ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 14 (2024) No. 4, pp. 466-475
DOI:10.33168/JSMS.2024.0430

Predicting Stock Price by Bidirectional-Lstm with Dual Attention Network

Agus Chandra, Abba Suganda Girsang

Computer Science Department, Binus Graduate Program, Master of Computer Science Bina Nusantara University, Jakarta, Indonesia, 11480

agus.chandra001@binus.ac.id (Corresponding Author)

Abstract. Predicting stock prices is one of the challenges in the field of financial markets. Due to the high volatility of stock prices, traditional data prediction methods cannot identify the most relevant and critical market data for predicting stock prices. This study discusses the use of Bidirectional-LSTM and Dual Attention Network in the process of predicting stock prices using stock data of state-owned banks. from historical Yahoo finance. The purpose of this study is to build a model that applies the two methods of Bidirectional-LSTM and Dual attention Network. The features used in the prediction process are open, high, low, volume, adj close. The method proposed by this research is compared with the LSTM and BI-LSTM models. The results show that the performance of the proposed method is the best, MAE, MSE, RMSE are the smallest with the results MAE = 0.079, MSE = 0.012, RMSE = 0.110.

Keywords: Bidirectional-LSTM, Dual Attention Network, Stock prediction

1. Introduction

A stock is a piece of paper that shows the right of the to obtain a stock of the prospects or wealth of the organization that issues the security and various conditions that allow these investors to exercise their rights (Adrison & Flukeria, 2016). Stocks are one of the most popular financial market instruments. Stocks are an investment tool that many investors choose because stocks are able to provide an attractive level of profit (Novita, 2009). However, many investors are still hesitant to invest, because the fluctuations in the stock price index are very fast and it is feared that they will not meet expectations(Izzah & Widyastuti, 2017). A stock index is a statistical measure that reflects the overall price movement of a group of stocks that are selected based on certain criteria and methodologies and are evaluated periodically. The index is a tool to measure market sentiment or investor confidence. Changes in value reflected in one index can be used as an indicator that reflects the collective opinion of all market participants. Thus, predicting or forecasting a stock price becomes an important tool for more effective and efficient planning so that investors can easily measure market sentiment, both predictions using machine learning(Chowdhury et al., 2020) and deep learning (Liu & Long, 2020).

In a study conducted by (Chaudhari et al., 2022), predictions were made using recurrent neural network techniques such as LSTM and BI-LSTM, which were trained based on historical data from various stock pools. In this study they tested the performance of the LSTM and BI-LSTM approaches in terms of how they generate results for specific stocks to identify optimal algorithms with high accuracy. They used over 50 stocks to train their models, and trained them for over 45 days to improve model training. The average accuracy for the LSTM model is 95.90, while the average accuracy for the BILSTM model is 93.30, indicating that the LSTM model is far superior to the BILSTM model in terms of accuracy. In addition, the time required for the Lstm model is greater than that required for the Bilstm. The average time needed for Lstm is 4.2 and for BiLstm it is 3.5.Furthermore, (Wiranda & Sadikin, 2019) employed LSTM to forecast product sales, yielding promising results and conducted data training and modelling, demonstrating LSTM's capability to accurately predict product sales in the time series data for 2017-2019, with a value of 13,762,154.00 for RMSE in rupiah and a MAPE of 12%. In conclusion, the LSTM algorithm exhibits excellent performance in computations involving time series data. Additionally, the Bidirectional LSTM (Bi-LSTM), commonly used in natural language processing, allows for information flow in both directions and is adept at capturing sequential dependencies between words and phrases. It serves as a powerful tool in modelling sequential data in both forward and backward directions. Thus, further research will be carried out using the dual attention network bidirectional LSTM method. In this research, several trial stages will be carried out, including pre-processing, DA-BILSTM method modelling, model training, results evaluation.

2. Literature Review

Stock are certificates that serve as tangible evidence of owning a company, granting shareholders the right to claim profits and assets. They represent securities that signify partial ownership in a company, indicating that investors who purchase stocks are acquiring a portion of the company. As a result, investors are entitled to receive dividends, which represent the company's profits. Stocks are issued by companies in the form of limited liability entities, often referred to as issuers. The ownership of stocks establishes shareholders as partial owners of the company. Therefore, when an investor purchases stocks, they effectively become a shareholder and hold a stake in the company's ownership (Nurhayati, 2016).

