high confidence in sentiment polarity labeling, and the sentiment label is used in opinion term extraction. This research used 17,525 news data, separated into 15,113 as aspect investment, 2,021 as aspect industry, and 391 as finance aspect. The statistics of the news dataset are shown in Table. 3.

The best model for aspect categorization and sentiment polarity classification is chosen to generate aspect and opinion terms. The indobert-base-p2 model beat the others in the model performance comparison of aspect classification, as Table 4 illustrates. In the aspect classification model, the indobert-base-p2 achieved 0.90 accuracy, 0.77 precision, 0.72 recall, and 0.74 F1-score. The second place has a different F1-score of 0.03. Table 5 shows the sentiment categorization model performance comparison. Once more, the indobert-base-p2 model performed better than the others. The sentiment classification model's indobert-base-p2 achieved 0.94 F1-score, 0.94 accuracy, 0.92 recall, 0.93 precision, and 0.92 recall. The Indobert-Base-P2 model enhances the aspect-sentiment quadruplet by extracting aspect and opinion terms using the Captum library.

TABLE 8. Comparison of generative ABSA in pair aspect-sentiment classification results.

Model	Pre	Rec	F1	Acc
Bart-Base	0.800	0.780	0.790	0.780
T5-Small	0.770	0.770	0.770	0.770
T5-Base	0.770	0.780	0.750	0.780
FlanT5-Small	0.780	0.790	0.770	0.790
FlanT5-Base	**0.810**	**0.800**	**0.800**	**0.800**
FlanT5-Base*	0.620	0.450	0.470	0.810

Best-performing model in generative ABSA
*Conventional classification using a generative model

TABLE 9. Comparison of generative ABSA model results.

Model	BLEU	ROUGE-1	ROUGE-L	WER
Bart-Base	0.522	0.789	0.780	0.184
T5-Small	0.623	0.827	0.824	0.147
T5-Base	0.623	0.825	0.823	0.148
FlanT5-Small	0.623	0.829	0.826	0.145
FlanT5-Base	**0.638**	**0.836**	**0.833**	**0.140**

B. GENERATIVE ASPECT-BASED SENTIMENT ANALYSIS

The generative ABSA processes news titles to aspect-sentiment quadruplets in natural language form with a token separator to parse the element during inferences. The generative ABSA uses a paraphrase generation framework to produce aspect category, sentiment polarity, aspect term, and opinion term. Various scenarios are evaluated to analyze the performance of the generative model. Table. 6 compares the generative ABSA model in aspect categorization. The FlanT5-Base outperformed the others; even though the accuracy of FlanT5-Base is equal with FlanT5-Small and T5-Small, the F1-score FlanT5-Base is the best. The FlanT5-Base in aspect categorization has 0.90 precision, 0.89 recall, 0.90 F1-score, and 0.89 accuracy.

In sentiment classification, the best generative ABSA is FlanT5-Base with 0.89 precision, 0.90 recall, 0.89 F1-score, and 0.90 accuracy. The comparison of the generative ABSA model in sentiment classification is shown in Table. 7. As shown in Table. 8, comparison of generative ABSA model in pair aspect-sentiment classification, present FlanT5-Base as the best model. The FlanT5-Base in pair aspect-sentiment classification has 0.81 precision, 0.80 recall, 0.80 F1-score, and 0.80 accuracy. The evaluation pair aspect-sentiment classification is conducted to reveal the relation aspect category and sentiment classification as a pair.

Another FlanT5-Base scenario is added to compare conventional classification and generative ABSA. The conventional classification uses a training dataset, testing dataset, and epoch with early stopping equally to generative ABSA. As shown in Table. 6 to Table. 8, the FlanT5-Base conventional classification performance not even close with the FlanT5-Base generative ABSA. The FlanT5-Base conventional classification is only better than the FlanT5-Base

TABLE 10. Comparison of unified stock price forecasting results without ABSA sentiment score.

Architecture	R^2	MAE	MSE
CEM	0.99947	69.34156	19,017.38782
REM	0.99947	69.34156	19,017.38782
FEM	0.99946	70.90242	19,372.35638
XGBoost	0.99808	105.58838	69,422.77859
LightGBM	0.99834	98.23507	60,047.07194
CatBoost	0.99825	105.31080	63,425.15953
MLP	0.99932	84.28837	24,677.97658
LSTM	0.99974	67.39176	9,279.41634
GRU	0.99986	49.43634	5,188.58184
CNN-LSTM	0.99983	49.80499	6,313.34780
Transformer	0.99992	42.54592	3,075.38631
FormerLSTM	0.99978	56.80479	8,139.42993
FormerGRU	0.99968	72.48244	11,782.88129
TELSTM	0.99986	45.30870	5,101.96885
TEGRU	0.99975	71.85916	9,015.89583
TKAN	0.99998	20.09218	900.64249
TKANAtt	0.99081	522.29964	333,226.07953
pTKAN	**0.99999**	**11.08959**	**450.02682**
pTKANAtt	0.99202	513.26394	289,165.51839

Architecture	sMAPE	F1	MWRE@0.75
CEM	0.00744	0.32963	0.50463
REM	0.00744	0.32963	0.50463
FEM	0.00786	0.32161	0.51076
XGBoost	0.00864	0.32217	0.51053
LightGBM	0.00834	0.32235	0.51032
CatBoost	0.00963	0.32410	0.50933
MLP	0.00973	0.32837	0.50615
LSTM	0.01468	0.42043	0.43834
GRU	0.01050	0.45378	0.41229
CNN-LSTM	0.00692	0.47730	0.39376
Transformer	0.00644	0.46445	0.40327
FormerLSTM	0.00827	0.47528	0.39560
FormerGRU	0.00979	0.43037	0.42967
TELSTM	0.00690	0.49520	0.38032
TEGRU	0.01594	0.42135	0.43797
TKAN	0.00403	0.55428	0.33530
TKANAtt	0.07658	0.24293	0.58695
pTKAN	**0.00143**	**0.60539**	**0.29631**
pTKANAtt	0.10297	0.28467	0.56224

Best-performing stock price forecasting architecture
Second best-performing stock price forecasting architecture

TABLE 11. Comparison of unified stock price forecasting results with ABSA sentiment score.

