

Prophet Vs ARIMA: Utilising Machine Learning Model in Forecasting Mortality Trend in Malaysia during COVID-19 Pandemic

Muhd Zulfadli Hafiz bin Ismail, Tan Lay Kim, Evi Diana Omar, Zamtira Seman, Nik Noor Syamimi Ismail, Shubash Shander Ganapathy, Mohd Azahadi Omar

Sector for Biostatistics & Data Repository, Office of NIH Manager, National Institutes of Health (NIH), Ministry of Health Malaysia; Institute for Public Health, National Institutes of Health (NIH), Ministry of Health Malaysia

ABSTRACT:

INTRODUCTION: The COVID-19 pandemic has resulted in considerable strain on the health systems, delays in acute care and exacerbation of chronic diseases, leading to mortality beyond that directly due to COVID-19. Forecasting the trend of mortality due to the pandemic among the Malaysians is crucial in guiding the appropriate measures not only in handling the pandemic but also to assess the readiness of the health system for continuous improvement. This can be achieved by utilizing Machine Learning in producing appropriate forecasting models using local data. Hence, we aimed to compare Facebook 's Prophet and ARIMA Forecasting Model 's performance in forecasting the mortality trend among the Malaysians.

METHODS: We utilised the study materials from Epidemiological Study on Mortality and Selected Morbidity Rates Pre and Post COVID-19 Vaccination in Malaysia. A total of 1,070,925 data from the National Registration Department between March 2015 and June 2021 was used in data splitting for training, validating and testing both models. The mortality trend was forecasted from February 2020 onwards. The models ' performances were evaluated using Mean Absolute Percentage Error (MAPE), obtained by comparing forecasted and actual values validation set period from March 2020 till June 2021. Lower MAPE score showed smaller difference on average between the actual and the forecasted values and signified the better performance model.

RESULTS: Similar forecasting pattern was seen in both models during the validation period, in which forecasted values were generally higher than expected from March 2020 till December 2020, before showing lower forecasted values than actual figure from January until June 2021. However, using ARIMA model with ARIMA (3,1,0) (2,1,1)[52] demonstrated MAPE value of 4.62% as compared to Prophet 's MAPE of 4.92%.

DISCUSSION/CONCLUSION: Both models appear to be performing well in forecasting mortality trend, although ARIMA was slightly outperforming Prophet 's Model based on the smaller MAPE value. The superiority of ARIMA in time series were well established and demonstrated in this analysis as well. As Prophet 's strength was more apparent for time series with obvious seasonality, the absence of such pattern in this particular dataset potentially reduces its performance. The findings indicate the suitability of ARIMA in forecasting the mortality pattern, although the models should be evaluated further as more data made available, in ensuring the validity of the forecasted pattern in the future.

KEYWORDS: Machine learning, ARIMA, Prophet, Mortality