Available online at http://www.iorajournal.org/index.php/ijgor/index

International Journal of Global Operations Research

e-ISSN: 2722-1016
p-ISSN: 2723-1739

Vol. 6, No. 3, pp. 220-226, 2025

Comparative Analysis of Activation Functions in LSTM Models for Predicting Bank BNI Stock Prices

Astrid Sulistya Azahra[1*], Moch Panji Agung Saputra[2], Rizki Apriva Hidayana[3]

[1,3]Doctoral of Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, Indonesia

[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, Indonesia

*Corresponding author email: astrid23002@mail.unpad.ac.id

Abstract

The Indonesian capital market has experienced rapid development in the last two decades, with the banking sector as one of the main drivers. Stock price prediction is a crucial aspect for investors and market players to minimize risk and optimize investment strategies. Price fluctuations influenced by fundamental factors, market sentiment, and external conditions make prediction a complex challenge. This study aims to compare the performance of four activation functions: Rectified Linear Unit (ReLU), hyperbolic tangent (Tanh), Sigmoid, and Softplus, in the Long Short-Term Memory (LSTM) model in predicting the stock price of Bank Negara Indonesia (BNI). The method used is a quantitative approach with experiments, using historical data of BNI's closing stock prices for the period May 1, 2020, to April 30, 2025, obtained from Yahoo Finance. The data is processed through cleaning, normalization, transformation into a supervised learning format, and division into training data (80%) and test data (20%). Performance evaluation is carried out using RMSE, MAE, MAPE, and R^2 metrics. The results showed that the Softplus activation function produced the best performance with RMSE 128.714, MAE 101.815, MAPE 2.358%, and R^2 0.924, followed by ReLU which had competitive performance and more efficient training time. The Tanh activation function was in the middle position, while Sigmoid showed the lowest performance. These findings indicate that Softplus and ReLU are optimal choices for BNI stock price prediction using LSTM, with Softplus excelling in accuracy and ReLU providing a balance between performance and efficiency.

Keywords: Stock prediction, BNI bank, activation function, LSTM, deep learning.

1. Introduction

Indonesia's capital markets have experienced rapid development over the past two decades, driven by increased participation from both domestic and foreign investors. As a key pillar of the national economy, the capital market plays a crucial role in providing companies with access to funding and providing an investment platform for the public (Syazali et al., 2024). Among the various sectors listed on the Indonesia Stock Exchange (IDX), the banking sector plays a particularly dominant role, driving economic activity through its financial intermediation, fundraising, and credit distribution functions (Felle & Santioso 2024).

Stock price prediction analysis is crucial for investors, traders, and financial managers in formulating appropriate investment strategies. Information about potential future price movements allows market participants to minimize risk, optimize portfolio allocation, and maximize profit opportunities (Sebastian & Tantita, 2024). For long-term investors, stock price predictions can help determine the right moment to buy or sell assets, while for traders, this analysis provides the basis for quick and accurate short-term decision-making (Shaban et al., 2024).

Stock price movements are determined not only by a company's fundamental performance but also by market sentiment, global economic conditions, government policies, exchange rate fluctuations, and political and social events (Raza et al., 2023). This combination of factors creates high uncertainty, making price movement patterns often difficult to accurately predict. The rapidly changing nature of the market requires predictive models to adapt to the latest data and capture non-linear relationships (Darwish et al., 2025). This requires the use of more sophisticated analytical approaches than traditional statistical methods, including the use of machine learning and deep learning models to improve prediction accuracy.

The use of deep learning methods to model time series data has been widely used by previous researchers. One prominent architecture is Long Short-Term Memory (LSTM), a variant of Recurrent Neural Network (RNN) designed to address the vanishing gradient problem and capture long-term dependency patterns in historical data (Malashin et al., 2024). LSTM's superiority in processing sequential data makes it widely used in stock price prediction. Selecting the right activation function can be a determining factor in the model's success in generating accurate predictions on volatile stock market data.

Common activation functions used in LSTM modeling include Rectified Linear Unit (ReLU), hyperbolic tangent (tanh), sigmoid, and softplus, each with distinct advantages in managing the flow of information within a neural network. The primary challenge lies in determining which activation function is most suitable for predicting banking stock price movements, given the dynamic nature of market data and its sensitivity to changes in external factors. The goal of this approach is to find the activation function that produces the most accurate predictions, thus supporting more informed, empirically-based investment decision-making.

