JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115

COMPARISON OF ACTIVATION AND OPTIMIZER PERFORMANCE IN LSTM MODEL FOR PURE BEEF PRICE PREDICTION

Dasril Aldo[1*]

School of Computing, Informatics[1]
Telkom University, Purwokerto, Indonesia[1]
telkomuniversity.ac.id[1]
dasrilaldo@telkomuniversity.ac.id[1*]

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *One of the primary factors impacting the economy is the ability to forecast the prices of commodities such as beef. This paper aims to evaluate the effectiveness of various activation functions and optimization strategies when integrated into the LSTM (Long Short-Term Memory) architecture model in predicting the price of lean beef in Aceh. The data sample utilized was obtained from the Indonesian National Food Agency panel, which shows daily prices for beef within the time frame of July 14th, 2022, to July 31st, 2024. As for the conducted research, the process of preparation data preprocessing, partitioning data into training, validation and test sets and the actual execution of the LSTM model which was trained using four different types of activation functions: tanh, ReLU, sigmoid and PReLU together with three different optimizers: Adam, Nadam and RMSprop for 50, 70, 100 and 200 training iterations. The evaluation metrics employed were Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R-squared). The best performance was recorded at 200 epochs with the combination of PReLU activation function and Nadam optimizer, which had the best performance with RMSE = 2.56, MAPE = 0.65% and R^2 = 0.104. This combination was more effective than others since it depicted better overall performance in identifying complex non-linear relationships that existed in the price data. Further on, Nadam seems to have benefits in terms of allowing the model to converge faster and making the training more stable. This work stresses the selection of activation functions and optimization methods when building LSTM models aimed at forecasting prices of commodities with large volatility. It will be very helpful in developing better predictive models and decision-making processes in the agro-business. Another way to enhance predictive performance could be changing the model architecture or using different techniques, such as attention mechanisms.*

Keywords*: activation, beef price prediction, LSTM, optimization.*

Intisari— *Prediksi harga komoditas, seperti daging sapi murni, merupakan isu penting yang memengaruhi stabilitas ekonomi. Penelitian ini bertujuan untuk mengevaluasi kinerja berbagai fungsi aktivasi dan teknik optimasi dalam model Long Short-Term Memory (LSTM) untuk memprediksi harga daging sapi murni di Aceh. Data harga harian dari 14 Juli 2022 hingga 31 Juli 2024 diperoleh dari panel Badan Pangan Nasional Indonesia. Tahapan penelitian meliputi normalisasi data, pembagian dataset menjadi set pelatihan, validasi, dan pengujian, serta pelatihan model LSTM menggunakan empat fungsi aktivasi (tanh, ReLU, sigmoid, PReLU) yang dikombinasikan dengan tiga optimasi (Adam, Nadam, RMSprop) untuk 50, 70, 100, dan 200 epoch. Evaluasi model dilakukan menggunakan metrik Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), dan koefisien determinasi (R^2). Hasil penelitian menunjukkan bahwa kombinasi fungsi aktivasi PReLU dan optimasi Nadam pada epoch 200 menghasilkan kinerja terbaik dengan nilai RMSE terendah sebesar 2.56, MAPE terendah sebesar 0.65%, dan nilai R^2 positif sebesar 0.104. Kombinasi ini menunjukkan kemampuan superior dalam menangkap pola non-linear yang kompleks pada data harga dibandingkan dengan fungsi aktivasi lainnya. Selain itu, optimasi Nadam menunjukkan keunggulan dalam mempercepat konvergensi dan menstabilkan pelatihan model. Penelitian ini menekankan pentingnya pemilihan fungsi* *aktivasi dan teknik optimasi yang tepat dalam membangun model LSTM untuk prediksi harga komoditas yang volatil. Temuan ini memberikan wawasan berharga untuk mengembangkan model prediksi yang lebih akurat dan strategi pengambilan keputusan di sektor pertanian. Penelitian lebih lanjut dapat mengeksplorasi variasi arsitektur model atau mengintegrasikan teknik lanjutan, seperti mekanisme perhatian (attention mechanism), untuk meningkatkan akurasi prediksi.*

Kata Kunci*: aktivasi, LSTM, Optimasi, Prediksi Harga Daging Sapi.*

INTRODUCTION

Proper nutrition for communities should not be undermined and controlled on the interaction of it with beef and overall food prices since it is directly impacting economics and commercialization [1][2]. Beef especially in the areas of Aceh in Indonesia remains a stable evacuation commodity across all seasons [3]. Others include uncertainty for consumers and business actors are price fluctuations in pure beef which often becomes a complex issue [4]. This information comes from The Food Price Panel, that claimed that the price of pure beef in Aceh between 14 July 2022 and July 31 2024, fluctuated from day to day in accordance with economic conditions, shifts in supply and demand as well as government policies. Bringing this problem to the forefront is the price which is often difficult to control and ensure the purchasing capacity of the community remains intact.

Aceh Province was chosen as the location for the study because it has the highest average beef price in Indonesia, both before and during the Covid-19 pandemic, reaching Rp 165,850/kg at the peak of the pandemic. This is influenced by the traditions of the Acehnese people, such as meugang, which increases the demand for local beef for religious events. In addition, the limited local production and the distance from Aceh cattle production centers are other factors that cause high beef prices in this region compared to other provinces [5]. These characteristics make Aceh a very relevant location to study beef price fluctuations using predictive models such as LSTM. To address these challenges, advanced computational methods are needed to analyze and predict the dynamic patterns of beef price fluctuations.

For that purpose, Long Short-Term Memory (LSTM) [6][7][8], which is categorized as the architecture of an artificial neural network in deep learning, is considered to be useful. LSTM architecture was specifically invented to address the long-term dependencies in time-series data [9][10]. But then again, the effectiveness of LSTM depends a lot on the choice of the activation function on that neural network [11]. The choice of activation functions is critical as it has a bearing on how the neuron of a neural network will receive processes information and thus determines the output of the prediction [12]. With regard to the pure beef price forecasting, the model may get bias due to inappropriate activation function, leading to unsatisfactory result. LSTM was chosen because of its ability to capture non-linear and complex patterns in time series data. This architecture is designed to retain relevant information over the long term, making it ideal for dynamic and fluctuating pure beef price analysis.

Recent research shows the importance of combining activation functions and optimization techniques in LSTM models to improve the accuracy of price predictions. The ReLU activation function is proven to provide the best performance in predicting the closing price of natural gas stocks with the highest R^2 value compared to the Sigmoid and Tanh activation functions [13]. Uning of the ReLU activation function and the Nadam optimizer on the Bi-LSTM model also gives the best results in BTCL stock price prediction [14]. In addition, the hyperbolic tan activation function shows better results than sigmoid in stock price prediction on the Indonesia Stock Exchange [15]. In addition, the optimization of the LSTM model through Principal Component Analysis (PCA) has succeeded in improving the accuracy and efficiency of model training [16]. These studies confirm that the right combination of activation functions, optimizers, and tuning techniques is essential to improve the performance of LSTM models in various price prediction domains.

