ANALYTICS OF INVENTORY PRIORITY LEVEL FOR DATA DRIVEN LOGISTICS AND SUPPLY CHAIN MANAGEMENT DECISIONS

DURKAHPUVANESVARI A/P CHANDRAN

A project report submitted in fulfilment of the
requirements for the award of the degree of
Master of Science (Data Science)

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

MARCH 2021

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor, Dr. Nor Erne Nazira binti Bazin, for her encouragement, guidance, critics and friendship. I am also very thankful to my industrial supervisor, Mr. Shawn Kwek for his guidance, advices and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.

Next, I would like to express my gratitude to my parents and my brother for their constant moral support and encouragement throughout this study period. I would not have got where I am today without them.

Lastly, I would like to thank to all my colleagues and those who contributed to this research directly or indirectly.

ABSTRACT

The purpose of this study is to implement analytics techniques in data driven logistics and supply chain management decisions as a measure of inventory control. Descriptive analytics is applied to make an informed decision on the service level requirement of the inventory based on its priority level. The priority level of the inventory is analysed by performing integrated ABC and FSN analysis. The priority level is determined through the ABC–FSN matrix and which later is visualised through a scatter plot using Tableau to show the distributions of the inventory with different priority levels so that it is easy to give some advice to the company on the inventory that needs higher service level. High service level, 99.9% availability and the greatest attention is proposed for the inventory classified as priority 1 since those item are classified as high risk items which frequently runs out of stock due to its demand. In order to keep updated as well as to predict the future status of an item in term of its priority level, predictive analytics is crucial. Hence, supervised machine learning technique, classification algorithm is applied in this study. Synthetic minority oversampling technique (SMOTE) is required to avoid the drawback of undersampling issue. A robust and optimised analytical model, Random Forest with the specific parameters of *“max_depth”* = 7, *“min_samples_leaf”*= 6, *“min_samples_split”* = 2 and *“max_features”*= auto is built upon holdout validation method and hyperparameter tuning. The performance measure of the model in classifying the inventory priority level is evaluated. The model accuracy obtained is 98.53% with 2685 instances correctly classified and 40 instances incorrectly classified. The weighted average of the model precision, recall and F1-score has a very good score of 0.99.

ABSTRAK

Tujuan kajian ini adalah untuk menerapkan teknik analisis dalam keputusan logistik dan pengurusan rantaian bekalan utuk kawalan inventori. Analisis deskriptif digunakan untuk membuat keputusan yang tepat mengenai keperluan tahap perkhidmatan inventori berdasarkan tahap keutamaannya. Tahap keutamaan inventori dianalisis dengan melakukan analisis ABC dan FSN bersepadu. Tahap keutamaan ditentukan melalui matriks ABC–FSN kemudiannya, diperhatikan melalui plot berselerak menggunakan Tableau untuk menunjukkan klasifikasi inventori dengan tahap keutamaan yang berbeza untuk memudahkan pemberian nasihat kepada syarikat mengenai inventori yang memerlukan tahap perkhidmatan yang lebih tinggi. Tahap perkhidmatan yang tinggi, 99.9% dicadangkan dan tumpuan khas diperlukan untuk inventori yang diklasifikasikan sebagai keutamaan 1 kerana inventori tersebut diklasifikasikan sebagai inventori yang berisiko tinggi kehabisan stok oleh kerana permintaannya. Untuk memastikan status inventori sentiasa dikemas kini dan juga untuk meramalkan status inventori dari masa ke semasa berdasarkan tahap keutamaannya, analisis ramalan sangat penting. Oleh itu, teknik pembelajaran mesin yang diselia, algoritma klasifikasi diterapkan dalam kajian ini. Teknik *SMOTE (synthetic minority oversampling technique)* diperlukan untuk mengelakkan masalah kekurangan sampel. Model analitik yang kukuh dan optimum dibina dengan parameter "*max_depth*" = 7, "*min_samples_leaf*"= 6, "*min_samples_split*" = 2 and "*max_features*"= auto berdasarkan kaedah *holdout* dan penalaan *hyperparameter*. Prestasi ketepatan model dalam mengklasifikasikan tahap keutamaan inventori dinilai. Ketepatan model yang diperolehi adalah 98.53% dengan 2685 data dikelaskan sebagai betul dan 40 data dikelaskan sebagai tidak betul. Purata wajaran model "*precision*", "*recall*" dan "*F1-score*" mempunyai skor 0.99 yang sangat baik.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABC	-	Always, Better, Control
DT	-	Decision Tree
ETL	-	Extract, Transform and Load
FSN	-	Fast, Slow, Non-Moving
LR	-	Logistic Regression
ML	-	Machine Learning
RF	-	Random Forest
SKUs	-	Stock Keeping Units
SC	-	Supply Chain
SCM	-	Supply Chain Management
SMOTE	-	Synthetic Minority Oversampling Technique
SVM	-	Super Vector Machine

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

This chapter discusses the introduction of data analytics application as a solution provider in supply chain logistics operation. Furthermore, this chapter describes the problem statement, project question, objective, scope, significance of the project and the thesis organizations.

