Comparison of Machine Learning Techniques in Demand Forecasting for Imported Food Items in Retail Industry: A Study on a Supermarket Chain in Sri Lanka

HENADEERA APPUHAMILAGE DONA LAKNA LAMALI GUNASEKERA

MSc

2022

Abstract

Managing consistent product availability of imported food items while cost-effectively managing inventory is a challenge for supermarket chains. This has been a real challenge with the number of macroeconomic factors had happened during the past couple of years. Reduction in stock-out situations has a direct correlation to an increase in customer satisfaction. Hence, accurate demand prediction has gained attention in literature whilst industry experts strongly believe that consistent product availability is key to retaining a loyal customer base and improving customer satisfaction as well. Through an intensive review of literature, it was recognized that the comparison of machine learning-based demand prediction for imported food items has gained less attention. Hence, this research has been carried out to perform a comparison on machine learning-based demand prediction models for imported food items and suggest the business entity consider developing a strong prediction model for the demand of imported food items. Moreover, this research discusses the business strategies which would help in the implementation of the proposed machine learning-based prediction model for the real-world business case.

The study includes a detailed explanation of the acquisition of the dataset and the features used as there are some unique features when considering the demand for imported food items, especially when deciding the lead times to make the product available at the store level. Python programming language has been used to carry out descriptive statistics and tests to check the stationarity of the dataset along with the correlation of features. The time series model and the machine learning-based models have been developed by using the Orange data mining and visualizing tool. The result of the study shows that the time series model has less relevance since there are complex features involved in deciding sales for an imported food item and the most appropriate machine learning method is the Gradient Boosting technique among other four machine learning techniques used to carry out the study, namely Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Support Vector Model (SVM), and Random Forest. As a future study, the models can be further evaluated by considering the seasonality factor of substitutable brands as a result of using improved features for the study. The proposed machine learning technique could be used to develop a strong demand prediction model for the imported food category after considering the real

need of maintaining consistent availability at the store level and to reduce lost sales for the business entity.