Architecture	R^2	MAE	MSE
CEM	0.99947	69.33875	19,016.86043
REM	0.99947	69.33875	19,016.86043
FEM	0.99946	70.89649	19,371.68468
XGBoost	0.99801	105.27605	72,041.30418
LightGBM	0.99843	96.76132	56,754.18872
CatBoost	0.99834	104.06354	60,254.51239
MLP	0.99932	94.16094	24,536.74073
LSTM	0.99981	48.35726	6,771.18137
GRU	0.99986	45.81115	5,031.56097
CNN-LSTM	0.99980	49.07813	7,139.87430
Transformer	0.99987	55.06966	4,870.30382
FormerLSTM	0.99962	82.42948	13,648.45649
FormerGRU	0.99965	88.12011	12,663.93928
TELSTM	0.99986	45.82395	4,987.97065
TEGRU	0.99960	78.71838	14,403.81592
TKAN	**0.99997**	**18.32822**	**916.76957**
TKANAtt	0.99221	433.95192	282,537.74491
pTKAN	0.99996	20.53953	1,280.51007
pTKANAtt	0.99804	187.01466	71,021.63010

Architecture	sMAPE	F1	MWRE@0.75
CEM	0.00744	0.33052	0.50397
REM	0.00744	0.33052	0.50397
FEM	0.00786	0.32151	0.51083
XGBoost	0.00861	0.32725	0.50672
LightGBM	0.00833	0.32185	0.51070
CatBoost	0.00949	0.32397	0.50940
MLP	0.01287	0.27760	0.54502
LSTM	0.00764	0.47820	0.39326
GRU	0.00931	0.46577	0.40300
CNN-LSTM	0.00730	0.46429	0.40361
Transformer	0.00720	0.43283	0.42717
FormerLSTM	0.01751	0.42342	0.43681
FormerGRU	0.01333	0.38578	0.46400
TELSTM	0.00662	0.48474	0.38810
TEGRU	0.01201	0.41884	0.43887
TKAN	**0.00295**	**0.56794**	**0.32478**
TKANAtt	0.07041	0.27096	0.56438
pTKAN	0.00298	0.55627	0.33354
pTKANAtt	0.01923	0.29366	0.53456

Best-performing stock price forecasting architecture
Second best-performing stock price forecasting architecture

generative ABSA in accuracy with a different 0.01 in aspect categorization and pair aspect-sentiment classification. The SCL loss in generative ABSA successfully distinguishes word representation based on aspect category, sentiment polarity, and opinion term.

The predicted aspect-sentiment quadruplet is evaluated using generative methods, which are BLUE, ROUGE-1, ROUGE-L, and WER. As shown in Table. 9, the FlanT5-Base is the best model to produce aspect-sentiment quadruplet with 0.638 BLEU, 0.836 ROUGE-1, 0.833 ROUGE-L, and 0.140 WER. The highest ROUGE-1 and ROUGE-L achieved by FlanT5-Base indicate strong performance capturing overlap of unigram and long matching sequences in predicted and reference text. The Bart-Base is unsuitable for our dataset because it gets the lowest performance in all scenario evaluations. The FlanT5-Base is the best model in all scenario evaluations because the FlanT5-Base has the biggest parameter compared to the others. The FlanT5-Base is selected to produce aspect-sentiment quadruplets on all news datasets, and then daily sentiment scores are produced, as mentioned in the section. III-C.

C. UNIFIED STOCK PRICE FORECASTING

The R^2, MAE, MSE, sMAPE, F1, and MWRE@0.75 are used to evaluate unified stock price forecasting. The MWRE@0.75 means the MWRE evaluation method using α 0.75 to emphasize movement direction error. Forecasting value in the financial domain not only discusses regression errors but also movement direction errors, which impact financial loss. This section compares unified stock price forecasting results with or without the ABSA sentiment score.

1) STOCK FORECASTING WITHOUT ABSA SENTIMENT

Unified stock forecasting uses open, high, low, and close prices and volume in previous transactions to predict prices one step ahead. Our proposed architecture, pTKAN, outperformed the traditional, ML, and other DL in 6 of 6 metrics evaluations, as shown in Table. 10 The pTKAN performance has 0.99999 R^2, 11.08959 MAE, 450.02682 MSE, 0.00143 sMAPE, 0.60539 F1-score, and 0.29631 MWRE@0.75. In second place filled by TKAN architecture with 0.99998 R^2, 20.09218 MAE, 900.64249 MSE, 0.00403 sMAPE, 0.55428 F1-score, and 0.33530 MWRE@0.75. The pTKAN has 50.03% lower MSE than TKAN, which is the most significant difference indicating pTKAN success lowering error above one. In MWRE evaluation, the pTKAN is 11.62% lower than TKAN, which means the pTKAN more accurately predicts stock price movement direction than TKAN.

