Seminar Presentation

December 31th, 2024

Engineering Intelligence: Developing an Accurate Prediction Model

Presenter: **Thang Le Quoc**,

Logistics Automation Laboratory

Department of Logistics, Korea Maritime & Ocean University

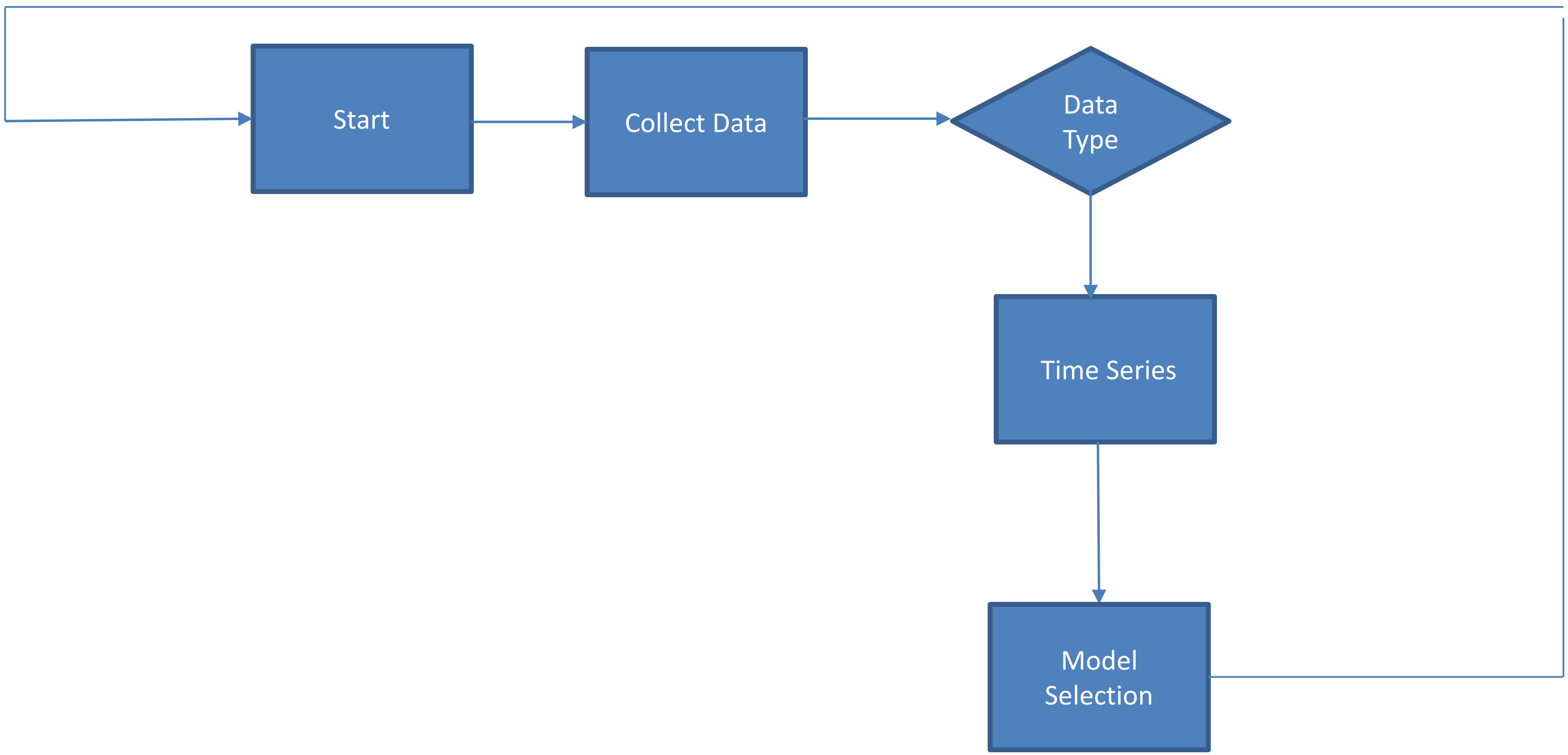
Contents

1. Introduction

"Forecasting Model Development"

By using a forecasting model, companies may better plan, manage resources more efficiently, and make better decisions by predicting future trends and uncertainties.

Performance Workflow



1. Introduction

Background

- The earliest forms of forecasting were rudimentary and based on astrology and patterns observed in nature. Farmers predict the Nile's flooding based on environmental cues.

Motivation

- Developing forecasting models stems from the inherent need to predict future events, allowing for risk management and resource optimization.

