Research Exhibition in Mathematics and Computer Sciences (REMACS 6.0)/
ISBN: 978-629-97440-5-4

IMPROVE FORECAST ACCURACY BY USING REPEATED TIME SERIES CROSS VALIDATION (RTS-CV): A CASE STUDY OF DIGI SHARE PRICE

Nurul Izzanie Muhammad Zaki and Azlan Abdul Aziz
College of Computing, Informatics and Mathematics,
Universiti Teknologi MARA, Perlis Branch
zakiezanie@gmail.com and azlan172@uitm.edu.my

ABSTRACT - Forecasting is the entire process of creating the essential techniques to produce future values that can then be utilised as inputs suited to the aims and objectives of the company (Alias, L. 2011). Based on data from prior experience, forecasters can make precise choices for the near future. The goal of this study is to make predictions about the share price of Digi Telecommunications Sdn. Bhd. (DIGI) on the stock market. Azlan Abdul Aziz (2021) acted as the representative for this approach. Given that there are many instances of erroneous predicting, this strategy is an improvement over the prior one. Inaccurate predictions of future values will result in poor decision-making and, even worse, might cause investors and stockholders to become fearful. The models that produce the lowest error measures were collected and compared to decide the most excellent predictions for this study. The next step is to distinguish the most excellent performance from the five models utilized in this consideration. This study was conducted to predict Digi's share price on a daily, weekly, and monthly basis for data high and low only from May 9, 2006 to May 2, 2023. Outcomes for univariate time-series investigation models, such as the Naive, Mean, Single Exponential Smoothing (SES), Holt's, and ARIMA models, were analyzed. Five sets of data splits (high and low) were used for each model to ensure the accuracy of the predicted values. Moreover, the smallest error for instance RMSE, MAE, MAPE, and MASE, are vital in deciding the demonstrative execution, with lower values showing more productive prescient model.

Keywords: Improve forecast accuracy, Repeated Time Series Cross Validation (RTS-CV), share price, digi telecommunications, R Studio

1. INTRODUCTION

In achieving the desired accuracy, time-series forecasting has become a binding domain and developing a forecasting framework with a high level of accuracy is one of the most challenging tasks in the field. In this study, the focus is on predicting the stock market's share price of Digi Telecommunications Sdn. Bhd. (DIGI). Accurate forecasting is important to investors, especially to domestic and international investors. Poor forecast value estimates lead to losses to investors, especially to major shareholders of the companies listed on Bursa Malaysia. Among the popular criteria that can be used to measure the performance of the models stated is to see the lowest error measurement value. Apart from the use of the correct model and criteria, another method that can be used in producing accurate forecasts is to use the Repeated Time Series Cross-Validation. Then, five time-series models, Naïve, Mean, SES (Single Exponential Smoothing), Holt's and ARIMA, were analysed into the R Studio software. However, the "winning" model or set of data splits is chosen based on the model that produces the smallest error measure in the evaluation part (M.A. Lazim, 2011).

2. METHODOLOGY

This method is represented by Azlan Abdul Aziz (2021). It is a better version of the previous method, since there is numerous false forecasting. Inaccurate forecasts of future values will lead to incorrect decision-making and even worse, may induce fear among investors and shareholders. The process began with data cleaning. Data cleaning is to make certain the data is free from missing values or outliers that can affect the accuracy of the forecast value. Each stage was conducted until the best model to predict short-term forecast of DIGI daily share price could be determined. The forecast values were compared to the actual data to see how well the model performs on unseen data and how accurate it is.

3. RESULTS AND DISCUSSION

Table 1. Forecast and accuracy of Digi's share price on a daily basis in data high and low

	DAILY(HIGH)	MEAN 90	ACTUAL DATA	ACCURACY MEAN 90 (%)
1	4/4/2022	3.9	3.98	97.673
2	5/4/2022	3.9	3.96	98.167
3	6/4/2022	3.9	3.89	99.933
4	7/4/2022	3.9	3.94	98.665
5	8/4/2022	3.9	3.95	98.415

	DAILY(LOW)	MEAN 90	ACTUAL DATA	ACCURACY MEAN 90 (%)
1	4/4/2022	3.8	3.9	98.033
2	5/4/2022	3.8	3.88	98.538
3	6/4/2022	3.8	3.84	99.565
4	7/4/2022	3.8	3.85	99.306
5	8/4/2022	3.8	3.87	98.793

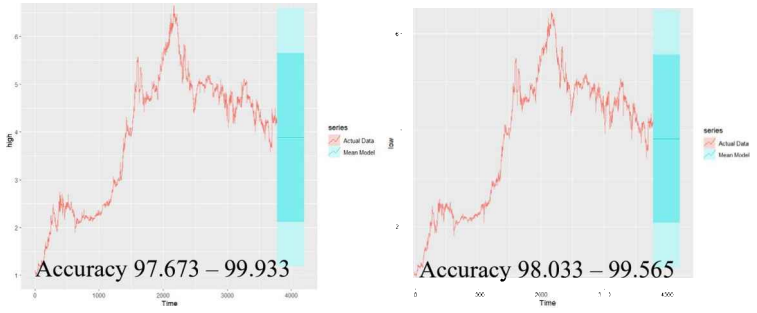

Table 1 shows the forecast and accuracy of the five-step ahead of Digi's share price on a daily basis in data high and low generated by the Mean Model. After comparing it with actual data obtained from Yahoo Finance, the forecast accuracy is between 97.673 percent to 99.933 percent for data high. Meanwhile, for data low, the five-step ahead of Digi's share price on a daily basis forecast accuracy is between 98.033 percent to 99.565 percent.

4. NOVELTY OF RESEARCH / PRODUCT

Time series cross-validation offers a novel and effective approach to improve forecast accuracy by considering the temporal dependencies inherent in time series data. Unlike traditional cross-validation methods, time series cross-validation techniques leverage the sequential nature of the data to create train-test splits that reflect real-world forecasting scenarios. By implementing rolling window cross-validation or recursive multi-step forecasting, models can be evaluated across different time periods, capturing evolving patterns and dynamics. Additionally, using time series-specific splitting techniques such as time series split or sliding window validation ensures that the temporal order is preserved during model evaluation. This enables a more accurate assessment of the model's ability to forecast into the future. Furthermore, incorporating specialized time series evaluation metrics and leveraging cross-validation for ensembling and model selection allows for more informed decision-making. This iterative process of model diagnostics and refinement, facilitated by time series cross-validation, leads to enhanced forecast accuracy and a deeper understanding of the data's temporal patterns. Overall, the application of time series cross-validation provides a valuable and innovative framework for improving forecasting performance.

5. CONCLUSION

Analysis and results were generated for univariate time-series analysis models and five sets of data splits (high and low) were used for each model to ensure the accuracy of the predicted values. Future studies should use long-term datasets with larger sample sizes, use cross-validation techniques to address data splits, and compare the performance and accuracy of predictive models.

REFERENCES

Alias, L. (2011). Introductory Business Forecasting a practical approach 3rd edition. (pp.4). UiTM Press.

Improve predictive accuracy of COVID-19 using time series analysis with repeated crossvalidation. Malaysia Technology Expo. (n.d.). Retrieved January 15, 2023, from https://mte.org.my/improve-predictive-accuracy-of-covid-19-using-time-seriesanalysis-with-repeated-cross-validation/

Research & Analysis. https://www.brightworkresearch.com/naive-forecast/