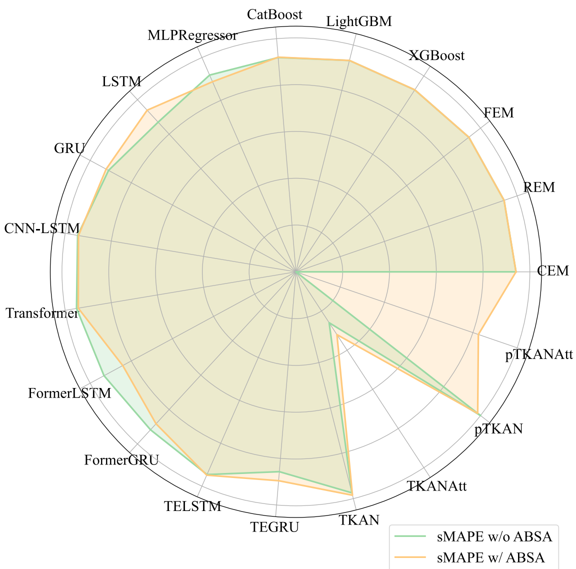


FIGURE 11. Scaled sMAPE comparison with and without ABSA sentiment score.

The best traditional architecture is CEM and REM with the same evaluation score, which are 0.99947 R^2, 63.34156 MAE, 19,017.38782 MSE, 0.00744 sMAPE, 0.32963 F1-score, and 0.50463 MWRE@0.75. The best ML architecture is LightGBM with evaluation scores are 0.99834 R^2, 98.23507 MAE, 60,047.07194 MSE, 0.00834 sMAPE, 0.32235 F1-score, and 0.51032 MWRE@0.75. The CEM and REM perform better than LightGBM. The best DL architectures without the Transformer variant and TKAN variant are GRU and CNN-LSTM. The GRU has best evaluation score in R^2, MAE, and MSE. The CNN-LSTM has the best evaluation score in sMAPE, F1-score, and MWRE@0.75. The GRU and CNN-LSTM perform better than CEM and REM. In the Transformer and TKAN variants, the vanilla Transformer has the best performance in 4 of 5 evaluation metrics with evaluation scores of 0.99992 R^2, 42.54592 MAE, 3,075.38631 MSE, 0.00644 sMAPE, 0.46445 F1-score, and 0.40327 MWRE@0.75. The TELSTM captures stock price movement better than the vanilla Transformer, with evaluation scores of 0.46445 F1-score and 0.40327 MWRE@0.75. Our architecture, which combined the TKAN variant and Transformer Encoder, fails to learn the dataset because it has the largest MSE and MWRE.

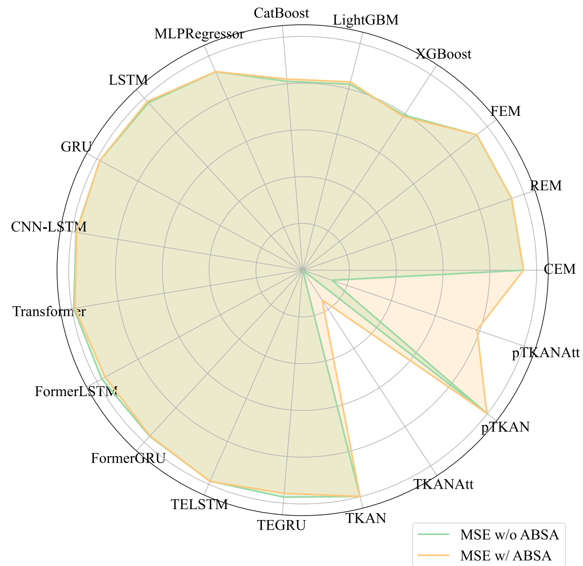


FIGURE 12. Scaled MSE comparison with and without ABSA sentiment score.

2) STOCK FORECASTING WITH ABSA SENTIMENT

The daily ABSA sentiment score is added in this scenario and then tested with 19 architectures. As shown in Table. 11, the pTKAN architecture as our proposed is filled second place best architecture with evaluation score is 0.99996 R^2, 20.53953 MAE, 1,280.51007 MSE, 0.00298 sMAPE, 0.55627 F1-score, and 0.33354 MWRE@0.75. The first place filled by TKAN architecture with evaluation score are 0.99997 R^2, 18.32822 MAE, 916.76957 MSE, 0.00295 sMAPE, 0.56794 F1-score, and 0.32478 MWRE@0.75. The TKAN has a better capture pattern of ABSA sentiment score than the pTKAN. The pTKAN architecture shows a 28.40% reduction in MSE compared to TKAN, indicating that pTKAN successfully reduced the error by more than one. In the MWRE evaluation, pTKAN outperforms TKAN by 2.62%, which more accurately predicts stock price movement direction.

The top traditional architectures, CEM and REM, achieved identical evaluation scores: 0.99947 R^2, 69.33875 MAE, 19,016.86043 MSE, 0.00744 sMAPE, 0.33052 F1-score, and 0.50397 MWRE@0.75. The best-performing machine learning model was LightGBM, with evaluation scores of 0.99843 R^2, 96.76132 MAE, 56,754.18872 MSE, 0.00833 sMAPE, 0.32185 F1-score, and 0.51070 MWRE@0.75. CEM and REM outperformed LightGBM. Among MLP, LSTM, GRU, and CNN-LSTM, GRU delivered the best performance in 3 out of 6 metrics, with scores of 0.99986 R^2, 45.81115 MAE, and 5,031.56097 MSE. CNN-LSTM excelled in the sMAPE metric, while LSTM surpassed MLP, GRU, and CNN-LSTM in F1-score and MWRE.

