STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **13**, January 2025, pp 368–377.
Published online in International Academic Press (www.IAPress.org)

Forecasting Nonstationary Time Series Based on Dicrete Hilbert Transform

Wahyuni Ekasasmita [1,*], Khaera Tunnisa [2], Muh. Tri Aditya[2]

[1] *Department of Mathematics, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia*
[2] *Department of Information System, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia*

Abstract Various predictive methods have been applied to predict the value of stocks. The purpose of this research is to implement the discrete Hilbert transform in stock returns. The ability to predict stock price movements has big implications for investors. Traditional methods are often limited in capturing the complexity of market dynamics. It was found that the proposed method obtained an average of MAE, RMSE and MAPE values of 0.02055, 0.02237, and 0.012985 which is lower than the conventional LSTM method. This research provides a new understanding of the application of discrete Hilbert transform in a dynamic global financial context.

Keywords Hilbert Transform, Forecasting, Time Series, Machine Learning

AMS 2010 subject classifications 26A33, 37M10, 60B11

DOI: 10.19139/soic-2310-5070-2060

1. Introduction

Share price prediction and financial risk determination are closely linked in the context of financial analysis and portfolio management. Share price forecasting involves an attempt to project the future value or trend of a stock based on historical information and market factors. The purpose of stock price predictions is to help investors and analysts make better investment decisions, such as when to buy or sell stocks. Financial risk assessment involves the evaluation and measurement of potential losses or volatility in investments or portfolios ([20]).

Stock price predictions provide important input in the determination of financial risk. Predictions of the stock price give an idea that a stock has a high volatility potential or a significant risk, this information is taken into account in the financial risk analysis ([21]). Accurate predictions can help portfolio managers and investors more effectively identify and manage potential risks in their investments. If there is a significant change in the stock price forecast, risk management measures can be used to mitigate the impact.

Time cycle data is a quantitative measurement of objects normally in the same sequence in a time period ([1]). Data scientists most often deal with this time cycle information (cite42). Many techniques, from basic to most technical, statistics to soft computing, have been used to understand the underlying structure or properties of time-specific data ([3]). The new field artificial intelegence has been used more precisely recently. A larger group of machine learning that is built on simulated neural networks and learning representations include deep learning ([4]).

One of the methods for predicting time cycle data is using deep neural networks ([5]; [6]). Deep Learning refers to the development of neural netways in machine learning by incorporating deeper layer structures. It has been used in many fields to solve a variety of problems and has produced remarkable results, especially in the fields of deep learning, natural language processing, and image classification ([7]). Recurrent Neural Networks (RNN) is a

*Correspondence to: Wahyuni Ekasasmita (Email: wahyuni.ekasasmita@ith.ac.id). Department of Mathematics, Institut Teknologi Bacharuddin Jusuf Habibie. Parepare, South Sulawesi, Indonesia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

unique technique that has been widely used mainly for regression such as predicting future values from time-long data ([8]). The shortcomings of RNN include long-term dependency problems ([9]) and lost or expanded gradient problems addressed by introducing Long Short Term Memory ([10]).

The quaternion approach in portfolio optimization has been developed by ([17]) to minimize portfolio risk under certain constraints as the extension of the complex portfolio diversification of risk proposed by ([22]) has shown that the proposed method goes beyond the conventional method of diversifying portfolio complex risk. In this study, the DHT-LSTM method was used to predict time around which could significantly improve prediction of time around data using historical stock data.

2. Long Short-Term Memory

Recurrent Neural Network (RNN) is one of the types of algorithms that apply sequential approaches. To overcome the weakness of recurrent neural network, Long Short-Term Memory (LSTM) is introduced, which is a special form of RNN ([12]). LSTM is the name of an RNN technique that has gained widespread acceptance and has been applied in a variety of contexts, especially those involving sequential data.

Compared to RNN, LSTM uses three gate in the cell blocks. Gate is known as input gate (I_g), output gate (O_g), and forget gate (F_g). In LSTM, each port adds values and spreads information. The LSTM calculation cell involves six equations, on the equation (1) to (6) ([11]).

$$F_g = \sigma(W_F h_{g-1} + U_F x_g + b_g). \tag{1}$$
$$I_g = \sigma(W_I h_{g-1} + U_F x_g + b_I). \tag{2}$$
$$\tilde{C}_g = tanh(W_C h_{g-1} + U_C x_g + b_C). \tag{3}$$
$$C_g = f_g \odot C_{g-1} + i_g \odot \tilde{C}_g. \tag{4}$$
$$O_g = \sigma(W_O \odot h_{g-1} + U_O x_g + b_O). \tag{5}$$
$$h_g = O_g \odot tanh(C_g). \tag{6}$$

Cell LSTM uses two activation functions (FA): the sigmoid function (σ) and the hyperbolic tangent function (tanh). Corresponding values for input gates, output gates, forget gates and candidate cell are W_F, W_I, W_C and U_f, U_I, U_O. Similarly, the default values of forget gate, input gate and output gate are b_g, b_i, b_o, b_C. $\tilde{C}_g$ specifies candicate cell, C_g specifies the current space of cell, h_g at the moment the value of the operator is x at the time the current state of the cell is specified. Hyperbolic (tanh) and sigmoid (σ) tangents are activation functions used in cell LSTM ([15]).

