International Journal of Information Management Data Insights 4 (2024) 100293

Contents lists available at ScienceDirect

International Journal of Information Management Data Insights

journal homepage: www.elsevier.com/locate/jjimei

How can we predict transportation stock prices using artificial intelligence? Findings from experiments with Long Short-Term Memory based algorithms

Dinar Ajeng Kristiyanti *, Willibrordus Bayu Nova Pramudya, Samuel Ady Sanjaya

Departement Information System, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Jl. Scientia Boulevard, Tangerang 15111, Indonesia

ARTICLE INFO

Keywords:
Activation
Long short-term memory
Optimizer
Prediction
Transportation stock

ABSTRACT

Inflation growth in Indonesia and other countries impacts the currency value and investors' purchasing power, particularly in the transportation sector. This research explores the impact of inflation growth in Indonesia and comparable nations on currency valuation and the purchasing power of investors, with a focus on the transportation sector. Data collection was carried out from April to October 2023 by scraping stock data from several transportation stocks such as: AKSI.JK, CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK, and WEHA. The research primarily aims to forecast stock prices in Indonesia's transportation sector, utilizing data mining techniques within the Cross Industry Standard Process for Data Mining (CRISP-DM) framework, which includes stages such as business understanding, data preparation, modeling, evaluation, and deployment. It employs the Long Short-Term Memory (LSTM) algorithm, assessing different hyperparameter activation functions (linear, ReLU, sigmoid, tanh) and optimizers (ADAM, ADAGRAD, NADAM, RMSPROP, ADADELTA, SGD, ADAMAX) to refine prediction accuracy. Findings demonstrate the ReLU activation function and ADAM optimizer's effectiveness, highlighted by evaluation metrics such as Mean Absolute Error (MAE) of 0.0092918, Mean Absolute Percentage Error (MAPE) of 0.06422, and R-Squared of 96 %. The study notably identifies significant growth in Temas (TMAS.JK) stock from April to October 2023, surpassing other sector stocks. Additionally, a web-based application for predicting transportation stock prices has been developed, facilitating user inputs like ticker, activation-optimizer choice, and date range.

1. Introduction

Growth of global capital markets has become one of the alternatives for investors to make short-term and long-term investments (Ritika et al., 2022). One of the forms or tangible evidence from investment in the capital market itself is stock (Rabinovich, 2023). Stock is a form of securities that represents an individual, institution, or organization's ownership in a company (Chen et al., 2021). When investors invest in stocks, they pay attention to the stock market to find out whether the stock price in a particular market is decreasing or increasing (Ballinari et al., 2022). Stock price is a value that represents the price of a stock during a certain period determined by market participants and the corresponding supply and demand in the stock market (Albahli et al., 2022). Decreasing or increasing stock prices are influenced by various factors, such as changes in interest rates, inflation, and others (Thampanya et al., 2020).

One of the right solutions for investors to make the right decision in understanding the stock market, especially in the transportation sector, is the role of data mining as a reference in predicting future stock prices, as has been shown in previous studies that successfully predict stocks using data mining methods (Kaur & Dharni, 2022). Analyzing transportation stocks is crucial as this sector is highly sensitive to economic changes, technological advancements, and regulatory policies. It serves as a barometer for overall economic health, influencing and being influenced by global trade flows and consumer behavior. Data mining methods applied to transportation stocks can uncover trends and patterns, offering investors insights for making informed decisions. This is especially valuable in a sector where stock performance can signal shifts in economic activity and consumer confidence. Data mining techniques have been used to predict the stock index of a Spanish company that engaged in infrastruc ture management sector is Acciona, Linear Regression (LR), Support Vector Regression (SVR), and Long Short-Term

* Corresponding author.
E-mail address: dinar.kristiyanti@umn.ac.id (D.A. Kristiyanti).

https://doi.org/10.1016/j.jjimei.2024.100293

Available online 28 September 2024

Memory (LSTM) algorithms. The lowest Root Mean Square Error (RMSE) metric evaluation was obtained with activation parameters linear and tanh, as well as adamax and adam optimisers of 0.0151 (Rana et al., 2019). National Stock Exchange (NSE) of India, LSTM and Multi-Layered Perceptron (MLP) algorithm model method was used with the lowest Mean Squared Error (MSE) metric evaluation of 0.58 % at 60 timestamps, using the activation relu parameter and the rmsprop's optimizer (Banyal* et al., 2020). Stock Exchange from Bangladesh, Hongkong, India, Indonesia, Japan, Malaysia, and Thailand used Long Short-Term Memory (LSTM) with several parameters of activation functions such as ELU, ReLU, and tanh along with number of epochs by 50–100. This study has found an accuracy from activation of tanh by 80 % compared to another activation function of ELU and ReLU to predict stock price prediction from several stock exchanges (Sami et al., 2023). Financial industries at Pakistan, Long Short-Term Memory (LSTM) algorithm obtained best accuracy score of 92 % from selu's activation and the adamax's optimizer with combining several parameters including 10 neurons, learning rate of 0.002, and 3000 epochs (Ali et al., 2019). Shanghai Composite Stock Index (CSI 300), Light Gradient Boosting Machine-Long Short-Term Memory (LightGBM-LSTM), Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU) algorithms along with activation and optimizer settings obtained lowest evaluation metric values for Root Mean Squared Error (RMSE) of 0.01108 and Mean Absolute Error (MAE) of 0.00770 from softsign's activation and rmsprop's optimizer in terms of training error. Besides that, testing error resulted in values of 0.01076 for RMSE and 0.00651 for MAE (Lv et al., 2021). Technology stock prediction is Apple Inc used Long Bilinear-Long Short-Term Memory (LBL-LSTM) to predict Apple Inc. This paper proposed the model has efficient performance compared to previous study. The comparative analysis reveals that LBL-LSTM performs better, achieving a 94 % accuracy rate, even when taking into account changing external market conditions, such as fluctuations in quarterly lending rates of banks. The previous research focuses on critical developments in construction management, predictive analytics, and engineering project management (Risan et al., 2024). The obstacles in construction phases and critical success factors in regional development programs provide foundational knowledge in understanding and improving project management practices. Besides the project management, the proposed previous research also has an approach on predictive analytics for river water suspended sediment using neural networks advances environmental management strategies (Q. Khan et al., 2023). These studies collectively underscore the importance of integrating advanced computational models and structural equation modeling in tackling complex engineering challenges, highlighting a multidisciplinary approach to enhance efficiency and sustainability in the sector. Further exploration into forecasting and determining cost performance indices for tunnel projects through artificial neural networks, Predict schedule performance indicators in highway projects underscores the pivotal role of machine learning in enhancing predictive accuracy and decision-making in complex engineering projects (AL-Zwainy & Aidan, 2017). This body of work collectively demonstrates a trend towards the adoption of AI and machine learning techniques to address traditional and emerging challenges within civil engineering and project management, pointing towards a future where data-driven approaches are standard in achieving project efficiency and sustainability.

Additionally, the reduced Root Mean Squared Relative Error (RMSRE) is only 0.00418, indicating a very low value (Gurav & Kotrappa, 2020). NIFTY 50 index stocks from financial, technology, energy, and consumer sectors used Reccurent Neural Network (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network along with activation functions such as ReLU, sigmoid, and softmax to predict several stocks from NIFTY 50 index. This study has performance of Mean Absolute Percentage Error (MAPE) from acitvation of ReLU by 0.4705 with Long Short-Term Memory (LSTM) algorithm compared to another activation functions such as sigmoid and softmax to predict stocks (Fathali et al., 2022). S&P500 stock index from United States of America, Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Convolutional Neural Network (CNN) algorithms obtained lowest evaluation metric value of Mean Absolute Error (MAE) was 0.0178 for validation data and 0.0260 for test data, using the rmsprop's optimizer (Kamalov et al., 2020). Banking stock predictions from Bank Rakyat Indonesia (BRI), Bank Nasional Indonesia (BNI), Bank Tabungan Negara (BTN), and Mandiri using Long Short-Term Memory (LSTM) algorithm with several parameter from optimizers such as SGD, adam, and RMSProp along with number of epochs by 25, 50, 75, and 100. This study has lower performance of Root Mean Squared Error (RMSE) by 48.32 from optimizer of adam with accuracy score by 95 % to predict Bank Rakyat Indonesia (BRI) stock price (Mutmainah, Marfuah, Nopianti & Tri Panudju, 2022).

Based on previous research, stock prediction using the LSTM algorithm model, as well as the comparison of activation and optimizer, has been carried out. However, the results of evaluation and activation metrics and optimisers are not fully utilised. Based on several studies that have been conducted, the Long Short-Term Memory (LSTM) algorithm modelling only uses several activation hyperparameters such as linear, relu, sigmoid, or tanh as well as the optimiser adam, adagrad, nadam, rmsprop, adadelta, sgd, or adamax optimisers only. In addition, the modelling only uses several metric evaluations such as Mean Absolute Error (MAE), Mean Squared Error (MSE), R2, Root Mean Squared Error (RMSE), or Mean Absolute Percentage Error (MAPE) and the use of streamlit also without based on modelling the LSTM algorithm with hyperparameter activation and optimizers. Therefore, this study aims using 6 transportation datasets in Indonesia. Data mining method or framework used is CRISP-DM for implementing conventional of Long Short-Term Memory (LSTM) algorithm modelling with the LSTM model from the comparison of hyperparameter comparisons of all activations namely linear, ReLU, sigmoid, and tanh, as well as all optimizers namely adam, adagrad, nadam, RMSProp, adadelta, SGD, and adamax. The evaluation elaboration uses metrics chosen are MAPE, MAE, RMSE, MSE, R2, and elapsed time in measuring the training model as well as the statistical test of the shapiro-wilk test. The contributions of this paper are as follows:

- To predict transportation stock prices using the Long Short-Term Memory (LSTM) algorithm model with various activation comparisons such as linear, ReLU (Rectified Linear Unit), tanh, and sigmoid, and optimizers such as adam, adagrad (Adaptive Gradient), nadam, RMSProp (Root Mean Squared Propagation), adadelta, SGD (Stochastic Gradient Descent), and adamax.
- To evaluate the success of the various activation comparisons and optimizers used to measure performance MAPE, MAE, RMSE, MSE, R-squared, elapsed time, and statistical tests.
- To design and develop a web application to predict transportation stocks price in Indonesia.

The structure of this paper is organised as follows. Section 2 presents of Materials and Method, starting with CRISP-DM framework implementation until LSTM model architecture using activation function and optimizer parameters. The CRISP-DM is used for process related to big data, such as IoT, cloud computing, Hadoop, and data centers. The study covers the entire value chain of big data, including data acquisition, generation, analysis, and storage, discussing technical aspects and recent advances. It also examines applications of big data in fields like IoT, enterprise management, medical applications, online social networks, smart grid, and collective intelligence (Rasheed et al., 2021). Section 3, the results and discussion of model performance from LSTM conventional model and LSTM using activation function and optimizer parameters along with the prediction of 6 transportation stocks with web-based information system from development of LSTM model based on activation function and optimizer parameters. Section 4, the paper concludes with model comparison limitations and some perspectives for future work

2. Literature review

National inflation occurred in October 2022. Year-on year inflation in Indonesia is approaching an increase of 5.71 % (Limanseto, 2022). This causes the transportation sector to be more impacted than the food, beverage, and tobaco, clothing and footwear, housing, household equipment, health, education, and other sectors at 16.03 % of this inflation (BPS Statistics, 2024). The growth of inflation in several international countries and Indonesia can cause the purchasing power of shares for investors to decrease (Purwono et al., 2020). Some experts and academics contend that the stock market is still unpredictable, even with the abundance of techniques for anticipating it. One of the right solutions for making decisions for investors is to understand the stock market, especially transportation, the role of artificial intelligence can be a reference in predicting transportation stock prices that will occur in the future (Fathali et al., 2022).

In a number of industries, including banking and transportation, artificial intelligence (AI) has become a potent instrument for stock price prediction. Because of the stock market's volatility and the impact of economic factors like inflation, academics and industry professionals are investigating AI methods in an effort to improve forecast accuracy. AI uses sophisticated algorithms to examine large datasets and spot patterns and trends that human analysis could miss (Chopra & Sharma, 2021). A common use of AI is the prediction of stock prices by machine learning. To forecast future patterns, it makes use of algorithms that are trained on historical data. Among the popular machine learning methods used in this field are random forests, support vector machines (SVM), and linear regression (Rana et al., 2019). For example, linear regression provides a simple way to make predictions by modeling the relationship between a dependent variable (stock price) and one or more independent variables (past stock data) (Qiu & Song, 2016). Deep learning has demonstrated great potential in managing big and complicated datasets, especially when it comes to multi-layer neural networks. For time-series data, such as stock prices, recurrent neural networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, operate very well (Banyal* et al., 2020). Because they can learn long-term dependencies, LSTM networks are useful for forecasting changes in stock prices based on past data (Selvin et al., 2017).

The application of AI in stock market predictions offers several advantages. AI systems can process and analyze large volumes of data more quickly and accurately than human analysts. This efficiency allows for real-time predictions and insights, which are crucial for making timely investment decisions. Additionally, AI algorithms can adapt to new data and changing market conditions, continuously improving their prediction accuracy over time (Feng et al., 2019). While AI holds significant potential for enhancing stock price predictions, it is essential to address the associated challenges to ensure reliable and accurate outcomes. As technology advances, AI-driven stock price prediction is likely to become an indispensable tool for investors, providing them with deeper insights and more informed decision-making capabilities. Future research should focus on improving AI models' robustness and integrating more diverse datasets to enhance prediction accuracy further.

The application of AI in stock price prediction represents a significant advancement in the field of information systems, particularly within the framework of design science research. Design science seeks to create and evaluate artifacts intended to solve identified organizational problems. This section explores how the current work on AI-driven stock price prediction contributes to the theoretical and practical foundations of design science in information systems.

