PRODUCT BRAND PREDICTION IN RETAIL INDUSTRY

NOR AFIFA BINTI ISMAIL

A dissertation submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Computer Science

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

OCTOBER 2018

“My dearest son, *Muhammad Harraz, Family and Friends”*

This is for all of you

ACKNOWLEDGEMENT

In the name of Allah, Most Gracious, Most Merciful. Alhamdulillah, thanks to Allah SWT because with the abundance of His grace, I can provide this dissertation. I would like to express my appreciation to my Supervisor, **Prof. Dr. Siti Mariyam binti Shamsuddin** and Co-Supervisor, **Dr. Zuriahati binti Mohd Yunos** for the guidance, encouragement, criticisms and thoughts when preparing for this dissertation. I'm so grateful to have found these two supervisors who are very helpful to me while preparing for this dissertation. Without them, it is unlikely that this dissertation can be produced.

I am also grateful to UTM for giving me the opportunity to continue my studies at this university. Thanks to the support staff for working with me to make it easier for me to get the materials for this dissertation.

Last but not least, I would like to express my appreciation to my father, mother and sister for being my spine to continue my studies. Without them, I can hardly get the blessing that allows me to reach this stage. Thanks also to my husband for helping me to give an idea and understand my commitment right now. Indeed, only Allah SWT can repay to all of you.

ABSTRACT

Nowadays, there are some challenges regarding to sales and retailing issues. There are some factors that impact to the retailer influences to their customer. It is including the competition from online businesses. Sometimes it is hard to achieve the consumer demands expectation. At the same time, the retailer may need to overcome the financing pressure and marketing challenges just to ensure that their brand could be survive in business matter. The retailing challenges that always facing by company including finding the financing to stay in business. The only ways to solve the problem is by making prediction to investigate their product sales by using the data in year 2015 and 2016. The objective is to study the existing data sales based on type of brand, to develop prediction model for each type of brand, and to evaluate the performance of Artificial Neural Network (ANN) and Exponential Smoothing (ES) model in selecting the best model. The measurement performance used to analyse the data are using RMSE, MAE and MAPE. The comparison for the best model is based on the actual and predicted data that helps to get the result for profit sales and loss sales for each brand. Result obtained shows that MLP model is better in prediction model compared to the ES model.

ABSTRAK

Hari ini, terdapat beberapa cabaran berkaitan dengan isu peruncitan dan penjualan. Terdapat beberapa faktor yang memberi kesan kepada peruncit kepada pelanggan mereka termasuklah saingan daripada perniagaan secara dalam talian. Kadangkala sukar untuk mencapai kehendak pelanggan. Pada masa yang sama, peruncit berkemungkinan berhadapan dengan masalah tekanan terhadap kewangan dan cabaran pasaran hanya untuk memastikan produk mereka terus bertahan. Terdapat cara untuk menangani masalah ini dengan membuat ramalan terhadap jualan produk dengan menggunakan data pada tahun 2015 dan 2016. Objektif kajian yang dijalankan ini ialah untuk mempelajari data jualan yang sedia ada berdasarkan jenama tertentu, untuk membangunkan model jangkaan untuk jenama barangan sukan serta untuk menilai prestasi model *Artificial Neural Network* (ANN) dan *Exponential Smoothing* (ES) untuk memilih model yang terbaik. Pengukuran prestasi yang digunakan untuk menganalisis data adalah dengan menggunakan *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE) dan *Mean Absoloute Percentage Error* (MAPE). Perbandingan model terbaik berdasarkan data nilai sebenar dan nilai ramalan membantu untuk mendapatkan keputusan jualan keuntungan dan kerugian bagi setiap jenama. Berdasarkan keputusan, model MLP adalah lebih baik untuk peramalan berbanding model ES.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	-	Artificial Neural Network
ES	-	Exponential Smoothing
LR	-	Learning Rate
M	-	Momentum
MA	-	Moving Average
MAE	-	Mean Absolute Error
MAPE	-	Mean Absolute Percentage Error
MLP	-	Multilayer Perceptron
RMSE	-	Root Mean Squared Error
WEKA	-	Waikato Environment for Knowledge Analysis

LIST OF SYMBOLS

μ	-	Mean
$f_i x_i$	-	Sum of products
n	-	Number of data values
X	-	Normalization
x_0	-	Current value
x_{min}	-	Minimum value
x_{max}	-	Maximum value
x	-	Actual data
y	-	Predicted data
$\lvert\,\rvert$	-	Absolute value

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the modern age, there are a lot of dumped data that had been collected by the industry. It is included gained something more valuable. The retailer may compare pricing data from the other competitor and the retailer can be compromised to adjust the price of its own product. Absolutely retailer wanted to develop a good reputation and became the selection of the customer wanted a lower price and a good experience.