These studies provide important insights into the influence of activation functions and optimization techniques in ESG models to maximize the accuracy of price predictions, as well as become a key reference in this field. One study comparing activation functions such as ReLU, ELU, and TanH on 25 stocks from various stock exchanges, found that the hyperbolic tan function produced the best accuracy for stock price prediction using the multivariable time series method [17]. Another study used a combination of Genetic Algorithm (GA) and LSTM algorithms to optimize model parameters, including tuning optimizers, which showed that Adam provided stable and low MAPE and RMSE values [18]. Additional research

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115

comparing the multilayer LSTM model with the single-layer model for stock price prediction, found that the single-layer model produced better performance, demonstrating the importance of optimal parameter tuning [19]. A study also explores the use of Stacked Bidirectional LSTM (StacBi LSTM) to improve the accuracy of stock price predictions, with results demonstrating the power of the model in effectively capturing stock price trends[20]. This research is the main reference in understanding and improving the performance of the LSTM model for more accurate price prediction.

A key challenge when managing pure beef prices has been the lack of conventional methods to model such complex and non-linear patterns of fluctuations. Over say, the long time periods, traditional methods such as linear regression or ARIMA (Autoregressive Integrated Moving Average Models) are not powerful enough to accurately explain the changing behaviors of food prices [21][22][23]. Such drawbacks suggest a change in strategies towards improved ones that can adequately fit the non-linear tendencies in the data of commodity prices. This gap highlights the need for methods like LSTM, which are designed to handle non-linear patterns more effectively while accommodating the dynamic nature of commodity price fluctuations.

Past researchers have utilized LSTM in an effort to forecast prices of other goods, but do not highlight how much it matters to use different activation functions. For example, Wang and Li [24] have applied the usage of LSTM to predict the prices of food in China, yet did not focus on the best choice of activation within the neural network. Zhang et al. [25] focused also on free cash flows for stock predictions using a Stacked LSTM setup that incorporated a regression linear unit (ReLU) to reduce some overfitting effects in model training, however, market structures for stocks and food items are not the same. In their paper Kayim and Yılmaz [26] also reported a case on beef and lamb prices, they proposed an enhancement to the LSTM model whereby a volatility activation function was used to assist in the prediction of complicated time series data. They argue that this novel activation function is more effective than the commonly used activations of sigmoid and ReLU, which are often subjected to issues of data convergence and vanishing gradient problems. Wang et al. [27] were very emphatic on the proper selection of activation function while using LSTM for electrical energy forecasting, stating that it had a major bearing on the outcome. While there have been several studies looking at LSTM, there is still a void in the work aimed at performing the hybrid prediction of beef price using different specific activations.

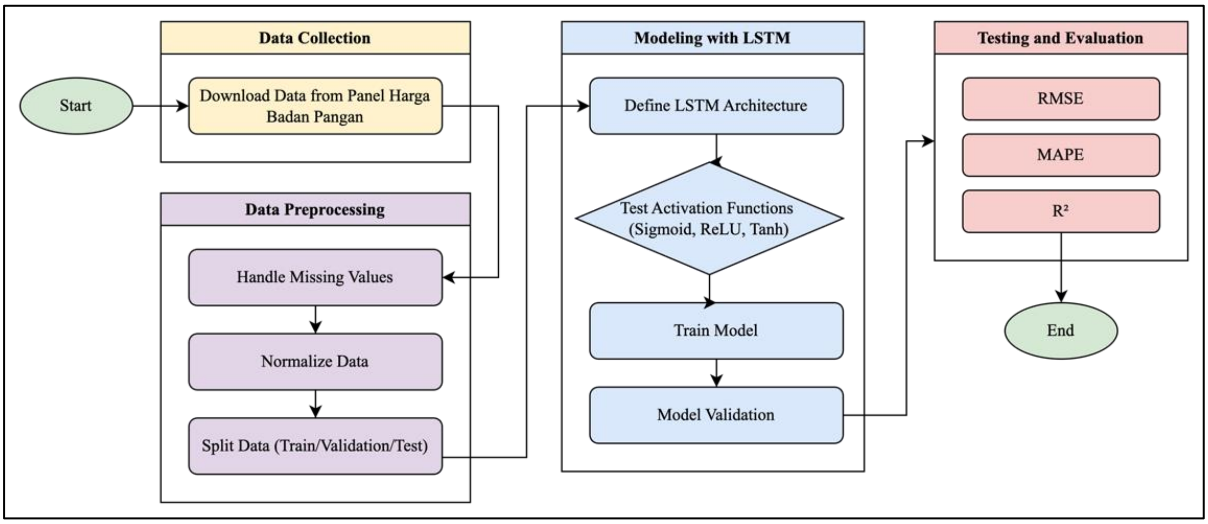
This study aims to evaluate how various activation functions (tanh, ReLU, sigmoid, and PReLU) affect the performance of the Long Short-Term Memory (LSTM) model in predicting the price of pure beef in Aceh. By exploring the combination of activation functions and optimization methods (Adam, Nadam, and RMSprop), this study presents a novelty by identifying the best configurations to improve the model's accuracy in capturing complex and non-linear price patterns. In contrast to previous studies that were more common in the application of LSTMs without focusing on configuration analysis, this study specifically makes a new contribution in the selection of the most effective activation function for predicting commodity prices with high volatility. These findings not only enrich the deep learning literature, especially related to the NGO network for food price forecasting, but also provide relevant policy recommendations to maintain the stability of beef prices in Aceh.

MATERIALS AND METHODS

This research uses a quantitative approach with a Long Short-Term Memory (LSTM) model to predict pure beef. Stages of the study include data collection, preprocessing, model development, validation as well as testing and evaluation of the model which are graphically represented in Figure 1. Based on the information in Figure 1, this research framework has importance articulating factors that are pivotal:

Data Collection

At this stage, the authors decided to go with raw beef wholesale prices that they sourced from the National Food Agency's Food Price Panel site. The reason for choosing this particular dataset is based of several key criteria. To begin with, the relevance of the data is very critical as the data being used needs to portray the changes in wholesale raw beef markets accurately to assist in investors analysis and future price predictions. Second, the time period of the data, spanning from July 14, 2022, to July 31, 2024, was specifically chosen to capture the presence of seasonal patterns, long-term trends, and other significant variations. Third, the data source which is used is the Food Security Agency's Food Price Panel which is esteemed to be an official agency where accurate and reliable daily price data is available. Last, daily prices data is particularly suitable for the



Source: (Research Results, 2024)

Figure 1. Research Methodology

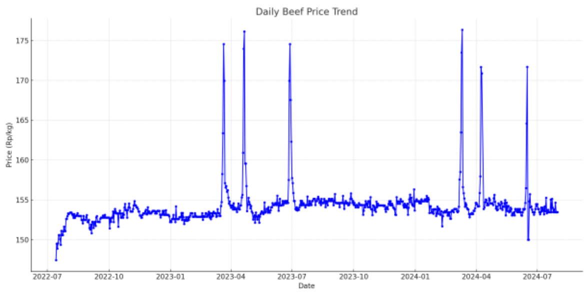
application of LSTM model which requires data with regular time intervals. Keeping these criteria in mind, the selected dataset is considered to be appropriate to cover the goals of this study. Table 1 presents extracts from the databases used in the study, including the dates and prices of beef per kilogram.