Design science research (DSR) involves the creation and evaluation of artifacts to address specific problems, with a focus on both the utility of these artifacts and the advancement of knowledge. DSR in information systems encompasses building, evaluating, and theorizing about the artifacts created. This research contributes to DSR by developing AI models as artifacts for stock price prediction and by evaluating their effectiveness in real-world scenarios (Hevner et al., 2004). This research creates AI-driven models for stock price prediction, which serve as artifacts. These models are designed to address the unpredictability of stock markets, especially in sectors significantly impacted by inflation, such as transportation. The creation process involves identifying relevant data sources, selecting appropriate AI techniques (e.g., machine learning, deep learning, NLP), and developing algorithms that can analyze and predict stock prices with high accuracy. The evaluation of these artifacts is a critical component of design science research. The AI models are assessed based on their predictive accuracy, robustness, and adaptability to changing market conditions. This evaluation is conducted through backtesting with historical data and real-time testing in current market conditions, aligning with the DSR guideline of rigorous evaluation (Hevner et al., 2004)

The iterative nature of developing and refining AI models aligns with the principles of action design research (ADR), as described by Sein et al., and Lindgren (2011). ADR emphasizes the importance of collaboration between researchers and practitioners in the design and implementation process. This research involves continuous feedback loops with financial experts to refine the AI models, ensuring they meet the practical needs of the industry while also contributing to theoretical advancements. The integration of AI for stock price prediction within the framework of design science research highlights the symbiotic relationship between theory and practice in information systems. By creating, evaluating, and refining AI artifacts, this research not only addresses the practical challenge of market unpredictability but also enriches the theoretical discourse on the application of AI in financial forecasting. This dual contribution underscores the importance of design science as a methodological approach in information systems research, fostering innovation and practical solutions.

3. Materials and method

3.1. Proposed research

The research workflow in Fig 3.1 shows an overview for transportation stock price predictions using data mining technique, namely CRISP-DM. First, business understanding explain inflation that happened from international and national countries. Data understanding was applied 6 transportation stock datasets from AKSI.JK, CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK, and WEHA.JK are collected and predicted from April until October 2023. Data preparation was applied including data pre-processing to replace missing values using mean imputation and splitting dataset randomly. From each datasets are used sequences for train the model using several activation function and optimizer combinations. The results were recorded based on metric performances in web-based system.

3.2. Business understanding

Business understanding process shown in Fig 3.1, International inflation is the inflation that occurred in the top 10 countries in 2022, categorized by the percentage increase The country that experienced the highest increase in inflation was Venezuela by 310 % compared to other countries such as Zimbabwe by 244 %, Argentina by 95 %, Sudan by 87 %, Turkey by 64 %, Sri Lanka by 57 %, Suriname by 55 %, Ghana by 54 %, Iran by 50 %, and South Sudan experienced the lowest inflation by 41 % (Jarne & Shveda, 2023). In addition, year-on-year inflation in Indonesia inflation in Indonesia has increased from 2.18 % in January 2022 to highest inflation in September 2022 of 5.95 %, but the inflation trend decreased and constantly from September 2022 of 5.95 % to 5.51 % in December 2022 (Hadi, 2023). Growth of inflation in some international countries, including Indonesia, may lead to a decline in the purchasing power of shares for investors (Chikwira, 2023).

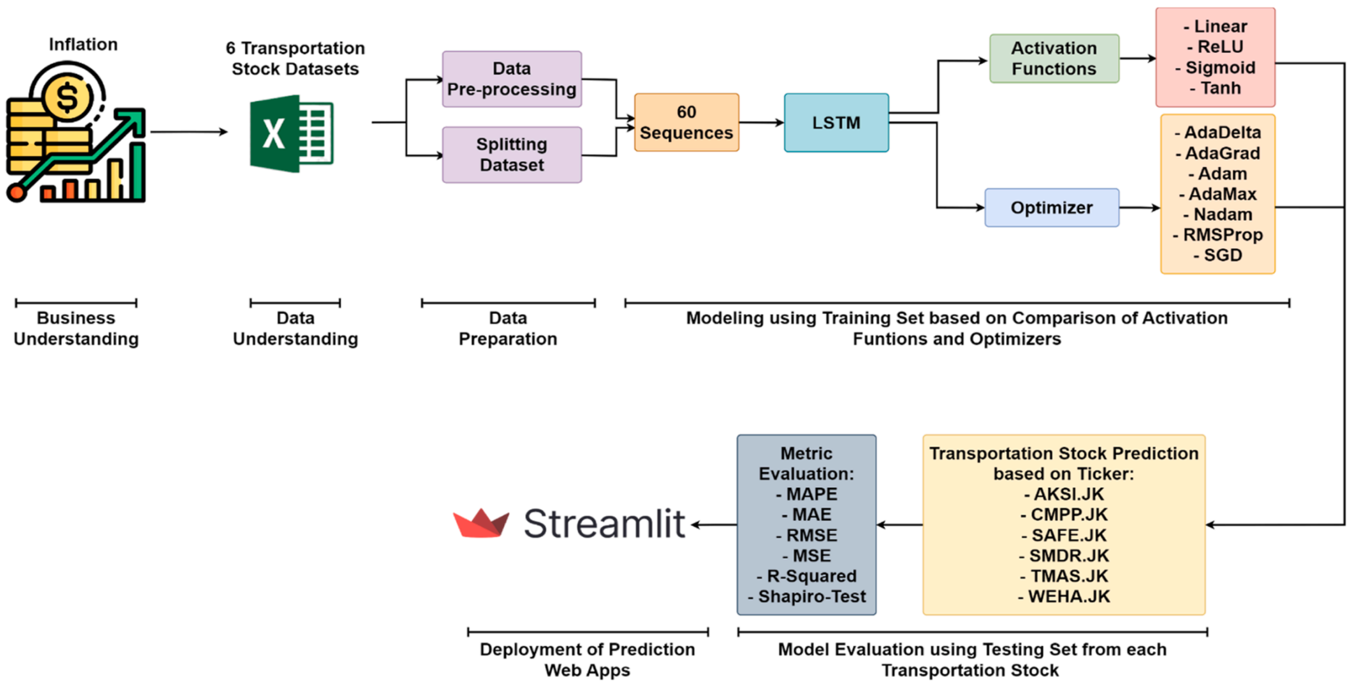


Fig. 3.1. Research workflow.

3.3. Data understanding

Initial data from data understanding process show in Fig 3.1 above, was collected from six transportation stocks namely Mineral Sumberdaya Mandiri (AKSI.JK), Air Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK) from Yahoo Finance in 1st April to 2011 to 1st April to 2023 (Yahoo Finance, 2023). All transportation stock datasets have six attributes such as Open, High, Low, Close, Adj. Close, and Volume along with obtained have same number of datasets are 2.980 show in Table 3.1.

3.4. Data preparation

Data understanding is process to prepare all transportation stock datasets such as Mineral Sumberdaya Mandiri, Air Asia Indonesia, Steady Safe, Samudera Indonesia, Temas, and WEHA Transportasi Indonesia from Yahoo Finance website to several sheet in excel file. Storage of all transportation stock datasets from periods April 1, 2011 until April 1, 2023 will be prepared into 2 processes area data preprocessing and splitting dataset. The following are several stages of data preparation, among others:

3.4.1. Data pre-processing

Data preprocessing is an important step on data analysis process. Some proposed approach and method to enhance the performance of Bangla news article classification addresses imbalances and interpretation challenges, showcasing the effectiveness of tailored strategies in natural language processing (NLP) (Hasib et al., 2023). Their investigation into depression detection from social networks data highlights the potential of computational methods in mental health assessment. Moreover, their development of BMNet-5, aimed at classifying Bengali music genres, and their novel approach to sentiment analysis of Twitter data regarding US airline services, illustrate the broad applicability of deep learning models in analyzing complex datasets across different fields, including healthcare, music classification, and customer service evaluation (Hasib et al., 2022). The proposed MCNN-LSTM to classify multi-class text in imbalanced news data pushes the boundary of combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. This research signifies a leap towards understanding and processing imbalanced datasets, particularly in the news domain, highlighting their commitment to addressing real-world problems through advanced computational techniques. Their work is a testament to the versatility and depth of machine learning applications, from social media analytics to music genre classification, showcasing a wide array of possibilities for future research and application development (Hasib et al., 2023), (Hasib et al., 2023).

Data Preparation stage shown in Fig 3.1 above is the process of preparing data by adjusting the dataset to suit the modelling needs. Data Pre-processing is the process of preparing raw data for modelling purposes. In the process of data pre-processing, there are several stages starting from checking for missing values for NULL values shown in Table 3.2.

Total missing values for each attribute from Table 3.2 above across from 6 transportation stock datasets. SAFE.JK have 8 more than missing values from another transportation stock datasets such as AKSI.JK,

Table 3.1
Transportation stock dataset info.

Dataset	Total Attribute	Total Rows
AKSI.JK	6	2.980
CMPP.JK	6	2.980
SAFE.JK	6	2.980
SMDR.JK	6	2.980
TMAS.JK	6	2.980
WEHA.JK	6	2.980

Table 3.2
Total missing values from datasets.

Dataset	Total Missing Values	Total Rows
AKSI.JK	1	2.979
CMPP.JK	1	2.979
SAFE.JK	8	2.972
SMDR.JK	0	2.980
TMAS.JK	0	2.980
WEHA.JK	1	2.979

CMPP.JK, WEHA.JK, SMDR.JK, and TMAS.JK. AKSI.JK, WEHA.JK, and CMPP.JK only has 1 missing value from each dataset, while SMDR.JK and TMAS.JK don't have missing values from each dataset.

Four datasets of SAFE.JK, AKSI.JK, CMPP.JK, and WEHA.JK that have missing values, requires further data processing from mean imputation. Mean Imputation is one of the most commonly used single imputation methods to replace missing data with the sample mean or median, mode, or distribution of the data (Jadhav et al., 2019). The process of replacing missing data ignores the correlation between each data point as its value is considered independent of any predictor value, leading to a bias in variance (Emmanuel et al., 2021). This method can also ignore the computation time required to replace missing data and significantly reduce the predictive performance of the model (Fletcher Mercaldo & Blume, 2020). This method will be implemented to handle missing values for each transportation stock datasets are AKSI.JK, CMPP.JK, SAFE.JK, and WEHA.JK.

3.4.2. Splitting dataset

Splitting dataset from Fig 3.1 above, training datasets plays a role in building the algorithm model, while testing datasets plays a role in measuring or evaluating the performance of the algorithm model used from each six transportation stock datasets such as AKSI.JK, CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK, and WEHA.JK. Target attribute from each transportation stock datasets is "close" attribute, that used to build a prediction algorithm model with a training data ratio of 90 % and a test data ratio of 10 % (Rana et al., 2019). Example from comparison of training and testing datasets obtained 2.682 samples from training set and 298 from testing set of each transportation stock datasets are AKSI.JK, CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK, WEHA.JK show in Table 3.3 that have been data cleaned using mean imputation.

Data distributions from training and testing set in transportation stock datasets are CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK, and WEHA.JK will be processed with the same distribution of AKSI.JK dataset show in Table 3.3 above. Next process, all transportation stock datasets will be normalized using MinMaxScaler. MinMaxScaler is data preprocessing technique used to normalize data and ensure that all data has the same range of values (Deepa & Ramesh, 2022). It adjusts and shifts each feature independently based on the range specified in the training set (Pires et al., 2020). The normalization process can be expressed to implemented from each stock datasets are AKSI.JK, CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK and WEHA.JK before modeling process, where X_i represents the value of an observation in the feature, and X_{min} and X_{max} denote the minimum and maximum values of the feature, respectively (Mohd Faizal et al., 2023).

$$Z = \frac{X_i \; -- X_{min}}{X_{max} \; -- X_{min}} \tag{1}$$

3.5. Modeling

Modeling process shown in Fig 3.1 above, specific methods and algorithms to construct models as solutions to fulfill the business understanding. Modeling process will utilize Visual Studio Code, employing the Python programming language and other frameworks such as Keras and Tensorflow. Python programming language is chosen because there have been numerous previous studies (Banyal* et al., 2020; Rana et al., 2019; Sami et al., 2023). have used Python, and it has been proven that Python is a more efficient programming language. Stages of modeling in deep learning algorithmic modeling with 60 sequences with activation function comparison and optimization of 240 sequences, including (Kamalov et al., 2020):

Table 3.3
Example of training and testing set from AKSI.JK dataset.

	Ratio	No. Samples
Training	90 %	2.682
Testing	10 %	298

3.5.1. Long short-term memory

LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) that focuses on long-term memory (Nan, 2023). It is a suitable option for analyzing, handling, and making forecasts using time series data due to the potential presence of an uncertain lack of time intervals between significant events (Kristiyanti & Hardani, 2023). Additionally, LSTM utilizes a specialized hidden state and three controlling gates to regulate and manage both of long-term and short-term memory (Huang et al., 2023). LSTM network, being a distinct type of RNN, consists of an input layer, a hidden layer, and an output layer, with the structure of the LSTM cell depicted in Fig 3.2 (Filho et al., 2020).

The process of sending the information initially starts from the forget gate layer, as shown in Fig 3.2 above. The function of the forget gate is to retain significant data in the cell state while discarding irrelevant information (Zaheer et al., 2023). It can be proven by the calculation following below (Kratzert et al., 2018).

$$f_t = \sigma\,(Wf[ht - 1,\ xt] + bf) \tag{2}$$

f_t is the resultant vector containing values between 0 and 1, $\sigma(.)$ denotes the logistic sigmoid function, and Wf, Uf and bf consists of a set of adjustable parameters that determine the weight matrix and bias vector for forgetting gates. Hidden states, referred to as ht are initialized in the first step using a vector of zeros with a length specified by the user. Next, the potential update vector for the cell state is calculated in the next step using the current input xt and the previous hidden state $ht \; -- 1$ like the following below (Zaheer et al., 2023).

$$\widehat{C}t = tanh\,(Wc\,.\,[ht - 1,\ xt] + bc) \tag{3}$$

Wc represents the weights, $ht - 1$ indicates the previous state value, xt represents the input at the current time step, bc represents the bias, and $\widehat{C}t$ represents the cell state function value at a given time step. The result of cell function and output of the input gate are combined by multiplying and adding them to the cell state like the following below (Sen & Raghunathan, 2018).

$$Ct = ft \odot Ct - 1 + it \odot \widehat{C}t \tag{4}$$

The input vector and the candidate memory cell vector are combined together to modify the previous memory cell $Ct \; -- 1$ and $\odot$ represents element-wise multiplication. The output gate layer holds the results of computation or processing of information from the early stage layers, including forget gate and update gate. In this layer, the overall calculation result is obtained through two computational stages to get the final result like following below (Sen & Raghunathan, 2018).