Moreover, from predictive analysis the retailer may get an accurate prediction from the data gathered. The collection of data on customer's measurement and their item preferences allowed the retailer to offer a better experience to the customer. From the data also the retailer may know further about the styles of the product and colour item that are popular to choose from the customer. This might help the retailer to keep the right products in stock and negotiated the best rates possible with the supplier (Mallon, 2016).

1.2 Background of the Problem

Nowadays, there are some challenges regarding to sales and retailing issues. There are some factors that impact to the retailer is influenced to their customer. It included the competition from online businesses. Sometimes it is hard to achieve the consumer demand expectation. At the same time, the retailer may need to overcome the financial pressure and marketing challenges just to ensure that their brand could be surviving in business matters. The retailing is challenging that always facing by companies included found the financing to stay in business. Most retailers made a loan to ensure they could be run the business frequently.

Besides that, the retailers need to maintain their inventory product by discover a way to access from various locations of outlets is essential for maximizing the sales. If the retailer could not organize in inventory process very well, they might lose their sales or delays in fulfilled orders. The best way to stay in business demand is ensured the product also could be produced by online marketing. By marketing from social media and user friendly from website retailer could reach a wide customer to the product. The retailer must ensure that they had marketing campaign online by producing a banner and sales promotion (Hecht, 2016).

Other challenges could be faced by retailers is data silos. Data silos are a Big Data's kryptonite, which the data is store separated and isolated. From the data silos, the retailer could produce a monthly sales report. Data silos could be the sales and marketing team are not worked together. Data silos also make the customer looked for another brand since they thought their needs are not being met and smaller. Not only that, inaccurate data also led to ineffective on an operational level. To overcome the problem encountered, the data need to be integrated. Since the technology moves so fast day by day, the retailer sometimes preferred to keep their databases in premises. If data can be obtained and analysed quickly, surface actionable insights and driven them to the operational systems, then the retailer could affected events as they are still unfolding.

Lack of skills workers also is the reason they are unable to use their data properly. The retailer needed to find the workers that understand data from a data science perspective, understand the business flow and customer demands, and how data findings apply directly to them (Rombaut, 2016).

This study is for Malaysia Retailer Sports namely as Company XYZ Sdn. Bhd. This company was established in 1993 with a single store. Initially, this company is stated as sole proprietorship enterprise, but after for a long time the company had evolved into a Private Limited Company and had more than 100 outlets in Peninsular Malaysia.

After for a long time, there some challenges that this company needs to face. The biggest factor is another new company that also provided the same services like this company; which are provided the sports stuff like shoes, jersey and sports equipment. The competition of multi company made this company to find the solution to stand still with their brand. One of the ways is by making the prediction to investigate their product sales for few years' back which is between 2015 and 2016.

The usage of the MLP model helps in predicting to reach the arbitrary input and output mapping (Zhang *et al.,* 1997) instead of learning based on the examples and captured the relationship among the data although the relationships are unknown to describe. Hence, the MLP model is suitable to observe the data. (White, 1989; Ripley, 1993; Cheng and Titterington, 1994). While for Exponential Smoothing model, even though it is one of the simplest models, it is powerful for prediction and widely used for business matter especially for inventories (Gardner, 1985). Thus, the data raw from the Company XYZ Sdn. Bhd will be predicted by used these two models to compare which model could be the best for prediction.

1.3 Problem Statement

Prediction is an important part in our lives, and sales prediction also played a major role for retailers or enterprises in making business plans more accurate and gained competitive advantage. Most studies had depended on historical sales data in predicting sales, but the sales are also affected by the contents of the business products. However, the important issue in the prediction is how to develop a model that can be produced accurate prediction results.

In this research, the aim is to develop a prediction model based on the types of brand. Based on the prediction model, it can be predicted and identified which brand gets the highest sales and the lowest sales. However, the constraints to conduct the prediction happened when the data is incomplete and had the missing values. The dataset also had a small scale to make the prediction where the data provided are only for two years ago.

