Comparative Study on Stock Price Forecasting Using Deep Learning Method Based on Combination Dataset

Yhudha Juwono
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
6025231055@student.its.ac.id

Riyanarto Sarno
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
riyanarto@if.its.ac.id

Ratih Nur Esti Anggraini
Department of informatics
Institut Teknologi Sepuluh November
Surabaya, Indonesia
ratih_nea@if.its.ac.id

Agus Tri Haryono
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
7025231020@student.its.ac.id

Abdullah Faqih Septiyanto
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
7025221022@student.its.ac.id

2024 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS) | 979-8-3503-5052-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/AIMS61812.2024.10513288

Abstract—Stock forecasting is the process of employing various analysis methods and mathematical models, including deep learning techniques, to predict future stock price movements based on historical data and relevant market factors. This paper aims to contribute to the field of stock price prediction by introducing a comprehensive forecasting model. The model integrates OHLCV, technical indicators, macroeconomic variables, and fundamental dataset, leveraging a multifaceted dataset approach. Through the incorporation of these diverse datasets, the proposed model seeks to enhance the accuracy and robustness of stock price forecasts, providing a more holistic understanding of market dynamics for investors and researchers alike. In conclusion, the test results indicate that the application of a combined dataset using feature selection, along with the utilization of the TFGRU model, yielded positive results. The model achieved an RMSE of 16.19, a MAPE of 2.65, and an AcMAPE of 0.85. Lower RMSE and MAPE values suggest enhanced performance, and the relatively low AcMAPE, considering both accuracy and percentage error, further underscores a favorable outcome.

Keywords— stock price forecasting, technical indicator, fundamental, macro-economics, combined dataset

I. INTRODUCTION

Stock price prediction involves employing various analytical methods, mathematical models, and statistical techniques to project the future value or movement of stock prices. This approach incorporates the utilization of historical data, technical indicators, as well as fundamental and macroeconomic factors to comprehend and anticipate changes in the value of a company's stock. The primary goal of stock price prediction is to offer insights to investors, traders, and financial decision-makers, enabling them to make more informed decisions grounded in risk analysis. Commonly used methods in stock price prediction encompass technical analysis, fundamental analysis, and, more recently, artificial intelligence-based approaches such as deep learning.

Predicting stock prices through deep learning and utilizing a combination of datasets is crucial, as this approach allows for the integration of information from various sources, such as historical data, technical indicators, and fundamental factors. Deep learning, as a branch of artificial intelligence, can handle high-dimensional data complexity and extract patterns that are challenging to discern using conventional methods. By employing a combination of datasets, a more holistic understanding of market dynamics can be obtained, encompassing historical aspects, macroeconomics, and other factors influencing stock price movements. This enables the model to make more accurate and reliable predictions, providing a significant advantage for investors and market participants to make better-informed decisions.

In this research, the primary focus lies in predicting stock prices through the utilization of a combination of datasets, with a specific emphasis on implementing preprocessing methods, feature selection and dimension reduction. Furthermore, this study conducts a performance comparison with various deep learning methods, aiming to identify the most effective approach in enhancing the accuracy of stock price predictions.

Within the scope of this research, we compare deep learning methods for predicting stock prices by leveraging a combination of datasets gathered from various sources. This endeavor not only emphasizes the development of predictive models but also involves an in-depth analysis of the relative contributions of each dataset element in enhancing the accuracy of stock price predictions. The second section discusses prior works, while the research methodology is elucidated in the third section, followed by results and discussions in the fourth section. The fifth section concludes our research and provides recommendations for further studies.

II. RELATED WORKS

In the context of stock price prediction, prior research has explored diverse methodologies, spanning traditional analyses like technical and fundamental approaches to contemporary paradigms, including machine learning and deep learning. The integration of various datasets, encompassing historical stock prices, technical indicators, and macroeconomic factors, has been a focal point, aiming for a comprehensive understanding of market dynamics. Additionally, investigations into feature selection, dimension reduction, and the evaluation of different deep learning architectures underscore ongoing efforts to enhance predictive accuracy in financial forecasting. This review navigates through these related works, contributing to the continual refinement of methodologies in the field of stock price prediction.