TABLE 12. Comparison of unified stock price forecasting results on each stock issuer.

Ticker	TKAN w/o ABSA			TKAN w/ ABSA			pTKAN w/o ABSA			pTKAN w/ ABSA		
	MAE	MSE	MWRE	MAE	MSE	MWRE	MAE	MSE	MWRE	MAE	MSE	MWRE
ACES	28.61213	911.37121	0.49499	14.45530	283.42823	0.37610	**6.75254**	**66.12386**	**0.36739**	6.95531	77.03580	0.36760
ADRO	13.23133	258.74527	0.33684	11.24207	202.09484	0.32462	**4.47067**	**30.52005**	**0.28662**	8.52817	108.45942	0.31839
AKRA	17.34742	365.31862	0.39919	10.82180	162.22087	0.35353	**8.56258**	**89.21619**	**0.31929**	14.28129	218.79907	0.39314
ANTM	13.17605	232.22604	0.37735	9.09244	123.58378	0.33771	**3.53751**	**17.32638**	**0.29326**	13.67711	208.35466	0.39280
ASII	11.21474	196.75827	**0.29428**	12.95194	266.90264	0.30782	**10.22869**	**160.57067**	0.29501	22.69524	710.16622	0.32448
BBCA	**25.50486**	**835.54053**	**0.32337**	43.18358	2,529.68523	0.35714	56.75573	4,283.05874	0.38733	78.07456	10,401.39641	0.41475
BBNI	11.44219	213.56669	0.29756	10.62606	181.16583	0.29629	**8.12134**	**105.17404**	**0.28671**	21.68102	601.97811	0.33794
BBRI	15.57831	378.80459	0.29078	**11.90047**	**218.12878**	**0.28845**	15.60121	986.99053	0.29924	28.50517	1,172.43576	0.33598
BMRI	14.95299	349.86946	0.29841	11.64698	222.04599	0.30260	**9.74811**	**221.63821**	**0.29736**	26.85032	996.99138	0.35126
BRPT	21.32481	675.91614	0.39747	15.73185	498.85355	0.34341	**6.80115**	**102.44322**	**0.29415**	11.84288	221.76682	0.32216
CTRA	14.92489	294.81520	0.39603	8.87924	125.60074	0.32994	**2.95814**	**14.33650**	**0.27249**	10.96461	130.45107	0.37978
HMSP	23.39427	584.59334	0.55389	12.33972	185.70488	0.46040	**2.78070**	**11.27154**	**0.32336**	7.08188	60.74710	0.42747
ICBP	20.83976	612.48649	0.29132	17.94779	517.08575	0.29414	**11.32460**	**232.65704**	**0.27849**	18.70603	679.31333	0.28581
INCO	17.01994	493.75290	0.29880	16.32818	466.40379	0.29237	**8.08986**	**100.04317**	**0.28541**	17.53946	492.25326	0.29864
INTP	12.96895	275.19651	0.28697	14.22415	308.89647	0.28976	**9.02774**	**154.14614**	**0.28427**	16.98962	574.06304	0.29812
ITMG	61.83025	6,550.52697	0.28340	68.91092	8,472.07221	0.28565	**33.76457**	**1,921.27664**	**0.26986**	43.71950	3,035.61277	0.27563
JSMR	16.80613	460.69010	0.29373	17.40690	470.60228	0.28345	**7.37903**	**96.70462**	**0.27553**	12.86233	245.00592	0.29801
KLBF	12.68581	235.27781	0.37833	12.19572	220.59438	0.37829	**4.26934**	**27.51839**	**0.30226**	11.18515	148.55899	0.34971
MDKA	17.51662	545.58032	0.30386	17.21926	523.85635	0.31706	**5.51229**	**52.67033**	**0.28033**	19.48867	500.61208	0.31578
MEDC	16.60318	455.04315	0.35973	14.02147	337.33618	0.33347	**5.05479**	**45.74275**	**0.29181**	11.70212	174.90922	0.33175
PGAS	13.07415	229.18480	0.46616	8.95585	118.83448	0.42525	**2.46841**	**10.40109**	**0.29987**	10.16829	118.83046	0.38251
PTBA	9.85465	235.66138	0.28830	12.47829	290.51387	0.34126	**3.82563**	**30.93475**	**0.28281**	14.60920	249.92174	0.32784
SMGR	13.81163	306.26509	0.29401	16.71148	403.21228	0.30163	**6.97181**	**85.33788**	**0.29136**	19.99651	567.14691	0.30266
TLKM	10.82935	183.35959	0.31416	9.86884	164.44480	0.30905	**6.06544**	**63.22198**	**0.28876**	7.53124	91.65892	0.29093
TPIA	29.38924	2,319.57803	**0.30469**	28.85787	2,360.43165	0.32440	**24.89323**	**1,876.37742**	0.30875	44.50040	9,205.62648	0.31729
UNTR	63.13909	5,738.63942	0.28245	52.36029	4,748.89622	0.28326	**29.13058**	**1,323.77630**	**0.26541**	44.59811	3,434.27512	0.27023
UNVR	15.41678	378.72299	0.30839	14.50368	350.28640	0.30123	**5.32371**	**41.27645**	**0.28113**	9.83022	147.25292	0.29736

Best-performing stock price forecasting architecture
Second best-performing stock price forecasting architecture