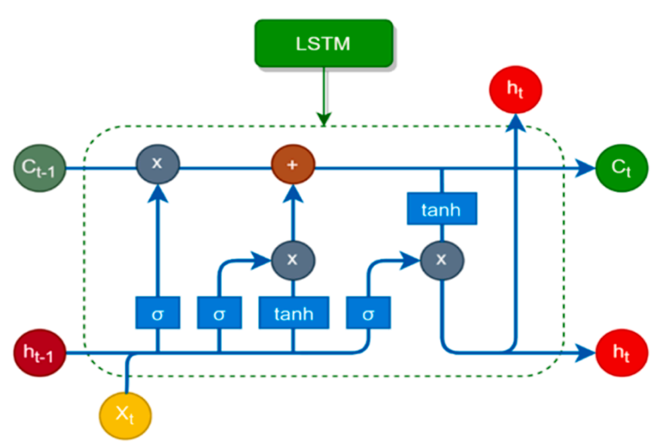


Fig. 3.2. LSTM architecture.

$$Ot = \sigma\,(Wo\,.\,[ht - 1,\ xt] + bo) \tag{5}$$

Ot represents the output gate, known as the final gate, organize transfer of information from all current memory cell to the current hidden state. This hidden state is also the output of the LSTM at the current time step. An LSTM architecture involves several calculation procedures within its layers to produce results, and these procedures are detailed in Table 3.4, an outline of the calculation process for each layer in the LSTM.

3.5.2. Activation techniques

Activation is a function used to determine whether a node will be activated or not in artificial neural network modelling (Brito da Silva et al., 2019). It can also be applied to the neuron's output to obtain more specific results in the form of a nonlinear transformation of the input signal (Zhao et al., 2021). Activation that were used in the experiments are Linear, ReLU (Rectified Linear Unit), Tanh, and Sigmoid.

Linear's activation is a function that has a gradient that is directly proportional to the input entered in neural network modelling (Sharma et al., 2020). ReLU's activation is a simple function that acts as an identity function for positive input values and returns 0 for negative input values (Dubey et al., 2022). Tanh's activation is a simple computational function, the derivative of one used in the back-propagation process of Recurrent Neural Network (RNN) modelling (Poornima & Pushpalatha, 2019). Sigmoid's activation is a function designed using a non-linear resistive load that takes current as an input value and provides voltage as an output value to the neurons (Venugopal & Vigneswaran, 2019). The formula for Linear, ReLU (Rectified Linear Unit), Tanh, and Sigmoid activation's are given in Eqs. (6–9).

$$f(x_i) = kx_i \tag{6}$$

$$ReLU(x) = max(x,\ 0) \tag{7}$$

$$Tanh\ (x) = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{8}$$

$$f(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

3.5.3. Optimization techniques

An optimizer is a mechanism that aims to improve processing or learning efficiency in a prediction algorithm model based on improving accuracy scores in a prediction task (Hassan et al., 2023). In other words, optimizer by the optimiser can shape and improve the accuracy score by utilising the weights to achieve the most accurate prediction possible (Lee & Kim, 2023). Optimizer that were used in the experiments are Adam, AdaGrad (Adaptive Gradients), Nadam, RMSProp (Root Mean Square Propagation), AdaDelta, SGD (Stochastic Gradient Descent), and AdaMax.

Adam's Optimizer is an algorithm optimization method based on stochastic gradients that is well suited for direct implementation in any model, in particular for large data sets and parameters (Ilboudo et al., 2022). AdaGrad's Optimizer is a variant of the Stochastic Gradient Descent optimisation algorithm that adjusts the learning rate dynamically for different parameter (Tokgoz et al., 2023). Nadam's optimizer is an extension of Adam's algorithm that combines Nesterov momentum with gradient descent (Fofana et al., 2021). RMSProp's optimizer stands for Root Mean Square Propagation, and has been developed to achieve better results and prevent convergence by not reducing the learning rate (Olanipekun et al., 2022). AdaDelta's optimizer is an extended optimization algorithm of AdaGrad that solves the problem of decreasing learning rate but not accumulating squared gradient values (Yaqub et al., 2020). SGD's optimizer stands for Stochastic Gradient Descent, a popular choice of optimization algorithm for many recent advances in deep learning across domains (Abdulkadirov et al., 2023). AdaMax's optimizer is a variant of Adam's optimization algorithm that simplifies the learning rate range and has been used to train Convolutional Neural Network (CNN) models (Weiss et al., 2022). The formula for Adam, AdaGrad (Adaptive Gradients), Nadam, RMSProp (Root Mean Square Propagation), AdaDelta, SGD (Stochastic Gradient Descent), and AdaMax optimizer's are given in Eqs. (10–16).

$$\theta_{t+1} = \theta_t - \frac{n}{\sqrt{\hat{v}_t} + \in}\,\hat{m}_t \tag{10}$$

$$W_t = W_{t-1} - \lambda'_t\,\frac{\partial\,\nabla}{\partial\,W_{t-1}} \tag{11}$$

$$w^i_t = w^i_{t-1} - \frac{\eta}{\sqrt{v_t} + \varepsilon}\,.\,\tilde{m}_t \tag{12}$$

$$G = \nabla_W C(w_t) \tag{13}$$

$$g'_t = \frac{\sqrt{u_{t-1} + \varepsilon}}{\sqrt{s_t}}\,g_t \tag{14}$$

$$W_{i+1} = W_i - \eta\,\frac{\partial L}{\partial W_i} \tag{15}$$

$$v_t = \beta_2\,v_{t-1} + (1 - \beta_2)|g_t|^2 \tag{16}$$

3.6. Evaluation metrics

Evaluation process shown in Fig 3.1 above, processes or stages carried out to assess the performance of the applied modeling. This evaluation process serves to determine how well the predictions made using the model built to answer the business understanding stage of each transportation stock prediction from Mineral Sumberdaya Mandiri, Air Asia Indonesia, Steady Safe, Samudera Indonesia, Temas, and WEHA Transportasi Indonesia. Evaluation metrics that were used in the experiments are Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), r-squared, elapsed time, and shapiro-wilk test.

Mean Absolute Percentage Error (MAPE) calculates the prediction error by dividing the absolute error of each period by the actual observed value for that period (Hayder et al., 2023). Mean Absolute Error (MAE) is one of the evaluation metrics that can be used when outliers on each attribute resubmit dirty data (Chicco et al., 2021). Root

Table 3.4
Pseudocode LSTM algorithm.

Algorithm 1 Pseudocode LSTM Algorithm
1: **Input:** $x = [x_1, \ldots\ldots, x_n],\ x_t \in \mathbb{R}^m$
2: **Given parameters:** $\mathbf{W}_f, \mathbf{U}_f, \boldsymbol{b}_f, \mathbf{W}_{\tilde{c}}, \mathbf{U}_{\tilde{c}}, \boldsymbol{b}_{\tilde{c}}, \mathbf{W}_i, \mathbf{U}_i, \boldsymbol{b}_i, \mathbf{W}_o, \mathbf{U}_0, \boldsymbol{b}_0$
3: **Initiliaze** $\boldsymbol{h}_0, \boldsymbol{c}_0 = \vec{0}$ of length p
4: **for** $t = 1, \ldots, n$ **do**
5: **Calculate** $\mathbf{f}_t, \hat{\mathbf{c}}_t, \mathbf{i}_t$
6: **Update cell state** $\mathbf{c}_t$
7: **Calculate** $\mathbf{o}_t$
8: **end for**
9: **Output:** $h = [h_1, \ldots\ldots, h_n],\ h_t \in \mathbb{R}^p$

Mean Squared Error (RMSE) is a metric that measures the distribution of residuals, indicating the extent to that data is spread around the optimal line of fit (Murugesan & Jung, 2021). Mean Squared Error (MSE) is a widely used approach in neural network models to determine error and precision (Lyu et al., 2022). R-squared proves useful when comparing different models and evaluating their ability to capture fundamental patterns in the data (P. W. Khan et al., 2020). Elapsed Time refers to the average time taken from when a data point is generated by SenSE until the performance result is received (Kamienski et al., 2019). Shapiro-Wilk test is the most widely used statistical hypothesis testing method to test the normality of data (Mishra et al., 2019). The formula for Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), r-squared, elapsed time, and shapiro-wilk test are given in Eqs. (17-23).

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|yl - y|}{yi} \tag{17}$$

$$MAE = \frac{\sum_{n=1}^{N}|\hat{r}_n - r_n|}{N} \tag{18}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{19}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{20}$$

$$R - Squared = 1 - \frac{\sum_{i=1}^{n}(X_i - y_i)^2}{\sum_{i=1}^{n}(\overline{Y} - Y_i)^2} \tag{21}$$

$$T = t_n - t_1 \tag{22}$$

$$W_X = \frac{\left(\sum_{i=1}^{n} a_i X_{(i)}\right)^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{23}$$

3.7. Deployment

Deployment process shown in Fig 3.1 above, the process of developing algorithmic models with various activation comparisons namely linear, ReLU, tanh, and sigmoid, as well as optimizations namely adam, adagrad, nadam, RMSProp, adadelta, SGD, and adamax into a web-based application using streamlit with input parameters will be carried out. The deployment stage begins with saving all models that have been trained based on linear, ReLU, tanh, and sigmoid activation comparisons, as well as optimization of adam, adagrad, nadam, RMSProp, adadelta, SGD, and adamax into .h5 format. The stored models will be implemented in a web-based application to display transportation stock predictions as well as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and R-squared evaluation metrics.

4. Result and discussion

This section are the results and discussion of performance from Long Short-Term Memory (LSTM) conventional model and comparison LSTM model using activation functions and optimizers to predict transportation stocks in Indonesia namely Mineral Sumberdaya Mandiri, Air Asia Indonesia, Steady Safe, Samudera Indonesia, Temas, and WEHA Transportasi Indonesia.

4.1. Comparison algorithm and evaluation metrics

In this section, Long Short-Term Memory (LSTM) algorithm will compared based on conventional model and LSTM with hyperparameter from activation functions and optimizers. Evaluation metrics that used to evaluate the model from each transportation stock datasets are MAE, MAPE, MSE, R-Squared, and RMSE along with elapsed time from training time model in LSTM conventional.

4.1.1. LSTM

This section presents the performance results of the conventional LSTM model in predicting the transportation stock dataset, specifically Mineral Sumberdaya Mandiri (AKSI.JK), Air Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK). Table 4.1 and Fig 4.1 show the performance of the conventional LSTM model for the transportation stock dataset. The time taken analysis shows that the TMAS.JK, SMDR.JK, and WEHA.JK stock datasets exhibit slower training speed for the LSTM algorithm compared to the CMPP.JK, SAFE.JK, and AKSI.JK stock datasets. Among the three datasets, TMAS.JK has the shortest training time of 9.56 min, followed by SMDR.JK and WEHA.JK. On the other hand, the CMPP.JK, SAFE.JK, and AKSI.JK datasets require longer time compared to the TMAS.JK, SMDR.JK, and WEHA.JK datasets. In particular, the AKSI.JK stock dataset has the longest time of 21.12 min, exceeding the time taken by the SAFE.JK and CMPP.JK datasets.

MAE analysis shows that the AKSI.JK, CMPP.JK, and SA-FE.JK stock datasets have lower performance errors compared to the SMDR.JK, TMAS.JK, and WEHA.JK stock datasets. In particular, the AKSI.JK stock dataset has a smaller Mean Absolute Error (MAE) of 0.0116605 compared to the stock datasets. On the other hand, the SMDR.JK, TMAS.JK, and WEHA.JK stock datasets have higher MAE values than the SAFE.JK, CMPP.JK, and AKSI.JK. SAFE.JK, AKSI.JK, and WEHA.JK stock datasets based on MAPE analysis, have lower performance errors compared to the TMAS.JK, SMDR.JK, and CMPP.JK stock datasets. In particular, the SAFE.JK stock dataset has a smaller Mean Absolute Error (MAE) of 0.04884 compared to AKSI.JK and WEHA.JK. On the other hand, the TMAS.JK, SMDR.JK, and CMPP.JK stock datasets have higher MAE values compared to the SAFE.JK, AKSI.JK, and WEHA.JK stock datasets. In particular, the CMPP.JK stock dataset has a smaller Mean Absolute Error of 0.10927 compared to the SMDR.JK and TMAS.JK stock datasets.

Specifically, the performance from LSTM conventional model from each transportation stock datasets, Air Asia Indonesia (CMPP.JK) has best performance from MAE of 0.0146438, MAPE of 0.10927, MSE of 0.0003563, R-Squared of 93 %, RMSE of 0.01888. Although this stock has a longer LSTM model training of 12.14 min14 minutes and higher error performance compared to the SAFE.JK, WEHA.JK, TMAS.JK, and SMDR.JK stock datasets, there is something interesting about this dataset, it has a coefficient of determination value that reaches 93 % greater than the other stock datasets. This result indicates that the predicted value results have a greater chance of explaining the dependent variable or a very accurate prediction value and the prediction error performance of MAE, MAPE, MSE, and RMSE is smaller than the AKSI.JK, SAFE.JK, WEHA.JK, TMAS.JK, and SMDR.JK stock datasets.

4.1.2. LSTM with activation functions & optimizers

This section presents the performance of LSTM models based on different hyperparameters, including activation functions and optimizers. The activation functions evaluated for the LSTM model are linear, ReLU, tanh, and sigmoid. The optimizers used are Adam, Adagrad, Nadam, RMSProp, Adadelta, SGD, and Adamax. The performance of these activation functions and optimizers is demonstrated on a dataset of transportation stocks, which include Mineral Sumberdaya Mandiri (AKSI.JK), Air Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK).