A. Stock Price Analysis

Stock price analysis is a methodology rooted in the study of historical price data and patterns with the aim of forecasting future stock prices. This practice is a fundamental aspect of technical analysis, a field of study in financial markets that emphasizes the use of historical market data, charts, and various technical indicators to make predictions about future price movements. Unlike fundamental analysis, which delves into a company's financial health and intrinsic value, price analysis focuses primarily on the stock's open, high, low, close, and volume [1].

Price analysis is particularly favored by short-term traders and technical analysts who seek to exploit short-term price fluctuations. It is important to note that while price analysis provides valuable insights into market behavior, it comes with risks, as historical patterns may not always accurately predict future outcomes.

B. Technical Indicator Analysis

In the realm of price stock analysis, technical indicators play a pivotal role, serving as a methodology that merges statistical computations and chart analysis. This approach is employed to uncover intricate patterns and trends within stock prices and trading volumes. At its essence, technical indicators involve the analysis and computation of historical data, aiming to discern specific patterns in both price and volume. The goal is to predict pattern and stock price movement [2].

C. Macroeconomic Analysis

Macroeconomics is a branch of economics that examines the overall functioning and behavior of the entire economy. Macroeconomic analysis in stock price prediction involves examining and assessing the broader economic factors that can impact the overall performance of financial markets and, consequently, influence the prices of individual stocks. Although macroeconomic data often has longer time periods, such as monthly or quarterly, it still exhibits strong significance and can be utilized to enhance stock price predictions [3].

D. Fundamental Data Analysis

Fundamental analysis is commonly employed to forecast stock prices over the mid-term and long-term periods[4]. Fundamental analysis is a method used in stock price prediction that involves evaluating a company's intrinsic value by examining various financial and economic factors. Unlike technical analysis, which focuses on historical price movements and trading patterns, fundamental analysis looks at the underlying health and performance of a company to determine its potential for future growth and profitability.

E. Feature Selection

The purpose of feature selection in deep learning is to enhance the efficiency, effectiveness, and interpretability of the models [5]. Selecting features that are significant can enhance the accuracy of predictions [6]. These techniques are employed to mitigate the challenges associated with high-dimensional data and to improve the overall performance of deep learning models.

Various approaches are undertaken in relation to feature selection or dimensional reduction. Shen et al. in [7] proposed a comprehensive solution during the preparation phase, encompassing essential steps such as feature expansion, feature selection, and the ultimate application of dimension reduction through PCA. Implementing adaptive feature selection on technical indicators enhances the model [8]. Ultimately, the implementation of feature selection demonstrates an improvement in accuracy in stock price prediction.

F. Application of Deep Learning in Stock Price Forecasting

The implementation of deep learning techniques in stock price forecasting involves the utilization of sophisticated neural network architectures to model and predict the complex patterns inherent in financial time series data. What makes deep learning particularly relevant in this context is its ability to automatically learn features from raw input data, capturing intricate dependencies that traditional models might struggle to discern. Deep learning models, such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs), are well-suited for handling sequential data like stock prices [9], as they can effectively capture temporal dependencies and patterns that evolve over time.

The significance of implementing deep learning in stock price forecasting lies in its potential to enhance predictive accuracy. Deep learning models excel at capturing nonlinear relationships and intricate temporal dependencies, allowing them to uncover hidden patterns and exploit information embedded in historical price movements. This can lead to more accurate predictions, aiding investors, traders, and financial analysts in making informed decisions and mitigating risks in dynamic financial markets.