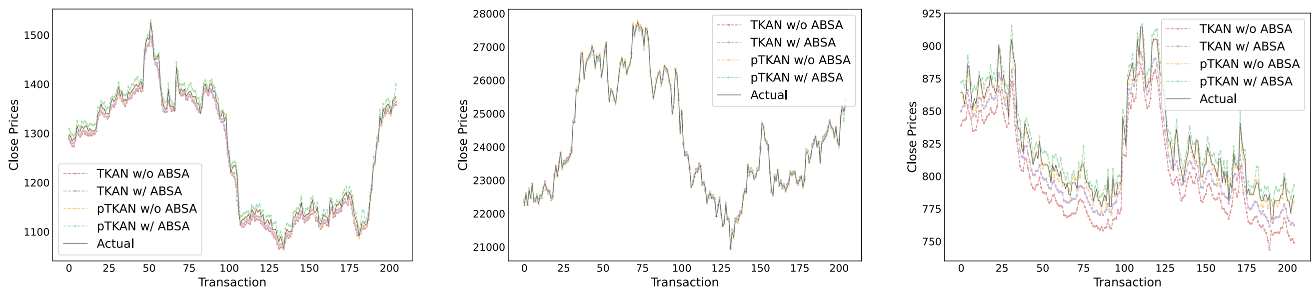
Among Transformer and TKAN variants, TELSTM led in 4 out of 5 metrics, with scores of 0.99986 R^2, 45.82395 MAE, 4,987.97065 MSE, 0.00662 sMAPE, 0.48474 F1-score, and 0.38810 MWRE@0.75. However, the vanilla Transformer was more effective in reducing significant errors, with an MSE of 4,870.30382. Our combined TKAN variant and Transformer Encoder architecture failed to learn the dataset, showing the worst MSE and MWRE.

V. DISCUSSION

The influence of daily ABSA sentiment score in unified stock price forecasting was analyzed using sMAPE and MSE evaluation metrics. The sMAPE represents a symmetric absolute percentage error that can be reduced by the daily ABSA sentiment score. The MSE represents the large error the daily ABSA sentiment score can reduce. Fig. 11 and Fig. 12 use error normalization scaling to show different performances with and without daily ABSA sentiment scores in various unified stock price forecasting. The higher value is better performance. As shown in Fig. 11 using scaled sMAPE, the daily ABSA sentiment score positively influences 13 of 19 architectures. The scaled MSE in Fig. 12 reveals the CEM, REM, FEM, LightGBM, CatBoost, MLP, LSTM, GRU, TELSTM, TKANAtt, and pTKANAtt success in lowering significant error while using daily ABSA sentiment score. The right combination of architecture and features is vital to produce the highest performance in stock price forecasting.

The influence analysis of the daily ABSA sentiment score is detailed in each stock issuer using TKAN and pTKAN architecture. The evaluation metrics used are MAE, MSE, and MWRE@0.75 to determine the best architectures and features in stock price forecasting. As shown in Table. 12, pTKAN without daily ABSA sentiment score outperformed the other scenario, except in BBRI and BBCA stock issuer. The most immense amount of news data is owned by the BBRI stock issuer, which succeeds in increasing the performance of BBRI stock price forecasting. The second largest news data, BBCA, is still not enough to achieve better performance while using daily sentiment score. The daily sentiment score increased performance while using the TKAN architecture, indicated by the amount of the second place best performance on TKAN with a daily ABSA sentiment score is more than TKAN with the stock transactions only.

According to the MAE evaluation, the best performance was achieved by the PGAS stock using pTKAN without daily ABSA sentiment, with an MAE of 2.46841. In contrast, the ITMG stock using TKAN with daily ABSA sentiment had the worst performance, with an MAE of 68.91092. For the MSE metric, PGAS again performed the best using pTKAN without daily ABSA sentiment, registering an MSE of 10.40109, while the worst performance came from the BBCA stock, which had an MSE of 10,401.39641 when using pTKAN with daily ABSA sentiment. Regarding MWRE@0.75, the UNTR stock performed best using pTKAN without daily ABSA sentiment, with an MWRE of 0.26541. On the other



(a) PGAS (b) UNTR (c) HMSP

FIGURE 13. Comparison of the predicted close price using unified stock price forecasting in the stock issuer.

hand, the worst performance for MWRE@0.75 was from HMSP using pTKAN without daily ABSA sentiment, with a value of 0.55389. The best performance of each stock issuer is influenced by appropriate architecture and features. The pTKAN architecture is suitable for unified stock price forecasting in almost all stock issuers. The daily ABSA sentiment score can increase performance with the right stock price forecasting architecture, which is TKAN.

The predicted close price from PGAS, UNTR, and HMSP stock issuers is chosen to prove the lower regression error is not always better. Predicted close prices from PGAS with TKAN and daily ABSA sentiment score have a low MAE of 8.95585 but a high MWRE@0.75. As shown in Fig. 13, the predicted close prices from PGAS with TKAN and daily ABSA sentiment score are close to the actual prices, but sometimes the predicted price under the actual prices, which indicates the daily movement direction is wrong. The line of predicted stock prices UNTR from all TKAN and pTKAN architecture in Fig. 13 follows the line of the actual prices, in which the daily movement direction of stock prices is almost similar to actual prices, even though the MSE is above ten. The worst movement direction is represented by the HMSP stock issuer. The predicted close prices from the HMSP stock issuer are far from actual prices and lower than actual prices, especially when using TKAN architecture without a daily ABSA sentiment score. The movement direction in stock price forecasting is crucial to take action in the stock market. The wrong action in the stock market can inflict financial loss on investors. The MWRE evaluation method can present error regression and movement direction errors in one metric. The MWRE is a general method because parameters can be changed to emphasize one error or distribute the weight evenly.