2. Methodology

2.2. Research Design

This study uses a quantitative approach with experimental methods to test and compare the performance of various activation functions in the Long Short-Term Memory (LSTM) model in predicting stock price movements of Bank Negara Indonesia (BNI). The approach used is a deep learning model focused on time series forecasting. The experimental process involves designing an LSTM model using various activation functions, then evaluating the prediction performance based on relevant accuracy and error metrics to determine the most appropriate activation function for the characteristics of banking stock data.

2.2 Data Collection

Data collection in this study was conducted using data mining techniques using the yfinance library to obtain historical stock price data from the Yahoo Finance platform. The data used covered the period from May 1, 2020, to April 30, 2025, with a total of 1,200 daily data points. The closing price (close) was chosen as the target variable because it represents the final trading value, which is often used as a reference in stock price analysis and forecasting. The closing price of the stock will be shown in figure 1.

Figure 1: BNI stock closing price chart

2.3 Data Preprocessing

Preprocessing is carried out to ensure the data is ready for use in LSTM model training. The steps include:

1) Data Cleaning
   Removing missing values from data to maintain data quality and prevent biased predictions from the model.

2) Data Normalization
   Using StandardScaler to standardize all features to the same scale, thereby speeding up the training process and improving model performance (Raju et al., 2020).

3) Data Transformation
   Transforming time series data into a supervised learning format by adding a lag variable as input to the LSTM model.
4) Data Splitting
   Splitting the data into a training set (80%) and a testing set (20%) of the total data.

2.4 Model Development

Long Short-Term Memory (LSTM) is a Recurrent Neural Network (RNN)-based artificial neural network architecture designed to learn long-term dependencies in sequential data. LSTM uses a gating mechanism that allows important information to be retained over long periods of time while discarding irrelevant information (Le et al., 2019). The main structure of an LSTM consists of an Input Gate, a Forget Gate, a Cell State, a New Candidate, an Output Gate, and a Hidden State. Its mathematical formulation is as follows:

1) Input Gate

$$i_t = \sigma\ (W_i \times X_i + U_i \times h_{t-1} + b_i \tag{1}$$

2) Forgate Gate

$$f_t = \sigma\ (W_f \times X_t + U_f \times h_{t-1} + b_f \tag{2}$$

3) Cell State

$$c_t = f_t \odot\ c_{t-1} + i_t + \odot\ \acute{c}_t \tag{3}$$

4) New Candidate

$$\acute{c}_t = tanh(W_c \times X_c + U_c \times h_{t-1} + b_c \tag{4}$$

5) Output gate

$$o_t = \sigma(W_0 \times X_i + U_o \times h_{t-1} + b_o \tag{5}$$

6) Hiddent State

$$h_t = o_t \odot\ \sigma(c_t) \tag{6}$$

where

X_t is the input at time t
h_{t-1} is the previous hidden state
c_t is the cell state at time t
σ is the sigmoid activation function
$\odot$ is the element-by-element multiplication (Hadamard product)
W and U are weight matrices, b is the bias vector

The LSTM model was trained using the Adam optimizer with 100 epochs. To prevent overfitting, early stopping was used with validation loss monitoring. Training was performed using a loop to compare the performance of four activation functions:

1. Sigmoid converts input values into probabilities between 0 and 1. Large positive values approach 1, large negative values approach 0 (Montesinos López et al., 2022).

$$f(x) = \frac{1}{1 + e^{-1}} \tag{7}$$

2. Tanh is similar to Sigmoid but is centered at 0, making positive and negative outputs balanced.

$$f(x) = \frac{1 - e^{2x}}{1 + e^{2x}} \tag{8}$$

3. Relu works If the input is positive, the function returns its value immediately; if negative, it returns 0.

$$f(x) = \max{(0, x)} \tag{9}$$

4. Softplus is a smoother version of ReLU. Instead of clipping negatives to zero, it provides a smooth transition from negative to positive values (Hammad, 2024).

$$f(x) = In\,(1 + e^{x}) \tag{10}$$

2.5 Model Evaluation

Model evaluation aims to measure the accuracy and reliability of prediction results relative to actual data. In this study, model performance was measured using four evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2).