The price of stocks in the capital market is largely determined by the strength of demand and supply. The more investors who buy stocks, the higher the price of these stocks. In trading and investing, stock price refers to the current stock price in stock trading. The stock price indicator illustrates many things that are currently happening between buyers and sellers. This stock price indicator not only describes the market price, but also describes the party currently in control of the capital market. The latest information that enters the capital market will cause investors to buy or sell stocks, this causes price

movements. Stock prices can change up or down in very fast calculations, in just a matter of minutes, even seconds.

Attention is a complex cognitive function that is very necessary for humans (Corbetta & Shulman, 2002; Rensink, 2000). Perception possesses a crucial characteristic wherein humans typically do not process all available information simultaneously. Instead, humans tend to selectively focus on specific information when and where it is necessary, disregarding other information that can be comprehended concurrently. For instance, when visually observing things, humans do not typically take in the entire scene from beginning to end. Rather, they observe and pay attention to specific parts as required. When humans frequently encounter a scene containing something of interest in a particular section, they learn to concentrate on that section when similar scenes reappear, allocating more attention to the relevant portion. This ability enables humans to swiftly extract valuable information from vast amounts of data, utilizing limited cognitive resources. The attention mechanism significantly enhances the efficiency and accuracy of perceptual information processing. In situations where computing power is constrained, it allows for the processing of more important information using limited resources. Consequently, some researchers are directing their attention to the field of computer vision.(Itti et al., 2005) introduced a salience-based visual attention model that identifies salient potential areas through the extraction of local low-level visual features. In the realm of neural networks,(Mnih et al., 2014) employed attention mechanisms in recurrent neural network models for image classification.(Bahdanau et al., 2015) utilized attention mechanisms to simultaneously perform translation and alignment in machine translation tasks.

There are various types of attention models, including global attention and local attention, as described(Luong et al., 2015). In the case of global attention, input is received from each source state (encoder) and the decoder status prior to the current state, which is then used to calculate the output. Figure 1 illustrates the structure of the global attention model. In this model, the alignment weight or attention weight (a<t>) is computed based on each encoder step and the previous decoder step (h<t>). Subsequently, the context vector is obtained by taking the product of the global alignment weights and each encoder step (a<t>). This context vector is then fed into the RNN cell to determine the decoder output.

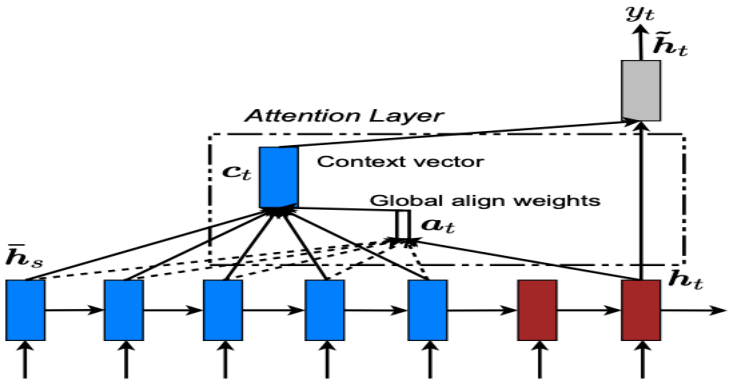

Fig 1: Global Attention Model

As for local attention, it differs from the Global Attention Model in the way that in the Local attention model, only several positions from the source (encoder) are used to calculate the alignment weight (a<t>). Below is a diagram for the Local Attention model.