VI. CONCLUSION AND FUTURE WORKS

This research proposed the Permuted Temporal Kolmogorov-Arnold Network (pTKAN) architecture for unified stock price forecasting using generative aspect-based sentiment analysis (ABSA). The generative ABSA outperforms the conventional classification in the pair aspect-sentiment classification task. The FlanT5-Base performed better than the other pre-trained model to produce aspect-sentiment quadruplet in natural form. The aspect-sentiment quadruplet quantifies daily ABSA sentiment score using the Loughran and McDonald Sentiment Lexicon with out-of-vocabulary handling. The daily ABSA sentiment score increases performance while using TKAN architecture. The pTKAN architecture excels in 25 of 27 stock issuers and is compared to 19 architectures, which include the traditional, ML, and DL architecture. The Movement-Weighted Regression Error evaluation method successfully represents movement direction error and regression error with general weighted in one metric.

This research still needs to resolve several issues, such as managing the news dataset containing multiple aspect-sentiment quadruplets in a single source text. Enhance the performance of pTKAN by utilizing daily ABSA sentiment scores for price forecasting in other stock markets to generalize the proposed method. Also, construct a stock price prediction model using MWRE as a loss function in the training model to reduce directional movement errors.

REFERENCES

[1] L. Nemes and A. Kiss, "Prediction of stock values changes using sentiment analysis of stock news headlines," *J. Inf. Telecommun.*, vol. 5, no. 3, pp. 375–394, Jul. 2021, doi: 10.1080/24751839.2021.1874252.

[2] Y.-J. Chen, Y.-M. Chen, and C. L. Lu, "Enhancement of stock market forecasting using an improved fundamental analysis-based approach," *Soft Comput.*, vol. 21, no. 13, pp. 3735–3757, Jul. 2017, doi: 10.1007/s00500-016-2028-y.

[3] P. Mehta, S. Pandya, and K. Kotecha, "Harvesting social media sentiment analysis to enhance stock market prediction using deep learning," *PeerJ Comput. Sci.*, vol. 7, p. e476, Apr. 2021, doi: 10.7717/peerj-cs.476.

[4] P.-Y. Hao, C.-F. Kung, C.-Y. Chang, and J.-B. Ou, "Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane," *Appl. Soft Comput.*, vol. 98, Jan. 2021, Art. no. 106806, doi: 10.1016/j.asoc.2020.106806.

[5] A. T. Haryono, R. Sarno, and R. Abdullah, "Aspect-based sentiment analysis of financial headlines and microblogs using semantic similarity and bidirectional long short-term memory," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 3, pp. 233–241, Mar. 2022. [Online]. Available: https://inass.org/wp-content/uploads/2021/12/2022063020-2.pdf

[6] A. N. Sutranggono and R. Sarno, "Detection and sentiment analysis based on mental disorders aspects using bidirectional gated recurrent unit and semantic similarity," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 4, pp. 1–15, 2024, doi: 10.22266/ijies2024.0831.01.

[7] F. Nurifan, R. Sarno, and K. Sungkono, "Aspect based sentiment analysis for restaurant reviews using hybrid ELMoWikipedia and hybrid expanded opinion lexicon-SentiCircle," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 6, pp. 47–58, Dec. 2019, doi: 10.22266/ijies2019.1231.05.

[8] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *J. Mach. Learn. Res.*, vol. 21, no. 1, Jan. 2020. [Online]. Available: http://jmlr.org/papers/v21/20-074.html

[9] A. Prakash, S. A. Hasan, K. Lee, V. Datla, A. Qadir, J. Liu, and O. Farri, "Neural paraphrase generation with stacked residual LSTM networks," in *Proc. COLING 26th Int. Conf. Comput. Linguistics, Tech. Papers*, Dec. 2016, pp. 2923–2934. [Online]. Available: https://aclanthology.org/C16-1275

[10] Y. Fu, Y. Feng, and J. P. Cunningham, "Paraphrase generation with latent bag of words," in *Proc. Adv. Neural Inf. Process. Syst.* Newry, U.K.: Curran Associates, vol. 32, 2019, pp. 1–12.

[11] J. Sun, X. Ma, and N. Peng, "AESOP: Paraphrase generation with adaptive syntactic control," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2021, pp. 5176–5189. [Online]. Available: https://aclanthology.org/2021.emnlp-main.420

[12] W. Zhang, Y. Deng, X. Li, Y. Yuan, L. Bing, and W. Lam, "Aspect sentiment quad prediction as paraphrase generation," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2021, pp. 9209–9219. [Online]. Available: https://aclanthology.org/2021.emnlp-main.726

[13] J. Peper and L. Wang, "Generative aspect-based sentiment analysis with contrastive learning and expressive structure," in *Proc. Findings Assoc. Comput. Linguistics, EMNLP*, 2022, pp. 6089–6095. [Online]. Available: https://aclanthology.org/2022.findings-emnlp.451

[14] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction," *Expert Syst. Appl.*, vol. 135, pp. 60–70, Nov. 2019, doi: 10.1016/j.eswa.2019.06.014.

[15] B. Fazlija and P. Harder, "Using financial news sentiment for stock price direction prediction," *Mathematics*, vol. 10, no. 13, p. 2156, Jun. 2022, doi: 10.3390/math10132156.

[16] 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," *Knowl.-Based Syst.*, vol. 247, Jul. 2022, Art. no. 108742, doi: 10.1016/j.knosys.2022.108742.

[17] N. Jing, Z. Wu, and H. Wang, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction," *Expert Syst. Appl.*, vol. 178, Sep. 2021, Art. no. 115019, doi: 10.1016/j.eswa.2021.115019.