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |f_t - y_t| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \tilde{y}_i)^2} \tag{12}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \tilde{y}_i}{y_i}\right| \times 100\% \tag{13}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{14}$$

3. Results and Discussion

3.1. Evaluation of activation function performance

Table 1: Activation Function Performance

Activation	RMSE	MAE	MAPE (%)	R^2	Train Time (ms)	Pred Time (ms)
relu	128.519	103.363	2.407	0.924	105,496,150	438,836
tanh	129.692	102.146	2.387	0.923	104,498,321	713,850
sigmoid	147.220	117.302	2.739	0.900	46,113,803	686,981
softplus	128.714	101.815	2.358	0.924	113,633,627	429,862

The results in table 1 show that the softplus activation function produces the best overall prediction performance with an RMSE of 128.714, MAE of 101.815, and MAPE of 2.358%, accompanied by an R^2 value of 0.924, which indicates the model's ability to explain variations in actual data is quite high. ReLU activation also shows competitive performance with an RMSE of 128.519, MAE of 103.363, MAPE of 2.407%, and R^2 of 0.924, although it is slightly inferior in absolute error compared to softplus. Meanwhile, tanh activation produces an RMSE of 129.692, MAE of 102.146, and MAPE of 2.387% with an R^2 of 0.923, which is in the middle of ReLU and softplus in terms of performance.

The sigmoid activation function showed the lowest performance with the highest RMSE of 147.220, MAE of 117.302, MAPE of 2.739%, and R^2 of 0.900, indicating a relatively lower accuracy level compared to other activation functions. In terms of training time, sigmoid required the shortest time (46,113 ms) while softplus required the longest time (113,634 ms). For prediction time, softplus also excelled with the fastest prediction (429,862 ms) compared to tanh which required the longest prediction time (713,850 ms). These findings indicate that softplus and ReLU are more optimal choices for BNI stock price prediction using LSTM in the context of this study.

3.2. Comparison of activation functions during training

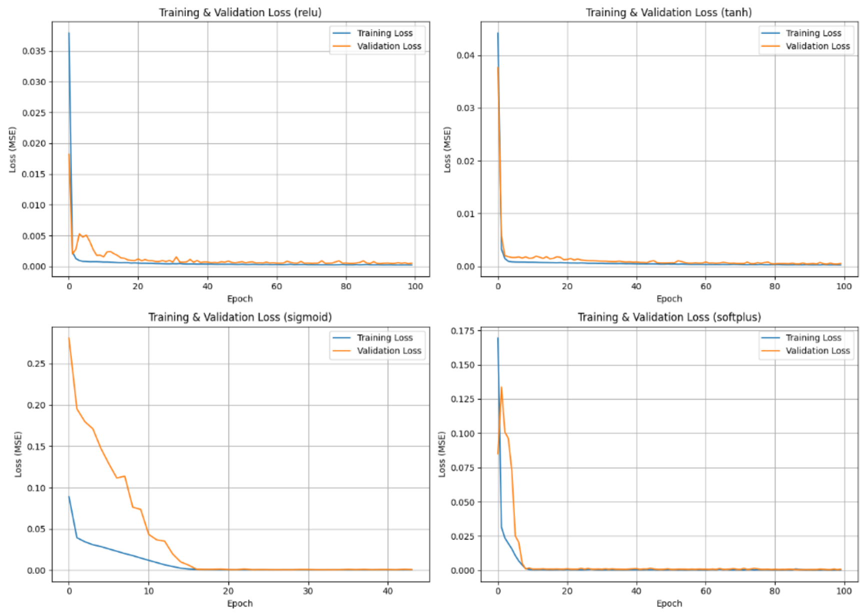

Figure 2: Comparison graph during the training process

Loss decreases rapidly in the ReLU and tanh graphs at the beginning of training, then stabilizes after about 10 epochs. Validation loss values tend to follow training loss consistently, indicating that these two activation functions are able to maintain model generalization well and do not show significant indications of overfitting. In the sigmoid graph, validation loss decreases more slowly than the other activation functions, with a relatively high initial loss. This is consistent with the results of the quantitative evaluation where sigmoid provides the lowest performance. The softplus graph shows a rapid convergence process in the first 10 epochs with a stable loss in the subsequent stages. The consistency between training loss and validation loss in softplus indicates the stability of the model in predicting unseen data, and also explains why softplus produces the best prediction performance among all activation functions. This training and validation curve analysis confirms that softplus and ReLU have good convergence and stability characteristics, while sigmoid tends to be less optimal for forecasting BNI stock prices in this study.