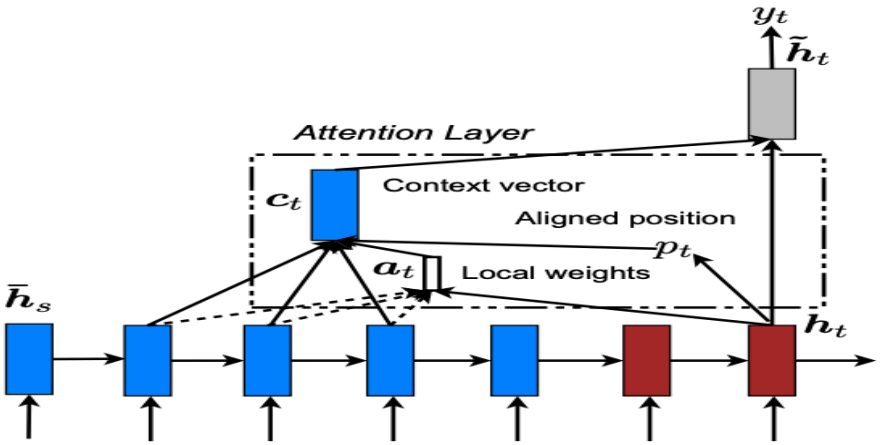

Fig 2: Local Attention Model

As can be seen from the figure 2, first the single-aligned position (p<t>) is found then the word window from the layer source (encoder) together with (h<t>) is used to calculate the align weights and the context vector. Local Attention-It consists of two types of Monotonic alignment and Predictive alignment. In monotone alignment, we just set the position (p<t>) as "t" whereas in predictive alignment, position (p<t>) instead of just thinking of it as "t", it is predicted by the predictive model.

LSTM is one of the RNN architectures. LSTM was created by(Hochreiter & Schmidhuber, 1997) and then developed and popularized by many researchers. Like RNNs, LSTM networks also consist of modules with iterative processing. In short, (Gers & Schmidhuber, 2001) LSTM adds a selection process in the control box (cell) so that it can select which information is appropriate to be forwarded, as well as being a solution to the vanishing gradient problem. Architectural drawings from RNN to LSTM, the hidden layer in the RNN will be shaped like a cell that functions to remember past events. By paying more attention to what is happening in the hidden layer, it is necessary to dissect the contents of the Xt perceptron. There is also a new notation It (read as input at time t) which is the input for the Xt perceptron. The activation function used can be sigmoid or hyperbolic (tanh). (Gers et al., 1999) LSTM is divided into several gates including Forget Gate, Input Gate, Cell Gate and Output Gate which have their respective duties.

The Bidirectional LSTM (Bi-LSTM) is a type of recurrent neural network predominantly employed in natural language processing tasks. Unlike standard LSTMs, the Bi-LSTM allows for bidirectional input flow, enabling the utilization of information from both preceding and succeeding elements in the sequence. It serves as a potent tool for modeling sequential dependencies among words and phrases in both forward and backward directions. To elaborate, the Bi-LSTM incorporates an additional LSTM layer that reverses the direction of information flow. Consequently, the input sequence is processed in reverse order within this supplementary LSTM layer. The outputs from both the standard and reversed LSTM layers are then combined through various methods, such as averaging, summing, multiplying, or other types of combination techniques.

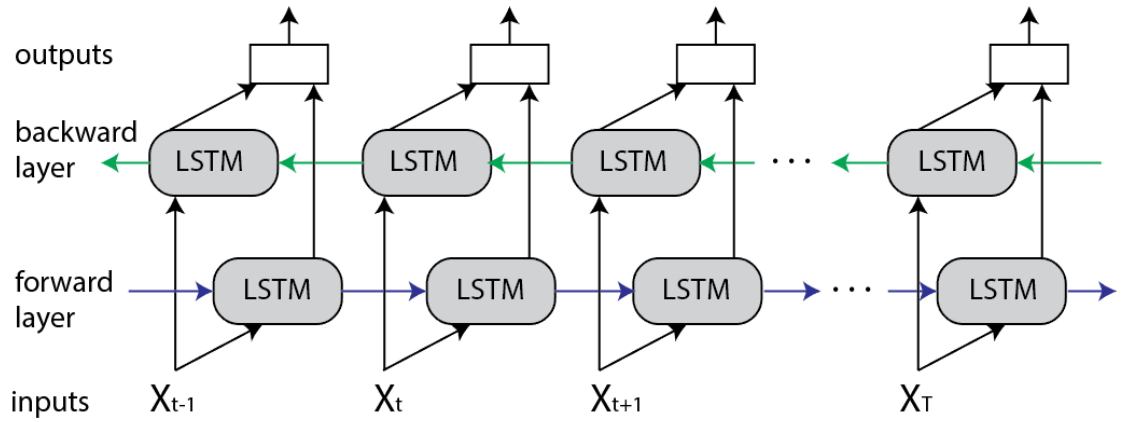

Fig 3: Bidirectional-LSTM Model

3. Methodology

The description of our research methodology is provided in this section. Section 3.1 is the description of the datasets used in our study. The proposed method is described in section 3.2. Section 3.3 is the description of evaluation matrix used to compare the result between the proposed method and other methods.