However, the adoption of deep learning in stock price forecasting comes with both advantages and challenges. On the positive side, deep learning models can handle vast amounts of data and learn complex patterns, providing a more nuanced understanding of market dynamics. They are also capable of adapting to changing market conditions, offering flexibility in capturing evolving trends. Nevertheless, challenges include the need for substantial computational resources, potential overfitting concerns, and the inherent difficulty of interpreting the internal workings of complex neural networks. Despite these challenges, the potential benefits of leveraging deep learning in stock price forecasting make it a compelling area of research and application in the finance domain.

III. Methodology

The primary aim of this research is to conduct a comprehensive comparative analysis focused on forecasting stock prices through the integration of multiple datasets. What sets this research apart is the intentional integration of multiple datasets, each contributing unique dimensions to the predictive model. This comprehensive dataset combination is expected to enrich the learning process of the GRU model, providing a more holistic understanding of the intricate patterns within stock price movements. To achieve this, the methodology employed follows the structure and guidance outlined in Fig 1. This Fig serves as a visual representation of the step-by-step approach taken in the research process, illustrating how the combination of datasets and the implementation the GRU models are systematically carried out and assessed.

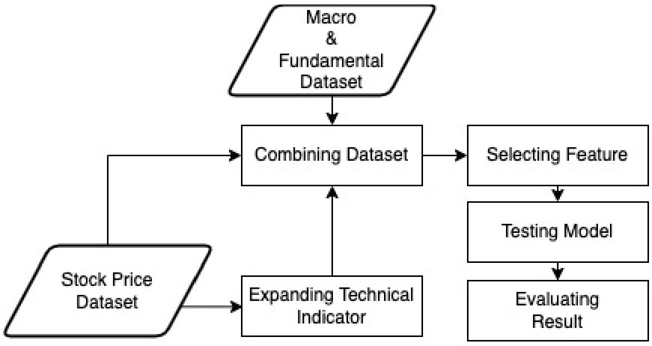

Fig 1. Experiment diagram

By leveraging a comparative analysis framework, the study aims to contribute valuable insights into the effectiveness of diverse data combinations for stock price forecasting, facilitating a more informed understanding of the factors influencing prediction accuracy in financial markets.

A. Dataset

This research was conducted using four types of datasets collected from various sources. OLHCV (Open, Low, High, Close, Volume), technical indicator, Macro-economics, and Fundamental data were chosen due to their significant impact on predicting future stock prices. We collected data for 8 Indonesian stocks, covering a timeframe of 16 to 23 years, corresponding to the Initial Public Offering (IPO) period. The data includes daily frequency records for OLHCV, Technical, and macro-economic data, along with quarterly frequency data for fundamental information.

1) Dataset of Stock Price

This research utilizes stock price datasets from eight prominent issuers in Indonesia, sourced from Yahoo Finance. The collected dataset is presented daily, segmented into columns for open, high, low, close, volume, dividend, and stock split.

TABLE I. STOCK DATA STATISTIC

Issuer	First Date	Last Date	Count	Min	Max	Mean
ASII	17/10/2000	11/12/2023	5,786	42.83	6997.28	3242.71
BBCA	08/06/2004	11/12/2023	4,837	105.65	9,355.36	2,826.44
BBRI	10/112003	11/12/2023	4,987	53.34	5,700	1,724.18
BMRI	14/07/2003	11/12/2023	5,073	89.19	10,225	1,752.58
HITS	17/10/2005	11/12/2023	4,124	177	2,891.12	539.39
INDX	28/05/2001	11/12/2023	5,627	35	18,000	801.27
SMMT	04/12/2007	11/12/2023	3,795	10.18	5,169.98	433.22
TLKM	28/09/2004	11/12/2023	4,756	465.68	4,558.16	2,123.02

The statistical data on prices, as presented in Table 1, illustrates the variance across 8 issuers, spanning from the IPO date to December 11, 2023.

2) Dataset of Technical Indicator

The technical indicator dataset was derived through an expanding transformation of stock price data using mathematical computational methods[10]. The implementation of technical indicators demonstrates more accurate results in financial market [11][12]. This approach involved employing mathematical computations to extract technical indicators from the original stock price dataset, capturing and enhancing underlying patterns and trends. The utilization of this expansion technique aligns with mathematical modeling principles, contributing to a rigorous and systematic analysis for stock price prediction.