3.3. Comparison of Predictions and Actuals

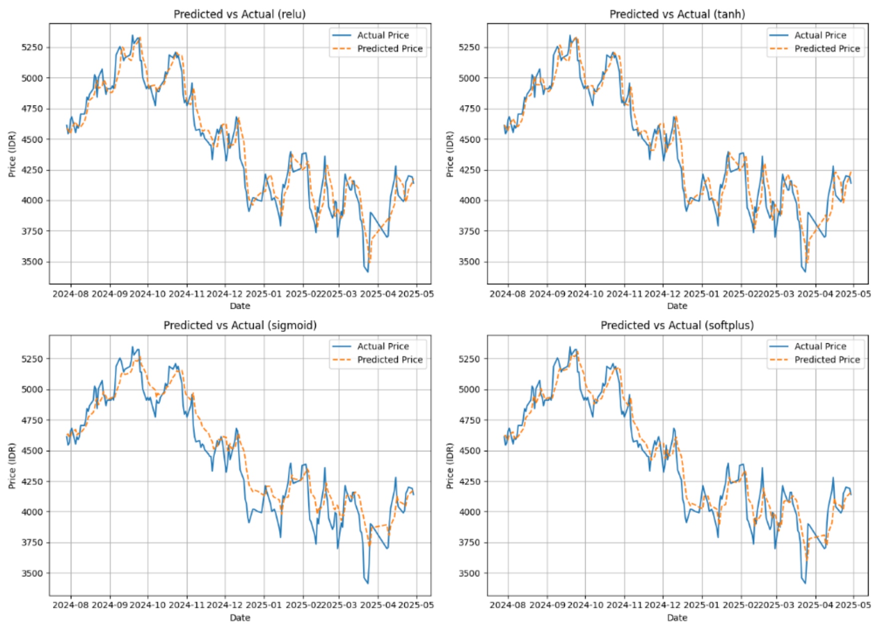

Figure 3: Comparison graph of prediction and actual

In the ReLU and Tanh activation functions, the prediction curves (dotted lines) are almost always parallel and close to the actual curve (solid line), indicating high accuracy and relatively small prediction errors across most timescales. This is consistent with the high R^2 values for both functions. The Sigmoid activation function shows a slight lag in predictions at some price peaks and troughs, although the overall trend follows the actual data. This difference indicates a slight limitation of the sigmoid in capturing the dynamics of extreme price changes. For the Softplus activation function, the prediction results are also quite close to the actual values, although there are slight deviations during periods of sharp fluctuations. This model remains capable of predicting trend direction well, reflecting its ability to represent long-term data patterns.

Date	Actual	Pred_relu	Pred_tanh	Pred_sigmoid	Pred_softplus
2024-07-29	4611.892578	4585.532227	4611.351074	4625.799316	4617.097656
2024-07-30	4543.058594	4579.793457	4601.488281	4634.273926	4622.863770
2024-07-31	4561.414062	4552.538086	4575.724609	4608.460938	4587.108398
2024-08-01	4657.782227	4544.435059	4565.675781	4609.579102	4585.943848
2024-08-02	4680.727051	4589.635742	4610.427246	4649.153320	4629.406738
2024-08-05	4552.236328	4640.374512	4666.588867	4669.873047	4650.368164
2024-08-06	4611.892578	4615.158691	4642.207520	4624.775879	4595.789062
2024-08-07	4588.947754	4610.450195	4623.118652	4639.884766	4615.499023
2024-08-08	4611.892578	4597.047363	4602.438965	4633.648438	4604.870117
2024-08-09	4703.671387	4597.752441	4602.804688	4641.885742	4613.510254
2024-08-12	4703.671387	4643.165039	4654.335449	4682.684082	4660.861328
2024-08-13	4749.561035	4680.108398	4702.615723	4695.397949	4674.276855
2024-08-14	4841.339844	4721.914062	4747.954590	4722.041504	4707.033203
2024-08-15	4818.395508	4787.001953	4810.894531	4771.918945	4766.723145
2024-08-16	4864.284668	4816.122070	4838.865234	4781.704590	4777.606445
2024-08-19	4910.174316	4843.174805	4860.176270	4810.909180	4813.859375
2024-08-20	5024.897949	4874.255859	4888.931641	4844.461914	4853.099609
2024-08-21	5001.953125	4942.691406	4960.684570	4908.569824	4928.810059
2024-08-22	4841.339844	4977.375000	5005.189941	4924.976074	4946.260742
2024-08-23	5001.953125	4913.783203	4943.065430	4872.735840	4883.295410
2024-08-26	5070.787598	4928.678711	4942.666992	4932.526855	4955.426270
2024-08-27	4956.063965	4981.958984	4999.264648	4980.543457	5005.993652
2024-08-28	4933.119141	4971.610352	5001.219727	4952.817383	4968.714844
2024-08-29	4864.284668	4944.204590	4970.783691	4942.942871	4956.654785
2024-08-30	4910.174316	4891.492676	4909.001465	4914.703613	4918.658691
2024-09-02	4910.174316	4878.410156	4887.013184	4927.182617	4929.627930
2024-09-03	4933.119141	4882.141113	4893.038086	4929.869141	4927.177246
2024-09-04	4910.174316	4904.048340	4919.981934	4940.566406	4936.738281
2024-09-05	5001.953125	4913.641602	4931.731934	4934.630371	4926.989258
2024-09-06	5185.511230	4961.011230	4973.560547	4972.354004	4971.321289