The problem statement for this research is how to develop a model that can be produced accurate prediction results. The followings are some research questions that will be addressed to answer the above problem statement:

1. What is the most brands that customers looked for?
2. How to predict the data with smaller scales?
3. What are the model and methods used to predict the data?
4. Which classifier between Artificial Neural Network (ANN) and Exponential Smoothing (ES) provided better performance for brand sales prediction?

1.4 Research Objectives

The aim of this research is to develop a prediction model that will be improved prediction accuracy, able to deal with insufficient and incomplete data. The objectives of this project are:

i. To study the existed data sales based on type of brand.
ii. To develop the Artificial Neural Network (ANN) Exponential Smoothing (ES) model.
iii. To measure and analyze the performance of ANN and ES for selected the best model.

1.5 Research Scope and Assumptions

The scope of the research is limited to the following:

i. MLP and ES are respectively chosen as prediction model.

ii. The experiment is tested using MLP with different value of the learning rate and momentum.

iii. The experiment is tested using ES with different value of the damping factor.

iv. The data consisted of five types of brand and covered only for twenty four months, which is data from January until December 2015 and the data started from January until September 2016. Only 3 inputs from MLP had been tested for this experiment.

v. For evaluation measurement, MAE, RMSE and MAPE had been used by compared each other.

1.6 Significance of Research

The significance of this research is the usage of MLP and ES model that can help the decision makers to make the necessary decision based on accurate prediction results. They are able to plan effectively and gave sound decision to achieve the desired results even the data size is limited.

1.7 Dissertation Outline

This dissertation is organized as follows:

Chapter 1 explained about the introduction of predicted sales, background study of the related research regarded to the predict sales and soft computing problem, problem statement, aim, objective and scope of the study.

Chapter 2 described further the related literature review of predicted sales, types of prediction techniques, model development, soft computing techniques and measurement performance used to evaluate the result of datasets.

Chapter 3 clarified the methodology and architecture used for this study. It also included data collection, data processing, data curation, data modelling and analysis and evaluation of the models.

Chapter 4 discussed about the model development, which are consisted of Multilayer Perceptron and Exponential Smoothing. The feature undergoing the process of model development, the evaluation from the measurement performance and the selection of the best model according to the comparison of both model developments.

Chapter 5 indicated the comparative prediction result from the two model developments. The comparative models included the best model in the training data and test data. The result of comparative will go through for the analysis.

Chapter 6 summarised and concluded the entire of the study and recommendations for future works

REFERENCES

Adhikari, R. & Agrawal, R.K., (2013). An Introductory Study on Time Series Modelling and Forecasting.

Al Shalabi, L. & Shaaban Z., (2006). Normalization as a Preprocessing Engine for Data Mining and the Approach of Preference Matrix. Proc. International in Dependability of Computer Systems (DepCos-RELCOMEX ‘06), pp. 207-214.

Amimi, J., (2008). Optimum learning rate in back propagation network for classification of satellite images (IRS- 1D), Scientia Iranica, Vol 15 No. 6 PP 558567, Dec 2008 .

Big Data Top 5 Problems Big Data. Available at: http://www.business2community.com/big-data/top-5-problems-big-data-solve01597918#sBs40KxWDjBC7z2D.97 [Accessed 29 April 2017].

Boyko, J., (2013). Handling Data with Three Types of Missing Values. University of Connecticut.

Brown, R.G., (1959). Less Risk in Inventory Estimates. HBR July- August 1959, page 104.

Blocher, J. D. et al., (2004). Forecasting: Including an Introduction to Forecasting using the SAP R/3 System. United States of America: Indiana University Kelley School of Business.

Bullard, B., (2017). Style and Statistics: The Art of Retail Analytics. Canada: John Wiley & Sons, Inc.

Byrne, R. F., (2012). Beyond Traditional Time-Series: Using Demand Sensing to Improve Forecasts in Volatile Times. *Journal of Business Forecasting, 31,* 13.

Caglar, G., et al., (2016). Noisy activation functions. In International Conference on Machine Learning, pp. 3059–3068.

Chambers, J.C., Mullick S.K. & Smith D. D., (1971). How to Choose the Right Forecasting Technique.