3.1. Datasets and Data Pre-Processing

Datasets was obtained from yahoo finance, including data on stocks of Indonesia state-owned banks including BRI, BNI, BTN and Mandiri. Stock data was taken from 1 January 2012 – 31 December 2022. The stock data collected consisted of various columns, namely date, open, close, high, low, adj close, and volume. The date column represents the specific date, month, and year when the stock market opened for trading. The open column indicates the stock price at the beginning of the trading day when the first transaction occurs. High and Low reflect the daily price fluctuations, guiding rational decisions for buying or selling stocks. The close column signifies the stock price at the end of the trading day when all transactions on the stock exchange conclude. Adj Close represents the adjusted closing price, which considers factors such as dividends and stock splits. Volume corresponds to the total number of stocks or lots traded within a specific period. For this study, the focus lies on the closing stock prices, serving as the reference or output data for training and testing purposes. In the data processing process, the existing data will be normalized or commonly known as scaling. Where data is changed using Standard Scaler which is a preprocessing technique in machine learning that is used to standardize features by removing averages and scaling to unit variances. This changes the feature such that it has a mean of zero and a standard deviation of 1. Furthermore, the data will be segmented or grouped data from the data retrieval process on yahoo finance. Data taken as many as 2732 rows and 7 columns based on banking data. Then the data is divided into 3 datasets, namely the training dataset, the validation dataset, the testing dataset. The proportion between the training dataset and the validation datesets is 70% for the training dataset, 20% for the validation dataset and 10% for the testing dataset (Borovkova & Tsiamas, 2019). Furthermore, the dataset that has been divided is then converted into a supervised learning dataset, namely in the form of sequential data in the form of input and output. The input data consists of open, high, low, volume, adj close which are used to predict the value of stock closes. The process of this change is carried out using the rolling window method, namely partitioning the time series data from the dataset into a small window so that it can produce more accurate predictions (Khalis Sofi et al., 2021).

3.2. BI-LSTM Dual Attention Network Model

In this study, we propose a new two-stage attention-based recurrent neural network (DA-BILSTM) for time series prediction. In the first stage, an input attention mechanism is introduced to adaptively extract the relevant driver set (aka, input feature) at each time step with reference to the hidden state of the previous encoder. In the second stage, a temporal attention mechanism is introduced to select the relevant encoder hidden states across all time steps (Qin et al., 2017).