Table 4.2 below presents the model performance of the activation function and optimizer of the LSTM model from the AKSI.JK and CMPP.JK stock datasets. The AKSI.JK dataset has 5 best performances based on activation function and optimizer, namely ReLU, linear, and tanh

Table 4.1
Model performances of LSTM conventional model from transportation stocks.

Dataset	Elapsed Time (min)	MAE (%)	MAPE (%)	MSE (%)	R-Squared (%)	RMSE (%)
AKSI.JK	21.12	0.0116605	0.07697	0.0002741	0.329813	0.01656
CMPP.JK	**12.14**	**0.0146438**	**0.10927**	**0.0003563**	**0.939591**	**0.01888**
SAFE.JK	14.7	0.0191644	0.04884	0.000599	0.40595	0.02447
SMDR.JK	10.15	0.6891372	0.30152	0.607971	−2.377048	0.77972
TMAS.JK	9.56	0.5825399	0.36527	0.3841733	−3.518707	0.61982
WEHA.JK	11.65	0.0253701	0.08396	0.00114	0.891818	0.03376

(a) Elapsed Time (min) from Transportation Stock

(b) Mean Absolute Error from Transportation Stock

c) Mean Absolute Percentage Error from Transportation Stock

(d) Mean Squared Error from Transportation Stock

(e) R-Squared Error from Transportation Stock

(f) Root Mean Squared Error from Transportation Stock

Fig. 4.1. Model performances of LSTM conventional model from transportation stocks.

activation with ad-am, nadam, RMSProp, and adamax optimizers. Based on the performance of the top 5 models, there is something interesting about this dataset. ReLU activation with nadam optimizer has a lower performance in terms of MSE of 0.0002039, RMSE of 0.01428, and a higher coefficient of determination from R-Squared of 50 % when compared to other activations, namely linear and tanh with adam, RMSProp, and adamax. The CMPP.JK dataset has 5 best performances based on activation function and optimization such as ReLU, tanh, and

Table 4.2
Model performances of LSTM model by activation functions and optimizers from AKSI.JK and CMPP.JK stock datasets.

Dataset	Activation	Optimizer	Elapsed Time (min)	MAE (%)	MAPE (%)	MSE (%)	R-Squared (%)	RMSE (%)
AKSI.JK	Linear	AdaDelta	114.08	0.1411001	0.93828	0.0203102	−48.658497	0.14251
		AdaGrad	120.22	0.0407017	0.26025	0.0020034	−3.898347	0.04476
		Adam	125.25	0.013728	0.08885	0.0003509	0.141948	0.01873
		AdaMax	128.11	0.0147873	0.09769	0.0004355	−0.064861	0.02087
		Nadam	124.63	0.0093744	0.06011	0.0002329	0.43057	0.01526
		RMSProp	122.99	0.0108179	0.06903	0.0002787	0.318511	0.01669
		SGD	115.98	0.032883	0.20994	0.001365	−2.337351	0.03695
	ReLU	AdaDelta	113.75	0.1463257	0.97308	0.0218322	−52.379844	0.14776
		AdaGrad	118.91	0.0564243	0.36556	0.003572	−7.733444	0.05977
		Adam	119.66	0.0083481	0.05275	0.0002139	0.476992	0.01463
		AdaMax	111.45	0.0166904	0.11066	0.0005104	−0.247851	0.02259
		Nadam	**127.26**	**0.0088794**	**0.05601**	**0.0002039**	**0.501388**	**0.01428**
		RMSProp	120.75	0.0102312	0.06505	0.0003208	0.21568	0.01791
		SGD	117.2	0.0358757	0.22994	0.00157	−2.838739	0.03962
	Sigmoid	AdaDelta	107.27	0.115355	0.79716	0.0137251	−32.557885	0.11715
		AdaGrad	122.5	0.1210827	0.83573	0.0150717	−35.850371	0.12277
		Adam	101.21	0.0110326	0.0704	0.0003012	0.263464	0.01736
		AdaMax	109.99	0.0133187	0.08828	0.0005293	−0.294086	0.02301
		Nadam	121.19	0.0109521	0.07034	0.0003428	0.161754	0.01851
		RMSProp	126.1	0.0279379	0.18246	0.0010325	−1.524418	0.03213
		SGD	108.9	0.0444058	0.28396	0.0023744	−4.80547	0.04873
	Tanh	AdaDelta	117.94	0.146567	0.97489	0.0219018	−52.550061	0.14799
		AdaGrad	110.03	0.0443995	0.28459	0.0023436	−4.73006	0.04841
		Adam	108.86	0.0192328	0.12722	0.0005197	−0.270645	0.0228
		AdaMax	117.74	0.009806	0.06264	0.0002964	0.275403	0.01722
		Nadam	116.83	0.014555	0.09527	0.0003581	0.124507	0.01892
		RMSProp	117.03	0.0213475	0.14184	0.0006638	−0.622879	0.02576
		SGD	115.3	0.0366961	0.23549	0.0016294	−2.983821	0.04037
CMPP.JK	Linear	AdaDelta	141.16	0.1189503	0.84968	0.0192562	−2.265067	0.13877
		AdaGrad	117.75	0.0457831	0.2762	0.00065805	−0.115791	0.08112
		Adam	116.18	0.0121662	0.09417	0.0002672	0.954698	0.01635
		AdaMax	118.2	0.0168838	0.13477	0.0004791	0.918757	0.02189
		Nadam	135.91	0.0119413	0.08811	0.000268	0.954561	0.01637
		RMSProp	146.07	0.0181908	0.14688	0.0004841	0.91791	0.022
		SGD	147.05	0.0364132	0.22087	0.004005	0.320908	0.06329
	ReLU	AdaDelta	112.97	0.1294893	0.93195	0.0224197	−2.801469	0.14973
		AdaGrad	121.72	0.0530037	0.29859	0.0078665	−0.333831	0.08869
		Adam	**104.35**	**0.0092918**	**0.06422**	**0.0002123**	**0.963999**	**0.01457**
		AdaMax	120.25	0.0122212	0.09055	0.0003171	0.946239	0.01781
		Nadam	111.58	0.0093115	0.06567	0.000216	0.963374	0.0147
		RMSProp	121.12	0.0184314	0.15146	0.0004958	0.915934	0.02227
		SGD	98.96	0.0372071	0.22589	0.0042278	0.283142	0.06502
	Sigmoid	AdaDelta	115.49	0.2384949	2.26155	0.0605894	−9.273472	0.24615
		AdaGrad	106.56	0.0472191	0.2799	0.0069161	0.172683	0.08316
		Adam	126.7	0.0113837	0.08164	0.0002794	0.952617	0.01672
		AdaMax	82.5	0.0172547	0.12295	0.0005945	0.899198	0.02438
		Nadam	110.99	0.0262671	0.1897	0.0013371	0.773279	0.03657
		RMSProp	119.66	0.0332332	0.27776	0.001394	0.763629	0.03734
		SGD	121.86	0.0471479	0.27893	0.0068557	−0.162437	0.0828
	Tanh	AdaDelta	91.82	0.1373486	0.98181	0.0256706	0.6484	0.16022
		AdaGrad	89.88	0.0369543	0.22637	0.0041433	0.297473	0.06437
		Adam	77.37	0.0143202	0.0922	0.000507	0.914033	0.02252
		AdaMax	86.5	0.0098673	0.06988	0.0002425	0.958882	0.01557
		Nadam	87.29	0.0120772	0.09287	0.0002585	0.956175	0.01608
		RMSProp	96.59	0.0130281	0.0974	0.0003032	0.948582	0.01741
		SGD	84.86	0.0327754	0.20161	0.0032411	0.450447	0.05693

linear activation with adam, nadam, and adamax optimization. Based on the top 5 model performance, there is something interesting about this dataset. ReLU activation with adam optimizer has a lower performance in terms of MSE of 0.0002123, RMSE of 0.01457, and a higher coefficient of determination from R- Squared of 96 % compared to other activations, namely linear and tanh with adam, nadam, and adamax optimizers. The performance of LSTM model is further shown in Table 4.3 below which displays the SAFE.JK and SMDR.JK stock datasets. The SAFE.JK dataset has 5 best performances based on activation functions and optimizations such as tanh, sigmoid, and ReLU activations with nadam and adam optimizations. Based on the top 5 best performing models, there is something interesting about this dataset. Tanh activation with nadam optimizer has a lower performance in terms of MSE of 0.0004659, RMSE of 0.02158, and a higher coefficient of determination from R-Squared of 53 % compared to sigmoid and ReLU activation. The SMDR.JK dataset has 5 best performances based on activation function and optimization such as linear, tanh, and ReLU activation with SGD, adamax, and adam optimization. Based on the top 5 model performance, there is something interesting about this dataset. The tanh activation with adamax optimizer has a lower performance in terms of MSE of 0.0880576, RMSE of 0.29675, and a higher coefficient of determination from R-Squared of 51 % compared to other activations, namely linear and tanh with SGD and adam optimizers.

Table 4.4 below presents the model performance of the activation function and optimizer of the LSTM model from the TMAS.JK and WEHA.JK stock datasets. The TMAS.JK dataset has 5 best performances based on activation function and optimizer, namely ReLU, linear, and tanh activation with ad-amax and nadam optimizers. Based on the top 5

Table 4.3
Model performances of LSTM model by activation functions and optimizers from SAFE.JK and SMDR.JK stock dataset.

Dataset	Activation	Optimizer	Elapsed Time (min)	MAE (%)	MAPE (%)	MSE (%)	R-Squared (%)	RMSE (%)
SAFE.JK	Linear	AdaDelta	71.65	0.3891133	0.96892	0.1524513	−150.19068	0.39045
		AdaGrad	69.15	0.0948563	0.23246	0.0097508	−8.670196	0.09875
		Adam	59.36	0.0199288	0.05077	0.0006262	0.378951	0.02502
		AdaMax	73.79	0.0183383	0.0456	0.0006443	0.361067	0.02538
		Nadam	72.88	0.0211785	0.05391	0.0006727	0.332891	0.02594
		RMSProp	73.28	0.0225846	0.05784	0.0007769	0.229568	0.02787
		SGD	76.14	0.0329282	0.07959	0.0015712	−0.558246	0.03964
	ReLU	AdaDelta	70.21	0.3496615	0.87044	0.1231789	−121.160323	0.35097
		AdaGrad	72.3	0.1523978	0.37608	0.0241184	−22.919016	0.1553
		Adam	65.61	0.0184185	0.04551	0.0006138	0.391229	0.02477
		AdaMax	78.57	0.0187016	0.04709	0.0006586	0.346848	0.02566
		Nadam	74.36	0.016943	0.04197	0.000543	0.461457	0.0233
		RMSProp	77.64	0.0296531	0.07295	0.0012285	−0.218327	0.03505
		SGD	72.31	0.0458683	0.11119	0.0026298	−1.608041	0.05128
	Sigmoid	AdaDelta	69.85	0.4281985	1.06717	0.1843068	−181.782786	0.42931
		AdaGrad	76.42	0.1581624	0.39018	0.0260116	−24.796492	0.16128
		Adam	79.59	0.0211831	0.05229	0.0008193	0.187497	0.02862
		AdaMax	85.06	0.0346803	0.08257	0.0018996	−0.883877	0.04358
		Nadam	82.98	0.0163909	0.04056	0.0005087	0.495545	0.02255
		RMSProp	74.82	0.0264128	0.06493	0.0010606	−0.051861	0.03257
		SGD	67.68	0.157868	0.38957	0.0258794	−24.66537	0.16087
	Tanh	AdaDelta	76.08	0.3340162	0.83136	0.1124454	−110.515619	0.33533
		AdaGrad	107.67	0.0945184	0.23162	0.0096887	−8.608627	0.09843
		Adam	100.28	0.0160239	0.03985	0.0005089	0.495301	0.02256
		AdaMax	75.47	0.0183708	0.0464	0.0006471	0.358257	0.02544
		Nadam	**87.06**	**0.0158117**	**0.03984**	**0.0004659**	**0.537937**	**0.02158**
		RMSProp	76.88	0.0395254	0.09723	0.0019361	−0.920052	0.044
		SGD	87.56	0.0406081	0.09852	0.0021445	−1.12678	0.04631
SMDR.JK	Linear	AdaDelta	114.54	1.9216183	0.8891	3.8524586	−20.398944	1.96277
		AdaGrad	91.07	1.8465863	0.8513	3.5877094	−18.928362	1.89412
		Adam	89.82	0.2800096	0.12416	0.1098653	0.38974	0.33146
		AdaMax	125.05	0.3157966	0.13806	0.1429296	0.20608	0.37806
		Nadam	136.34	0.3781864	0.16957	0.1800447	−8e-05	0.42432
		RMSProp	111.17	0.3444461	0.14917	0.175245	0.02658	0.41862
		SGD	76.14	0.0329282	0.07959	0.0015712	−0.558246	0.03964
	ReLU	AdaDelta	173.14	1.9350504	0.89484	3.9116207	−20.727567	1.97778
		AdaGrad	119.61	2.0320714	0.93917	4.3175704	−22.982463	2.07788
		Adam	127.87	0.3489477	0.15544	0.1585388	0.119377	0.39817
		AdaMax	177.01	0.5362993	0.23549	0.3676068	−1.041916	0.60631
		Nadam	110.22	0.4412922	0.19799	0.2385993	−0.325328	0.48847
		RMSProp	72.56	0.4054026	0.17519	0.2361109	−0.311507	0.48591
		SGD	193.23	1.783136	0.82222	3.3430151	−17.569179	1.82839
	Sigmoid	AdaDelta	178.63	0.6316804	0.2673	0.5868883	−2.259942	0.76609
		AdaGrad	138.71	2.061405	0.95381	4.4313475	−23.614451	2.10508
		Adam	176.3	1.1191406	0.50405	1.4135866	−6.851936	1.18894
		AdaMax	175.23	1.4587871	0.6665	2.2933001	−11.73841	1.51436
		Nadam	174.26	0.8827884	0.39446	0.9128652	−4.07062	0.95544
		RMSProp	175.78	1.0499303	0.47232	1.2513875	−5.950982	1.11865
		SGD	178.08	1.9433511	0.89779	3.9528967	−20.956839	1.98819
	Tanh	AdaDelta	172.61	2.093085	0.96907	4.5629419	−24.345408	2.1361
		AdaGrad	173.93	1.7444241	0.80444	3.2005391	−16.77778	1.78901
		Adam	183.67	0.3569744	0.15756	0.1681515	0.065982	0.41006
		AdaMax	**173.18**	**0.2101362**	**0.08806**	**0.0880576**	**0.510873**	**0.29675**
		Nadam	190.53	0.4370276	0.19432	0.2381563	−0.322868	0.48801
		RMSProp	176.31	0.5245328	0.22956	0.360404	−1.001907	0.60034
		SGD	172.5	1.4444561	0.66291	2.2245312	−11.356426	1.49149

model performance, there is something interesting about this dataset. ReLU activation with adamax optimizer has a lower performance in terms of MSE of 0.0404941, RMSE of 0.20123, and a higher coefficient of determination from R-Squared of 52 % compared to other activations, namely linear and tanh with adam optimizer. The WEHA.JK dataset has 5 best performances based on activation functions and optimizers such as linear activation and tanh with adam, nadam, and adamax optimizers. Based on the top 5 model performance, there is something interesting about this dataset. Linear activation with adam optimizer has a lower performance in terms of MSE of 0.0010775, RMSE of 0.03283, and a higher coefficient of determination from R-Squared of 89 % compared to other activations, namely tanh with nadam and adamax optimizers.