2.1. Hilbert Transform

Hilbert transformation of real signal $z(t)$ to $t \in [0, \infty)$ is defined as

$$H[z(t)] = \frac{1}{\pi} \int_0^\pi \frac{z(\tau)}{t - \tau} d\tau. \tag{7}$$

Discreet Hilbert transformation is defined as

$$H_D[z_k] = -isgn\left(k - \frac{N}{2}\right) \sum_{n=0}^{N-1} z_n e^{i\frac{2\pi n}{N}} \tag{8}$$

with $sgn(.)$ signum function ([19]).

2.2. Prosedur

We propose a novel approach by incorporating DHT within Deep Learning methods, particularly LSTM, for time series forecasting. DHT could acs as a transform data technique that contibutes to the overall forecasting results of time series data with high volatility.

Figure 1 shows the schematic diagram of the proposed methods, namely DHT-LSTM. The time series data will be preprocessed to handle any missing data. DHT is applied to the cleaned data before data splitting process with 80% data train and 20% data test. To handle any missing values from DHT calculation, xero data imputation is applied to the train set before deep learning model development.

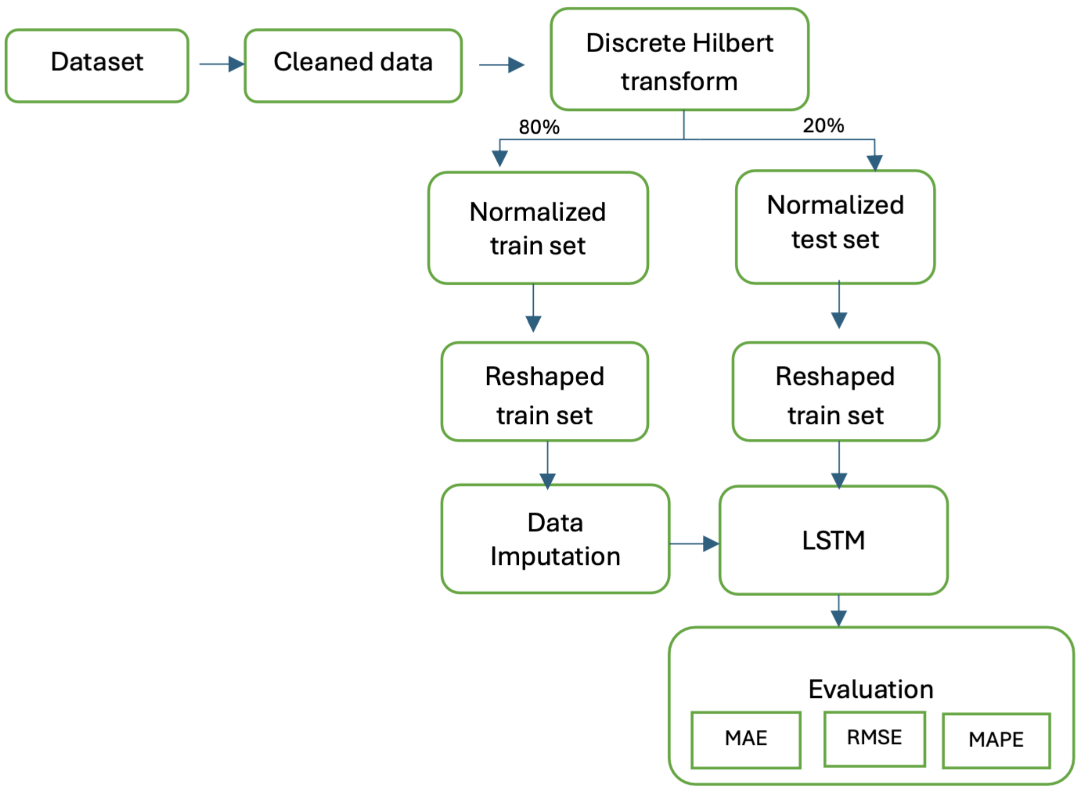

Figure 1. **Schematic diagram of proposed methods**

In this section, the DHT-LSTM procedure is reviewed using the following steps.

1. ${p_t^{(m)}}_{0<t<T}$; stock price m^{th};
2. $r_t^{(m)} = \frac{p_{t+1}^{(m)} - p_t^{(m)}}{p_t^{(m)}}$; return stock price m^{th};
3. Applying discret Hilbert transform to portfolio returns using equations (8) and obtain analytical signals that are then converted to equation signals $x_t = r_t + iH_D[z_t]$;
4. Use LSTM to predict the return on a portfolio.