Cheng, B. & Titterington, D.M., (1994). Neural networks: A review from a statistical perspective. Statistical Science 9 (1), 2–54.

Chopra, S. & Meindl, P., (2013). Supply Chain Management: Strategy, Planning and Operation, 5th Edition, Prentice Hall, Page 195, Chapter 7.

Dean, J., (2014). Big Data, Data Mining and Machine Learning: Value Creation for Business Leaders and Practitioners. United States of America: Wiley.

Frank, E., Hall, M. A. & Witten I.H., (2016). Data Mining: Practical Machine Learning Tools and Techniques. Fourth Edition. United States of America:Morgan Kaufmann Publications.

Gardner, E. S., (1985). Exponential smoothing: The state of the art. Journal of Forecasting, 4, 1– 28.

Gheyas, I. A. & Smith, L.S., (2010). A neural network-based framework for the reconstruction of incomplete data sets. University of Stirling, Department of Computing Science and Mathematics, Stirling, Scotland FK9 4LA, UK.

Gonzalez, J., (2017). Sales Forecasting and the Role of Predictive Analytics. Available at: http://vortini.com/blog/forecasting-predictive-analytics/ [Accessed 29 April 2017].

Gor, R. M., (2005). Industrial Statistics and Operational Management, Chapter 6: Forecasting Techniques. ICFAI Business School, Bodakdev.

Hadley, G., (1968). Introduction to Business Statistics (San Francisco, Holden-Day, Inc.

Harmon, M. & Klabjan, D., (2017). Activation Ensembles for Deep Neural Networks, arXiv:1702.07790v1 [stat.ML] 24 Feb 2017.

Hassani, H. & Silva, E., (2015). Forecasting with big data: A review. Annals of Data Science, (1),5–19.

Heaton Research. (2017, June 1). The Number of Hidden Layers. Available at: http://www.heatonresearch.com/2017/06/01/hidden-layers.html

Hecht, J., (2016). 5 Obstacles Facing Retail Small Businesses. Available at: http://www.huffingtonpost.com/jared-hecht/5-obstacles- facingretail_ b_ 7881350.html [Accessed 29 April 2017].

Hecht-Nielsen, R., (1990). Neurocomputing. Addison-Wesley, Menlo Park, CA.

Heizer, J. & Render, B., (2011). Operations Management, 10th Edition, Prentice Hall. Page 113, Chapter 4.

Hipel, K.W., McLeod, A.I., (1994). "Time Series Modelling of Water Resourc Environmental Systems", Amsterdam, Elsevier.

Hurwitz, J.S, Nugent, A.F, Halper & F.,Kaufman, M.A.(2013). Big Data for Dummies. Canada: John Wiley & Sons, Inc.

Husin, N. A., Salim, N. & Ahmad, A.R., (2008). Modelling of dengue outbreak prediction in Malaysia: A comparison of Neural Network and Nonlinear Regression Model, Proc. International Symposium in Information Technology (ITSim 2008), 2008, pp. 1-4.

Hyndman, R. J. & Athanasopoulos, G., (2012). Forecasting: principles and practice. OTexts. Available at: http://otexts.com/fpp [Accessed 29 April 2017].

Johnston, L.R., (2017). Curating Research Data. Volume One: Practical Strategies for Your Digital Repository. United States of America: Association of College and Research Libraries, A division of the American Library Association Chicago, Illinois.

Kang, S., (1991). An Investigation of the Use of Feedforward Neural Networks for Forecasting. Ph.D. Thesis, Kent State University.

Karsoliya, S., (2012). Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture. *International Journal of Engineering Trends and Technology- Volume3Issue6- 2012*. Department of Computer Science & Engineering, Maulana Azad National Institute of Technology, Bhopal, India.

Kihoro, J. M., Otieno, R.O. & Wafula, C., (2004). "Seasonal Time Series Forecasting: A Comparative Study of ARIMA and ANN Models", African Journal of Science and Technology (AJST) Science and Engineering Series Vol. 5, No. 2, pages: 41-49.

Kotsiantis, S. B., Kanellopoulos, D. & Pintelas, P. E., (2006). Data preprocessing for Supervised Learning, Computer Science, vol. 1, 1306-4428.