Several technical models employed in this study include the Simple Moving Average (SMA), Exponential Moving Average (EMA), Average Directional Movement (ADX), Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), and Bollinger Bands. These models were selected for their effectiveness in analyzing stock price trends and providing valuable insights into market dynamics.

3) Dataset of Fundamental

The fundamental dataset explains essential factors related to a company's health. Various indicators within a company are crucial for assessing its stock condition. In this study we utilized ten key ratio indicators, including metrics such as Price to Book Value, Book Value Per Share, Total Equity, Total Debt, Net Debt, and Working Capital. These indicators were selected for their significance in providing comprehensive insights into the financial health and performance of the companies under examination.

4) Dataset of Macroeconomics

The integration of macroeconomic datasets has a significant impact in enhancing the precision of stock forecasting [3]. These datasets encompass a spectrum of economic indicators such as inflation rates, interest rates, GDP growth, and employment Figs, offering a broader contextual understanding of the market forces at play. By incorporating macro-economic variables into forecasting models, analysts can capture the influence of broader economic trends on stock prices, allowing for more informed predictions. For instance, changes in interest rates may impact borrowing costs for companies, influencing their profitability and, consequently, stock prices. Similarly, shifts in GDP growth can indicate the overall health of the economy, affecting various sectors and, in turn, stock market performance. The inclusion of macro-economic datasets provides a holistic perspective, enabling forecasting models to navigate and adapt to the dynamic interplay between economic indicators and stock market movements.

B. Data Preprocessing

1) Data Merging

Data merging is performed to combine information from multiple sources that can enrich the dataset. In this context, stock price data is integrated with technical, macroeconomic, and fundamental company data. This merging process aims to enhance the depth and diversity of information available for the model.

2) Data Scaling

Data scaling is applied to ensure that all variables have a similar range of values, avoiding the dominance of variables with large scales. In the context of stock price prediction, we use the Min-Max scaling method to convert variable values into the range [0, 1].

C. Feature Selection

Feature selection is a crucial process in machine learning that involves choosing a subset of relevant variables from a dataset to enhance model performance. The primary goal is to reduce the dimensionality of the data by retaining the most impactful information, thereby improving the quality and efficiency of the model. One of the key benefits of feature selection is its ability to mitigate overfitting. By eliminating irrelevant or redundant features, the risk of the model memorizing the training data and failing to generalize well to new data is minimized. This not only ensures more robust model performance but also promotes better adaptability to diverse datasets. The following are the options used:

1) LASSO Regression

Also known as Least Absolute Shrinkage and Selection Operator, is a regularization method in linear regression designed to address overfitting issues and perform automatic feature selection. The primary goal of Lasso Regression is to drive some regression coefficients to zero, effectively retaining only a small subset of features with significant impact.

The application of Lasso Regression for feature selection is particularly valuable when dealing with many features, not all of which are essential for building a robust model [13]. By applying Lasso Regression, we can identify the most relevant subset of features and disregard less significant ones [14]. This helps reduce the dimensionality of the dataset, enhances model interpretability, and often results in a model more resilient to overfitting.

D. Models

1) Gated Recurrent Unit (GRU)

GRU is a type of recurrent neural network (RNN) architecture designed to address some of the challenges faced by traditional RNNs, such as difficulties in capturing long-term dependencies in sequential data [15]. The GRU shows better compared to other RNN such as LSTM models in predicting stock prices [16], this also demonstrates better in the Indonesian stock market [17]. Unlike LSTM, GRU has a simplified architecture with two gates: the reset gate and the update gate. The reset gate determines which information from the previous time step should be discarded, while the update gate regulates the amount of new information to be stored in the current state. GRU's design allows it to efficiently capture and propagate relevant information across sequential data, making it computationally less expensive and often yielding comparable performance to more complex models.

