Master thesis: Smooth Sailing with Neural Networks Improving Accuracy and Transparency in Freight Premium Forecasting

Abstract:

This thesis investigates how neural networks can be used to improve the accuracy and transparency of regional freight rate premium predictions in the Handysize dry bulk shipping market. While previous studies have applied neural networks to forecast movements in the Baltic Handysize Index (BHSI), limited attention has been given to predicting the difference between the index and regional spot rates, also referred to as the regional premium. Predicting the premium is challenging because the same events that influence the index often affect regional spot rates, though not always in the same way. Building on earlier work that found FFNNs outperformed both statistical and traditional machine learning models in this context, this thesis extends the approach through a case study in collaboration with Navi Merchants. We focus on a single Handysize trade route from Europe to the US Gulf, exploring a broader set of explanatory features and specialized neural network architectures for sequential data, including LSTM, GRU, and one-dimensional CNNs. Model performance is benchmarked against a simple FFNN baseline to evaluate the benefits of sequenceaware architectures. The aim is to improve predictive performance and enhance transparency by using Monte Carlo Dropout to estimate model uncertainty and Integrated Gradients to identify how input features influence predictions. These contributions are intended to support ship brokers in making more informed and confident vessel allocation decisions. Our best-performing model was a one-dimensional CNN trained on a combined dataset of macroeconomic indicators, ship location data, and geopolitical risk measures. It achieved a test RMSE of 1,343, representing a substantial improvement over the baseline. Monte Carlo Dropout produced 95% prediction intervals with a typical range of 2,000 to 4,000 on either side and a coverage rate of 92.5% on the test set. Integrated Gradients revealed that AIS-based vessel location data were the most influential features, followed by geopolitical risk indicators. Sequential neural network architectures show promise, although LSTM and GRU models tended to overfit, likely due to their complexity. The superior performance of CNNs suggests that short-term patterns are more important for premium prediction than longer-term dependencies. Integrated Gradients provided valuable insight into how each feature influenced predictions, helping to mitigate the black-box nature of neural networks. In contrast, the wide uncertainty bands produced by Monte Carlo Dropout, combined with a near-zero premium throughout the test period, made it difficult to identify any clearly profitable allocation windows for the selected route.

conclusion:

This thesis set out to investigate how regional freight rate premiums can be predicted accurately and transparently under volatile market conditions. To address this question, we conducted a case study on the Europe to US Gulf Handysize route and evaluated a range of neural network architectures suited for sequential data. The work aimed to improve both predictive performance and model transparency. To meet the first objective, we trained several models and compared their performance to a simple feedforward neural network baseline. The best results came from a one-dimensional convolutional neural network trained on macroeconomic indicators, AIS vessel location data, and geopolitical risk measures. This model achieved a test RMSE of 1343, which marked a clear improvement over the baseline. The CNN outperformed more complex models like LSTM and GRU, which showed stronger signs of overfitting. This suggests that short-term patterns may be more useful for this task than long-term dependencies. To improve transparency, we applied Integrated Gradients and Monte Carlo Dropout. Integrated Gradients provided meaningful explanations by showing how each feature influenced the model's predictions. AIS data and geopolitical risk were identified as the most important inputs. This helped reduce the black-box nature of the model. Monte Carlo Dropout produced wide prediction intervals, in some cases up to 6,000. Since the premium stayed close to zero throughout the test period, it was difficult to identify any clearly profitable vessel allocation windows. Still, changes in the uncertainty bands gave useful signals about whether the model was seeing familiar patterns in the data. While this thesis shows that neural networks, combined with interpretability methods, can improve premium forecasting for a specific Handysize route, the findings are 149 Master Thesis subject to important limitations. The small dataset and periods of extreme volatility made it difficult for models to generalize across time, and overfitting remained a challenge despite regularization. Still, the results offer Navi Merchants a clearer view of which features drive predictions and how model performance varies under different market conditions. This work also provides a foundation for future research into more targeted data sources, broader route generalization, and the development of decision support tools.