1.2 Problem Background

In this globalization era, analytics plays an important role in the process of assisting management to make supply chain operation decisions, which often include demand planning, procurement, production, inventory, and logistics by anticipating trends and outcomes. The deployment of analytics is essential especially in the supply chain (SC) field, a large network comprising various entities namely suppliers, manufactures, logistics, warehouses and distribution centers, retailers and consumers. The application of analytics as a solution provider to the SC problems becomes very precious in data driven decision making and for prediction future outcomes. In data analytics, data from various sources and enterprise systems can be integrated, providing a unified data to manage inventory and attune to the changing customer behavior, supplier and market landscape, elevating inventory investment and ensuring continuous supply of inventory.

As the SC becomes more complex, firms are facing challenges in managing inventory effectively due to more and more uncertain demands. One of the most critical effects due to demand uncertainty is providing better service and maintaining the good customer relationship while ensuring no occurrences of excess stock and items run out of stock which might result in backorder situations is also becoming a challenging task. In order to manage the situation, making use of the analytics techniques in terms of visualization can help to improve the entire portion of a firm by improving the access to the firm's data and later, transforming the data into insights of their business processes, identifying the trends and patterns in inventory use. These insights will be used to make strategic business decisions that can increase the efficacy and efficiency like improve productivity, increase revenue and accelerate growth as well as profitability. Analytics techniques using machine learning (ML) as a tool can be used to classify the inventory materials based on their priority in order to identify which inventory needs more attention as well as optimizing the inventory level by obtaining optimal inventory required in an organization to reduce the overstocking costs thus improving the SC's effectiveness. In addition, predictions through analytics facilitates the future consumption or sales and reduces the supply disruption by identifying and anticipating the inventory turnover, lost sales and customers behavior, potential stock-outs allowing an organization to take decisions and act proactively.

Previously, there were a case in a logistics company where the company had to deal with spare parts movement of an automobile company. Some spare parts of the automobile products were noticed to be spoilt and need to be returning back to the company warehouse for mending it or to be prepared with a new spare parts. In order to have a proper inventory level of stock keeping units (SKUs) and avoid insufficiency, descriptive analytics approach was applied to the large movement table to classify the movement and demand category by performing ranking level inventory management analysis. This analysis were performed to identify the priority level of the SKUs so that, customers demand at item level can be planned and controlled to match supply and demand. As a result of the analysis, the visualization produce distributions with different priority level where it is easy to give some advice to the company on the inventory level that need to allocated with higher service level. This helps customers of the automobile company to experience a better customer service since the automobile company is able provide the sufficient inventories to be placed in the logistics company warehouse in a short lead time and the logistics company can deliver the product in timely manner.

As the logistic company has come up with the integrated inventory management analysis, they provide the service level required for each SKU based on that analysis. The big drawback felt by the company is they could not keep updated on the upcoming data and also predict the future priority level of the SKUs as they do not have a system which could do the prediction. Here comes the role of predictive analytics which plays an important role in SCM in order to provide the insights which ultimately predict the future event occurring through the historic data. Hence, the company planned to implement ML technique to predict the priority level of the inventory. The execution of prediction method in inventory control was operationally effective where the company was able to keep update on the current status of the inventory in term of its priority level and also predict the future status of the inventory through the demand forecasting data.

Hence, visualization and ML tools are required as a solution for issues in SC areas: identifying current inventory levels including SKUs. Visualization assist in understanding the behavior and make informed decisions. Interactive visualizations enables the view from a very general to in-depth information of the data ensuring the users are aware of serious issues like inventory level, stock uncertainty, back orders and other business factors. Besides, application of predictive analytics as classification, regression and clustering which can be implemented as a ML techniques is also very important in inventory control and planning. Other than predicting future trends such as sales demand, exchange rates and other important SC metrics, predictive analytics allows SC managers to determine detailed inventory requirements by region, location and annual consumption, service level and etc.

1.3 Problem Statement

Data analytics techniques can be deployed in a business operation that implements supply chain management (SCM) to maintain a profitable business by managing the growing organization data, analyze it and consequently make the right decision at the right time. Based on the case study of a logistics company, the new customer of the logistics company provide the traditionally utilize information generated from their accounting records, it is expected that some analytics study need to be performed. The data provided is also seems to be a large data and it is time consuming if traditional method is implemented. Descriptive analytics is required as a solution provider to make more informed decisions from the visualization generated in the form of a report.