Figure 4: Comparison of actual and predicted prices

Based on Figure 4, the four models with the ReLU, Tanh, Sigmoid, and Softplus activation functions were able to produce predictions very close to the actual data. The difference between the actual price and the predicted results was generally in the tens of points range, indicating a relatively low error rate. At the beginning of the period, all models provided predictions that differed only slightly from the actual value (approximately ±30 points), with a similar pattern across subsequent dates. The Tanh and ReLU activation functions consistently provided predictions closest to the actual data at almost all time points, particularly during periods of gradual downtrends or uptrends. The model with the Sigmoid activation function tended to provide slightly higher predictions at some peaks, while the Softplus model sometimes produced larger deviations, especially during sharp price fluctuations.

4. Conclusion

Based on the evaluation results using RMSE, MAE, MAPE, and R^2 metrics, it was found that Softplus provided the best overall performance with an RMSE of 128.714, MAE of 101.815, MAPE of 2.358%, and R^2 of 0.924, indicating high accuracy and good stability during the training process, although it required the longest training time. ReLU occupies a competitive position with performance close to Softplus, having fast convergence and good generalization capabilities, but slightly inferior in MAE value. The Tanh activation function is in an intermediate position with stable results but slightly lower accuracy than Softplus and ReLU. Meanwhile, Sigmoid shows the lowest performance with the largest error value (RMSE of 147.220) and the lowest R^2 (0.900), making it less than optimal for BNI stock price prediction. Overall, the results of this study conclude that Softplus and ReLU are the most optimal activation function choices for modeling and predicting BNI stock price movements using LSTM, where Softplus excels in terms of accuracy, while ReLU offers a balance between performance and training time efficiency.

References

Darwish, M., Hassanien, E. E., & Eissa, A. H. (2025). Stock Market Forecasting: From Traditional Predictive Models to Large Language Models: M. Darwish et al. *Computational Economics*, 1-45.

Felle, A. R., & Santioso, L. (2024). The Effect Of Credit Risk And Bank-Specific Factors On Financial Performance Of Banks Listed In Indonesia Stock Exchange (Idx). *International Journal of Application on Economics and Business*, *2*(2), 3392-3403.

Hammad, M. M. (2024). Deep learning activation functions: fixed-shape, parametric, adaptive, stochastic, miscellaneous, non-standard, ensemble. *arXiv preprint arXiv:2407.11090*.

Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of long short-term memory (LSTM) neural network for flood forecasting. *Water*, *11*(7), 1387.

Malashin, I., Tynchenko, V., Gantimurov, A., Nelyub, V., & Borodulin, A. (2024). Applications of Long Short-Term Memory (LSTM) Networks in Polymeric Sciences: A Review. *Polymers*, *16*(18), 2607.

Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Fundamentals of artificial neural networks and deep learning. In *Multivariate statistical machine learning methods for genomic prediction* (pp. 379-425). Cham: Springer International Publishing.

Raju, V. G., Lakshmi, K. P., Jain, V. M., Kalidindi, A., & Padma, V. (2020, August). Study the influence of normalization/transformation process on the accuracy of supervised classification. In *2020 third international conference on smart systems and inventive technology (icssit)* (pp. 729-735). IEEE.

Raza, S., Baiqing, S., Kay-Khine, P., & Ali Kemal, M. (2023). Uncovering the effect of news signals on daily stock market performance: An econometric analysis. *International Journal of Financial Studies*, *11*(3), 99.

Sebastian, A., & Tantia, V. (2024). Deep Learning for Stock Price Prediction and Portfolio Optimization. *International Journal of Advanced Computer Science & Applications*, *15*(9).

Shaban, W. M., Ashraf, E., & Slama, A. E. (2024). SMP-DL: a novel stock market prediction approach based on deep learning for effective trend forecasting. *Neural Computing and Applications*, *36*(4), 1849-1873.

Syazali, E. A., Najwan, J., & Muskibah, M. (2024). Legal Reform and Strengthening of The Financial Services Authority (OJK): Protecting Investors From Insider Trading in Indonesia. *As-Siyasi: Journal of Constitutional Law*, *4*(2), 184-209.