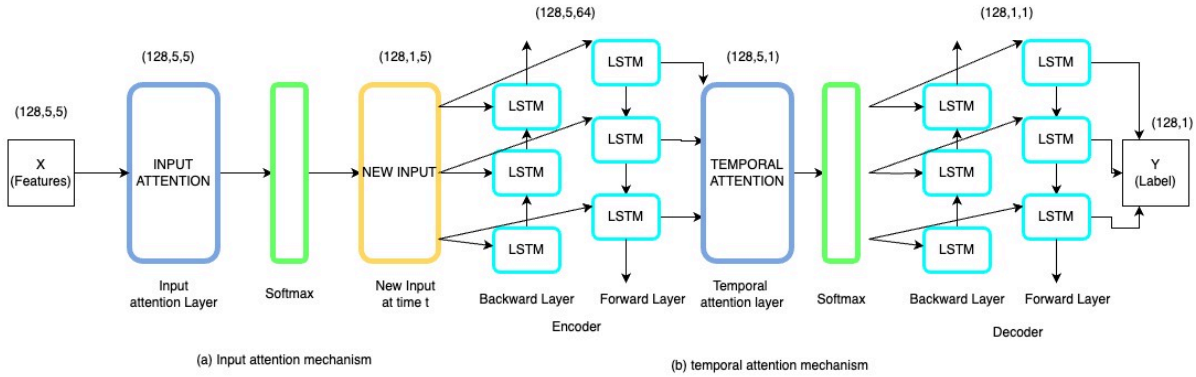

Fig 4: BI-LSTM Dual Attention Network Model

With the example Batch_size = 128, T = 5, Features = 5, the process that occurs when the dataset is entered into the model can be seen in Figure 4. The Input Attention layer calculates the attention weight for the Encoder at a given time step T. The layer applies the transformation using dense layers and generates the attention weight using the softmax function. The attention weight indicates the importance of each data at a given time step.The encoder layer generates `X_tilde` input for the encoder in this model. It consists of input `X` is processed for each time step. The `Alpha_t` attention weight is calculated using the`input_attention` layer. `X_tilde_t` is obtained by multiplying `Alpha_t` by the input in the `t` time step. The input is then passed through the BI-LSTM layer, resulting in an updated hidden state. Hidden states are merged and saved. Finally, the combined hidden state forms the encoder hidden state, which is returned. The encoder layer plays a critical role in capturing meaningful information from the input sequence using attention mechanisms and bidirectional-LSTM. In the Temporal Attention layer, the process that occurs is to calculate the attention weight for each encoder hidden state based on the provided hidden state, cell state, and encoder hidden state. It applies transformations and softmax functions to determine the relevance and importance of each encoder hidden state in relation to the current hidden state. This allows the model to focus on the relevant encoder state during sequence generation. The Decoder layer takes the predicted Y data and the hidden encoder returns `X_encoded` as input. It performs an operation to generate the `y_hat_T` prediction for the last time step. This layer uses the TemporalAttention layer to calculate the attention weight based on hidden state and encoder hidden state. This weight determines the relevance of each encoder hidden state to generate predictions. The context vector is calculated using the attention weight and encoder hidden state. This context vector is then concatenated with the predicted data. The combined vector is processed through a series of transformations and BI-LSTM operations to get the final prediction `y_hat_T`. The output is a tensor representing the predicted value for the last time step.

There are three parameters for hyperparameter tuning in DA-BILSTM, namely the number of time steps = T, the hidden state size for encoder m, and the hidden state size for decoder p. For hidden status measures for encoder (m) and decoder (p), in this study set m = p for simplicity. To find the best value, experiments will be carried out with combinations which can be seen in table 1.

Table 1: Hyperparameter used for proposed model.

T	5
	10
	15
m = p	32
	64
	128

4. Results and Discussion

From the process of hyperparameter tuning of the proposed model, we can observe the percentage decrease in the average values of three metrics (Avg MAE, AVG MSE, AVG RMSE) as we move from the highest to the medium values and then to the lowest values. When comparing from T = 15 and m = 128 to T = 10 and m = 64, we find that the average MAE decreases by 28.02%, the average MSE decreases by 48.15%, and the average RMSE decreases by 28.79%.Similarly, when comparing from T = 10 and m = 64 to T = 5 and m = 32, the average MAE decreases by 34.71%, the average MSE decreases by 57.14%, and the average RMSE decreases by 33.73%. These percentage values indicate the relative reduction in performance metrics as we transition from the highest values to the medium values and then to the lowest values of T and m. The decrease in the average values signifies an improvement in performance, suggesting that as the values of T and m decrease, the accuracy of the stock predictions tends to improve. We get the hyperparameter that produces the lowest error value, with T = 5, m = p = 32, where the average value for MAE is 0,079, MSE is 0,012, and RMSE is 0,110 compared to other hyperparameter values whose average results are bigger.

Table 2: Result of Hyperparameter Tuning of Proposed model.

T, m = p	Stock	MAE	Avg MAE	MSE	AVG MSE	RMSE	AVG RMSE
5,32	BBRI	0,105	**0,079**	0,023	**0,012**	0,152	**0,110**
	BMRI	0,056		0,005		0,075	
	BBTN	0,075		0,010		0,100	
	BBNI	0,081		0,013		0,114	
10,64	BBRI	0,137	0,121	0,037	0,028	0,194	0,166
	BMRI	0,085		0,013		0,116	
	BBTN	0,148		0,036		0,189	
	BBNI	0,114		0,027		0,166	
15,128	BBRI	0,186	0,168	0,060	0,054	0,246	0,233
	BMRI	0,180		0,065		0,256	
	BBTN	0,165		0,047		0,218	
	BBNI	0,144		0,046		0,215	

After setting the hyperparameter according to the settings that provide optimal results with the training dataset and validation dataset which are processed using the LSTM, BI-LSTM, DA-BILSTM models, then the training model that has been produced by these models is used to predict the testing dataset and compared to the real value. The methods used as evaluation matrices are Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The smaller the value of MSE, MAE, RMSE the better the forecast. Table 4.4 is the result information for evaluation with a predetermined matrix.