In SC, proper inventory management is very important. In order to provide optimum service level, inventory availability is always crucial and the space allocated for the inventories in the warehouse is also need to be always sufficient and organized well. So, this logistics company is currently facing the problem in organizing the space in its warehouse due to the increase in the inventory type and quantity. This happened as the logistics company had a new customer with various type of inventories and quite a lot of inventories. So, the logistics company would like to classify the inventory movement and demand category which subsequently will help the company to identify the fast mover and high in demand SKUs. The identification of the movement and demand category is necessary to have a proper inventory control. A proper inventory management is crucial to ensure the availability of the inventory in a SC logistics firm. Customers demand at item level must be planned and controlled to match supply and demand. Carrying too much safety stock will lead to high holding cost of inventory whereas carrying too little stock might lead to stock insufficiency. Both the situation will cause customers disappointment towards the service provided or in other words the customers satisfactory level will not be met. So, a constant attention is needed in safety stock management and control. Inventory is costly to store, so inventory management is required to offer a continuous, appropriate supply of products to customers, partners and retailers and maintain an optimum level of buffer stock.

Hence, the application of descriptive analytics potentially can be a beneficial approach in categorizing the demand and movement type of the inventories as a measure of inventory control. In addition, its potential to provide a visualization of the inventory categorization based on the priority level of the inventory eases the process of generating the report and give advice to the customer on the appropriate service level of the SKUs needed to maintain the ideal level of inventory. Thus, higher service level can be provided to the customers and the customers can be satisfied with better customer service experience. In meantime, by understanding the concept of inventory categorization, a predictive model which can predict the priority level using inventory demand forecast is needed. Implementing a predictive analytics approach in inventory categorization would increase the operational efficiency of an organization. A robust predictive model enable accurate prediction for inventory priority level where appropriate service level availability of the inventory can be determined to ensure satisfactory service level provided to the customers to meet their expectation.

1.4 Study Question

This study attempts to answer the following questions in order to address the problem:-

(a) Q1: How analytics techniques are applied as a solution provider in Supply Chain Management (SCM)?

(b) Q2: How to perform the data analysis for categorizing inventories to make the informed decision?

(c) Q3: How visualization portrays as a meaningful descriptive analytics product to generate useful information for making inventory control decisions particularly on SKUs?

(d) Q4: How classification machine learning (ML) algorithm is applied as predictive analytics measure for future predictions of inventory priority level.

1.5 Objectives

The objective of this project is as following:-

(a) To categorize the priority level of the inventory by performing the integrated Always, Better, Control (ABC) and Fast, Slow and Non- Moving (FSN) analysis.

(b) To visualize the scatter plot distribution of the inventory highlighting its priority level by plotting the annual consumption value against average order frequency.

(c) To evaluate and make the decision on the inventory service level requirement by generating the visualization report.

(d) To create a robust classification model for inventory priority level by assessing the model score of the inventory priority level classifications using holdout method and hyperparameter tuning as well as evaluating the model performance to classify the inventory priority level.

1.6 Project Scope

This project is focused to provide the optimal solution by determining the appropriate service level availability for the inventory based on its priority level as an inventory control measure. The data is obtained from the logistics company. The information provided to perform the analysis were company and user details, warehouse details, inventory details and movement details. Descriptive analytics is applied by performing inventory categorization analysis and the outcome is visualized using Tableau as well as the report is generated to make the decision. Besides, predictive analytics approach is applied to construct a robust classification model to be reused for future predictions.

REFERENCES

Aly, M. (2005). Survey on multiclass classification methods. *Neural Netw, 19*, 1-9.

Anitha, P., & Patil, M. M. (2018). A review on data analytics for supply chain management: a case study. *International Journal of Information Engineering and Electronic Business, 11*(5), 30.

Arya, V., Sharma, P., Singh, A., & De Silva, P. (2017). An exploratory study on supply chain analytics applied to spare parts supply chain. *Benchmarking: An International Journal*.

Banks, J. (2000). *Introduction to simulation.* Paper presented at the 2000 Winter Simulation Conference Proceedings (Cat. No. 00CH37165).

Bellinger, C., Jabbar, M. S. M., Zaïane, O., & Osornio-Vargas, A. (2017). A systematic review of data mining and machine learning for air pollution epidemiology. *BMC public health, 17*(1), 1-19.

Bilgen, B. (2010). Application of fuzzy mathematical programming approach to the production allocation and distribution supply chain network problem. *Expert Systems with Applications, 37*(6), 4488-4495.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Chase, R., Jacobs, F., & Aquilano, N. (2006). Operations Management for Competitive Advantage, McGraw-Hill. *Irwin, New York*.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research, 16*, 321-357.