Proposed method

- The predictive model will be developed and implemented using the Streamlit library, which provides a seamless and efficient framework for building interactive and user-friendly applications.

In this study, we propose the development of a comprehensive prediction system that leverages the powerful capabilities of the Streamlit library.

2. Methodology

"Parameters"

Input Layer	Units representing the initial input features to the network.
Weight Parameters	Coefficients that multiply the input variables and their interactions.
Bias Terms	Constants added to the weighted sum of inputs to adjust the output.
Activation Function	Non-linear functions applied to the weighted sums of inputs to introduce non-linearity.
Output Layers	Units representing the final output predictions of the network.
Learning Rate	Hyperparameter controlling the step size during the optimization process.
Regularization Parameters	Techniques to prevent overfitting by adding constraints or penalties to the model.
Loss Function	Measure of the discrepancy between the predicted values and the actual values

2. Methodology

“ HOON function ”

- High-order neural networks (HOONs) provide a powerful extension to traditional neural networks by incorporating higher-order combinations of input features. This allows them to capture more complex relationships within the data, potentially leading to better performance on tasks with intricate dependencies. However, this increased power comes with additional computational costs and the risk of overfitting, necessitating careful design and regularization.

2. Methodology

"Model Formulation" – The HOON model:

- Weighted Sum with Bias (z_i):

$$z_i^{(1)} = \Sigma_{i=1}^{n} w_{ij}^{(1)} w_j + a_j^{(1)}$$

- High-Order Terms:

$$z_i^{(2)} = \Sigma_{i=1}^{n} \Sigma_{k=j}^{n} w_{ijk}^{(2)} x_j x_k$$

$$z_i^{(3)} = \Sigma_{i=1}^{n} \Sigma_{k=j}^{n} \Sigma_{l=k}^{n} w_{ijkl}^{(3)} x_j x_k x_l$$

- Hidden Layer (l)

$$z_i^{(l)} = \Sigma_j w_{ij}^{(l)} a_j^{(l-1)} + b_i^{(l)} + \Sigma_{j,k} w_{ijk}^{(l)} a_j^{(l-1)} a_k^{(l-}$$

- Activation Function (a_i):