Table 1. Daily Price Data for Pure Beef

No.	Date	Price
1	14/07/2022	147.44
2	15/07/2022	149.5
3	16/07/2022	148.87
4	17/07/2022	149.53
5	18/07/2022	150.57
6	19/07/2022	150.57
7	20/07/2022	150.04
8	21/07/2022	149.35
9	22/07/2022	151.11
10	23/07/2022	150.57
11	24/07/2022	150.57
12	25/07/2022	151.06
13	26/07/2022	151.62
14	27/07/2022	151.08
15	28/07/2022	151.08
...	...	...
...	...	...
747	29/07/2024	153.45
748	30/07/2024	153.45
749	31/07/2024	153.45

Source: (Research Results, 2024)

To complete this analysis, we would also include graphs that will provide an effective, visual image of the data in tables, portraying the trend of the prices of pure beef per day throughout the said timeline. Such charts will highlight the upward and downward movements and assist in determining most significant changes in price patterns, such as the increases and decreases that are clearly visible in Figure 2.



Source: (Research Results, 2024)

Figure 2. Data Visualization of Daily Prices for Pure Beef

Figure 2 illustrates the graphical representation of daily price changes for pure cattle beef from July 14, 2022, to July 31, 2024. The X-axis represents the dates, while the Y-axis shows the price in rupiah per kilogram. Based on the graph, the average price started at approximately Rp 147.44 per kilogram on July 14, 2022, and steadily increased to Rp 149.50 per kilogram on July 15 of the same year. Between July sixteenth and July seventeenth, the first kilogram of meat weighed one hundred and fourty eight and eighty seven cents and one hundred and forty nine and fifty three cents respectively. This reflects an increase until July 19, when the price per kilogram remained stable. From July 20 to July 22, the graph shows a continued upward trend in beef prices, reaching a peak on July 22 at Rp 151,57 per kilogram. On July 23, the price stabilized at Rp 151,57 per kilogram. However, this stabilization possibly impacted revenue negatively between July 24 and December 5.

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115

The price of beef remained relatively stable from July 25 to July 28, 2022, within the range of Rp 151.06 to Rp 151.08 per kilogram, before reaching its peak on July 26 at Rp 151.62 per kilogram. Towards the end of the observation period, the price trend displayed a gradual but steady increase. On July 29, 2024, the price rose to Rp 153.45 per kilogram and remained at this level for the next two days, July 30 and 31. This consistent price increase over the last two years could be attributed to factors such as inflation or rising production costs. The data also indicate that beef prices stabilized within a central range, particularly between Rp 150 and Rp 153 per kilogram, suggesting that market control was relatively effective during this period. However, occasional price spikes suggest variability in supply, possibly caused by sudden increases in demand or disruptions in logistics. This trend highlights the need for market players and policymakers to address these issues to ensure stable beef prices and a consistent supply in the market

Data Preprocessing

After the data is collected, the next step is preprocessing in order to prepare the data to be ready to be used in the model. The preprocessing step consists of a number of sub-steps, which include:

1. Handle Missing Values

Missing values have been known to bias the analysis[28]. In this study missing values were computed using linear interpolation technique:

$$\text{linear interpolation} = \frac{Y_{t+1} - Y_{t-1}}{2} \quad (1)$$

2. Normalize Data

Rather than employing the same mathematical processes employed in scaling, the Min-Max scaling strategy can be applied owing to its extreme value in consistently restricting data to certain limits, which in this situation is between 0 and 1[29]. Mathematically, this is expressed as, Data is frequently adjusted using Min-Max Scaling which shifts all data within a specified range. The Min-Max Scaling formula goes as follows:

$$X_{\text{normalized}} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

3. Split Data (Train/Validation/Test)

The above data normalization has been categorized into 3 subsets namely: Training Set: This is used to teach the model, mostly around 80% of the entire data set. Validation Set: This is set aside to help perfect the model so that there is no overfitting during the training phase. Typically accounts for 10% of the dataset. Testing Set: First is used to verify the performance of the model after training, which is about 10% of the dataset [30].

Modeling with LSTM

In this step, the Long Short-term Memory (LSTM) model is built and developed in order to make advance forecasting. The LSTM is chosen due to its ability to effectively model complex, nonlinear time series data [31].

1. Forget Gate

This equation represents the forget gate, which specifies the amount of information retained from the previous memory cell. The sigmoid activation function σ is employed to produce values ranging from 0 (completely forget) to 1 (fully retains all).

$$f_t = \sigma\left(W_f x_t + R_f h_{\{t-1\}} + b_f\right) \quad (3)$$

2. Input Gate

This equation describes the input gate which controls how much new information is to be introduced into the memory. The sigmoid activation σ is utilized to filter the new information which is deemed to be useful.

$$i_t = \sigma\left(W_i x_t + R_i h_{\{t-1\}} + b_i\right) \quad (4)$$

3. Output Gate

The equation provided illustrates an output gate which decides the amount of the cell state memory that is sent to the next time step. In this case, sigmoid is employed to obtain the output.

$$o_t = \sigma\left(W_o x_t + R_o h_{\{t-1\}} + b_o\right) \quad (5)$$

4. Candidate Cell State

This is the equation of the candidate cell state. The application of the function tanh is to scale the output to a range of -1 to +1 which will be added to the cell memory.

$$\tilde{c}_t = tanh\left(W_c x_t + R_c h_{\{t-1\}} + b_c\right) \quad (6)$$

5. Cell State Update

This equation depicts the state of the cell (a memory) with respect to its change. The information which is to be forgotten $f_t$from the earlier memory $C_{\{t-1\}}$ is combined with the new information which comes through candidate cell $\tilde{c}_t$.

$$C_t = f_t * C_{\{t-1\}} + i_t * \tilde{c}_t \quad (7)$$

6. Hidden State Update

The final equation describes the hidden state (output) of the LSTM. The output of the output gate o_t is multiplied with the new cell state C_t which has been activated by the activation function.

The steps in modeling are as follows: Define LSTM Architecture: The LSTM model structure is specified with a number of layers such as an input layer, several hidden layers and the output layer. The model architecture is structured to enable it to learn complex patterns and long term dependencies in the price data. Sest Activation Functions: Activation functions' retrieval is done in the model through iteration so as to see how the performance can be improved. The activation functions tested included sigmoid, ReLU, and tanh. Each of these activation functions has different characteristics in the processing of the input signal:

a. Sigmoid Function [32]:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (8)$$

b. ReLU Function [33]:

$$f(x) = \max(0, x) \quad (9)$$

c. Tanh Function [34]:

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \quad (10)$$

Train Model was trained using training data with Adam optimizer optimization technique and other training parameters such rate of learning, number of Epoch, and batch size. Model Validation: Trained model is tested against validation data to check whether the model is overfit or underfit.