[18] L. Xueling, X. Xiong, and S. Yucong, "Exchange rate market trend prediction based on sentiment analysis," *Comput. Electr. Eng.*, vol. 111, Oct. 2023, Art. no. 108901, doi: 10.1016/j.compeleceng.2023.108901.

[19] Y. Bi, H. Liu, R. Wang, and S. Li, "Predicting stock market movements through daily news headlines sentiment analysis: U.S. stock market," in *Proc. 2nd Int. Conf. Big Data Artif. Intell. Softw. Eng. (ICBASE)*, Sep. 2021, pp. 642–648, doi: 10.1109/ICBASE53849.2021.00127.

[20] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Transformer-gated recurrent unit method for predicting stock price based on news sentiments and technical indicators," *IEEE Access*, vol. 11, pp. 77132–77146, 2023, doi: 10.1109/ACCESS.2023.3298445.

[21] S. Ren, X. Wang, X. Zhou, and Y. Zhou, "A novel hybrid model for stock price forecasting integrating encoder forest and informer," *Expert Syst. Appl.*, vol. 234, Dec. 2023, Art. no. 121080, doi: 10.1016/j.eswa.2023.121080.

[22] Y. Yu, Y. Lin, X. Hou, and X. Zhang, "Novel optimization approach for realized volatility forecast of stock price index based on deep reinforcement learning model," *Expert Syst. Appl.*, vol. 233, Dec. 2023, Art. no. 120880, doi: 10.1016/j.eswa.2023.120880.

[23] J. Luo, G. Zhu, and H. Xiang, "Artificial intelligent based day-ahead stock market profit forecasting," *Comput. Electr. Eng.*, vol. 99, Apr. 2022, Art. no. 107837, doi: 10.1016/j.compeleceng.2022.107837.

[24] B. T. Khoa and T. T. Huynh, "Forecasting stock price movement direction by machine learning algorithm," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 6, pp. 6625–6634, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6625-6634.

[25] B. Pang, W. Wei, X. Li, X. Feng, and C. Li, "A representation-learning-based approach to predict stock price trend via dynamic spatiotemporal feature embedding," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 106849, doi: 10.1016/j.engappai.2023.106849.

[26] C. D. Kothapalli, G. Navya, U. Jaladhi, S. R. Sulthana, D. L. S. Kumar, and S. P. Praveen, "Predicting buy and sell signals for stocks using bollinger bands and MACD with the help of machine learning," in *Proc. Int. Conf. Sustain. Comput. Smart Syst. (ICSCSS)*, Jun. 2023, pp. 333–340, doi: 10.1109/ICSCSS57650.2023.10169500.

[27] W. Ye, J. Yang, and P. Chen, "Short-term stock price trend prediction with imaging high frequency limit order book data," *Int. J. Forecasting*, vol. 40, no. 3, pp. 1189–1205, Jul. 2024, doi: 10.1016/j.ijforecast.2023.10.008.

[28] J. Zhao, D. Zeng, S. Liang, H. Kang, and Q. Liu, "Prediction model for stock price trend based on recurrent neural network," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 1, pp. 745–753, Jan. 2021, doi: 10.1007/s12652-020-02057-0.

[29] C. Manjunath, B. Marimuthu, and B. Ghosh, "Deep learning for stock market index price movement forecasting using improved technical analysis," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 5, pp. 129–141, Oct. 2021, doi: 10.22266/ijies2021.1031.13.

[30] Z. Shi, Y. Hu, G. Mo, and J. Wu, "Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction," 2022, *arXiv:2204.02623*.

[31] H. Peng, K. Dong, and J. Yang, "Stock price movement prediction based on relation type guided graph convolutional network," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 106948, doi: 10.1016/j.engappai.2023.106948.

[32] F. Colasanto, L. Grilli, D. Santoro, and G. Villani, "AlBERTino for stock price prediction: A Gibbs sampling approach," *Inf. Sci.*, vol. 597, pp. 341–357, Jun. 2022, doi: 10.1016/j.ins.2022.03.051.

[33] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Stock price forecasting in Indonesia stock exchange using deep learning: A comparative study," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 14, no. 1, p. 861, Feb. 2024, doi: 10.11591/ijece.v14i1.pp861-869.

[34] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "KAN: Kolmogorov-Arnold networks," 2024, *arXiv:2404.19756*.

[35] R. Genet and H. Inzirillo, "TKAN: Temporal kolmogorov-Arnold networks," *SSRN Electron. J.*, pp. 1–6, May 2024, doi: 10.2139/ssrn.4825654.

[36] C. J. Vaca-Rubio, L. Blanco, R. Pereira, and M. Caus, "Kolmogorov-Arnold networks (KANs) for time series analysis," 2024, *arXiv:2405.08790*.

[37] K. Xu, L. Chen, and S. Wang, "Kolmogorov-Arnold networks for time series: Bridging predictive power and interpretability," 2024, *arXiv:2406.02496*.

[38] H. Cai, R. Xia, and J. Yu, "Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions," in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics, 11th Int. Joint Conf. Natural Lang. Process.* PA, USA: Association for Computational Linguistics, 2021, pp. 340–350. [Online]. Available: https://aclanthology.org/2021.acl-long.29

[39] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, Art. no. e0212320, doi: 10.1371/journal.pone.0212320.

[40] T. A. Assefa, O. A. Esqueda, and A. V. Mollick, "Stock returns and interest rates around the world: A panel data approach," *J. Econ. Bus.*, vol. 89, pp. 20–35, Jan. 2017, doi: 10.1016/j.jeconbus.2016.10.001.