Based on Fig 4.2, the prediction of transportation stocks using different activation functions and optimizers, along with performance evaluation metrics. The stocks analyzed for the period between April and October 2023 include Mineral Sumberdaya Mandiri (AKSI.JK), Air Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Te-mas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK). The evaluation metrics used are MAE, MAPE, MSE, R-Squared, and RMSE. Starting with Mineral Sumberdaya Mandiri stock, the first stock, there was no significant increase between April and May 2023, with prices ranging from 1.4 billion to 1.6 billion. However, in June to August, there was a significant increase in price, ranging from 1.4 billion to 2 billion, followed by a decrease from September to October to 1.2 billion. Predictions for this stock used the ReLU activation function and the Nadam optimizer Furthermore, Air Asia Indonesia, the second stock, followed a similar pattern to Mineral Sumberdaya

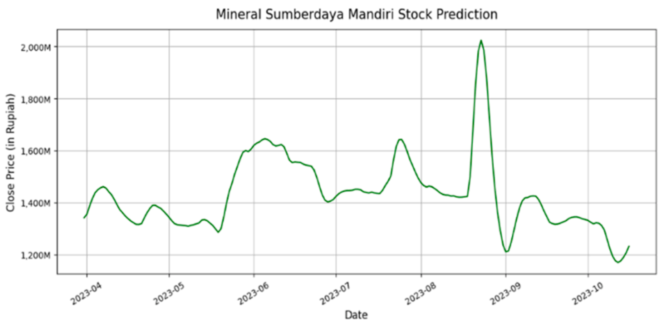
Table 4.4
Model performances of LSTM model by activation functions and optimizers from TMAS.JK and WEHA.JK stock datasets.

Dataset	Activation	Optimizer	Elapsed Time (min)	MAE (%)	MAPE (%)	MSE (%)	R-Squared (%)	RMSE (%)
TMAS.JK	Linear	AdaDelta	105.35	1.6988036	1.10194	2.9811393	−34.064634	1.7266
		AdaGrad	144.38	1.3594205	0.87749	1.9276772	−21.673645	1.38841
		Adam	72.91	0.1648907	0.10099	0.0413225	0.513958	0.20328
		AdaMax	129.16	0.1727892	0.10533	0.0471447	0.445476	0.21713
		Nadam	112.58	0.2766955	0.17442	0.0931482	−0.095623	0.3052
		RMSProp	135.32	0.377653	0.23375	0.1736341	−1.042312	0.41669
		SGD	138.87	1.3045373	0.8415	1.7777032	−19.909627	1.33331
	ReLU	AdaDelta	87.76	1.588086	1.02825	2.6142226	−29.7489	1.61686
		AdaGrad	153.03	1.2031804	0.77573	1.5141695	−16.809902	1.23052
		Adam	135.15	0.2298114	0.14208	0.0704677	0.171148	0.26546
		AdaMax	**69.0**	**0.1484958**	**0.09002**	**0.0404941**	**0.523703**	**0.20123**
		Nadam	133.87	0.2580699	0.16288	0.0830604	0.023031	0.2882
		RMSProp	87.05	0.2983446	0.18196	0.1174824	−0.381847	0.34276
		SGD	72.64	1.4147432	0.91228	2.091806	−23.604154	1.44631
	Sigmoid	AdaDelta	81.14	1.9527325	1.27434	3.8990916	−44.861733	1.97461
		AdaGrad	70.78	1.5624308	1.01335	2.5225199	−28.670279	1.58824
		Adam	63.9	1.0371461	0.6617	1.1529433	−12.561101	1.07375
		AdaMax	83.94	1.2344657	0.7924	1.6088571	−17.923633	1.26841
		Nadam	86.97	1.244571	0.79951	1.6324352	−18.200962	1.27767
		RMSProp	84.29	0.9123407	0.57994	0.9013942	−9.602341	0.94942
		SGD	85.25	1.4164581	0.91503	2.0891936	−23.573427	1.4454
	Tanh	AdaDelta	61.05	1.399788	0.90569	2.0341072	22.925492	1.42622
		AdaGrad	68.79	1.3858043	0.89492	2.0012696	−22.539251	1.41466
		Adam	84.72	0.2635969	0.16299	0.088622	−0.042387	0.29769
		AdaMax	60.99	0.2017587	0.12199	0.0610237	0.282229	0.24703
		Nadam	63.25	0.267067	0.16604	0.089608	−0.053984	0.29935
		RMSProp	65.1	0.3675875	0.22508	0.170889	−1.010024	0.41339
		SGD	63.0	1.2461414	0.803	1.6258883	−18.123956	1.2751
WEHA.JK	Linear	AdaDelta	64.08	0.2749349	0.88865	0.0845976	−7.028241	0.29086
		AdaGrad	70.43	0.0996529	0.39217	0.0117979	−0.119607	0.10862
		Adam	**62.05**	**0.0247843**	**0.08116**	**0.0010775**	**0.897742**	**0.03283**
		AdaMax	151.92	0.0269593	0.08949	0.001224	0.883845	0.03499
		Nadam	77.96	0.0254886	0.08243	0.0011212	0.893601	0.03348
		RMSProp	62.99	0.0376284	0.12874	0.0021684	0.794225	0.04657
		SGD	151.44	0.0479684	0.17974	0.0030352	0.711965	0.05509
	ReLU	AdaDelta	67.45	0.2712285	0.87701	0.0822644	−6.806816	0.28682
		AdaGrad	72.65	0.1044148	0.42142	0.0137194	−0.301959	0.11713
		Adam	139.2	0.0270935	0.08845	0.0012668	0.879783	0.03559
		AdaMax	70.0	0.0273615	0.09051	0.001263	0.880142	0.03554
		Nadam	71.62	0.0292362	0.09586	0.0015118	0.856534	0.03888
		RMSProp	70.42	0.0294683	0.10088	0.0014157	0.865647	0.03763
		SGD	73.97	0.0933411	0.36469	0.0102868	0.023787	0.10142
	Sigmoid	AdaDelta	62.75	0.4872805	1.64241	0.2470682	−22.446559	0.49706
		AdaGrad	71.42	0.0954021	0.34475	0.0103983	0.013213	0.10197
		Adam	73.09	0.0271899	0.08793	0.0012765	0.878858	0.03573
		AdaMax	70.2	0.0310937	0.10239	0.0017582	0.83315	0.04193
		Nadam	72.24	0.0558406	0.19467	0.0054107	0.486529	0.07356
		RMSProp	68.78	0.0561839	0.1949	0.06764	0.565823	0.06764
		SGD	69.89	0.1710757	0.6951	0.037233	−2.533382	0.19296
	Tanh	AdaDelta	65.84	0.2811222	0.90771	0.0886422	−7.412066	0.29773
		AdaGrad	71.78	0.1212396	0.48275	0.0176672	−0.676603	0.13292
		Adam	77.36	0.0256809	0.08505	0.0011522	0.890653	0.03394
		AdaMax	72.91	0.0269378	0.089	0.0012465	0.881711	0.03531
		Nadam	71.77	0.0255236	0.08348	0.0011487	0.890993	0.03389
		RMSProp	68.89	0.0345188	0.11929	0.001821	0.827186	0.04267
		SGD	71.24	0.0397785	0.13997	0.0024724	0.765371	0.04972

Mandiri. This stock ex perienced a significant decline from May to October, dropping from 1.6 billion to 600 million, after experiencing a price inc increase in April from 1 billion to 1.6 billion. The forecast for this stock also used the ReLU activation function and the Adam optimizer. Turning to Steady Safe, the third stock, there was a significant decline from April to June, dropping from 4.2 billion. to 3.4 billion, and then maintaining a constant price range of 3.4 billion to 4.2 billion from July to October. The prediction for this stock uses the Tanh activation function and the Nadam optimizer. For Samudera Indonesia, the fourth stock, there was a significant drop from 24 billion to 14 billion. Predictions for this stock used the Tanh activation function and the AdaMax optimizer. As for Temas, the fifth stock, there was a decline from April to June, down from 16 billion to 11 billion, followed by a significant increase from July to October, up from 11 billion to 17 billion. Predictions for this stock used the ReLU activation function and the AdaMax optimizer. Finally, the last stock is WEHA Transportasi Indonesia, experiencing a significant decline from April to July, dropping from 4.5 billion to 2 billion, and then maintaining a constant price in the range of 2 billion to 3 billion from July to October. The prediction for this stock uses Linear activation function and Adam's optimizer.

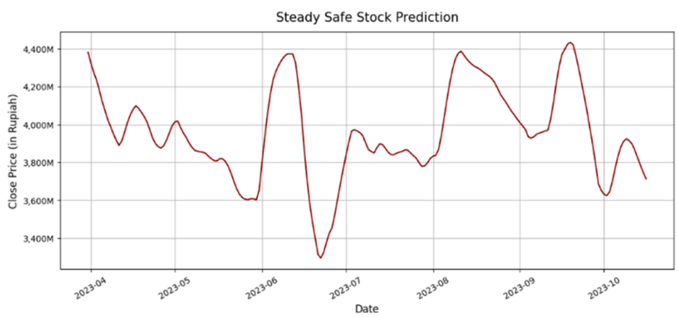
4.1.3. Statistical test to support the evaluation metrics results

This section presents statistical tests using the shapiro-wilk test to support the results of evaluation metrics based on activation functions and optimizers from transportation stock datasets such as Mineral Sumberdaya Mandiri (AKSI.JK), Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK). Based on Table 4.5 and Fig

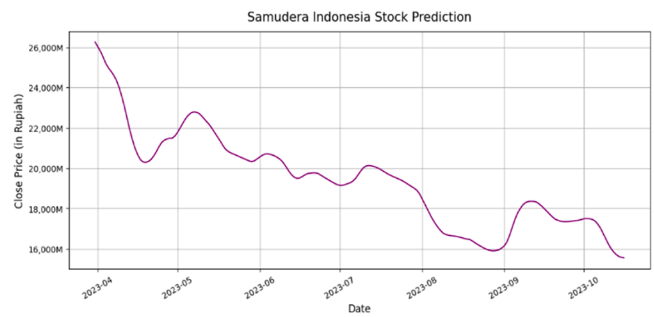


(a) Mineral Sumberdaya Mandiri Stock Price Prediction

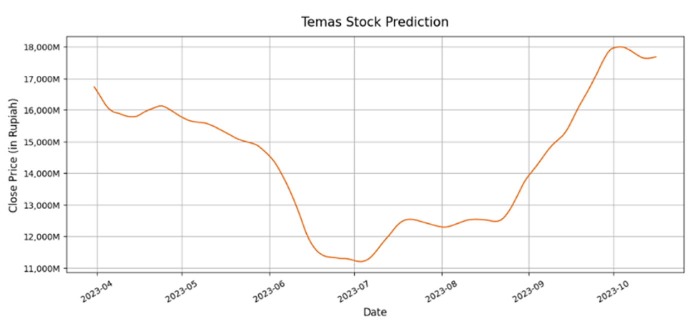
(b) Air Asia Indonesia Stock Price Prediction



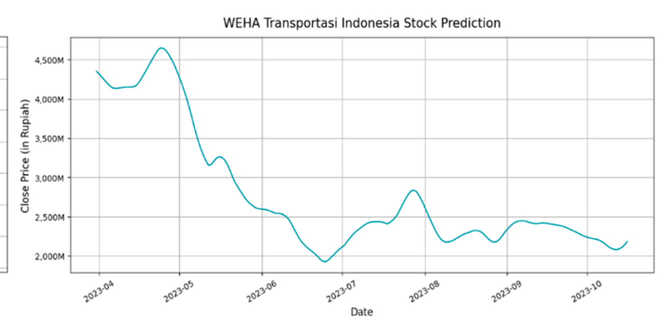
(c) Steady Safe Stock Price Prediction



(d) Samudera Indonesia Stock Price Prediction



(e) Temas Stock Price Prediction



(f) WEHA Transportasi Indonesia Stock Price Prediction

Fig. 4.2. Transportation stock predictions based on activations functions and optimizer.

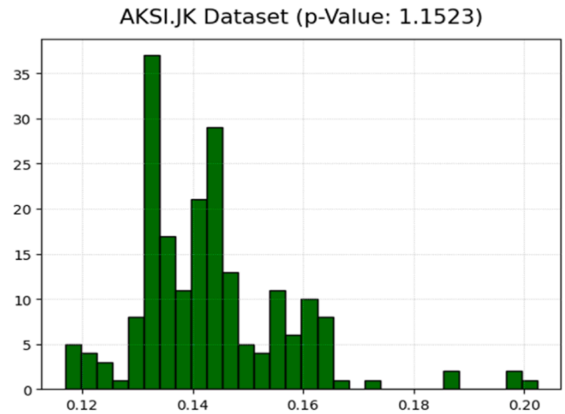
Table 4.5
Shapiro-Wilk test statistics of transportation stock dataset.

