FORECASTING SALES OF FAST-MOVING CONSUMER GOODS USING MACHINE LEARNING METHODS: A CASE STUDY

Adhikari Mudiyanselage Chandeep Jatila Bandara Adhikari

A dissertation submitted in partial fulfilment of the requirement for

Master of Science (Distinction) in Business Analytics

Department of Computing

Informatics Institute of Technology, Sri Lanka

in collaboration with

Robert Gordan University, Aberdeen Scotland

2022

Abstract

FMCG product sales are one of the fastest growing market segments with an equally fast offtake. As the name implies these are products which have a very short shelf life and a very high demand. The global FMCG market is continually growing due to the variety of products offered by a variety of competitor product brands. Most product in this segment are part of our day-to-day consumptions such as confectioneries, beverages, dairy product, cosmetics etc. Because of this nature the sales forecast of these products has become a very critical functionality to ensure the correct quantities are available to accommodate the consumer demand. In addition, the sales forecast is also the basis on which other critical functionalities in the supply chain such as logistics, production and supplies plan their activities. In more recent studies it has been highlighted that a lot of research has been directed toward using more data driven approaches such as machine learning to do the sales forecast. Especially in the regions of south Asia where there are several external factors such as holidays, geographic locations and other macro-economic factors which have made the sales forecasting even more challenging. Especially in more reason times with the Covid 19 pandemic the FMCG industry was one of the few industries that saw growth with the industry falling under essential goods across all global markets. This also created newer challenges in getting and accurate sales forecast for FMCG products. Hence the need for more data driven approaches to improve this sales forecast is even more evitable.

Data on FMCG sales in 3 regions of India are used in this project obtained from Kaggle.com. The data set is used to develop 5 independent machine learning models and 5 combined machine learning models. The 5 independent models were developed using unidirectional stacked LSTM, bidirectional LSTM, unidirectional stacked GRU, bidirectional GRU, and Support Vector Regression. The same five models were then optimized using K-means clustering to optimize the overall performance and device the combined models. The data set is split with two years of historical sales for

training and two months for testing of the models. The models were evaluated using the accuracy, RMSE, MAPE, Loss Validation plots and Sales against Forecast plots.

Though the models faced limitations due to the limited historical sales and features, the findings show that out of the models studied in this project GRU delivers the best results because of its optimized capability to successful update and rest its nodes during the training phase. When used with bidirectional modeling GRU is more optimal due to the availability to learn from future states as well. LSTM too is a viable alternative but is not as efficient as GRU. The overall results of SVR were unable to follow the sales trends as in the case of with the deep learning models. The SVR models require more features to perform a more optimal classification. Finally, the use of K-means helps optimize the results of the individual models for both deep learning and regression models.