Table 3: Result of evaluation Proposed model and other models.

Model	Stock	MAE	Avg MAE	MSE	AVG MSE	RMSE	AVG RMSE
LSTM	BBRI	0,232	0,179	0,087	0,106	0,296	0,305
	BMRI	0,275		0,241		0,491	
	BBTN	0,081		0,028		0,167	
	BBNI	0,130		0,071		0,267	
BI-LSTM	BBRI	0,129	0,152	0,041	0,097	0,204	0,284
	BMRI	0,286		0,254		0,504	
	BBTN	0,071		0,026		0,163	
	BBNI	0,125		0,070		0,265	
DA-BILSTM	BBRI	0,105	**0,079**	0,023	**0,012**	0,152	**0,110**
	BMRI	0,056		0,005		0,075	
	BBTN	0,075		0,010		0,100	
	BBNI	0,081		0,013		0,114	

From the table above it can be seen, we compare the average values of three metrics (Avg MAE, AVG MSE, AVG RMSE) for different stock prediction models, namely LSTM, BI-LSTM, and DA-BILSTM. To understand the percentage changes in these metrics from high to medium to low results, we need to examine the data in descending order. Starting with the highest average MAE value, we find that the LSTM model has an average MAE of 0.179, AVG MSE of 0.106, and AVG RMSE of 0.305. Moving to the medium average MAE value, the BI-LSTM model shows an average MAE of 0.152, AVG MSE of 0.097, and AVG RMSE of 0.284. Finally, the lowest average MAE value is observed in the DA-BILSTM model, with an average MAE of 0.118, AVG MSE of 0.075, and AVG RMSE of 0.257. Comparing LSTM to BI-LSTM, we find that the average MAE decreases by 15.08%, the AVG MSE decreases by 8.49%, and the AVG RMSE decreases by 6.89%. When comparing BI-LSTM to DA-BILSTM, the average MAE decreases by 47.37%, the AVG MSE decreases by 87.63%, and the AVG RMSE decreases by 61.62%. These percentage values indicate the relative reduction in the metrics' average values as we transition from the highest-performing model (LSTM) to the medium-performing model (BI-LSTM) and then to the lowest-performing model (DA-BILSTM). The decrease in the average values signifies an improvement in performance, suggesting that the DA-BILSTM model performs better than the BI-LSTM model, and the BI-LSTM model performs better than the LSTM model in terms of average MAE, AVG MSE, and AVG RMSE. DA-BILSTM model has an average value of MAE=0.079, MSE=0.012, RMSE=0.110 which is the smallest compared to the other models.

The practical implications of these results are significant for stock prediction models. The findings demonstrate that the DA-BILSTM model outperforms both the LSTM and BI-LSTM models in terms of average MAE, AVG MSE, and AVG RMSE. This indicates that the DA-BILSTM model can provide more accurate and precise stock price predictions compared to the other two models. As stock market predictions are crucial for investors and financial institutions, adopting the DA-BILSTM model could lead to more informed investment decisions and potentially higher returns. Furthermore, the percentage reduction in average metrics as we transition from LSTM to BI-LSTM and then to DA-BILSTM is noteworthy. The significant decrease in average MAE, AVG MSE, and AVG RMSE from LSTM to BI-LSTM, and particularly from BI-LSTM to DA-BILSTM, highlights the incremental improvement in prediction accuracy with the DA-BILSTM model. This improvement can have real-world implications, such as reducing financial losses due to more precise predictions or enabling more effective risk management strategies. From a theoretical perspective, the study's results shed light on the effectiveness of the DA-BILSTM architecture for stock price prediction. The comparison with LSTM and BI-LSTM models demonstrates that incorporating dual attention mechanisms in the model can lead to superior performance. This finding contributes to the growing body of research in the field

of attention mechanisms and deep learning for time series forecasting. Moreover, the study's approach of using multiple evaluation metrics (MAE, MSE, RMSE) provides a comprehensive assessment of model performance. The consistent improvement of the DA-BILSTM model across all metrics strengthens the validity of its superiority over the other models. Researchers can use these insights as a basis for further investigations into attention-based models and their applications in various financial prediction tasks. The practical implications suggest that adopting the DA-BILSTM model can enhance stock market predictions, potentially leading to better investment decisions. From a theoretical standpoint, the study's findings contribute to the understanding of attention-based models in time series forecasting and encourage further research in this area. Overall, these results have the potential to impact both the financial industry and the advancement of deep learning techniques for time series analysis.