E. Evaluation

1) Root Mean Square Error (RMSE)

RMSE is a widely employed metric for computing regression error, is the square root of the Mean Squared Error (MSE). It serves as an indicator of how well the data aligns with the regression line. A lower RMSE value is deemed optimal, signifying a more precise fit of the data around the regression line. The equation of RMSE is given in equation (2) as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{e}_{\mathrm{i}}^{2}} \quad (2)$$

2) Mean Absolute Percentage Error (MAPE)

MAPE is a metric that gauges the average percentage difference between predicted and actual values in a forecasting model. It provides a concise measure of prediction accuracy, with lower values indicating better model performance. The accuracy of MAPE is given equation (3) as:

$$\mathrm{MAPE} = [(1 - \frac{1}{\mathrm{N}}\sum\nolimits_{1=\mathrm{i}}^{\mathrm{n}} \left|\frac{\mathrm{e}_{\mathrm{i}}}{\mathrm{y}_{\mathrm{i}}}\right| * 100 * 100)] \quad (3)$$

3) Accuracy Mean Precentage Error (AcMAPE)

Accuracy is a critical metric used to evaluate the precision of predictive models. It measures the percentage of accurate predictions the model. The AcMAPE is a composite metric that integrates accuracy and normalized MAPE. The Accuracy and AcMAPE is given in equation (3) and (4) as:

$$\mathrm{Accuracy} = \frac{\mathrm{True\ Positive} + \mathrm{True\ Negative}}{\mathrm{Total\ Instance}} \quad (4)$$

$$\mathrm{AcMAPE} = \big((1 - \mathrm{accuracy}) \times 2\big) + (\frac{\mathrm{MAPE}}{100}) \quad (5)$$

The formula combines $(1 - \mathrm{accuracy}) \times 2$ to highlight the impact of accuracy deviations and $\mathrm{MAPE}/100$ for a normalized MAPE. This balanced approach provides a detailed evaluation, highlighting both accuracy and percentage error in predictions, making AcMAPE a comprehensive measure of forecasting reliability.

IV. Results And Discussion

Stock price forecasting using a combination of datasets is evaluated with three models that aggregate the average values of RMSE, MAPE, and AcMAPE with lookback parameters of 5, 10, 20, and 25. Subsequently, within each dataset combination, measurements are conducted and categorized into two types of new datasets: original dataset (O) and Feature-selected dataset (F).

A. Stock forecasting using stock price dataset

In the OLHCV experiment using the dataset containing open, low, high, close, volume, dividend, and stock split, as shown in Table 2, BBCA has the lowest MAPE value with an RMSE of 169.47. On the other hand, SMMT has the highest AcMAPE value with a MAPE of 20.51. Applying feature selection results in a decrease in almost all MAPE values, except for BBCA, HITS, SMMT, and TLKM, which experience an increase of MAPE score.

TABLE II. STOCK FORECASTING USING OLHCV DATA WITH FEATURE SELECTION.

Issuer	Features	Metrics RMSE	MAPE%	AcMAPE
ASII	O (6)	153.24	2.39	0.99
	F (4)	164.47	2.63	1.08
BBCA	O (6)	169.31	**1.83**	0.93
	F (4)	171.11	1.85	0.93
BBRI	O (6)	113.15	2.16	0.97
	F (4)	111.77	2.14	0.99
BMRI	O (6)	221.78	20.53	0.96
	F (3)	260.03	3.16	1.01
HITS	O (6)	34.31	4.55	1.02
	F (3)	37.60	5.00	0.96
INDX	O (6)	137.76	55.82	0.94
	F (3)	**33.20**	32.44	0.94
SMMT	O (6)	49.61	9.50	1.08
	F (4)	52.68	20.51	**1.21**
TLKM	O (6)	94.78	2.06	**0.91**
	F (4)	98.42	2.20	0.92

B. Stock forecasting using combination stock price and technical indicator dataset

The combination of stock price and technical indicator datasets resulted in a total of 21 original features, and after performing feature selection, the dataset changed to a minimum of 11 features for BMRI, INDX, and TLKM, and a maximum of 14 features for ASII and BBRI. In terms of stock price prediction results, as outlined in Table 3, TLKM has the lowest MAPE value with both the original dataset and feature selection, with RMSE values of 53.02 and 46.43, and AcMAPE values of 1.00 and 0.99, respectively.