[41] X. Yuan, J. Yuan, T. Jiang, and Q. U. Ain, "Integrated long-term stock selection models based on feature selection and machine learning algorithms for China stock market," *IEEE Access*, vol. 8, pp. 22672–22685, 2020, doi: 10.1109/ACCESS.2020.2969293.

[42] J. M. Chen, "An introduction to machine learning for panel data," *Int. Adv. Econ. Res.*, vol. 27, no. 1, pp. 1–16, Feb. 2021, doi: 10.1007/s11294-021-09815-6.

[43] H. Zhu, L. Bai, L. He, and Z. Liu, "Forecasting realized volatility with machine learning: Panel data perspective," *J. Empirical Finance*, vol. 73, pp. 251–271, Sep. 2023, doi: 10.1016/j.jempfin.2023.07.003.

[44] S. Erfanian, Y. Zhou, A. Razzaq, A. Abbas, A. A. Safeer, and T. Li, "Predicting bitcoin (BTC) price in the context of economic theories: A machine learning approach," *Entropy*, vol. 24, no. 10, p. 1487, Oct. 2022, doi: 10.3390/e24101487.

[45] M. Beniwal, A. Singh, and N. Kumar, "Forecasting long-term stock prices of global indices: A forward-validating genetic algorithm optimization approach for support vector regression," *Appl. Soft Comput.*, vol. 145, Sep. 2023, Art. no. 110566, doi: 10.1016/j.asoc.2023.110566.

[46] Y. Han, J. Kim, and D. Enke, "A machine learning trading system for the stock market based on N-period min-max labeling using XGBoost," *Expert Syst. Appl.*, vol. 211, Jan. 2023, Art. no. 118581, doi: 10.1016/j.eswa.2022.118581.

[47] H. Rezaei, H. Faaljou, and G. Mansourfar, "Stock price prediction using deep learning and frequency decomposition," *Expert Syst. Appl.*, vol. 169, May 2021, Art. no. 114332, doi: 10.1016/j.eswa.2020.114332.

[48] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a feeling: Accuracy and application of sentiment analysis," *Int. J. Res. Marketing*, vol. 40, no. 1, pp. 75–87, Mar. 2023, doi: 10.1016/j.ijresmar.2022.05.005.

[49] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," in *Proc. 28th Int. Conf. Comput. Linguistics*. Stroudsburg, PA, USA: International Committee on Computational Linguistics, 2020, pp. 757–770. [Online]. Available: https://www.aclweb.org/anthology/2020.coling-main.66

[50] X. Li and J. Li, "AoE: Angle-optimized embeddings for semantic textual similarity," in *Proc. 62nd Annu. Meeting Assoc. Comput. Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2024, pp. 1825–1839. [Online]. Available: https://aclanthology.org/2024.acl-long.101

[51] L. Wang, N. Yang, X. Huang, L. Yang, R. Majumder, and F. Wei, "Multilingual E5 text embeddings: A technical report," 2024, *arXiv:2402.05672*.

[52] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-networks," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2019, pp. 3982–3992. [Online]. Available: https://aclanthology.org/D19-1410

[53] S. Xiao, Z. Liu, P. Zhang, N. Muennighoff, D. Lian, and J.-Y. Nie, "C-pack: Packed resources for general Chinese embeddings," in *Proc. 47th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2024, pp. 641–649, doi: 10.1145/3626772.3657878.

[54] N. Kokhlikyan, V. Miglani, M. Martin, E. Wang, B. Alsallakh, J. Reynolds, A. Melnikov, N. Kliushkina, C. Araya, S. Yan, and O. Reblitz-Richardson, "Captum: A unified and generic model interpretability library for PyTorch," 2020, *arXiv:2009.07896*.

[55] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in *Proc. Int. Conf. Mach. Learn.*, vol. 70, 2017, pp. 3319–3328.

[56] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, 2020, pp. 7871–7880. [Online]. Available: https://aclanthology.org/2020.acl-main.703

[57] H. W. Chung et al., "Scaling instruction-finetuned language models," *J. Mach. Learn. Res.*, vol. 25, no. 70, pp. 1–53, 2024.

[58] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.* Stroudsburg, PA, USA: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: http://aclweb.org/anthology/N19-1423

AGUS TRI HARYONO was born in Balikpapan, Indonesia, in 1988. He received the master's degree in informatics engineering from Institut Teknologi Sepuluh Nopember (ITS), Indonesia, in 2023, where he is pursuing the degree in computer science. His research interests include financial time series, text mining, and machine learning.

RIYANARTO SARNO (Senior Member, IEEE) received the Ph.D. degree, in 1992. He is a Professor with the Informatics Department, Institut Teknologi Sepuluh Nopember (ITS). He was the author of more than five books and more than 300 scientific articles, incorporating him in the top 2% world ranking scientist by Standford University, in 2020. He has researched process mining for five years. His research interests include machine learning, the Internet of Things, knowledge engineering, enterprise computing, and information management.

RATIH NUR ESTI ANGGRAINI (Member, IEEE) received the Ph.D. degree in engineering mathematics from the University of Bristol, in 2021. She is currently a Lecturer with the Informatics Department, Institut Teknologi Sepuluh Nopember (ITS), Indonesia. Her research interests include text mining, artificial intelligence, and software engineering.

KELLY ROSSA SUNGKONO (Member, IEEE) was born in Surabaya, in June 1994. She received the Ph.D. degree in computer science, in 2024. She is currently a Lecturer with Institut Teknologi Sepuluh Nopember (ITS). Her research interests include databases, information systems, and machine learning.