Testing and Evaluation

Once the model is trained and validated the final step is to put the model to test using the testing set. The predicted values are now compared with the actual values, and the performance of the model is assessed with the aid of several evaluation metrics, namely the following ones:

1. Root Mean Square Error (RMSE)

Measure the mean square root of the squared deviation between the predicted value and the actual value. RMSE is defined as [35]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(y_i - \acute{y}_i)^2} \quad (11)$$

2. Mean Absolute Percentage Error (MAPE)

The formula for MAPE is as follows [36]:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\acute{y}_i}{y_i}\right| * 100\% \quad (12)$$

3. R-squared (R^2)

To measure how well a model explains the variability of the actual data, one can look at the coefficient of determination given by R^2. A value of R^2 close to one indicates a model which is very good at explaining the variability of the data set. The formula for R^2 is given by [37]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\acute{y}_i)2}{\sum_{i=1}^{n}(y_i-\acute{y})2} \quad (13)$$

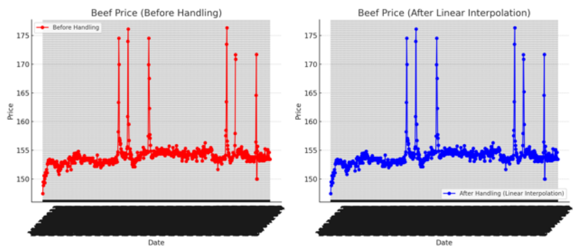
RESULTS AND DISCUSSION

Results

This section presents the results resulting from the implementation of the Long Short-Term Memory (LSTM) model in the prediction of pure beef daily prices based on the analyzed data. The obtained results will be compared to actual values in order to assess model performance and subsequently in looking overall into the price patterns which were found as well as the effect of using different activation functions on the accuracy of the predictions. The discussion also includes an interpretation of the findings and analysis of the causes that lead to price volatility.

1. Data Pre-processing

In this stage, the first step is to check whether or not there are any missing values through an analysis. The result is displayed in Figure 3.



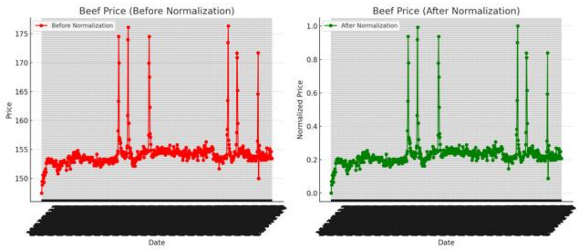
Source: (Research Results, 2024)

Figure 3. Results of Handling Missing Value

The lost values in this study were treated by linear interpolation method which fills in the missing value through a straight line drawn between two known points or data. The rows that underwent filling are the 21st, 637, and 638th rows. This method was selected so as to preserve the trend of the data from discontinuous changes. As a result, the data became more even and coherent which facilitates more precise analysis and price forecasting. Subsequently the data was normalized as shown in Figure 4.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115



Source: (Research Results, 2024)
Figure 4. Normalization Results

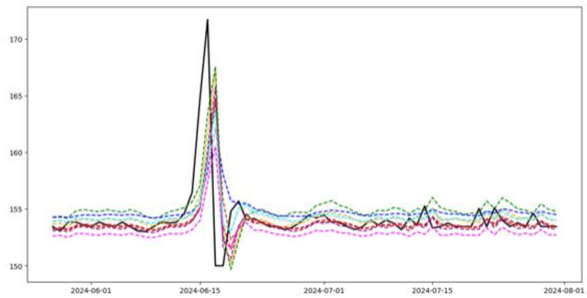
Data standardization is facilitated in this research through applying Min-Max Normalization technique. The purpose of this technique is to transform the original data values to a range between 0 and 1, so that all features are on the same scale. This is important because machine learning models are often sensitive to the scale differences between features which can affect the training oerformance. The results of the normalization showed that the data on the price of beef became more uniform and ordered as all values now fell within a uniform range which in turn aided in performing the models enhanced. Subsequently the next step was to perform data split such as shown in Figure 4.

Source: (Research Results, 2024)
Figure 5. Spliting Data

Figure 5 shows the distribution of data after the division into three main parts, namely the training data, the validation data, and the test data. The training data, indicated with blue color, consists of 80% of the overall data and is used to train the model in such a manner that it learns to identify patterns and relationships within the data. The validation data, indicated with orange color, consists of 10% of the overall data and aims to assess the performance of the model with respect to training processes, in such a way that the model does not suffer from over-optimization of the learning process. On the other hand, the test data indicated with red color, accounts for 10% of the overall data to be used for the final model evaluation. This separate allocation of data makes it possible for the model to be trained, validated and tested in an orderly and efficient manner so that the evaluation results are more accurate and representative.

2. Modeling with LSTM

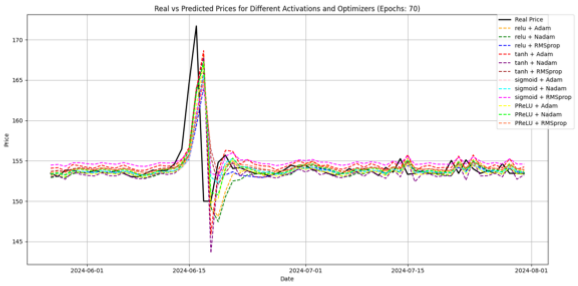
As the modeling tasks were completed using LSTM, various activation and optimization functions were utilized. The activations employed included ReLU, tanh, sigmoid, as well as PReLU, while optimizations included Adam, Nadam, and RMSprop. Number of epoch that were experimented on every combination were from 50 to 70, 100 and more to ensure the model was fully fitted.



Source: (Research Results, 2024)
Figure 6. Test Results with Epoch 50

LSTM modeling at epoch 50 is shown in Figure 6 indicating inconsistent performance of the model based on chosen combinations of activation and optimization for A) Adam optimization, tanh activation achieved a Final Epoch Loss of 0.00256 with an RMSE of 2.79, MAPE of 0.67%, and R2 of -0.0636. On the other hand ReLU activation with Adam had RMSE of 2.80, MAPE of 0.77% and R2 of -0.0747 which indicates the model has performed slightly worse than earlier mentioned tasks. Also, sigmoid activation produced Final Epoch Loss of 0.00312 with RMSE of 2.65 and MAPE of 0.83% and R2 positive of 0.0378 which indicates that the model is better able to explain the variability in the data. PReLU activation with Adam produced an RMSE of 2.82 and MAPE of 0.88%, however, R2 was unfavorable at -0.0861 which implies that more still means of more thorough explanation of variability in data exists.