TABLE III. STOCK FORECASTING USING OLHCV AND TECHNICAL INDICATOR DATASET WITH FEATURE SELECTION.

Issuer	Features	Metrics RMSE	MAPE %	AcMAPE
ASII	O (21)	205.68	3.15	0.89
	F (14)	182.58	2.94	1.05
BBCA	O (21)	250.33	2.61	0.96
	F (12)	242.93	2.63	0.96
BBRI	O (21)	137.02	2.46	0.98
	F (14)	134.32	2.53	1.01
BMRI	O (21)	269.92	3.29	1.00
	F (11)	237.91	2.94	1.01
HITS	O (21)	35.44	4.23	0.92
	F (12)	35.37	4.76	0.93
INDX	O (21)	213.47	177.68	2.49
	F (11)	168.20	170.02	2.34
SMMT	O (21)	53.02	9.51	1.13
	F (12)	46.43	18.53	1.26
TLKM	O (21)	103.44	**2.33**	1.00
	F (11)	97.15	**2.22**	0.99

C. Stock forecasting using combination of stock price and macro-economics dataset

The combination of stock price and macroeconomics datasets resulted in 90 original features. After performing feature selection, the dataset was reduced to a minimum of 26 features for HITS and 43 for ASII. The results from Table 4 explain that ASII has the lowest MAPE value for both the original dataset and the feature-selected data, with RMSE values of 195.94 and 277.67, and AcMAPE values of 1.05 and 1.03, respectively.

TABLE IV. STOCK FORECASTING USING OLHCV AND MACRO-ECONOMIC DATASET WITH FEATURE SELECTION.

Issuer	Features	Metrics RMSE	MAPE %	AcMAPE
ASII	O (90)	295.94	**4.82**	1.05
	F (43)	277.67	**4.54**	1.03
BBCA	O (90)	1,312.75	13.42	1.05
	F (40)	631.47	6.41	0.97
BBRI	O (90)	765.42	14.14	1.08
	F (40)	314.65	5.87	0.99
BMRI	O (90)	814.41	12.46	1.07
	F (38)	438.23	6.37	1.00
HITS	O (90)	80.28	8.70	1.09
	F (26)	59.08	6.33	0.94
INDX	O (90)	228.76	162.15	**2.77**
	F (33)	160.69	131.78	**2.54**
SMMT	O (90)	91.06	17.16	2.77
	F (29)	57.21	12.35	1.16
TLKM	O (90)	369.31	7.69	1.00
	F (40)	220.75	4.82	0.97

Meanwhile, the highest MAPE value is for the INDX, with MAPE score 162.15% for the original dataset and 131.78% for the feature-selected dataset.

D. Stock forecasting using combination of stock price and fundamental dataset

The combination of datasets involving stock prices and fundamental data results in 16 original features. After conducting feature selection, the dataset is reduced to 4 features for ASII, BBCA, and TLKM. Notably, for TLKM, no fundamental-related features are found after feature selection. This occurrence may stem from the significance of fundamental data in influencing TLKM stocks. As illustrated in Table 5, it is evident that, overall, the values of RMSE, MAPE, and AcMAPE are smaller compared to the combination of stock prices and technical indicator datasets. This consistent trend can be attributed to the smaller number of features. Additionally, the application of feature selection consistently improves evaluation metrics. For the original features, BBRI attains a MAPE value of 1.88, an RMSE of 103.39, and an AcMAPE of 0.96. Meanwhile, for the dataset with feature selection, BBCA, with 4 features, achieves a MAPE of 1.59, an RMSE of 155.22, and an AcMAPE of 0.88.