3. Error Criteria

Mean absolute error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage error (MAPE) are statistical tools used to measure the accuracy of a statistical model in making predictions or forecasts. To perform the triple calculation of this satisticus, use the equation (9) to (10).

$$MAE = \frac{1}{m} \sum_{t=1}^{m} |Ad_t - Fv_t| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{t=1}^{m}(Ad_t - Fv_t)^2} \tag{10}$$

$$MAPE = \left(\frac{1}{m}\sum_{t=1}^{m}\left|\frac{Ad_t - Fv_t}{Ad_t}\right|\right)100\% \tag{11}$$

With Ad_t is the actual data and Fv_t is the prediction data value.

4. Eksperiment

The time period data used is Yahoo Finance Statistics data. The information provided shows the actual data. Nucor Corporation, Cleveland-Cliffs Inc. and Silver Sep from January 2019 to January 2023. The stock data is given in the following table (1).

Table 1. **Descriptive Data**

Descriptive	CLF	NUE	SI=F
Count	1010	1010	1010
Mean	13.8105	76.8113	20.9317
Std	6.90822	36.4555	4.20206
Min	3.02152	26.0112	11.7350
25	7.69739	46.6178	17.3580
50	13.5950	53.5319	21.4980
75	19.6174	108.748	24.5677
max	33.0700	170.601	29.3980

Python is used to train and test all deep learning models in training and testing. The LSTM and HDT-LSTM methods are used to predict time period data. All deep learning techniques use the same design, which is a simple three-layer architecture, as described in the previous section. In addition, two different stock data, Nucor Corporation, Cleveland-Cliffs Inc and Silver Sep are used to run each test. The table (2) shows the calculations.

Table 2. **Performance Measurement Method of Stock Price Movement**

Metode	LSTM		
	CLF	NUE	SI=F
MAPE	0.05754	0.04371	0.61398
RMSE	1.44109	7.42675	0.75949
MAE	1.09478	5.62356	0.02966

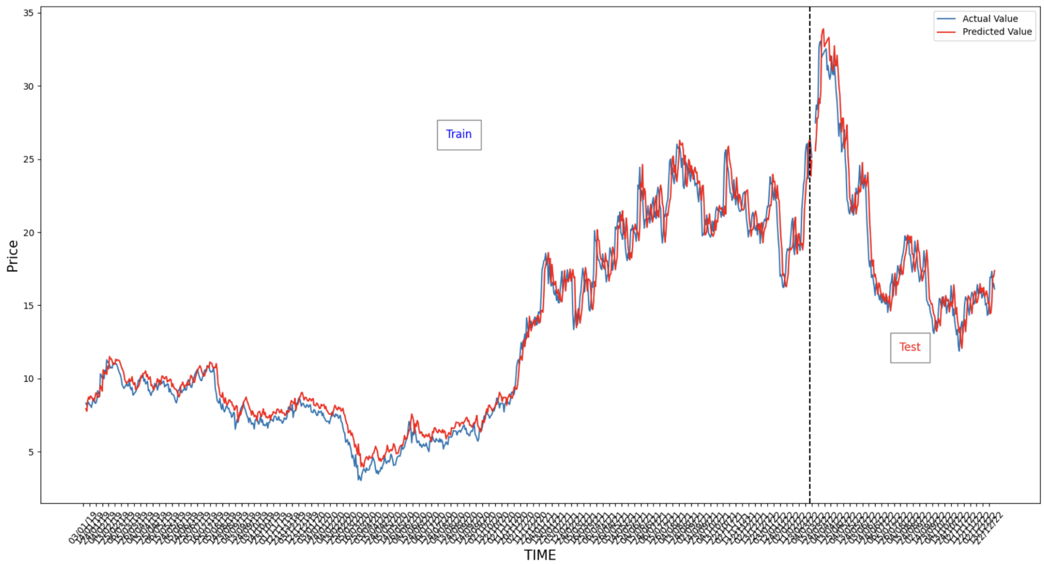

Figure 2. **Predict CLF stock price using LSTM**

Figure 2 illustrates the movement of CLF stock return using LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of CLF 1.09478, 1.44109, and 0.05754.