Dataset	Activation	Optimizer	MAE (%)	MAPE (%)	MSE (%)	R-Squared (%)	RMSE (%)	T-Statistic	P-Value
AKSI.JK	ReLU	Nadam	0.0088794	0.05601	0.0002039	0.501388	0.01428	0.8974	1.1523
CMPP.JK	ReLU	Adam	**0.0092918**	**0.06422**	**0.0002123**	**0.963999**	**0.01457**	**0.7102**	**4.7160**
SAFE.JK	Tanh	Nadam	0.0158117	0.03984	0.0004659	0.537937	0.02158	0.9761	0.0004
SMDR.JK	Tanh	AdaMax	0.2101362	0.08806	0.0880576	0.510873	0.29675	0.946	1.0190
TMAS.JK	ReLU	AdaMax	0.1484958	0.09002	0.0404941	0.523703	0.20123	0.9372	1.4691
WEHA.JK	Linear	Adam	0.0247843	0.08116	0.0010775	0.897742	0.03283	0.8278	1.5495

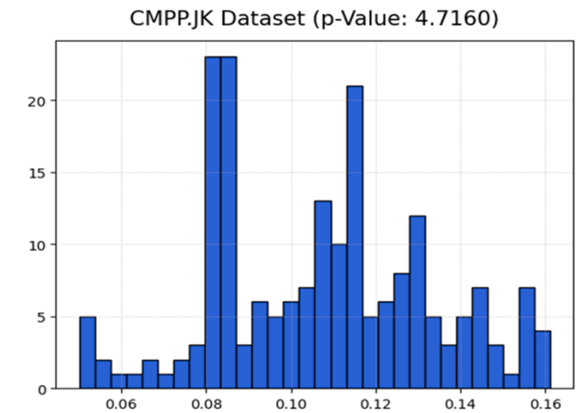
4.3, the transportation stock dataset has several normality distributions from the shapiro-wilk test. The CMPP.JK and SAFE.JK transportation stock datasets have a normally distri buted population compared to other transportation stock datasets such as the AKSI.JK, SMDR.JK, TMAS.JK, and WEHA.JK stock datasets. The normality distribution of the CMPP.JK stock dataset has a higher shapiro-wilk test with a p-value of 4.7160 compared to the SAFE.JK stock dataset in the shapiro-wilk test with a p-value of 0.0004. AKSI.JK, SMDR.JK, TMAS.JK, and WEHA.JK stock datasets have a skewed population distribution. The WEHA.JK stock dataset has a population with a higher skew distribution compared to other transportation stock datasets such as the AKSI.JK, SMDR.JK, and TMAS.JK transportation stock datasets. The WEHA.JK stock dataset has a skew distributed population with a p-value of 1.5495, while AKSI. JK has a p-value of 1.1523, SMDR.JK has a p-value of 1.0190, and TMAS. JK has a p-value of 1.0190, and TMAS.JK has a p-value of 1.4691 from the shapiro-wilk test statistics. There is something interesting about the CMPP.JK transportation stock dataset. This dataset has a higher shapiro-wilk test statistic compared to other MAPE, MSE, RMSE, and R-Squared. Specifically, the CMPP.JK transportation stock dataset has statistics from the shapiro-wilk test that support MAE of 0.0092918, MAPE of 0.06422, MSE of 0.0002123, R-Squared of 96 %, and RMSE of 0.01457.

4.1.4. Model implementation in web-based application

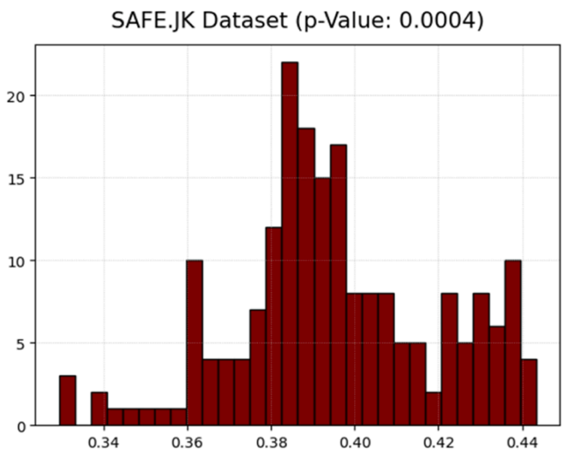
This section presents the implementation of the Long Short-Term Memory (LSTM) algorithm model into a web-based application for the prediction of transportation stock prices such as Mineral Sumberdaya



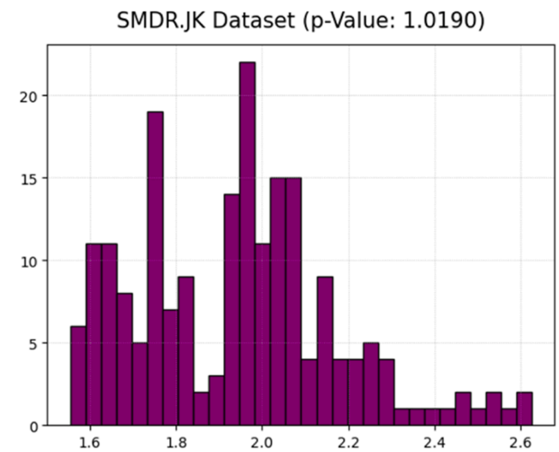
a) Mineral Sumberdaya Mandiri Stock Price Prediction Distribution



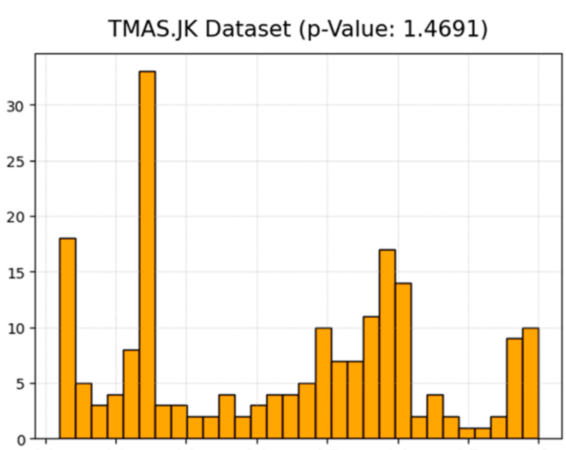
(b) Air Asia Indonesia Stock Price Prediction Distribution



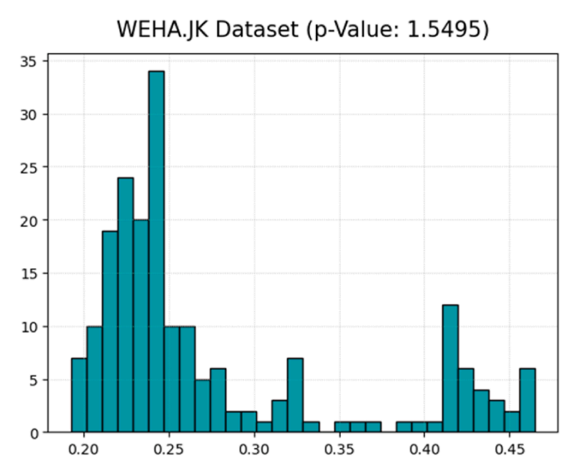
(c) Steady Safe Stock Price Prediction Distribution



(d) Samudera Indonesia Stock Price Prediction Distribution



(e) Temas Stock Price Prediction Distribution



(f) WEHA Transportasi Indonesia Stock Price Prediction Distribution

Fig. 4.3. Shapiro-Wilk Test Statistics of Transportation Stock Datasets.

Mandiri (AKSI.JK), Air Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK) show in Fig 4.4 and Fig 4.5. This prediction uses activation functions such as linear, ReLU, tanh, and sigmoid, and optimizers such as adam, adagrad, nadam, RMSProp, adadelta, SGD, and adamax. The web-based appli cation to generate prediction results based on 4 inputted parameters namely ticker, activation and optimizer, start date, and end date. The ticker parameter is a dataset of transportation stocks such as AKSI.JK, CMPP.JK, SAFE.JK, SMDR.JK, TMAS.JK, and WEHA.JK, while the start date and end date are the prediction start period to the selected prediction end period.

Fig 4.4 above shows the initial view of the web-based application for transportation stocks. Here, each transportation stock is displayed with several sections. There is the transportation stock dataset, the stock price history from April 1, 2011 to April 1, 2023, the current stock price prediction which is the actual price and stock price prediction from the

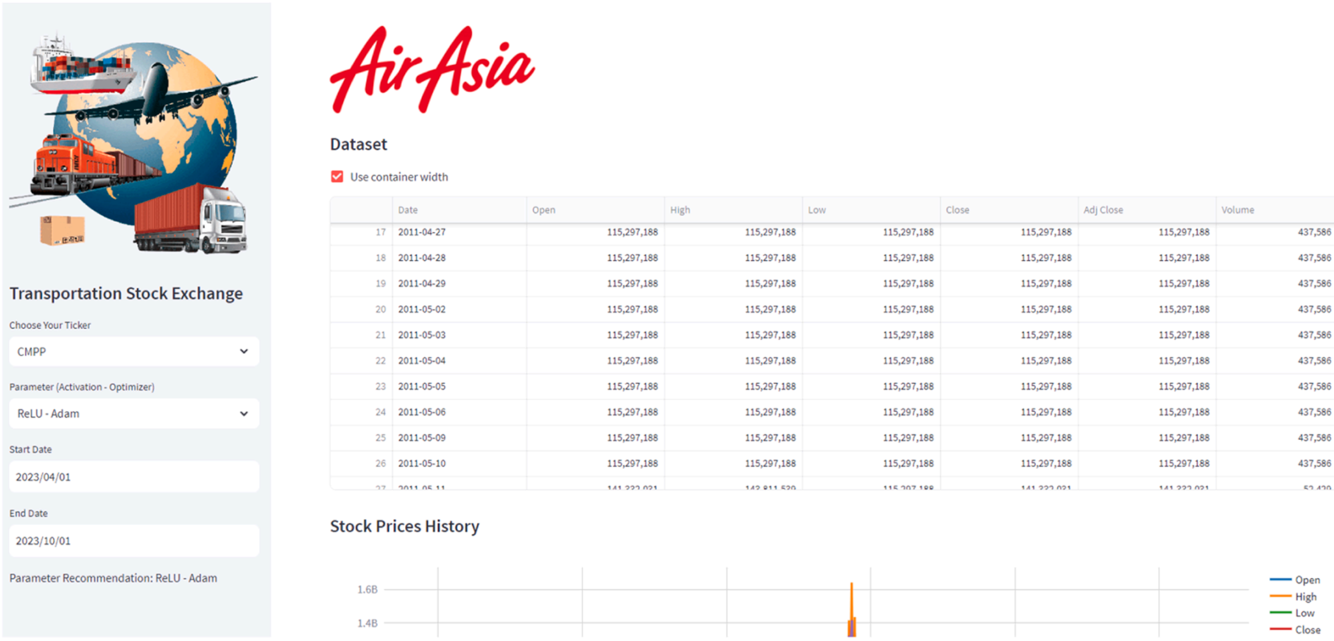


Fig. 4.4. Home screen of the web-based application for Transportation Stock.

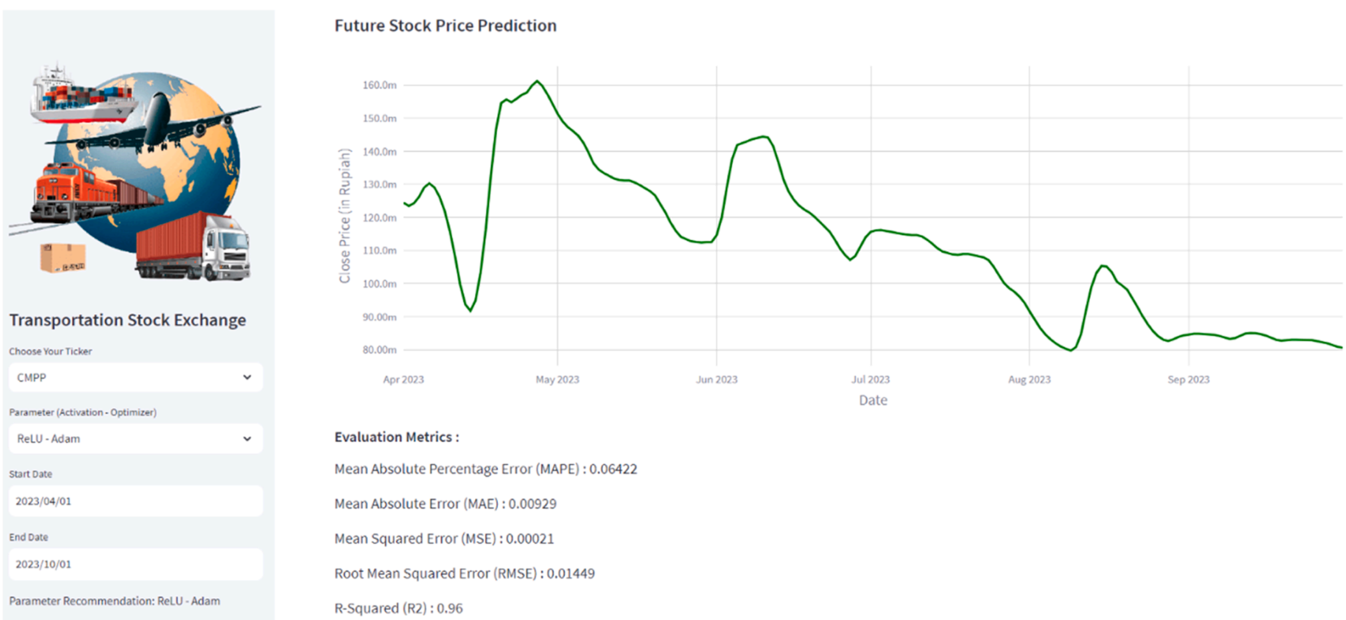


Fig. 4.5. Example result of web-based application for Transportation Stock.

closing price, and the future stock price prediction from April 2023 to August 2023 based on the MAPE, MAE, MSE, RMSE, and R-Squared evaluation metrics. An example of the results from transportation stocks shown in Fig 4.5 above, Air Asia Indonesia (CMPP.JK) transportation stocks have the best performance from MAE of 0.0092918, MAPE of 0.06422, MSE of 0.0002123, R-Squared of 96 %, and RMSE of 0. 01,457 compared to other transportation stock datasets such as Mineral Sumberdaya Mandiri (AKSI.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK). The purpose of implementing the model in a web-based application is to provide information about stock price trends to investors to evaluate several transportation stocks that have decreased or increased in price due to inflation. The web-based application for transportation stock price predictions can be assessed at streamlit web apps (bayunova28-transportation-stock-analysis-app-u94np4.streamlit.app/).