TABLE V. STOCK FORECASTING USING OLHCV AND FUNDAMENTAL DATASET WITH FEATURE SELECTION.

Issuer	Features	Metrics RMSE	MAPE %	AcMAPE
ASII	O (16)	161.45	2.53	0.98
	F (4)	138.69	2.07	0.91
BBCA	O (16)	277.92	3.01	0.95
	F (4)	155.22	**1.59**	0.88
BBRI	O (16)	103.39	**1.88**	0.96
	F (6)	93.80	1.68	0.90
BMRI	O (16)	233.45	2.16	0.97
	F (5)	233.17	2.06	0.99
HITS	O (16)	41.03	5.55	**1.15**
	F (7)	36.78	5.12	**1.16**
INDX	O (16)	**20.93**	5.25	0.87
	F (9)	**15.54**	4.31	0.92
SMMT	O (16)	68.84	13.61	1.08
	F (8)	61.29	11.45	1.19
TLKM	O (16)	99.46	2.19	0.93
	F (4)	89.48	1.97	0.91

E. Stock forecasting using combined dataset

The combination of all datasets, including stock prices, technical indicators, fundamental, and macroeconomic datasets, results in a total of 115 features. Subsequently, a feature selection process is conducted, yielding a new count of features for each issuer. Among them, INDX possesses the fewest features at 42, while BBCA has the highest with 56 features.

As shown in Table 6, it is evident that, overall, all metrics exhibit larger values for RMSE, MAPE, and AcMAPE when compared to several previous dataset combinations. From all experiments, it is discerned that these values correlate with the abundance of features within the dataset. For instance, INDX, with a feature-selected dataset comprising 24 features, has an RMSE value of 16.19. In contrast, BBCA, with the Original Combined dataset, records an RMSE of 1,036.73.

The best MAPE evaluation is achieved by ASII, with scores of 3.09 and 2.65 for the Original and feature-selected datasets, respectively. This trend is mirrored in the relatively small AcMAPE values, standing at 0.98 and 0.92, respectively. Notably, HITS attains the best AcMAPE for the feature-selected dataset, registering a value of 0.85.

TABLE VI. STOCK FORECASTING USING COMBINED DATASET WITH FEATURE SELECTION

Issuer	Features	Metrics		
		RMSE	MAPE %	AcMAPE
ASII	O (115)	209.41	**3.09**	0.98
	F (50)	199.03	**2.65**	0.92
BBCA	O (115)	1,036.73	10.28	1.00
	F (56)	734.31	7.46	0.96
BBRI	O (115)	469.18	8.35	1.01
	F (48)	308.15	5.39	0.98
BMRI	O (115)	652.92	9.35	1.05
	F (53)	538.49	7.62	1.03
HITS	O (115)	72.44	8.18	**1.06**
	F (42)	62.19	6.03	**0.85**
INDX	O (115)	**19.08**	7.51	1.09
	F (24)	**16.19**	5.38	1.07
SMMT	O (115)	115.18	26.47	1.30
	F (45)	79.42	14.86	1.13
TLKM	O (115)	325.12	7.03	0.98
	F (52)	244.32	5.33	0.93

V. CONCLUSION

Based on the comparative study of stock price forecasting using deep learning methods and a combination dataset, it is evident that the integration of stock price, technical indicators, macroeconomic, and fundamental dataset has the potential to enhance understanding regarding the utilization of dataset in stock price forecasting. The application of feature selection, along with the utilization of the TFGRU model, yielded positive results. The study also highlights the impact of feature selection on stock price prediction for various issuers and dataset combinations, demonstrating its potential to improve forecasting reliability. Additionally, the findings emphasize the consistent improvement in evaluation metrics with smaller feature sets, underscoring the importance of feature selection. Overall, this study provides valuable insights into the effectiveness of diverse data combinations and feature selection for stock price forecasting, contributing to the current knowledge base on predicting stock prices through the application of deep learning techniques.