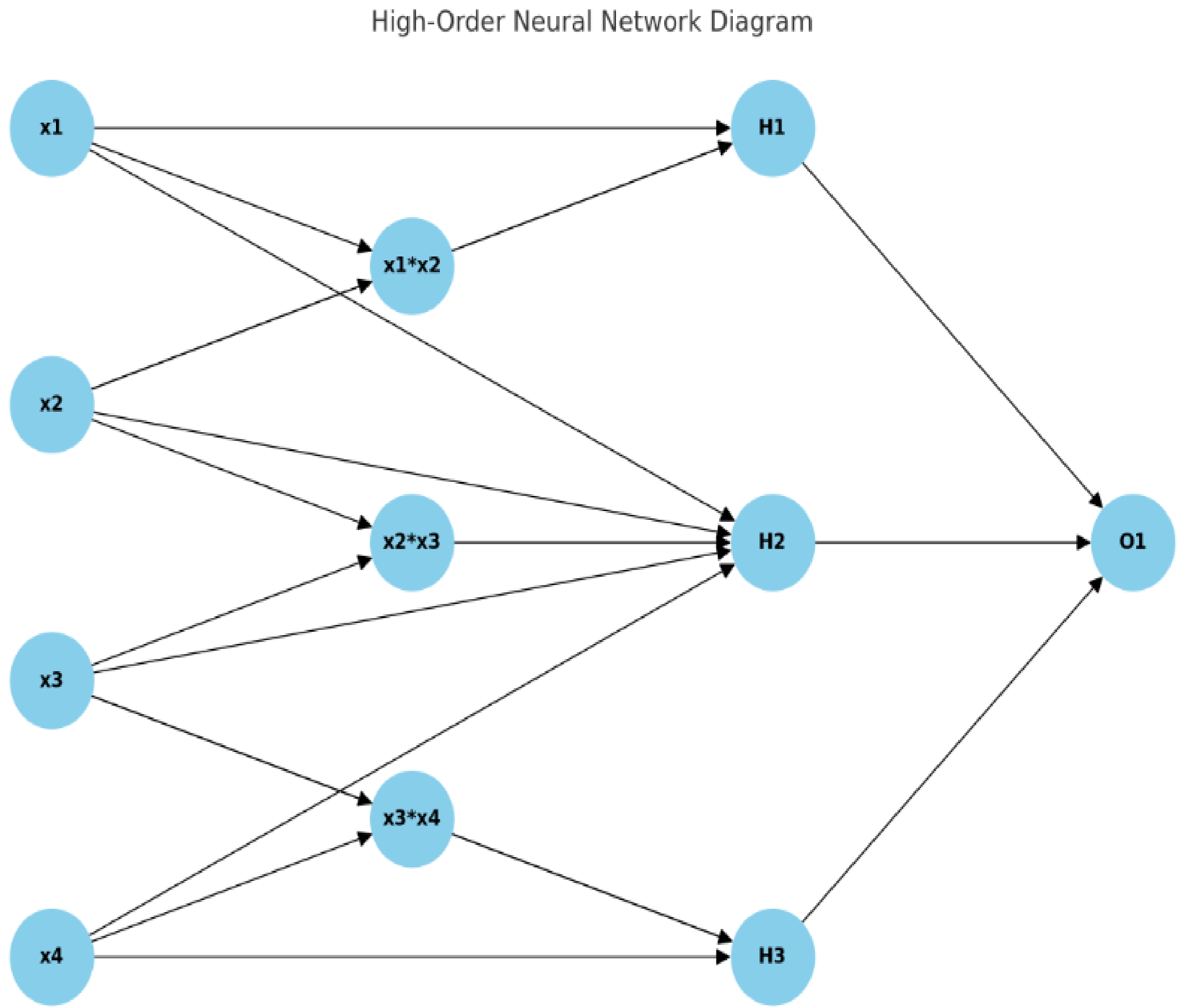
$$a_i = f(z_i)$$

- Output Layer (y)

$$y = \Sigma_i w_i^{(out)} a_i^{(L)} + b^{(out)}$$

- Regularization Terms (λ)

$$L_{total} = L + \lambda_1 \Sigma_i |w_i| + \lambda_2 \Sigma_i w_i^2$$



3. Stimulation

"Explanation data frame"

- The data used for this stock analysis is being sourced from Yahoo Finance, a comprehensive platform that provides a wide range of financial information, including historical stock prices, market trends, and other related metrics.

3. Stimulation

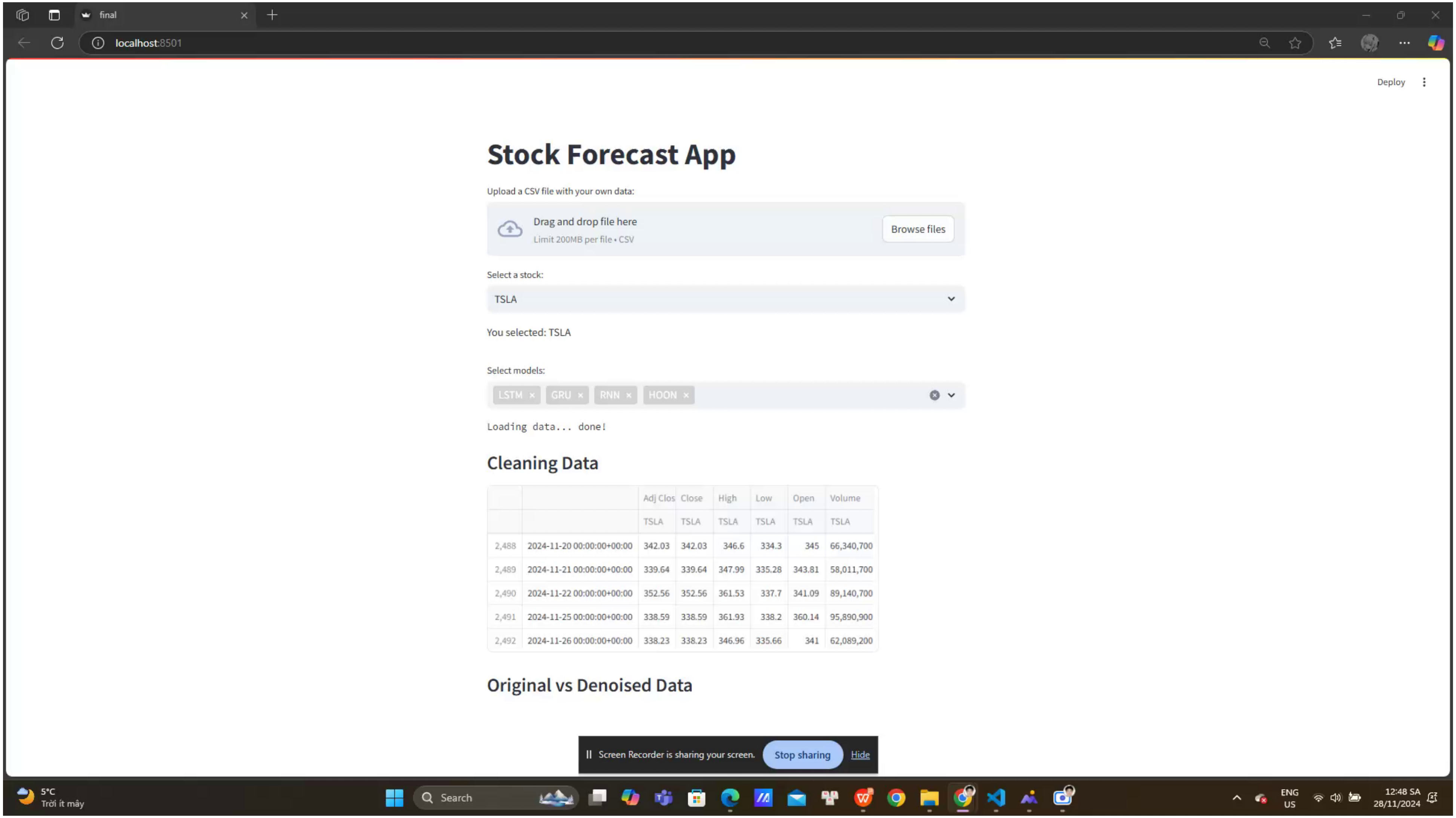
"Explanation training parameters"