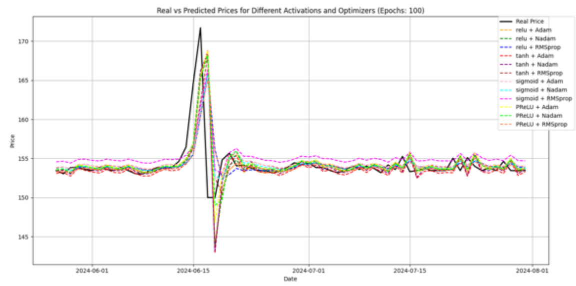
It can be observed that some activations performed better than others when using Nadam optimization. For example, a tanh activation resulted in approximately RMSE of 2.80, MAPE of 0.68%, and an R2 of 0.0695. On the other hand, ReLU performed poorly since the RMSE is estimated to be 3.16 with a MAPE of 1.34% and an R2 of -0.3600. Furthermore, the best results for the other combinations were achieved with the sigmoid activation which was characterized by RMSE of 2.60, MAPE of 0.69% and R2 of 0.0764. Moreover, the results alongside RMSprop optimization also differ. For tanh, the estimated RMSE is 2.80, the MAPE is 0.70%, and R2 is -0.0674. Moreover, activation using ReLU was somewhat better at RMSE of 2.75, MAPE of 0.68%, and an R2 of -0.0352. The most significant results came from the combination of sigmoid and RMSprop, which produced a RMSE of 2.64, a MAPE of 0.66%, and a positive R2 of 0.0496. Furthermore, when PReLU with RMSprop was utilized, empirical evidence suggests that the activation was able to achieve RMSE of 2.77, MAPE of 0.67%, and an R2 of -0.0471. Such weak results in R2 indicates that the model performance is still not excellent in explaining interpolation of the data.



Source: (Research Results, 2024)
Figure 7. Test Results with Epoch 70

Figure 7 shows the results of modeling an LSTM with an epoch of 70 and it is noted that the model's performance depends on a combination of activation and optimization. With the use of the optimizer by Adam, tanh activation gives a Final Epoch Loss of 0.00212 with RMSE metrics of 2.86, MAPE 0.74% and R squared of -0.1137. On the other hand, ReLU activation resulted in higher metrics of RMSE 2.92, MAPE 0.93% and R squared of -0.1616 which shows that complex models have explanatory power to variability in data. Using sigmoid activation and Adam optimizer, the Final Epoch Loss was 0.00286, the RMSE was at 2.76, the MAPE was 0.73% and the R squared figure was -0.0377 while the PReLU activation resulted in an RMSE of 2.76 and MAPE 0.71% while the R squared figure was -0.0414. All of this suggests that this combination of activation and optimization changes the model's capacity to capture patterns present in the data.

On the other hand, sigmoid activation with Nadam optimization yielded the best results with RMSE 2.72, MAPE 0.68%, and almost zero at -0.0112 for R2. With 0.71% MAPE and RMSE 2.82, tanh activation with Nadam optimization achieved an Epoch loss of 0.00233 and R2 - 0.0872. In the same case, using sigmoid activation under RMSprop optimization, better results were achieved with MAPE 0.65%, RMSE 2.64, and R2 0.0495 indicating that the model is quite good at explaining the variability in the data. PReLU combined with RMSprop optimization and recorded an RMSE of 2.75, MAPE of 0.68%, and an R-square value of -0.0314 indicating that this model is not yet optimal. Adam and Nadam's active optimizer with the use of sigmoid activation on the other hand showed better and consistent results compared to other optimization syndicates in epoch 70 and this is the main finding.



Source: (Research Results, 2024)
Figure 8. Test Results with Epoch 100

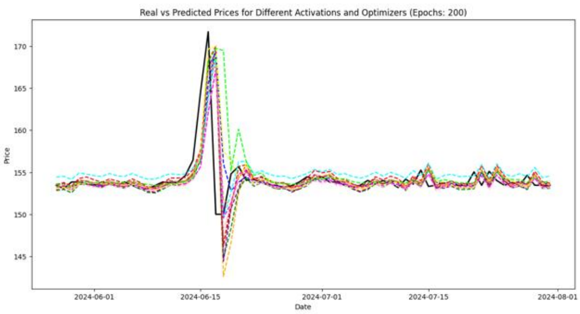
In Figure 8, the results of LSTM modeling with an epoch of 100 are obtained that a reasonable variation in performance has been obtained from the change in the combination of activation and optimization. In Adam's optimization, for tanh activation, the Final Epoch Loss was obtained of 0.00180 with a dimension value of RMSE 2.90, MAPE 0.84%, and R2 -0.1472, indicating that the model's ability can be said to be not good enough in explaining why there is variation in the data. ReLU activation with Adam showed slightly better performance with RMSE 2.81, MAPE 0.75 and R2 -0.0807. With the activation of sigmoid, the accuracy of the model becomes more consistent, indicated by an RMSE value of 2.74, MAPE 0.72% and R2 -0.0223 which is close to zero. PReLU activation has a large effect and shows an RMSE of 2.80 but has a higher MAPE of 0.97% and R2 -0.0731.

By utilizing the Nadam algorithm combined with the tanh activation function, it produces sub-optimal performance with a final epoch loss of 0.00186, RMSE of 2.92, MAPE at 0.78%, and R2 equal to -0.1647. An RMSE close to 0.7 with a value of 2.73 allows us to say that sigmoid activation combined with the Nadam algorithm gives the best results among all combinations, the values for MAPE and R2 are 0.67% and -0.0186 respectively. When RSMprop is used and the tanh function is applied together, the final epoch loss is lower than before, which is 0.00262 while RMSE, MAPE, and R2 are equal to 2.86, 0.72%, and -0.1187, respectively. The activation of the sigmoid function using the RMSprop algorithm was the most consistent

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115

achieving RMSE of 2.74, MAPE of 0.85% and R2 of -0.0279. After that, PreLeu is slightly better than tanh and ReLu by 20 basis points even though the activation function registers a negative R2 of -0.0682. Finally, there was a significant performance of the combination of ReLU+Nadam and sigmoid+adam with a greater degree of glucose absorption compared to the other with a value of 100.



