DEVELOPMENT OF AN ANALYTICAL TOOL TO FORECAST THE PROBABILITY OF DEFAULT IN BANKING AND FINANCE SECTOR

Jayani M W G Gurusinghe

A dissertation submitted in partial fulfilment of the requirement
for Master of Science in Big Data Analytics

School of Computing
Informatics Institute of Technology, Sri Lanka in collaboration with
Robert Gordon University, Aberdeen, Scotland

2023

Abstract

Credit Risk Management is a critical function in the field of Banking and Finance. With the several economic crises that took place in the world over the past few decades, the need to follow good credit risk management practices was heavily discussed among the various stakeholder groups. When a bank or NBFI faces an unforeseen loss because of non-payment or delays in payments by the borrowers, the company faces a major struggle. Probability of Default was a measure that was introduced to deal with this issue by estimating the future loss using this forecast along with several other estimated measures. Since PD plays a major role here, the regulatory accounting bodies have taken steps over the years to streamline the governance, computation and reporting of PD. Different methodologies have been introduced by various parties for this computation, out of which the most commonly used technique in the Sri Lankan context is scorecard-based models. However, the accuracy of the forecasts by these models may not always be satisfactory. PD is a measure that is affected by various macro-economic factors such as inflation, GDP, interest rate, employment rate, exchange rate etc. Therefore, a time series forecasting modelling approach for forecasting PD based on these macro-economic factors was developed. Several classical and machine learning techniques such as ARIMAX, TSLM, Linear models and XGBoost were applied, and the best model was chosen based on the accuracy measures such as RMSE and MAPE computed for the test data set. The tool was compiled into a R-based tool which could be utilized to visualize several analysis related PD and to generate forecasts for PD for the upcoming time periods. The tool considers inputs from the database for macro-economic factors and the forecasted PD figures are stored back in the database. Due to limitations such as lack of enough data points, the developed models did not show satisfactory levels of accuracy. Therefore, the developed models could be used as frameworks for the forecasting PD which could be further improved with the availability of data.

Keywords: Credit Risk Management, Bank, NBFI, PD, Inflation, Exchange Rate, Interest Rate, GDP, ARIMAX, TSLM, XGBoost