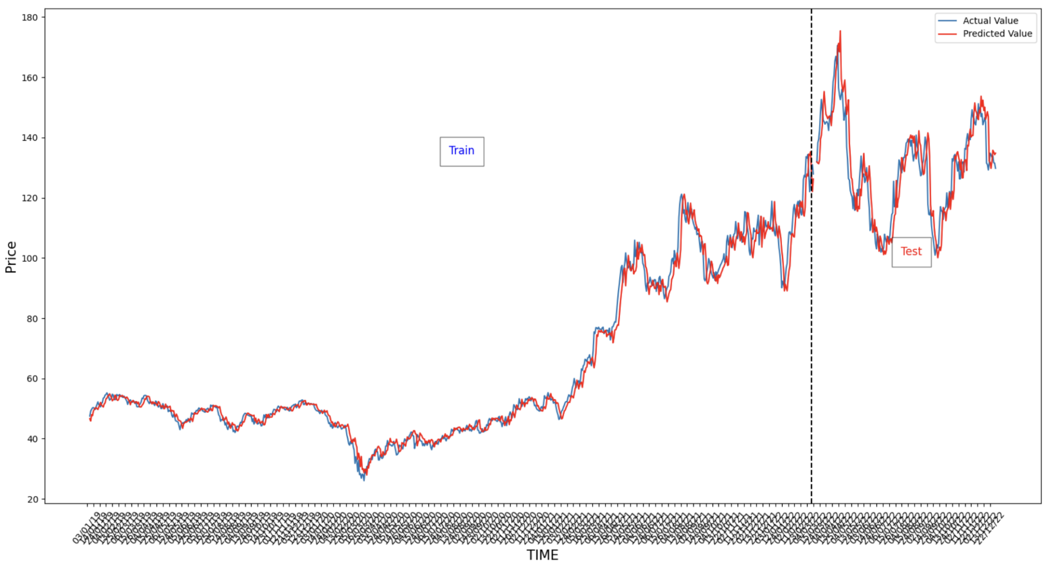

Figure 3. **Predict NUE stock price using LSTM**

Figure 3 illustrates the movement of NUE stock return using LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of NUE 5.62356, 7.42675, and 0.04371.

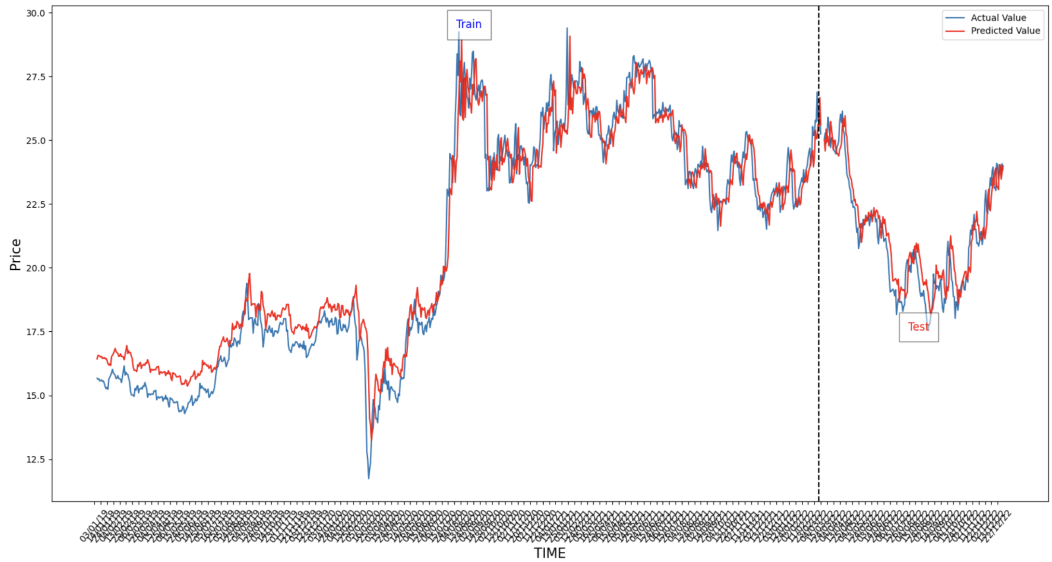

Figure 4. **Predict SI=F stock price using LSTM**

Figure 4 illustrates the movement of SI=F stock return using LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of SI=F 0.61398, 0.75949, and 0.02966.

Table 3. **Performance Measurement Method of Stock Return Movement**

Metode	LSTM			DHT-LSTM		
	CLF	NUE	SI=F	CLF	NUE	SI=F
MAPE	0.16040	1.32956	1.22311	0.02530	0.00067	0.00033
RMSE	0.04024	0.02942	0.75884	0.03213	0.01261	0.02013
MAE	0.03203	0.02258	0.61220	0.02702	0.01408	0.01550