Source: (Research Results, 2024)
Figure 9. Test Results with Epoch 200

From the results of the LSTM modeling in Figure 9 with an epoch of 200, we see that the model performance varies based on a combination of activation and optimization. In the case of optimization with Adam, tanh activation results in a final epoch loss value of 0.00137 and this loss value translates to RMSE 2.99, MAPE 0.77% and R2 -0.2257 which indicates that the model's performance is not efficient enough to account for variations in the data. ReLU activation combined with Adam's optimizer showed better performance with RMSE 2.67, MAPE 0.65% and R2 0.0248 which showed that the model was more effective in explaining the variability in the data. The activation of the sigmoid achieved a final loss epoch value of 0.00243, RMSE 2.79 and MAPE 0.75%, while PReLU activation gave the model RMSE 2.75, MAPE 0.74% and R2 -0.0333. With Nadam optimization, tanh was able to produce RMSE 2.94, MAPE 0.77%, and R2 -0.1793, but the performance was still below expectations. The activation of Nadam ReLU turned out to produce even lower numbers with RMSE 3.07, MAPE 0.87% and R2 -0.2891. However, Nadam PReLU activations performed best with values of 2.56, 0.65% and 0.1045 for RMSE, MAPE, and R2, respectively, indicating that the model was better able to identify patterns. On optimization with RMSprop, tanh activation results in a final epoch loss value of 0.00213. RMSE and MAPE yielded 2.91 and 0.75%, with an R2 of -0.1592.
The weather balance transformation function of sigmoid activation RMSE 2.79 and MAPE at 0.72%, while PReLU activation injury appears to be good with RMSE 2.69, MAPE 0.68%, and R2 positive 0.0122. In short, PReLU activation with Nadam optimizers and PReLU pose with Adam, while still within the positive extreme of all faces, felt less than optimal in other settings. Results on models with PReLU activation show a model that is better able to explain price variability. A narrow range describes a price with better variability capacity in a macro model setting that controls the empirical average of stock prices.
To conclude, the results of the LSTM comparison are shown in Table 2.

Table 2. Comparison Results of Activation and Optimization in LSTM

No	Epoch	Activation	Optimizer	Final Epoch Loss	RMSE	MAPE	R2
1	50	Tanh	Adam	0,0025563336	2,7907384777	0,0067236361	-0,0636419090
	70	Relu		0,0021486427	2,9163706028	0,0092505042	-0,1615625031
	100	Sigmoid		0,0025366852	2,7359641380	0,0072024963	-0,0222990595
	200	PReLU		0,0011497979	2,7506223357	0,0074031994	-0,0332825395
2	50	Tanh	NAdam	0,0025574833	2,7983532013	0,0067705677	-0,0694542695
	70	Relu		0,0021225926	2,8306489568	0,0074619817	-0,0942818286
	100	Sigmoid		0,0026400296	2,7310057012	0,0067352113	-0,0185969554
	200	PReLU		0,0009967067	2,5606960345	0,0065177937	0,1044842715
3	50	Tanh	RMSprop	0,0032121295	2,7956681574	0,0069857975	-0,0674029530
	70	Relu		0,0031411930	2,7794396945	0,0072381773	-0,0550466708
	100	Sigmoid		0,0031037224	2,7433988432	0,0085331380	-0,0278625965
	200	PReLU		0,0014036803	2,6894630794	0,0067694320	0,0121560745

Source: (Research Results, 2024)

In the recording of the evaluation results in theTable 2, four critical indicators can be identified that are involved: Final Epoch Loss, RMSE, MAPE, and R2 related to the combination of activation and optimization that gives the best results. First of all, the combination of ReLU activation and Adam's optimization at epoch 200 has the lowest Final Epoch Loss value of 0.00114 which confirms that the model successfully reduced the loss during training. However, the focus on the RMSE (Root Mean Squared Error) metric shows the combination of PReLU activation paired with the Nadam optimized model at epoch 200 as the most effective with the lowest RMSE value of 2.56. This suggests

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

that the model that uses this combination is the one that provides predictions closer to the actual value with the smallest mean error. Furthermore, in terms of MAPE (Mean Absolute Percentage Error), the use of sigmoid as an activation function with Nadam optimization in epoch 100 resulted in a MAPE of only 0.67%. This means that the actual value has a very low mean prediction error in percentage.

Other combinations apart from PReLU with NAdam at 200 epochs underperformed due to some technical issues. Tanh Activation for instance is limited as its output is within the scope of -1 to 1, making it less compatible with capturing complex data patterns that have a wide range of values. This resulted to higher loss, RMSE and MAPE values as well as a negative R^2 depicting the models failure to grasp the variability of the target data. ReLU activation suffers from dying ReLU issues where a neuron will stop activating due to giving a zero output for any negative input which causes the model to under learn. Meanwhile Sigmoid activation suffers from the backwards problem of increasing performance on the last layers while ineffective on the first making the predictions imprecise resulting in high values of RMSE and MAPE and negative R^2. Optimization wise, Adam and RMSprop do not perform to the same level as NAdam since they are both ineffective at speeding convergence and easily settle at local minima. On the other hand, with the implementation of Nesterov caused NAdam to be much more effective especially when trained on time series data which is the case in this study whereby any combination of PReLU and NAdam at 200 epochs is optimal.

In terms of R^2 metrics, the most promising combination is the combination of PReLU activation along with Nadam optimization at epoch time 200, which in this case has a positive R^2 value of 0.104. This shows that the combination is more effective in explaining the diversity of data than other combinations described. Thus, this would indicate that the combination of PReLU activation plus Nadam optimization at epoch time of 200 would be the most superior combination in terms of prediction accuracy as it has the smallest RMSE, reasonable MAPE, and positive R^2. Therefore, in this case, this combination can be considered the optimal choice for the LSTM model in this case.

Discussion

The results of this study show that the selection of activation functions and optimization techniques has a significant influence on the performance of the LSTM model in predicting beef prices. In particular, the combination of PReLU activation and Nadam optimization in epoch 200 proved to provide the best performance with the lowest RMSE value of 2.56, the lowest MAPE of 0.65% and a positive R^2 value of 0.104. The results show that the combination of LSTM and PReLU allows for better detection of non-linear patterns in price data compared to using other activation functions including tanh, ReLU, and sigmoid.

Generally, one of the problems in the application of ReLU and PReLU is the possibility of vanishing gradients in data modeling that have trends or non-linearity complexes. However, the results suggest that PReLU is superior to ReLU probably because PReLU's positive and negative slope units have lower RMSE and MAPE values, which allows for greater flexibility in handling negative values. This remains consistent with previous research that has shown that PReLU is effective in improving the performance of LSTM models applied to time series data with seasonal patterns or bulls and bears.

On the other hand, it is proven that the optimization of Nadam is better compared to the optimization of Adam and RMSprop in optimizing the parameters of the model. Nadam is a coping that incorporates adaptive momentum that will improve the quality of convergence and accelerate the learning process. This result is in line with previous research that has explained the advantages of Nadam in obtaining convergence time and reducing the final epoch loss in the neural network model.

However, it should be noted that the tanh and sigmoid components combined with RMSprop optimization managed to obtain relatively poor RMSE and R^2 values. The reason may be that these two activation functions are limited when facing situations where the prediction value changes and often experiences saturation problems. Therefore, although tanh and sigmoid are often used in LSTM, their use should be carefully observed according to the characteristics of the data to be analyzed.

