CDS 506:

Research, Consultancy and Professional Skills

Consultancy Project Proposal:

Real Private Consumption Time Series Prediction and Forecasting

Prepared by:

Muhammad Azzubair bin Azeman **P-COM0019/19**

Submitted for:

Dr Nasuha Lee Abdullah
Ts. Dr. Chew XinYing

Date of Submission:

26 June 2020

Masters of Da

School of Computer Science

Academic Session 2019 / 2020

Chapter 1: Introduction

1.1 Background

Level of economic advancement of one country to another differs by the main macroeconomics indicator which is the Gross Domestic Product (GDP). As such, world countries are categorised into three main categories which are Developed Economies, Economies in Transition, and Developing Economies, with additional Least Developed Countries listed (United Nations, 2020). In determining country classification, World Gross Product (WGP) (derived GDP) is included as one of the indicators.

Gross Domestic Product (GDP) is defined as "the market value of all final goods and services produced in an economy annually" (Hashim et. al, 2018). It is measured based on three main approaches which are the Production, Expenditure, and Income approaches (Department of Statistics Malaysia (DoSM), 2020). In terms of Production approach, it reflects on economic activities of individuals towards GDP as an overall; while for Expenditure approach, it determines the values of services and products consumed by consumers. As for Income approach, it includes all income sources and amounts gained in economy. Therefore, in order to determine the economic values of each approach, macroeconomic components are used as input to calculate the values of each approach.

Expenditure approach plays a crucial role in overall GDP as it contribute the most to the overall GDP since 2013 until 2018 (Asada et.al, 2019). This approach is dependent on five main macroeconomic expenditure indicators namely as Real Private Consumption (RPC), Real Government Consumption (RGC), Fixed Capital Formation (FCF), Changes in Inventories and Valuables (CIV), and Net Export (NE) (DoSM, 2020). Therefore, forecasting of econometrics indicators are crucial to be applied as reference for Ministry of Finance (MoF) Malaysia in making decisions for future financial planning. Currently, statistical techniques such as vector autoregression (VAR) model and Autoregressive Integrated Moving Average (ARIMA) model are commonly used models for econometrics forecasting (Razak, Khamis & Abdullah, 2017).

Econometrics forecasting using machine learning has been a major topic discussed in many literatures in the last two decades (Taieb, 2014). Several machine learning models such as Neural Network, Support Vector Machine, and K-Nearest Neighbour were proposed and discussed. However, machine learning are foreign among Malaysians until it was recommended by Minister of International Trade and Industry (MITI) (Bernama, 2018). As a result, machine mearning models and techniques are gradually being learned by Malaysians in many online courses recently (Fadzil, Latif, & Munira, 2015).

1.2 Problem Statement

Shifting from statistical models into machine learning models for econometrics forecasting definitely requires some time to be adopted. Plus, machine learning models have to compete with current statistical models used in government sectors. In particular, its ability to forecast future values is still questionable whether it could outperform current statistical models or not. Model evaluations such as model prediction accuracy, time taken for modelling and prediction, and model's complexity and transparency must be taken into account in weighing their reliability for econometrics forecasting. If machine learning is proved to be more reliable, then government sectors should reconsider in adapting to the new technology emerged nowadays. Otherwise, statistical models are remained.

1.3 Research Question

This research proposal makes an attempt to model and forecast time series of one of the econometrics which is the Real Private Consumption. This will consequently results in determining whether machine learning models or statistical models is better for econometrics forecasting. This project proposes machine learning models used for econometrics forecasting between Support Vector Machine (SVM) and Random Forest (RF) models with ARIMA and VAR models. This brings to the following research questions:

- How good are machine learning forecasting models compared to statistical models?
- What would be the RPC forecast of machine learning models for the third quarter of 2020?

Throughout this proposal, the research question answers will be reviewed from literatures to offer valid justifications, and to propose the best machine learning model for RPC forecasting.

1.4 Objectives

Therefore, the aim of this project is to propose machine learning techniques as a new approach to improve econometric forecasting in Malaysia. This project will be focusing on forecasting one of the econometric indicators which is the RPC published quarterly by DoSM. Typically, to achieve the aim of this project, there are two objectives listed below which are:

1. To model and forecast Real Private Consumption using SVM and RF.
2. To compare machine learning model performances with ARIMA, and VAR models.

1.5 Benefits of the project

This problem is actually a consultation project between DataMicron with Malaysia's Ministry of Finance (MoF). Therefore, this project will benefit DataMicron in providing proposed solution for their client's problem. In particular, this project will deliver the insights of previous RPC trends and will develop a reliable model for suggesting RPC forecasts as a reference for MoF's top management in making effective decisions.

1.6 Related Works

In this section, econometric literatures will be referred to explore the details of established and proposed methods in econometrics modelling. Published researches demonstrate on the theories and analysis behind every econometrics modelling and also some others include on proposed models potentially applicable in the future.

References

Asada, H., Kiang, T. K., Espinoza, R., and Vandeweyer, M. (2019). *OECD Economic Surveys 2019: Malaysia*. Retrieved on Oct. 8. 2020, from http://www.oecd.org/economy/surveys/Malaysia-2019-OECD-economic-survey-overview.pdf

Bernama (2018). *Azmin: Statistical Community needs to Embrace Digital Revolution*. The Edge Markets. Retrieved on Oct. 9, 2020, from https://www.theedgemarkets.com/article/azmin-statistical-community-needs-embrace-digital-revolution

Department of Statistics Malaysia (2020). *National Accounts FAQ.* Retrieved on Oct. 8, 2020 from https://www.dosm.gov.my/v1/index.php?r=column/cone&menu_id=dUtRR1JYWjk2TEJha1BrZml0REY4UT09

Fadzil, M., Latif, L. A., and Munira, T. A. (2015). MOOCsin Malaysia : A preliminary case study. In *MOOCs and Educational Challenges around Asia and Europe,* 1(6), pp. 65-86.

Hashim, E., Ramli, N. R., Romli, N., Jalil, N. A., Bakri, S. M., and Ron, N. W. (2018). Determinants of Real GDP in Malaysia. In *The Journal of Social Sciences Research*, No. 3, pp. 97-103.

Razak, N. A. A., Khamis, A., and Abdullah, M. A. A. (2017). ARIMA and VAR Modeling to Forecast Malaysian Economic Growth. In *Journal of Science and Technology: Special Issue on the Application of Science and Mathematics,* 9(3), pp. 16-24.

Taieb, S. B. (2014). *Machine learning Srategies For Multi-Step-Ahead Time Series Forecasting*. Université Libre de Bruxelles Belgium. Retrieved on Oct. 9, 2020 from http://souhaib-bentaieb.com/pdf/2014_phd.pdf

United Nations, (2020). *World Economic Situation and Prospects 2020*. United Nations Publication. Retrieved on Oct. 7, 2020 from https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/WESP2020_FullReport.pdf