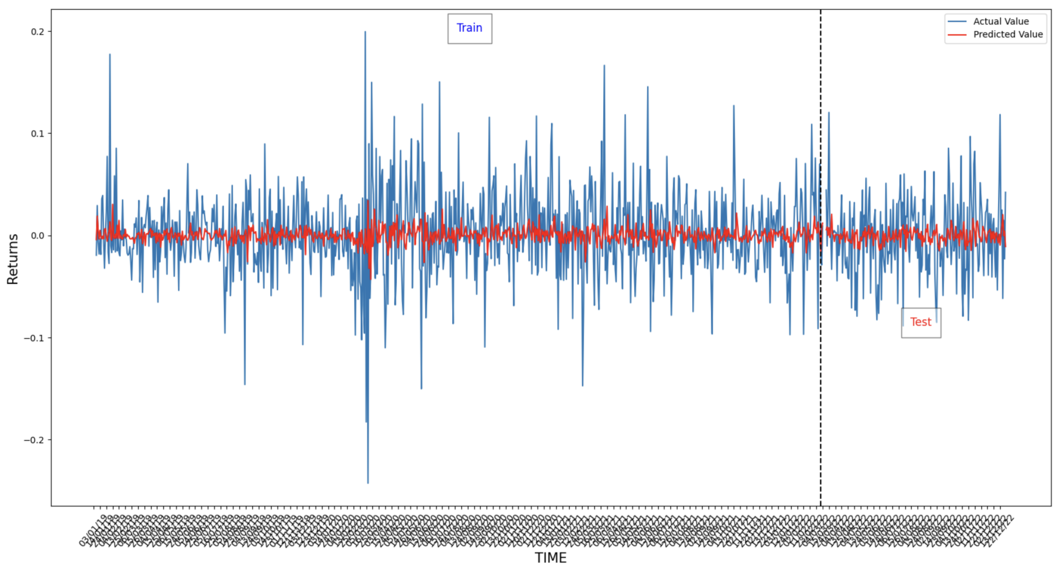

Figure 5. **Predict stock returns of CLF using LSTM**

Figure 5 figure out the movement of CLF stock return using LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of 0.03203, 0.04024, and 0.16040.

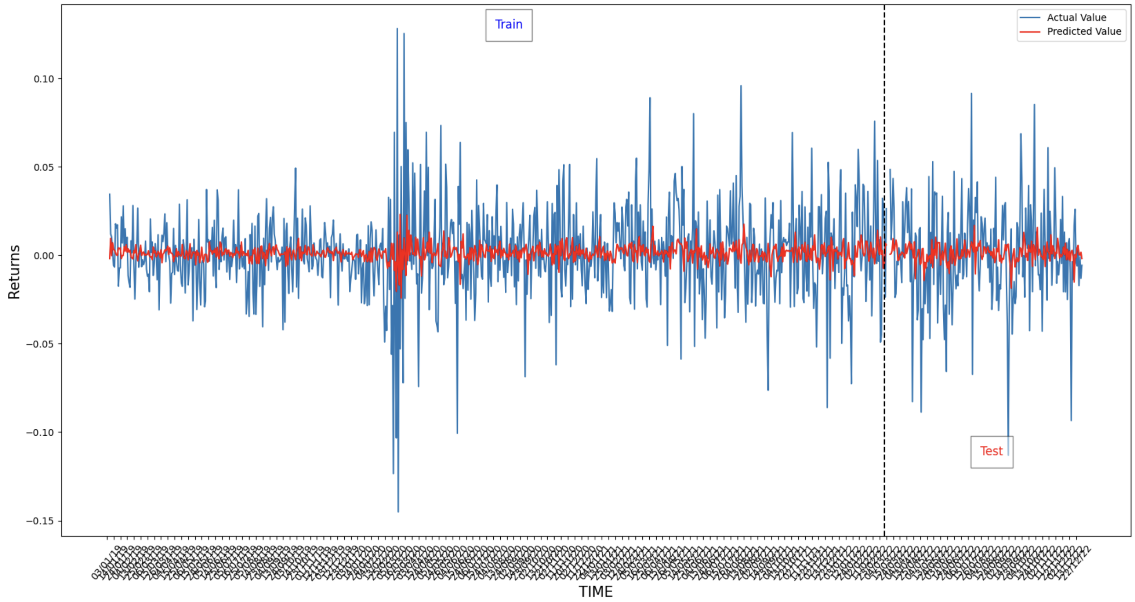

Figure 6. **Predict stock returns of NUE using LSTM**

Figure 6 figure out the movement of NUE stock return using LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of NUE is 0.02258, 0.02942, and 1.32956.