By paying attention to all the results obtained, the impetus from the results of this study considers that in the development of an LSTM model for predicting volatile commodity prices, the selection of the activation function is very important. The composition of PReLU and Nadam shows a better ability to overcome some of the complexities of price patterns. This innovation can be used as a reference in the development of similar predictive models for other commodities. The study also leaves it up to future researchers to delve deeper into the effects of variations in the number of neurons or LSTM layers in the model, as well as utilize other methods, such as attention

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115

mechanisms, to capture more complex patterns in price.

CONCLUSION

This research successfully evaluated the performance of various activation functions and optimization techniques on the LSTM model in predicting the price of pure beef in Aceh. The results show that the combination of PReLU activation and Nadam optimization in epoch 200 results in the best with the lowest RMSE of 2.56, the lowest MAPE of 0.65%, and the positive R^2 value of 0.104. These findings show that PReLU is more effective than other activation functions such as tanh, ReLU, and sigmoid in capturing complex price patterns.

Overall, the findings of this research highlight the importance of selecting appropriate activation functions and optimization techniques in the construction of the LSTM model to predict commodity prices that have high volatility. It is clear that Nadam optimization performs well in accelerating the convergence process and improving the stability of the model while PReLU activation offers better variation in price flexibility.

This research makes an important contribution to commodity price prediction modeling, especially in the context of highly volatile markets. The use of this combination of best models will help policymakers and industry players in designing more effective strategies to deal with price fluctuations. This research also opens up opportunities for further development such as the integration of other forecasting techniques, or the application of models to other commodities.

REFERENCE

[1] I. J. Liur, D. F. Souhoka, dan B. J. Papilaya, "ANALISIS KADAR AIR DAN KUALITAS FISIK DAGING SAPI YANG DIJUAL DI PASAR TRADISIONAL KOTA AMBON," *Agrinimal J. Ilmu Ternak Dan Tanam.*, vol. 10, no. 1, hlm. 45–50, Apr 2022, doi: 10.30598/ajitt.2022.10.1.45-50.

[2] Sri Hardianti Rosadi dan Fitry Purnamasari, "Transmisi Harga Bawang Merah Ditingkat Produsen dan Konsumen di Sulawesi Selatan: Price Transmission of Onion at Producer and Consumer Level in South Sulawesi," *Perbal J. Pertan. Berkelanjutan*, vol. 10, no. 2, hlm. 206–219, Jul 2022, doi: 10.30605/perbal.v10i2.1870.

[3] M. Masduqi, E. M. Sari, dan Mohd. A. Nashri Abdullah, "Identifikasi Sifat Kuantitatif dan Sifat Kualitatif pada Sapi Aceh Dalam Rangka Pelestarian Sumber Daya Genetik Ternak Lokal," *J. Agripet*, vol. 21, no. 2, hlm. 141–148, Okt 2021, doi: 10.17969/agripet.v21i2.21185.

[4] N. Helmiah dan Nasrudin, "SIMULASI KEBIJAKAN PADA IMPLEMENTASI PERJANJIAN KOMPREHENSIF INDONESIA-AUSTRALIA (IA-CEPA) TERHADAP PASAR DAGING SAPI DOMESTIK," *Bul. Ilm. Litbang Perdagang.*, vol. 15, no. 2, hlm. 157–180, Des 2021, doi: 10.30908/bilp.v15i2.633.

[5] A. I. Hasibuan, Sahara, dan S. Mulatsih, "Covid-19 dan Disparitas Harga Daging Sapi Indonesia," *Policy Brief Pertan. Kelaut. Dan Biosains Trop.*, vol. 4, no. 1, hlm. 175–178, Feb 2022, doi: 10.29244/agro-maritim.0401.175-178.

[6] K. Phoksawat, E. Phoksawat, dan B. Chanakot, "Forecasting smoked rubber sheets price based on a deep learning model with long short-term memory," *Int. J. Electr. Comput. Eng. IJECE*, vol. 13, no. 1, hlm. 688, Feb 2023, doi: 10.11591/ijece.v13i1.pp688-696.

[7] S. Sen, D. Sugiarto, dan A. Rochman, "Prediksi Harga Beras Menggunakan Metode Multilayer Perceptron (MLP) dan Long Short Term Memory (LSTM)," *Ultim. J. Tek. Inform.*, vol. 12, no. 1, hlm. 35–41, Jul 2020, doi: 10.31937/ti.v12i1.1572.

[8] R. M. S. Adi dan S. Sudianto, "Prediksi Harga Komoditas Pangan Menggunakan Algoritma Long Short-Term Memory (LSTM)," *Build. Inform. Technol. Sci. BITS*, vol. 4, no. 2, Sep 2022, doi: 10.47065/bits.v4i2.2229.

[9] A. Isah *dkk.*, "Digital Twins Temporal Dependencies-Based on Time Series Using Multivariate Long Short-Term Memory," *Electronics*, vol. 12, no. 19, hlm. 4187, Okt 2023, doi: 10.3390/electronics12194187.

[10] X. Wen dan W. Li, "Time Series Prediction Based on LSTM-Attention-LSTM Model," *IEEE Access*, vol. 11, hlm. 48322–48331, 2023, doi: 10.1109/ACCESS.2023.3276628.

[11] M. A. Mohamed, H. A. Hassan, M. H. Essai, H. Esmaiel, A. S. Mubarak, dan O. A. Omer, "Modified gate activation functions of Bi-LSTM-based SC-FDMA channel equalization," *J. Electr. Eng.*, vol. 74, no. 4, hlm. 256–266, Agu 2023, doi: 10.2478/jee-2023-0032.

[12] V. Sakellariou, V. Paliouras, I. Kouretas, H. Saleh, dan T. Stouraitis, "A High-performance RNS LSTM block," dalam *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, Austin, TX, USA: IEEE, Mei 2022, hlm. 1264–1268. doi: 10.1109/ISCAS48785.2022.9937633.

[13] B. Budiman, N. Alamsyah, dan R. Y. R. Alamsyah, "ACTIVATION FUNCTION IN LSTM FOR IMPROVED FORECASTING OF CLOSING NATURAL GAS STOCK PRICES," *JITK J. Ilmu Pengetah. Dan Teknol. Komput.*, vol. 10, no. 1, hlm. 100–107, Agu 2024, doi: 10.33480/jitk.v10i1.5258.

[14] Y. Loday, P. Apirukvorapinit, dan P. Vejjanugraha, "Stock Price Prediction Using Modified Bidirectional Long Short-Term Memory and Deep Learning Models: A Case Study of Bhutan Tourism Corporation Limited Stock Data," dalam *2023 8th International Conference on Business and Industrial Research (ICBIR)*, Bangkok, Thailand: IEEE, Mei 2023, hlm. 645–650. doi: 10.1109/ICBIR57571.2023.10147486.

[15] A. Wahab, A. Herdian, D. Wirawan, Y. Jumaryadi, S. Alam, dan A. Fiade, "Stock Prediction for Indonesia Stock Exchange with Long Short-Term Memory," *J. Ilm. FIFO*, vol. 16, no. 1, hlm. 96, Jun 2024, doi: 10.22441/fifo.2024.v16i1.010.