Dantzig, G., & Thapa, M. (1997). Linear Programmin 1: Introduction Springer. *New York*.

Dey, A. (2016). Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies, 7*(3), 1174-1179.

Eicker, T., & Cilliers, J. O. (2019). Managing logistical supply chain drivers: A predictor of small township retailers' odds of survival. *Acta Commercii, 19*(2), 1-12.

Ghasemzadeh, A., Hammit, B. E., Ahmed, M. M., & Young, R. K. (2018). Parametric ordinal logistic regression and non-parametric decision tree approaches for assessing the impact of weather conditions on driver speed selection using naturalistic driving data. *Transportation research record, 2672*(12), 137-147.

Grichi, Y., Beauregard, Y., & Dao, T. (2017). *A random forest method for obsolescence forecasting.* Paper presented at the 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM).

Harvey, H. B., & Sotardi, S. T. (2018). The pareto principle. *Journal of the American College of Radiology, 15*(6), 931.

Hugos, M. H. (2018). *Essentials of supply chain management*: John Wiley & Sons.

Huisman, D. O. (2015). *To What Extent do Predictive, Descriptive and Prescriptive Supply Chain Analytics Affect Organizational Performance?* , University of Twente.

Kibria, M. G., Nguyen, K., Villardi, G. P., Zhao, O., Ishizu, K., & Kojima, F. (2018). Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks. *IEEE access, 6*, 32328-32338.

Komarek, P. (2004). Logistic regression for data mining and high-dimensional classification.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news, 2*(3), 18-22.

Manivel, P., & Ranganathan, R. (2016). Prioritized ABC-FSN analysis of inventory management in private and hospital pharmacy followed by questionnaire. *International Research Journal of Pharmacy, 7*(12), 104-113.

May, B. I., Atkinson, M. P., & Ferrer, G. (2017). Applying inventory classification to a large inventory management system. *Journal of Operations and Supply Chain Management (JOSCM), 10*(1), 68-86.

Mishra, N., & Silakari, S. (2012). Predictive analytics: A survey, trends, applications, oppurtunities & challenges. *International Journal of Computer Science and Information Technologies, 3*(3), 4434-4438.

Murray, P. W., Agard, B., & Barajas, M. A. (2015). Forecasting supply chain demand by clustering customers. *IFAC-PapersOnLine, 48*(3), 1834-1839.

Pandya[1], B., & Thakkar, H. (2016). A review on inventory management control techniques: ABC-XYZ analysis. *REST Journal on Emerging trends in Modelling and Manufacturing, 2*, 15.

Parekh, S., Lee, J., & Kozman, T. A. (2008). A decision support system for inventory management. *URL: http://www. swdsi. org/swdsi08/paper/SWDSI% 20Proceedings% 20Paper% 20S206. pdf [Accesat la data de 3 iunie 2014]*.

Priore, P., Ponte, B., Rosillo, R., & de la Fuente, D. (2019). Applying machine learning to the dynamic selection of replenishment policies in fast-changing supply chain environments. *International Journal of Production Research, 57*(11), 3663-3677.

Quiroz, I. A., & Alférez, G. H. (2020). Image recognition of Legacy blueberries in a Chilean smart farm through deep learning. *Computers and Electronics in Agriculture, 168*, 105044.

Riadi, C. Z., Hendrawan, D., & Pratomo, R. M. H. (2020). The ABC-FSN Analysis and Determination of Forecasting Methods for Material Maintenance, Repair & Operations Inventory, Case Study at PT X. *Journal of Critical Reviews, 7*(17), 1738-1750.

RK., G., KK., G., Jain, B., & Garg, R. (2007). ABC and VED analysis in medical stores inventory control. *Medical Journal Armed Forces India, 63*(4), 325-327.

Selvakumar, S., & Ramya, C. (2017). Descriptive Analytics *Applying Predictive Analytics Within the Service Sector* (pp. 88-112): IGI Global.

Singh, A., Shukla, N., & Mishra, N. (2018). Social media data analytics to improve supply chain management in food industries. *Transportation Research Part E: Logistics and Transportation Review, 114*, 398-415.

Takigawa, I., Shimizu, K.-i., Tsuda, K., & Takakusagi, S. (2016). Machine-learning prediction of the d-band center for metals and bimetals. *RSC advances, 6*(58), 52587-52595.

Teichgraeber, H., & Brandt, A. R. (2019). Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. *Applied Energy, 239*, 1283-1293.

Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., Yu, H., . . . Yang, X. (2017). Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. *Remote Sensing, 9*(4), 309.