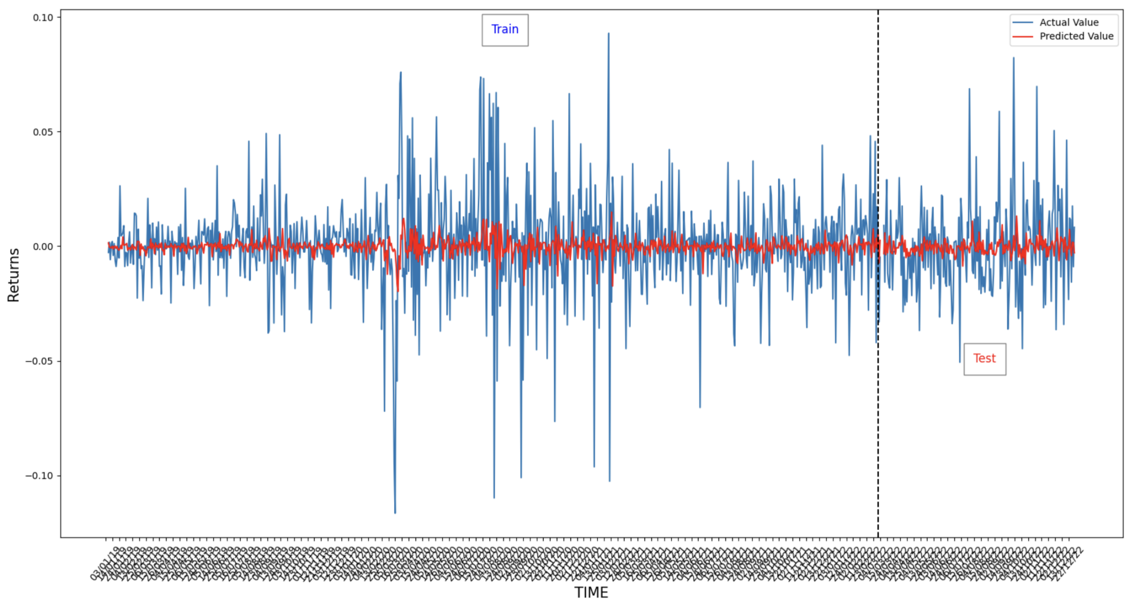

Figure 7. **Predict stock returns of SI=F using LSTM**

Figure 7 figure out the movement of SI=F stock return using LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of SI=F is 0.61220, 0.75884, and 1.22311.

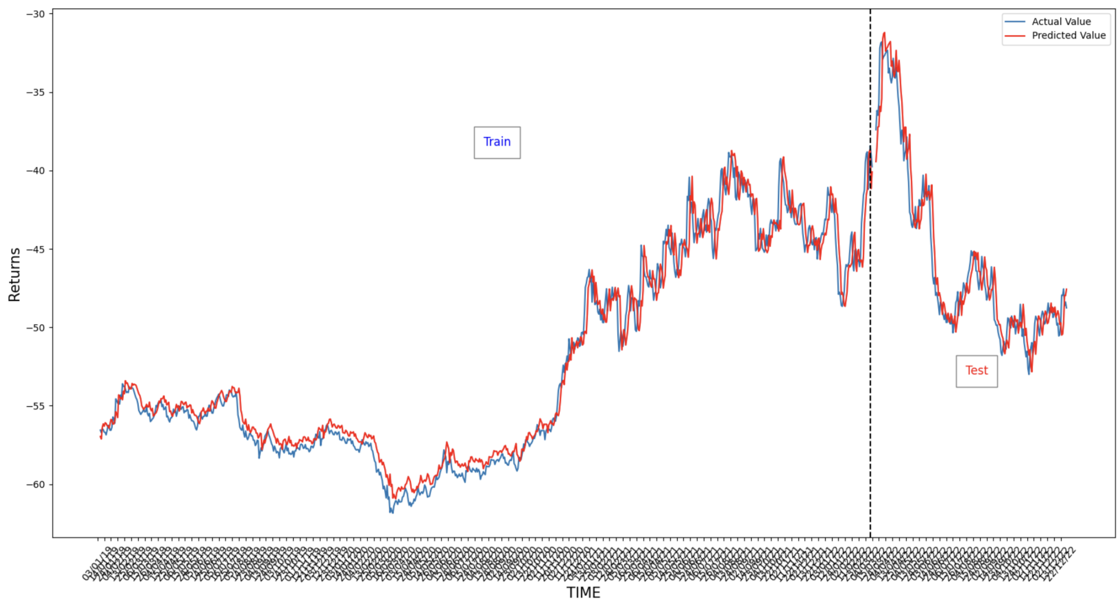

Figure 8. **Predict stock returns of CLF using DHT-LSTM**

Figure 8 figure out the movement of CLF stock return using DHT-LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of CLF is 0.02702, 0.03213, and 0.02530.