4.2. Discussion

4.2.1. Contributions to literature

This study contributes to stock prediction, particularly in the use of the Long Short-Term Memory (LSTM) algorithm for predicting Indonesian transportation stocks. The results of the study aim to predict the performance of six transportation stocks, specifically Mineral Sumberdaya Mandiri (AKSI.JK), Air Asia Indonesia (CMPP.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK). Performance evaluation using Long Short-Term Memory (LSTM) algorithm with comparison of activation functions such as linear, ReLU, tanh, and sigmoid, as well as optimizers such as adam, adagrad, nadam, RMSProp, adadelta, SGD, and adamax. Besides that, statistical test from shapiro-wilk test is very important to support evaluation metrics based on MAE, MAPE, MSE, RMSE, and R-Squared. Based on several of transportation stock datasets, CMPP.JK stock dataset has best performance from MAE of 0.0092918, MAPE of 0.06422, MSE of 0.0002123, R-Squared of 96 %, and RMSE of 0. 01,457, as well as training speed LSTM model from elapsed time of 104.35 min compared to other transportation stock datasets such as Mineral Sumberdaya Mandiri (AKSI.JK), Steady Safe (SAFE.JK), Samudera Indonesia (SMDR.JK), Temas (TMAS.JK), and WEHA Transportasi Indonesia (WEHA.JK). Besides that, there is something interesting about the CMPP.JK transportation stock dataset. This dataset has a higher shapiro-wilk test statistic compared to other transportation stock datasets with a p-value of 4.7160. This dataset supports evaluating the LSTM model algorithm based on the comparison of activation functions and optimizing hyperparameters, as well as the evaluation metrics of MAE, MAPE, MSE, RMSE, and R-Squared.

Compared to the previous study conducted by Rana et al. (2019), which used Linear Regression (LR), Support Vector Regression (SVR), and Long Short-Term Memory (LSTM) algorithms to predict Acciona stock, our study achieved superior results. The performance evaluation of the previous study showed an RMSE of 0.0151 using Linear and Tanh activations, along with the AdaMax optimizer, showing a higher error rate than this study with an RMSE value of 0.01457. It is important to note that both studies used the same dataset divided into a training set of 90 % and a testing set of 10 %, with batch size of 32, dropout of 0.2, and 10 epochs.

Furthermore, another study aims to predict the National Stock Exchange (NSE) stock index using the Long Short-Term Memory (LSTM) and MultiLayered Perceptron (MLP) algorithms. However, the Mean Squared Error (MSE) evaluation metric obtained in that study, MSE of 0.58, shows a relatively higher error performance compared to this study with an MSE value of 0.0002123 in ReLU's activation function and adam's optimizer with timestamp length used in their research is 60 (Banyal* et al., 2020).

Similarly, a study focused on predicting the stock exchange from Bangladesh, Hongkong, India, Indonesia, Japan, Malaysia, and Thailand used Long Short-Term Memory (LSTM) with several parameters of activation functions such as ELU, ReLU, and tanh along with number of epochs by 50–100. This study has found an accuracy from activation of tanh by 80 % compared to another activation function of ELU and ReLU to predict stock price prediction from several stock exchanges (Sami et al., 2023). These performance results exhibited lower performance of accuracy compared to this study, achieved R-squared value of 96 % with number of epochs by 10 on ReLU activation and adam optimizer.

Lastly, a separate study conducted by Kamalov et al. (2020), predicted banking stock from Bank Rakyat Indonesia (BRI), Bank Nasional Indonesia (BNI), Bank Tabungan Negara (BTN), and Mandiri using Long Short-Term Memory (LSTM) algorithm with several parameter from optimizers such as SGD, adam, and RMSProp along with number of epochs by 25, 50, 75, and 100. This research has a lower Root Mean Squared Error (RMSE) performance of 48.32 from adam optimizer with an accuracy value of 95 % to predict the stock price of Bank Rakyat Indonesia (BRI). The RMSE error performance is still higher than this research which has an RMSE of 0.01457 with ReLU activation and adam optimizer and gets an accuracy value of 96 %.

4.2.2. Practical implications

The practical implications of this study are significant for financial analysts, investors, and corporate stakeholders in the transportation sector. Our results demonstrate that the CMPP.JK stock performs exceptionally well under the LSTM model, with a notably low MAE of 0.0092918 and RMSE of 0.01457. Such predictive accuracy offers a valuable tool for investors looking to make data-driven decisions with higher confidence and reduced risk.

Moreover, the findings provide a framework for further optimization of machine learning models in other sectors. By demonstrating the effectiveness of specific activation functions and optimizers, this study equips industry practitioners with knowledge on how to fine-tune predictive models based on the characteristics of the stock data they are analyzing. This is particularly useful for decision-makers in volatile industries where accurate predictions can significantly impact strategic planning and financial performance.

This study also has implications for improving the efficiency of stock prediction models. The LSTM model used here, with only 10 epochs, achieves comparable or superior accuracy to other models that use 50–100 epochs. This reduction in computational time and resources makes it more feasible to implement machine learning models in real-world applications, especially where rapid decision-making is critical.

4.2.3. Future research directions

There are several avenues for future research based on the findings of this study. First, further exploration of stock prediction across different industries would be valuable. While this study focuses on the transportation sector, similar models could be applied to industries such as finance, energy, or technology. Expanding the application of LSTM to different sectors would provide a broader understanding of its strengths and limitations across various market conditions.

Second, future research could explore the integration of additional data sources, such as macroeconomic indicators or social media sentiment, to enhance the predictive power of LSTM models. The inclusion of real-time data streams or event-based triggers could further improve the model's ability to respond to market fluctuations in a timely manner.

Additionally, alternative machine learning approaches, such as hybrid models combining LSTM with attention mechanisms or ensemble methods like XGBoost, could be investigated for potential performance improvements. Advanced hyperparameter optimization techniques, such as Bayesian optimization or genetic algorithms, may also offer further enhancements to the model's performance.

Finally, real-world implementation challenges, such as computational efficiency and scalability, should be addressed. Future research could investigate lightweight versions of LSTM models or parallel processing techniques to improve model training times, making them more suitable for real-time stock prediction in high-frequency trading environments.

5. Conclusion

In this study, the prediction of transportation stock prices using the Long Short-Term Memory (LSTM) algorithm model by comparing several activations such as linear, ReLU, tanh, and sigmoid, as well as optimizers such as adam, adagrad, nadam, RMSProp, adadelta, SGD, and adamax. This study shows a decrease in the stock prices of Air Asia Indonesia (CMPP.JK), Samudera Indonesia (SMDR.JK), and WEHA Transportation Indonesia (WEHA.JK). However, the stock prices of Mineral Sumberdaya Mandiri (AKSI.JK) and Steady Safe (SAFE.JK) showed stable movements. In addition, there was a significant increase in the stock price of Temas (TMAS.JK) from April to October 2023. Besides that, ReLU activation with Adam optimizer is best performance to predict Air Asia Indonesia (CMPP.JK) transportation stock dataset from MAE of 0.0092918, MAPE of 0.06422, MSE of 0.00021230, R-Squared of 96 %, RMSE of 0.01457 and shapiro-wilk test on t-statistic of 0.7102 and p-value of 4.716007 with elapsed time of 104.35 min. This performance compared to another activation functions and optimizers from each transportation stock datasets such as AKSI.JK, SAFE.JK, SMDR.JK, TMAS.JK, and WEHA.JK.

However, it is important to note several limitations in this study. The data used consists of transportation stocks from April 1, 2011, to April 1, 2023. The transportation stocks included in the study are Mineral Sumberdaya Mandiri, Air Asia Indonesia, Steady Safe, Samudera Indonesia, Temas, and WEHA Transportasi Indonesia. The study utilized the Long Short-Term Memory (LSTM) algorithm with various activations such as linear, ReLU, tanh, and sigmoid, as well as optimizers such as Adam, AdaGrad, Nadam, RMSProp, AdaDelta, SGD, and AdaMax. In addition, the elapsed time process is very long when training the Long Short-Term Memory (LSTM) algorithm because the Graphics Processing Unit (GPU) is not used to speed up the training time of the Long Short-Term Memory (LSTM) algorithm.

For future research, we encourage researchers to explore modeling approaches with time series algorithms or other deep learning techniques such as Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), and Convolutional Neural Net-work (CNN) along with meta-heuristic optimization algorithms to speed up the training process and determine the best model for each model. In addition, they can compare different types of LSTM algorithms such as Vanilla, Stacked, and Bidirectional LSTM. Expanding the dataset to include stocks from sectors such as banking, technology, food and beverages, or other sectors, with longer time periods and more extensive data, will help avoid overfitting during the modeling process. Additionally, comparing the number of epochs and batch size to evaluate metrics such as Mean Forecast Error (MFE) and Mean Absolute Scaled Error (MASE), among other relevant metrics, can be included in the modeling process. In addition, researchers can explore the distribution of training and testing samples ranging from 70:30, 80:20, 90:10, or all of them as 70:30 and 80:20 performed poorly in this study.

Source of support

Any grants / equipment / drugs, and/ or other support that facilitated the conduct of research / writing of the manuscript (including AFMRC project details, if applicable)

CRediT authorship contribution statement

Dinar Ajeng Kristiyanti: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Willibrordus Bayu Nova Pramudya:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Samuel Ady Sanjaya:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article

Acknowledgements

(If any): We would like to acknowledge Universitas Multimedia Nusantara for providing support in this study.

References

Abdulkadirov, R., Lyakhov, P., & Nagornov, N. (2023). Survey of optimization algorithms in modern neural networks. *Mathematics, 11*(11), 1–37. https://doi.org/10.3390/math11112466

AL-Zwainy, F. M. S., & Aidan, I. A. A. (2017). Forecasting the cost of structure of infrastructure projects utilizing artificial neural network model (highway projects as case study). *Indian Journal of Science and Technology, 10*(20), 1–12. https://doi.org/10.17485/ijst/2017/v10i20/108567

Albahli, S., Nazir, T., Mehmood, A., Irtaza, A., Alkhalifah, A., & Albattah, W. (2022). AEI-DNET: A novel DenseNet model with an autoencoder for the stock market predictions using stock technical indicators. *Electronics (Switzerland), 11*(4). https://doi.org/10.3390/electronics11040611

Ali, M. M., Babar, M. I., Hamza, M., Jehanzeb, M., Habib, S., & Khan, M. S. (2019). Industrial financial forecasting using long short-term memory recurrent neural networks. *International Journal of Advanced Computer Science and Applications, 10*(4), 88–99. https://doi.org/10.14569/ijacsa.2019.0100410

Ballinari, D., Audrino, F., & Sigrist, F. (2022). When does attention matter? The effect of investor attention on stock market volatility around news releases. *International Review of Financial Analysis, 82*, Article 102185. https://doi.org/10.1016/j.irfa.2022.102185 (April).

Banyal, S., Goel, P., & Grover, D. (2020). Indian stock-market prediction using stacked LSTM and multi-layered perceptron. *International Journal of Innovative Technology and Exploring Engineering, 9*(3), 1051–1055. https://doi.org/10.35940/ijitee.c8026.019320

BPS Statistics, I. (2024). The year-on-year (y-on-y) inflation in October 2022 was 5.71 percent. The highest inflation saw in Tanjung Selor at 9.11 percent. https://www.bps.go.id/en/pressrelease/2022/11/01/1866/the-year-on-year–y-on-y–inflation-in-october-2022-was-5-71-percent–the-highest-inflation-saw-in-tanjung-selor-at-9-11-percent-.html.

Brito da Silva, L. E., Elnabarawy, I., & Wunsch, D. C. (2019). A survey of adaptive resonance theory neural network models for engineering applications. *Neural Networks, 120*, 167–203. https://doi.org/10.1016/j.neunet.2019.09.012 (xxxx).

Chen, W., Qu, S., Jiang, M., & Jiang, C. (2021). The construction of multilayer stock network model. *Physica A: Statistical Mechanics and Its Applications, 565*, Article 125608. https://doi.org/10.1016/j.physa.2020.125608

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, 1–24. https://doi.org/10.7717/PEERJ-CS.623

Chikwira, C. (2023). The impact of the stock market on liquidity and economic growth : Evidence of volatile market. *Economies, 11*(6), 155. https://doi.org/10.3390/economies11060155

Chopra, R., & Sharma, G. D. (2021). Application of artificial intelligence in stock market forecasting: A critique, review, and research agenda. *Journal of Risk and Financial Management, 14*(11). https://doi.org/10.3390/jrfm14110526

Deepa, B., & Ramesh, K. (2022). Epileptic seizure detection using deep learning through min max scaler normalization. *International Journal of Health Sciences, 6*, 10981–10996. https://doi.org/10.53730/ijhs.v6ns1.7801 (April).

Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing, 503*, 92–108. https://doi.org/10.1016/j.neucom.2022.06.111

Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., & Tabona, O. (2021). A survey on missing data in machine learning. *Journal of Big Data, 8*(1). https://doi.org/10.1186/s40537-021-00516-9

Fathali, Z., Kodia, Z., & Ben Said, L. (2022). Stock market prediction of NIFTY 50 index applying machine learning techniques. *Applied Artificial Intelligence, 36*(1). https://doi.org/10.1080/08839514.2022.2111134

Feng, F., He, X., Wang, X., Luo, C., Liu, Y., & Chua, T. S. (2019). Temporal relational ranking for stock prediction. *ACM Transactions on Information Systems, 37*(2), 1–27. https://doi.org/10.1145/3309547

Filho, H. C., de, C., Júnior, O. A., de, C., de Carvalho, O. L. F., de Bem, P. P., de Moura, R., dos, S., de Albuquerque, A. O., Silva, C. R., Ferreira, P. H. G., Guimarães, R. F., & Gomes, R. A. T. (2020). Rice crop detection using LSTM, Bi-LSTM, and machine learning models from sentinel-1 time series. *Remote Sensing, 12*(16), 1–25. https://doi.org/10.3390/RS12162655

Fletcher Mercaldo, S., & Blume, J. D. (2020). Missing data and prediction: The pattern submodel. *Biostatistics, 21*(2), 236–252. https://doi.org/10.1093/biostatistics/kxy040

Fofana, T., Ouattara, S., & Clement, A. (2021). Optimal flame detection of fires in videos based on deep learning and the use of various optimizers. *Open Journal of Applied Sciences, 11*(11), 1240–1255. https://doi.org/10.4236/ojapps.2021.1111094

Gurav, U., & Kotrappa, S. (2020). LBL - LSTM : Log bilinear and long short term memory based efficient stock forecasting model considering external fluctuating factor UMA. *International Journal of Engineering and Advanced Technology, 9*(4), 2057–2063. https://doi.org/10.35940/ijeat.d8680.049420

Hadi, A. (2023). Inflation in Indonesia brought under control faster than expected. *The Jakarta Post*, 5–9. https://asianews.network/inflation-in-indonesia-brought-under-control-faster-than-expected/.

Hasib, K. M., Azam, S., Karim, A., Marouf, A. Al, Shamrat, F. M. J. M., Montaha, S., Yeo, K. C., Jonkman, M., Alhajj, R., & Rokne, J. G. (2023). MCNN-LSTM: Combining CNN and LSTM to classify multi-class text in imbalanced news data. *IEEE Access, 11*, 93048–93063. https://doi.org/10.1109/ACCESS.2023.3309697 (August).

Hasib, K. M., Tanzim, A., Shin, J., Faruk, K. O., Mahmud, J. Al, & Mridha, M. F. (2022). BMNet-5: A novel approach of neural network to classify the genre of Bengali music based on audio features. *IEEE Access, 10*, 108545–108563. https://doi.org/10.1109/ACCESS.2022.3213818 (October).

Hasib, K. M., Towhid, N. A., Faruk, K. O., Al Mahmud, J., & Mridha, M. F. (2023). Strategies for enhancing the performance of news article classification in Bangla: Handling imbalance and interpretation. *Engineering Applications of Artificial Intelligence, 125*. https://doi.org/10.1016/j.engappai.2023.106688 (C).

Hassan, E., Shams, M. Y., Hikal, N. A., & Elmougy, S. (2023). The effect of choosing optimizer algorithms to improve computer vision tasks: A comparative study. *Multimedia Tools and Applications, 82*(11), 16591–16633. https://doi.org/10.1007/s11042-022-13820-0

Hayder, I. M., Al-Amiedy, T. A., Ghaban, W., Saeed, F., Nasser, M., Al-Ali, G. A., & Younis, H. A. (2023). An intelligent early flood forecasting and prediction leveraging machine and deep learning algorithms with advanced alert system. *Processes, 11*(2). https://doi.org/10.3390/pr11020481

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly: Management Information Systems, 28*(1), 75–105. https://doi.org/10.2307/25148625

Huang, S., Shen, J., Lv, Q., Zhou, Q., & Yong, B. (2023). A novel NODE approach combined with LSTM for short-term electricity load forecasting. *Future Internet, 15* (1), 1–20. https://doi.org/10.3390/fi15010022
Ilboudo, W. E. L., Kobayashi, T., & Sugimoto, K. (2022). Robust stochastic gradient descent with student-t distribution based first-order momentum. *IEEE Transactions on Neural Networks and Learning Systems, 33*(3), 1324–1337. https://doi.org/10.1109/TNNLS.2020.3041755
Jadhav, A., Pramod, D., & Ramanathan, K. (2019). Comparison of performance of data imputation methods for numeric dataset. *Applied Artificial Intelligence, 33*(10), 913–933. https://doi.org/10.1080/08839514.2019.1637138
Jarne, A., & Shveda, K. (2023). -If you think inflation in the US is high, look at the numbers in these countries. *CNN*, 1–10. https://ktvz.com/money/cnn-business-consumer/2023/07/07/if-you-think-inflation-in-the-us-is-high-look-at-the-numbers-in-these-countries/.
Kamalov, F., Smail, L., & Gurrib, I. (2020). Stock price forecast with deep learning. Proceedings of the 2020 International Conference on Decision Aid Sciences and Application (DASA), 1098–1102. https://doi.org/10.1109/DASA51403.2020.9317260.
Kamienski, C., Soininen, J., Taumberger, M., Dantas, R., Toscano, A., Cinotti, T. S., Maia, R. F., & Neto, T. (2019). Smart Water Management Platform : IoT-Based Precision Irrigation for Agriculture †. https://doi.org/10.3390/s19020276.
Kaur, J., & Dharni, K. (2022). Application and performance of data mining techniques in stock market: A review. Intelligent systems in accounting. *Finance and Management, 29*(4), 219–241. https://doi.org/10.1002/isaf.1518
Khan, P. W., Byun, Y. C., Lee, S. J., & Park, N. (2020). Machine learning based hybrid system for imputation and efficient energy demand forecasting. *Energies, 13*(11). https://doi.org/10.3390/en13112681
Khan, Q., Hayder, G., & Al-Zwainy, F. M. S. (2023). River water suspended sediment predictive analytics using artificial neural network and convolutional neural network approach: A review. *Advances in Science, Technology and Innovation*, 51–56. https://doi.org/10.1007/978-3-031-26580-8_10. May.
Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences, 22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018
Kristiyanti, D. A., & Hardani, S. (2023). Sentiment analysis of public acceptance of covid-19 vaccines types in indonesia using naïve bayes, support vector machine, and Long Short-Term Memory (LSTM). *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 7*(3), 722–732. https://doi.org/10.29207/RESTI.V7I3.4737
Lee, S., & Kim, T. (2023). Impact of deep learning optimizers and hyperparameter tuning on the performance of bearing fault diagnosis. *IEEE Access, 11*, 55046–55070. https://doi.org/10.1109/ACCESS.2023.3281910. June.
Limanseto, Haryo. (2022). Prices After the Fuel Price Adjustment Realization of Indonesian Inflation in November Declined, Supported by Stable Food. https://ekon.go.id/publikasi/detail/5031/realization-of-indonesian-inflation-in-november-declined-supported-by-stable-food-prices-after-the-fuel-price-adjustment.
Lv, J., Wang, C., Gao, W., & Zhao, Q. (2021). An economic forecasting method based on the LightGBM-optimized LSTM and time-series model. *Computational Intelligence and Neuroscience, 2021*. https://doi.org/10.1155/2021/8128879
Lyu, Z., Yu, Y., Samali, B., Rashidi, M., Mohammadi, M., Nguyen, T. N., & Nguyen, A. (2022). Back-Propagation neural network optimized by K-fold cross-validation for prediction of torsional strength of reinforced concrete beam. *Materials, 15*(4). https://doi.org/10.3390/ma15041477
Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia, 22*(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18
Mohd Faizal, A. S., Hon, W. Y., Thevarajah, T. M., Khor, S. M., & Chang, S. W. (2023). A biomarker discovery of acute myocardial infarction using feature selection and machine learning. *Medical and Biological Engineering and Computing, 2021*. https://doi.org/10.1007/s11517-023-02841-y
Murugesan, M., & Jung, D. W. (2021). Formability and failure evaluation of AA3003-H18 sheets in single-point incremental forming process through the design of experiments. *Materials, 14*(4), 1–26. https://doi.org/10.3390/ma14040808
Mutmainah, Marfuah, U., Nopianti, R., & Tri Panudju, A. (2022). LSTM algorithm analysis of banking sector stock price predictions. *International Journal of Advanced Research, 10*(01), 627–634. https://doi.org/10.21474/IJAR01/14082
Nan, H. (2023). Apply RF-LSTM to predicting future share price. Proceedings of the 2023 International Conference on Digital Economy and Management Science, 170, 0–3. https://doi.org/10.1051/shsconf/202317002012.
Olanipekun, A. T., Mashinini, P. M., Owojaiye, O. A., & Maledi, N. B. (2022). Applying a neural network-based machine learning to laser-welded spark plasma sintered steel: Predicting vickers micro-hardness. *Journal of Manufacturing and Materials Processing, 6*(5). https://doi.org/10.3390/jmmp6050091
Pires, I. M., Hussain, F., Garcia, N. M., Lameski, P., & Zdravevski, E. (2020). Homogeneous data normalization and deep learning: A case study in human activity classification. *Future Internet, 12*(11), 1–14. https://doi.org/10.3390/fi12110194
Poornima, S., & Pushpalatha, M. (2019). Prediction of rainfall using intensified LSTM based recurrent neural network with weighted linear units. *Atmosphere, 10*(11). https://doi.org/10.3390/atmos10110668
Purwono, R., Yasin, M. Z., & Mubin, M. K. (2020). Explaining regional inflation programmes in Indonesia: Does inflation rate converge? *Economic Change and Restructuring, 53*(4), 571–590. https://doi.org/10.1007/s10644-020-09264-x
Qiu, M., & Song, Y. (2016). Predicting the direction of stock market index movement using an optimized artificial neural network model. *PLoS ONE, 11*(5), 1–11. https://doi.org/10.1371/journal.pone.0155133
Rabinovich, J. (2023). Tangible and intangible investments and sales growth of US firms. *Structural Change and Economic Dynamics, 66*, 200–212. https://doi.org/10.1016/j.strueco.2023.05.001. September 2022.
Rana, M., Uddin, M. M., & Hoque, M. M. (2019). Effects of activation functions and optimizers on stock price prediction using LSTM recurrent networks. Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence, 12, 354–358. https://doi.org/10.1145/3374587.3374622.
Rasheed, R. T., Mohammed, M. A., & Tapus, N. (2021). Big data analysis. *Mesopotamian Journal of Big Data, 2021*, 485–501. https://doi.org/10.4324/9781003139850-37
Risan, H. K., Serhan, F. M., & Al-Azzawi, A. A. (2024). Management of a typical experiment in engineering and science. Proceedings of the AIP Conference Proceedings. https://doi.org/10.1063/5.0186079.
Ritika, Himanshu, & Kishor, N. (2022). Modeling of factors affecting investment behavior during the pandemic: A grey-DEMATEL approach. *Journal of Financial Services Marketing, 28*(2), 222–235. https://doi.org/10.1057/s41264-022-00141-4
Sami, H. M., Ahshan, K. A., & Rozario, P. N. (2023). Determining the best activation functions for predicting stock prices in different (stock exchanges) through multivariable time series forecasting of LSTM. *Australian Journal of Engineering and Innovative Technology, 7804*, 63–71. https://doi.org/10.34104/ajeit.023.063071
Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly: Management Information Systems, 35*(1), 37–56. https://doi.org/10.2307/23043488
Selvin, S., & R, V., E.A, G., Menon, V. K. (2017).& K.P, S. Stock price prediction using LSTM, RNN, and CNN-sliding window model. Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 132(December), 1643–1647. https://doi.org/10.1016/j.procs.2018.05.050.
Sen, S., & Raghunathan, A. (2018). Approximate computing for Long Short Term Memory (LSTM) neural networks. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 37*(11), 2266–2276. https://doi.org/10.1109/TCAD.2018.2858362
Sharma, S., Sharma, S., & Athaiya, A. (2020). Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology, 04*(12), 310–316. https://doi.org/10.33564/ijeast.2020.v04i12.054
Thampanya, N., Wu, J., Nasir, M. A., & Liu, J. (2020). Fundamental and behavioural determinants of stock return volatility in ASEAN-5 countries. *Journal of International Financial Markets, Institutions and Money, 65*, Article 101193. https://doi.org/10.1016/j.intfin.2020.101193
Tokgoz, E., Musafer, H., Faezipour, M., & Mahmood, A. (2023). Incorporating derivative-free convexity with trigonometric simplex designs for learning-rate estimation of stochastic gradient-descent method. *Electronics, 12*(2). https://doi.org/10.3390/electronics12020419
Venugopal, P., & Vigneswaran, T. (2019). State-of-Health estimation of Li-ion batteries in electric vehicle using IndRNN under variable load condition. *Energies, 12*(22). https://doi.org/10.3390/en12224338
Weiss, R., Karimijafarbigloo, S., Roggenbuck, D., & Rödiger, S. (2022). Applications of neural networks in biomedical data analysis. *Biomedicines, 10*(7), 1–30. https://doi.org/10.3390/biomedicines10071469
Yaqub, M., Jinchao, F., Zia, M. S., Arshid, K., Jia, K., Rehman, Z. U., & Mehmood, A. (2020). State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images. *Brain Sciences, 10*(7), 1–19. https://doi.org/10.3390/brainsci10070427
Zaheer, S., Anjum, N., Hussain, S., Algarni, A. D., Iqbal, J., Bourouis, S., & Ullah, S. S. (2023). A multi parameter forecasting for stock time series data using LSTM and deep learning model. *Mathematics, 11*(3), 1–24. https://doi.org/10.3390/math11030590
Zhao, M., Zhong, S., Fu, X., Tang, B., Dong, S., & Pecht, M. (2021). Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis. *IEEE Transactions on Industrial Electronics, 68*(3), 2587–2597. https://doi.org/10.1109/TIE.2020.2972458