[16] W. Zhai, "Stock Price Prediction Based on Optimized LSTM Model," dalam *Proceedings of the 2nd International Conference on Financial Innovation, FinTech and Information Technology, FFIT 2023, July 7–9, 2023, Chongqing, China*, Chongqing, People's Republic of China: EAI, 2023. doi: 10.4108/eai.7-7-2023.2338038.

[17] H. M. Sami, K. A. Ahsan, dan P. N. Rozario, "Determining the Best Activation Functions for Predicting Stock Prices in Different (Stock Exchanges) Through Multivariable Time Series Forecasting of LSTM," *Aust. J. Eng. Innov. Technol.*, hlm. 63–71, Apr 2023, doi: 10.34104/ajeit.023.063071.

[18] Y. L. Sukestiyarno, D. T. Wiyanti, L. Azizah, dan W. Widada, "Algorithm Optimizer in GA-LSTM for Stock Price Forecasting," *Contemp. Math.*, Jan 2024, doi: 10.37256/cm.5120243367.

[19] P. Paygude *dkk.*, "Optimizing Hyperparameters for Enhanced LSTM-Based Prediction System Performance," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. 10s, hlm. 203–213, Okt 2023, doi: 10.17762/ijritcc.v11i10s.7620.

[20] M. Diqi dan H. Hamzah, "Improving Stock Price Prediction Accuracy with StacBi LSTM," *JISKA J. Inform. Sunan Kalijaga*, vol. 9, no. 1, hlm. 10–26, Jan 2024, doi: 10.14421/jiska.2024.9.1.10-26.

[21] J. J. Pangaribuan, F. Fanny, O. P. Barus, dan R. Romindo, "Prediksi Penjualan Bisnis Rumah Properti Dengan Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA)," *J. Sist. Inf. Bisnis*, vol. 13, no. 2, hlm. 154–161, Okt 2023, doi: 10.21456/vol13iss2pp154-161.

[22] M. Ibrahim, H. R. Perwira Negara, dan S. Syaharuddin, "Prediction of Land Area Harvest, Production, Rice Productivity: A Accuracy Analysis of ARIMA Methods," *Protech Biosyst. J.*, vol. 1, no. 1, hlm. 1, Jun 2021, doi: 10.31764/protech.v1i1.4776.

[23] M. M. H. Khan, M. R. U. Mustafa, M. S. Hossain, S. Shams, dan A. D. Julius, "Short-Term and Long-Term Rainfall Forecasting Using ARIMA Model," *Int. J. Environ. Sci. Dev.*, vol. 14, no. 5, hlm. 292–298, 2023, doi: 10.18178/ijesd.2023.14.5.1447.

[24] Z. Wang dan Y. Li, "Price movement prediction using deep learning: a case study of the China futures market," dalam *Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)*, Y. Lu dan C. Cheng, Ed., Beijing, China: SPIE, Des 2022, hlm. 223. doi: 10.1117/12.2662524.

[25] X. Zhang, C. Li, K.-L. Chen, D. Chrysostomou, dan H. Yang, "Stock Prediction with Stacked-LSTM Neural Networks," dalam *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, Hainan, China: IEEE, Des 2021, hlm. 1119–1125. doi: 10.1109/QRS-C55045.2021.00166.

[26] F. Kayim dan A. Yilmaz, "Time Series Forecasting With Volatility Activation Function," *IEEE Access*, vol. 10, hlm. 104000–104010, 2022, doi: 10.1109/ACCESS.2022.3211312.

[27] J. Wang dan J. Wang, "A New Hybrid Forecasting Model Based on SW-LSTM and Wavelet Packet Decomposition: A Case Study of Oil Futures Prices," *Comput. Intell. Neurosci.*, vol. 2021, no. 1, hlm. 7653091, Jan 2021, doi: 10.1155/2021/7653091.

[28] G. Huang, "Missing data filling method based on linear interpolation and lightgbm," *J. Phys. Conf. Ser.*, vol. 1754, no. 1, hlm. 012187, Feb 2021, doi: 10.1088/1742-6596/1754/1/012187.

[29] M. Pagan, M. Zarlis, dan A. Candra, "Investigating the impact of data scaling on the k-nearest neighbor algorithm," *Comput. Sci. Inf. Technol.*, vol. 4, no. 2, hlm. 135–142, Jul 2023, doi: 10.11591/csit.v4i2.p135-142.

[30] H. Ghasemzadeh, R. E. Hillman, dan D. D. Mehta, "Toward Generalizable Machine Learning Models in Speech, Language, and Hearing Sciences: Estimating Sample Size and

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6115

Reducing Overfitting," *J. Speech Lang. Hear. Res.*, vol. 67, no. 3, hlm. 753–781, Mar 2024, doi: 10.1044/2023_JSLHR-23-00273.

[31] S. Feng dan Y. Feng, "A Dual-Staged Attention Based Conversion-Gated Long Short Term Memory for Multivariable Time Series Prediction," *IEEE Access*, vol. 10, hlm. 368–379, 2022, doi: 10.1109/ACCESS.2021.3136712.

[32] Z. Pan, Z. Gu, X. Jiang, G. Zhu, dan D. Ma, "A Modular Approximation Methodology for Efficient Fixed-Point Hardware Implementation of the Sigmoid Function," *IEEE Trans. Ind. Electron.*, vol. 69, no. 10, hlm. 10694–10703, Okt 2022, doi: 10.1109/TIE.2022.3146573.

[33] Y. Bai, "RELU-Function and Derived Function Review," *SHS Web Conf.*, vol. 144, hlm. 02006, 2022, doi: 10.1051/shsconf/202214402006.

[34] S. Sorayaasa dan M. Ahmadi, "Efficient Implementation of Tanh: A Comparative Study of New Results," dalam *Artificial Intelligence Advances*, Academy and Industry Research Collaboration Center (AIRCC), Apr 2023, hlm. 01–07. doi: 10.5121/csit.2023.130701.

[35] R. A. Hasan dan J. E. Jamaluddin, "Prediction of Covid-19 Cases for Malaysia, Egypt, and USA using Deep Learning Models," *Malays. J. Fundam. Appl. Sci.*, vol. 19, no. 3, hlm. 417–428, Mei 2023, doi: 10.11113/mjfas.v19n3.2992.

[36] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geosci. Model Dev.*, vol. 15, no. 14, hlm. 5481–5487, Jul 2022, doi: 10.5194/gmd-15-5481-2022.

[37] M. Bossé, E. Marland, G. Rhoads, J. A. Sanqui, dan Z. BeMent, "Generalizing R^2 for deming regressions," *Commun. Stat. - Theory Methods*, vol. 52, no. 21, hlm. 7731–7743, Nov 2023, doi: 10.1080/03610926.2022.2059678.