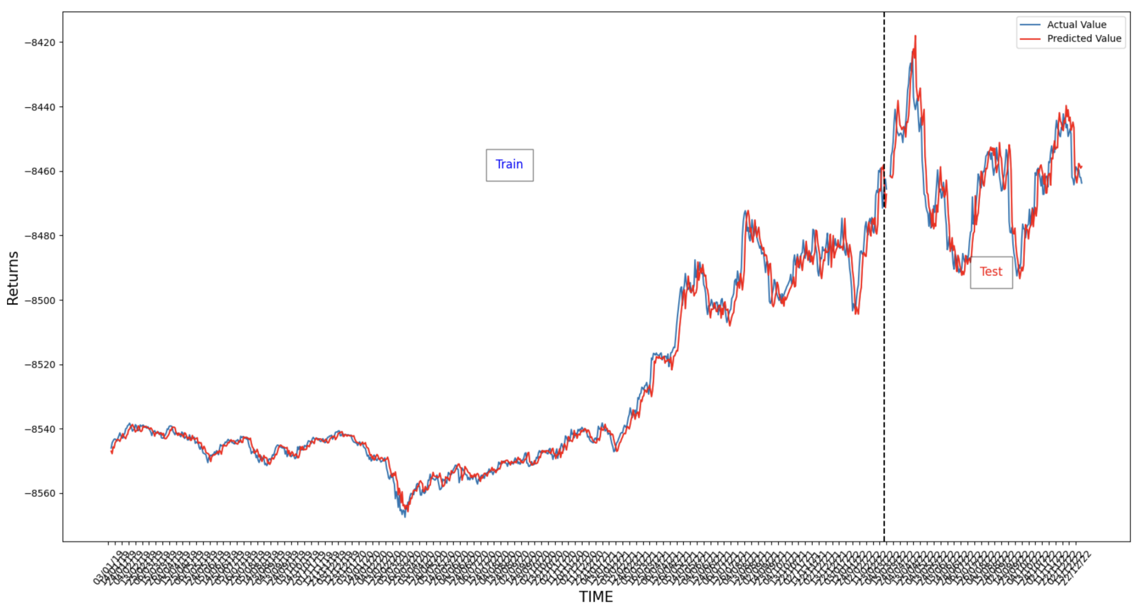

Figure 9. **Predict stock returns of NUE using DHT-LSTM**

Figure 9 figure out the movement of NUE stock return using DHT-LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of NUE is 0.01408, 0.01261, and 0,00067.

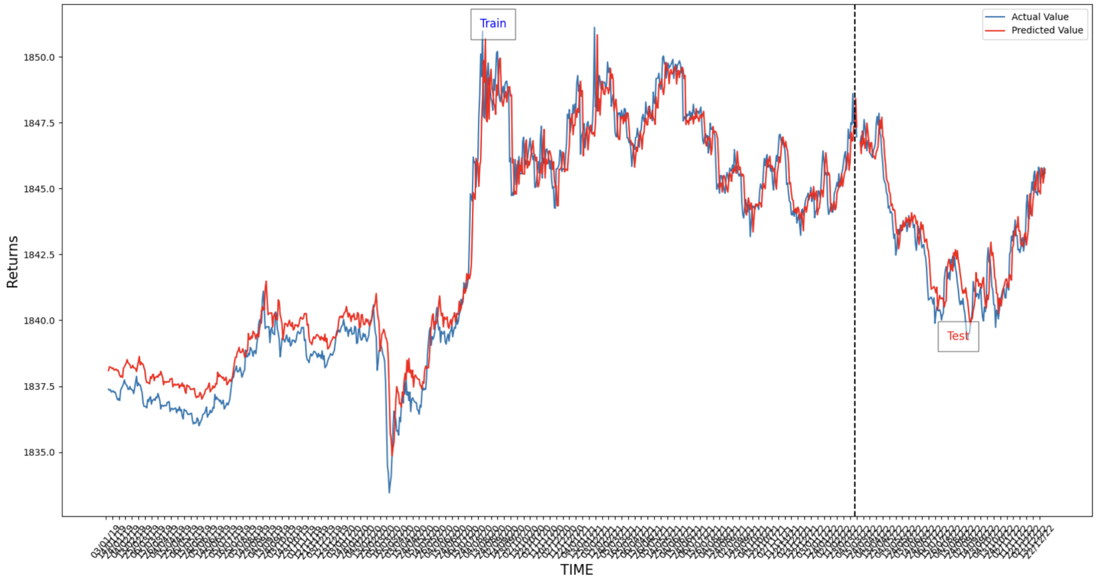

Figure 10. **Predict stock returns of SI=F using DHT-LSTM**

Figure 10 figure out the movement of SI=F stock return using DHT-LSTM displaying actual values and predictive values with MAE, RMSE, and MAPE values of SI=F is 0.01550, 0.02013 and 0,00033.

The proposed approach applied to the Nucor Corporation (NUE) and Cleveland-Cliffs Inc. (CLF), and Silver Sep (SI=F). in the table 2 had error values using DHT-LSTM obtained averages of MAE, RMSE, and MAPE being 0.01886, 0.02162, and 0.00876 lower than the LSTM used in this study with averages for MAE, RMSE, and MAPE being 0.22227, 0.27616, and 0.90435. The prediction results 5 and 8 showed that each deep learning used successfully followed the same data patterns as 7 and 9. Furthermore, given how little data was used in the experiment, the prediction could be said to be very good.

Google Colab is used for training and testing all deep learning models. The experimental configuration makes use of the Python programming language and several essential libraries, including Numpy, Matplotlib, Pandas, Keras, and sklearn. You may find the source code at https://github.com/jiunmath/dht-lstm/upload, the GitHub repository.