- Batch size is the number of samples processed before the model updates its internal parameters. Smaller batch sizes make updates more frequent, while larger batch sizes are more memory-efficient.
- Epoch is an one complete pass through the entire training dataset. Training typically involves multiple epochs to allow the model to learn effectively.
- Verbose is a setting that controls the level of detail displayed during training.

Figure 1. Performance Training Parameters

Training Parameters	Hyperparameters
Epochs	32
Batch Size	50
Verbose	1

3. Stimulation



4. Result

"Insights"

- The denoising process was effective in reducing noise while preserving the overall structure and trends of the data.
- The denoised data (red line) is likely better suited for use in machine learning models or other predictive tasks, as it removes extraneous noise that could negatively impact model performance.

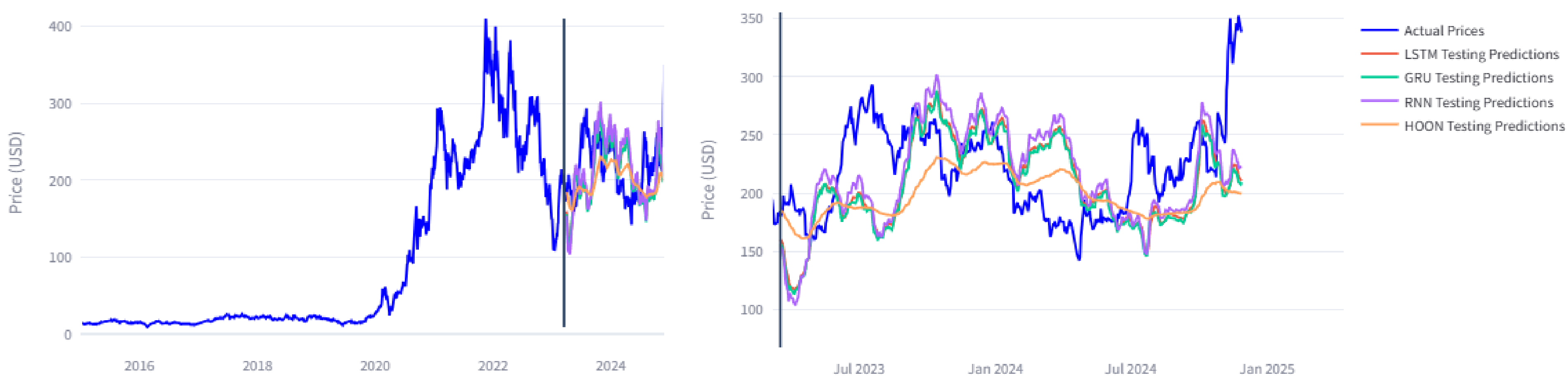
Figure 2. Comparing Denoised Data with Original Data

4. Result

“Insights”

- Based on the visual inspection, the HOON model (orange) seems to perform better, as it closely follows the actual prices.
- All models seem to capture the general upward or downward trends in the stock prices but struggle with finer fluctuations.
- The more closely a model follows the blue line (actual prices), the more reliable it is for prediction tasks.