REFERENCES

[1] Y. Ye, “Machine Learning in Stock Price Analysis,” Applied and Computational Engineering, vol. 8, no. 1, 2023, doi: 10.54254/2755-2721/8/20230252.

[2] M. Agrawal, P. K. Shukla, R. Nair, A. Nayyar, and M. Masud, “Stock prediction based on technical indicators using deep learning model,” Computers, Materials and Continua, vol. 70, no. 1, 2021, doi: 10.32604/cmc.2022.014637.

[3] S. J. Binoy and J. Jos, “Financial Market Forecasting using Macro-Economic Variables and RNN,” in 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2022, 2022. doi: 10.1109/ICACITE53722.2022.9823905.

[4] Y. Huang, L. F. Capretz, and D. Ho, “Machine Learning for Stock Prediction Based on Fundamental Analysis,” in 2021 IEEE Symposium Series on Computational Intelligence, SSCI 2021 - Proceedings, 2021. doi: 10.1109/SSCI50451.2021.9660134.

[5] B. Kumari, S. Patnaik, and T. Swarnkar, “Feature selection for stock price prediction: a critical review,” International Journal of Intelligent Enterprise, vol. 10, no. 1. 2022. doi: 10.1504/ijie.2023.127236.

[6] W. L. Yan, “Stock index futures price prediction using feature selection and deep learning,” North American Journal of Economics and Finance, vol. 64, 2023, doi: 10.1016/j.najef.2022.101867.

[7] J. Shen and M. O. Shafiq, “Short-term stock market price trend prediction using a comprehensive deep learning system,” J Big Data, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00333-6.

[8] G. Ji, J. Yu, K. Hu, J. Xie, and X. Ji, “An adaptive feature selection schema using improved technical indicators for predicting stock price movements,” Expert Syst Appl, vol. 200, 2022, doi: 10.1016/j.eswa.2022.116941.

[9] A. T. Haryono, R. Sarno, and K. R. Sungkono, “Stock price forecasting in Indonesia stock exchange using deep learning: a comparative study,” International Journal of Electrical and Computer Engineering (IJECE), vol. 14, no. 1, p. 861, Feb. 2024, doi: 10.11591/ijece.v14i1.pp861-869.

[10] 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.

[11] N. Parida, B. K. Balabantaray, R. Nayak, and J. K. Rout, “A Deep Learning based Approach to Stock Market Price Prediction using Technical indicators,” in 2023 International Conference on Advances in Intelligent Computing and Applications, AICAPS 2023, 2023. doi: 10.1109/AICAPS57044.2023.10074445.

[12] S. Y. Lin and H. Y. Lin, “Bond Price Prediction Using Technical Indicators and Machine Learning Techniques,” Journal of Information Science and Engineering, vol. 39, no. 2, 2023, doi: 10.6688/JISE.202303_39(2).0012.

[13] J. Yang, Y. Wang, and X. Li, “Prediction of stock price direction using the LASSO-LSTM model combines technical indicators and financial sentiment analysis,” PeerJ Comput Sci, vol. 8, 2022, doi: 10.7717/PEERJ-CS.1148.

[14] H. T. V. Ha, D. N. Hung, and N. T. Xuan, “Relationships between Accounting Information on the Business Financial Statements and the Stock Price: A Study with LASSO Method,” Quality - Access to Success, vol. 23, no. 186, 2022, doi: 10.47750/QAS/23.186.31.

[15] R. Jozefowicz, W. Zaremba, and I. Sutskever, “An empirical exploration of Recurrent Network architectures,” in 32nd International Conference on Machine Learning, ICML 2015, 2015.

[16] 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 Humaniz Comput, vol. 12, no. 1, 2021, doi: 10.1007/s12652-020-02057-0.

[17] 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, 2023, doi: 10.1109/ACCESS.2023.3298445.