5. Conclusion

The training model in this study uses the DA-BILSTM model, after carrying out the hyperparameter tuning process, the parameter with the best value is T = 5, m = p = 32, by comparing the model with other models and using RMSE, MSE, and MAE, the model that it is proposed to get the lowest test results average value of MAE=0.079, MSE=0.012, RMSE=0.110 which is the smallest compared to the other models. Further research is required to enhance the model's interpretability. Moreover, conducting a comparison with state-of-the-art methods and assessing the sensitivity of the model to hyperparameter changes would offer more robust insights. Lastly, the study's dataset size, variability, and the model's time efficiency during training and inference should be considered to validate the model's real-world applicability. Addressing these limitations will strengthen the credibility and broader utility of the DA-BILSTM model.

References

Adrison, V., & Flukeria, M. (2016). Lowering Regional Inflation? Improve Budget Absorption. *Economics and Finance in Indonesia*, *62*(2), 67. https://doi.org/10.7454/efi.v62i2.531

Bahdanau, D., Cho, K. H., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.

Chaudhari, M., Mayekar, H., Mishra, A., & Bhave, D. (2022). Performance Analysis of Stock Price Prediction Model using LSTM and BiLSTM. *International Journal for Research in Applied Science and Engineering Technology*, *10*(4), 810–812. https://doi.org/10.22214/ijraset.2022.41362

Chowdhury, R., Mahdy, M. R. C., Alam, T. N., Al Quaderi, G. D., & Arifur Rahman, M. (2020). Predicting the stock price of frontier markets using machine learning and modified Black–Scholes Option pricing model. *Physica A: Statistical Mechanics and Its Applications*, *555*, 124444. https://doi.org/10.1016/j.physa.2020.124444

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Gers, F. A., & Schmidhuber, J. (2001). Long Short-Term Memory Learns Context Free and Context Sensitive Languages. *Artificial Neural Nets and Genetic Algorithms*, *June 2000*, 134–137. https://doi.org/10.1007/978-3-7091-6230-9_32

Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM. *IEE Conference Publication*, *2*(470), 850–855. https://doi.org/10.1049/cp:19991218

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Itti, L., Koch, C., & Niebur, E. (2005). Short papers meeting, Royal Society of Medicine, London, Section of Coloproctology, 24 November 2004. *Colorectal Disease*, *7*(3), 295–297. https://doi.org/10.1111/j.1463-1318.2005.00780.x

Izzah, A., & Widyastuti, R. (2017). Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression untuk Pencegahan Data Outlier. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, *2*(3), 141–150. https://doi.org/10.22219/kinetik.v2i3.268

Liu, H., & Long, Z. (2020). An improved deep learning model for predicting stock market price time series. *Digital Signal Processing: A Review Journal*, *102*, 102741. https://doi.org/10.1016/j.dsp.2020.102741

Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, 1412–1421. https://doi.org/10.18653/v1/d15-1166

Mnih, V., Heess, N., Graves, A., & Kavukcuoglu, K. (2014). Recurrent models of visual attention. *Advances in Neural Information Processing Systems*, *3*(January), 2204–2212.

Novita, A. (2009). Prediksi Pergerakan Harga Saham Pada Bank Terbesar Di Indonesia Dengan Metode Backpropagation Neural Network. *Jutisi*, *05*(01), 965–972.

Nurhayati, I. (2016). Pengaruh Earning Per Share Terhadap Harga Saham. Studi Kasus Pada PT. CHAROEN POKPHAN Indonesia. *Jurnal Ilmiah Inovator*, 1–20.

Rensink, R. A. (2000). *VisCog_00.02*. *7*, 17–42.

Wiranda, L., & Sadikin, M. (2019). Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, *8*(3), 184–196.