Figure 3. Prediction Comparision



4. Result

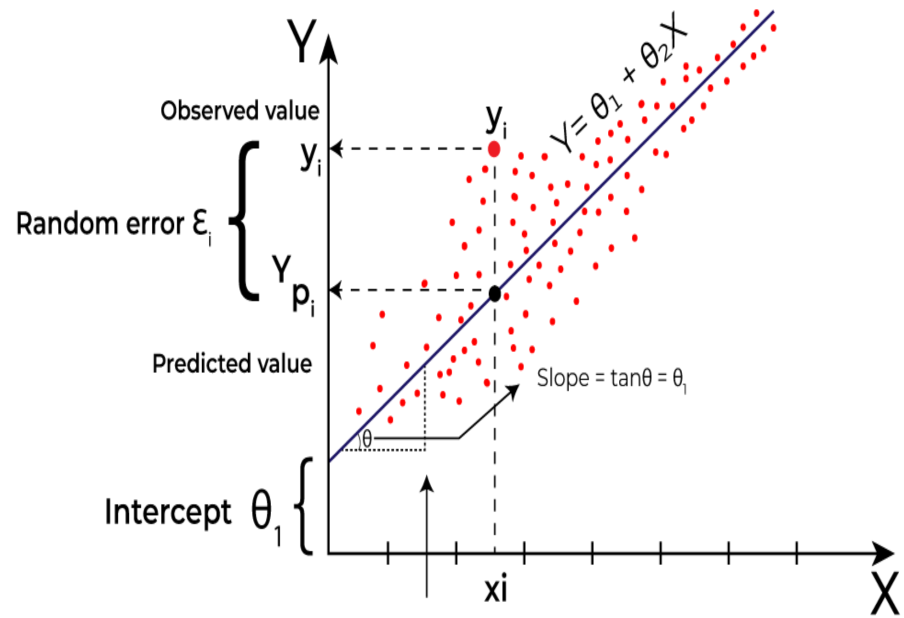
"Explanation formular"

- In this prediction, MAE, MSE, and RMSE are the parameters that show the rate of prediction can fail.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

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

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



4. Result

"Explanation chart visualization"

- As shown in Figure 4 Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) are popular metrics used to measure the accuracy of prediction models in statistics and machine learning. These metrics evaluate how close the predictions of a model are to the actual outcomes. It can be understood that the lower the number the more accurate the prediction.

Figure 4. Performance Model Metrics

Models	Metrics		
	MAE	MSE	RMSE
LSTM	213	47356	217
GRU	211	46495	215
RNN	221	50920	225
HOON	199	40183	200

5. Conclusion

"Summary of the Study"

- This article explores the performance of forecasting models using the Streamlit library, an open-source framework designed to transform data-driven scripts into interactive web applications.

"Discussion"

- The development of forecasting models plays a vital role across various industries, focusing on enhancing the accuracy of future prediction by using historical data.

"Future Work"

- Harnessing the potential of AI to continuously enhance the capabilities of forecasting models.

References

- Goriveau, Rafael; Medjaher, Kamal; Zerhouni, Noureddine (2016-11-14). From prognostics and health systems manage ment to predictive maintenance 1 : monitoring and prognostics. ISTE Ltd and John Wiley & Sons, Inc. ISBN 978-1-84821-937-3.
- Mobley, R. Keith (2002). An introduction to predictive maintenance (2nd ed.). Butterworth-Heinemann. pp. 4–6. ISBN 9 78-0-7506-7531-4.
- Wierstra, Daan; Foerster, Alexander; Peters, Jan; Schmidhuber, Juergen (2005). "Solving Deep Memory POMDPs with Re current Policy Gradients". International Conference on Artificial Neural Networks.
- Amruthnath, Nagdev; Gupta, Tarun (February 2018). "Fault Class Prediction in Unsupervised Learning using Model-Base d Clustering Approach".
- Mather, D. (2008). "The value of RCM". Plant Services
- Kennedy, Sheila (2006). "New tools for PdM". plantservices.com. Putman Media. Retrieved 19 Nov 2019.
- McCulloch, W; Pitts, W (1943). "A Logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathematical Bio physics. 5 (4): 115–133. doi:10.1007/BF02478259.
- Hebb, D.O. (1949). The Organization of Behavior. New York: Wiley & Sons.
- Bain (1873). Mind and Body: The Theories of Their Relation. New York: D. Appleton and Company.
- Levitan, Irwin; Kaczmarek, Leonard (August 19, 2015). "Intercellular communication". The Neuron: Cell and Molecular Biology (4th ed.). New York, NY: Oxford University Press. pp. 153–328. ISBN 978-0199773893.

Thank you for your attention!

QUESTIONS

Q&A

ANSWERS