5. Conclussions

The three error criteria applied in this study are MAE, RMSE and MAPE. Using DHT-LSTM obtained averages of MAE, RMSE, and MAPE being 0.01886, 0.02162, and 0.00876 are lower than the LSTM used in this study with averages for MAE, RMSE, and MAPE being 0.22227, 0.27616, and 0.90435. The lower the score of each criterion that shows a reduction in prediction error between actual and prediction values, the better the prediction system. Using discrete Hilbert transform with deep learning techniques can improve predictions.

Acknowledgement

This work was supported by Sains Department, Institut Teknologi Bacharuddin Jusuf Habibie, and LPPM-PM ITH (Research, Community Service, and Quality Assurance Institution).

REFERENCES

1. OECD. *Data and Metadata Reporting and Presentation Handbook*. Paris, France: OECD Publishing, 61. 2007.
2. Hansun, S., Charles,V., Indrati,C. R., and Subanar. *Revisiting the Holt-Winters' Additive Method for Better Forecasting*. International Journal of Enterprise Information Systems 15(2): 43–57. DOI: 10. 4018/IJEIS.2019040103. 2019.
3. Kureychick,V. M. and Kaplunov,T. G. *Time series forecasting method based on genetic algorithm for predicting the conditions of technical systems*. Journal of Physics: Conference Series 1333(3): 032046. DOI: 10.1088/1742-6596/1333/3/032046. 2019.
4. Adamson,A. S. and Smith, A. *Machine learning and health care disparities in dermatology*. JAMA Dermatology, vol. 154, no. 11, pp. 1247-1248. 2018.
5. Paras, S. M., Kumar, A., and Chandra, M. *A feature based neural network model for weather forecasting*. International Journal of Computational Intelligence, vol. 4, no. 3, pp. 209–216. 2009.
6. Liu S, Liao G, Ding Y. *Stock transaction prediction modeling and analysis based on LSTM*. In: 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA) [Internet]. Wuhan, China: IEEE; p. 2787–90. Available from: https://ieeexplore.ieee.org/document/8398183/. 2018. Abstr. Appl. Anal. 2013, 2013, 162769.
7. Lim,B. and Zohren,S. *Time-series forecasting with deep learning: a survey*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 379(2194): 20200209. DOI: 10.1098/rsta.2020.0209. 2021.
8. Hansun, S., and Young, J. C. *Predicting LQ45 Financial Sector Indices using RNN-LSTM*. Journal of Big Data 8(1): 104. DOI: 10.1186/s40537-021-00495-x. 2021.
9. Shahid, F. ,Zameer, A. , and Muneeb, M. *Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM*. Chaos, Solitons Fractals 140:110212. DOI: 10.1016/j.chaos.2020.110212. 2020.
10. Du J, Liu Q, Chen K, Wang J. *Forecasting stock prices in two ways based on LSTM neural network*. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) [Internet]. Chengdu, China: IEEE; pp. 1083–6. Available from: https://ieeexplore.ieee.org/document/8729026/. 2019.
11. Sepp, H and Jürgen, S. *Long short-term memory*. Neural computation, 9(8): 1735–1780. 1997.
12. S. Hochreiter and J.Schmidhuber. *Long Short Term Memory*. Neural Computation, 9(8) :1735-1780. 1997. DOI: 10.1162/neco.1997.9.8.1735.
13. Larry R. M. and Lakhmi J. J. *Recurrent neural networks*. Design and Applications, 5. 2001.
14. Minemoto,T., Isokawa,T., Nishimura, H., and Matsui, M. *Feed forward neural network with random quaternionic neurons*. Signal Processing, 136:59–68. 2017.
15. Titouan, P.,Mohamed, M, and Linarès, G. *Quaternion denoising encoder-decoder for theme identification of telephone conversations*. Proc. Interspeech, pp. 3325–3328, 2017a. 2017.
16. Ekasasmita, W., Bahri, M., Bachtiar, N., Rahim, A., Nur, M. *One-dimensional Wuaternion Fourier Transform with Application to Probability Theory*. Symmetry 2023, 15, 815, https://doi.org/10.3390/sym15040815. 2023.
17. Sugitomo, S. and Maeta, K. *Quaternion valued risk diversification*. Entropy, 22(4):390. 2020.
18. Titouan, P., Mirco, R., Mohamed, M., Georges, L., Chiheb, T., Renato, D.M, Yoshua, B. *Quaternion Recurrent Neural Networks*. ICLR. 2019.
19. Kak, S. *The discrete Hilbert transform*. Proc. IEEE 1970, 58, 585–586
20. Li, Y., Zhu, Z., Liu, Z., Zhou, Z., and Guo, Y. *Quantitative trading strategy and portfolio optimization analysis basen on LSTM mode*l. Highlights in Business, Economics and Management. 2023, 14.
21. Chen, M.,Shen, Z., and Xu, W. *Portfolio Construction Based on Combination of Individual Stocks and LSTM*. BCP Business Management. FIBA 2023, 44.
22. Yusuke, U., Takanori, K., and Kei, N. *Complex Valued Risk Diversification*. Entropy 2019, 21, 119.