Kulkarni, E. G. & Kulkarni, R.B., (2016). WEKA Powerful Tool in Data Mining. International Journal of Computer Applications (0975 – 8887). National Seminar on Recent Trends in Data Mining (RTDM 2016).

Kumar, R. & Balara, A., (2014). Time Series Forecasting of Nifty Stock Market using Weka. Gurgaon, India.

Lippmann, R.P., (1987). An introduction to computing with neural, nets, IEEE ASSP Magazine, April, 4–22.

Liu, L.M. et al., (1994). Forecasting and Time Series Analysis Using the SCA Statistical System. Volume 1. Scientific Computing Associate Corp. Chicago, USA.

Mallon, S., (2016). 5 Benefits of Big Data for E-Commerce Companies and Shoppers. Available at: http://www.smartdatacollective.com/seanmallonbizdaquk/410001/5-benefits-big data-e-commerce-companies-and-shoppers [Accessed 29 April 2017].

Moyo, V. & Sibanda, K., (2014). On the Optimal Learning Rate Size for the Generalization Ability of Artificial Neural Networks in Forecasting TCP/IP Traffic Trends, Department of Computer science, University of Fort Hare, South Africa. International Journal of Computer Applications (0975 – 8887) Volume 105 – No. 4, November 2014.

Mustaffa, Z. & Yusof, Y., (2010). A Comparison of Normalization Techniques in Predicting Dengue Outbreak, 2010 International Conference on Business and Economic Research, vol.1 (2011) © (2011) IACSIT Press, Kuala Lumpur, Malaysia.

OECD, "Time Series," 2006. Available at: http://stats.oecd.org/glossary/detail.asp?ID=2708 [Accessed 29 April 2017].

Panchal, F. S., Panchal, M., (2014). Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network. *IJCSMC, Vol. 3, Issue. 11, November 2014, pg.455 – 464.* Department of Computer Science & Engineering, Kalol Institute of Technology, Gujarat, India.

Patro, S. G. K. & Sahu, K. K., (2015). Normalization: A Preprocessing Stage. Department of CSE & IT, VSSUT, Burla, Odisha, India.

Peace, I.C. et al., (2015). Effect of Learning Rate on Artificial Neural Network. International Journal of Engineering Research & Technology (IJERT).Vol. 4 Issue 02, February-2015. University of Port Harcourt, Nigeria.

Peng, R.D., (2015). R Programming for Data Science. Lulu Press.

Ravinder, H.V., (2013). Determining the Optimal Values of ES Constants- Does Solver Really Work? American Journal of Business Education- May/ June 2013. Montclair State University, USA.

Ripley, B.D., (1993). Statistical aspects of neural networks. In: Barndorff-Nielsen, O.E., Jensen, J.L., Kendall, W.S. (Eds.), Networks and Chaos-Statistical and Probabilistic Aspects. Chapman and Hall, London, pp. 40–123.

Rombaut, V., (2016). Top 5 Problems with Big Data (and how to solve them).

Schroeder, R., Rungtusanatham, M. J., & Goldstein, S., 2013. Operations Management in the Supply Chain: Decisions and Cases, 6th Edition, McGraw-Hill Higher Education, Page 261, Chapter 11.

Shahin, M.A., Maier, H.R. & Jaksa, M. B., (2004). Data Division for Developing Neural Networks Applied to Geotechnical Engineering. Journal of Computing in Civil Engineering ASCE, 18 (2), 105-114. University of Adelaide.

Shalabi, L. A., Shaaban, Z. & Kasasbeh, B., (2006). Data Mining: A Preprocessing Engine, J. Comput, Sci., 2: 735-739.

Sibi, P., Jones, S. A. & Siddarth, P., (2013). Analysis Of Different Activation Functions Using Back Propagation Neural Networks. Journal of Theoretical and Applied Information Technology. 31st January 2013. Vol. 47 No.3. SASTRA University, Kumbakonam, India.

Tang, Z., & Fishwick, P.A., (1993). Feedforward neural nets as models for time series forecasting. ORSA Journal on Computing 5 (4), 374–385.

Thota, L. S. & Changalasetty, S. B., (2013). Optimum Learning Rate for Classification Problem With MLP In Data Mining. Department of Computer Science, King Khalid University, Abha, KSA. International Journal of Advances in Engineering & Technology. International Journal of Advances in Engineering & Technology.

Tibell, R., (2015). Training A Multilayer Perceptron To Predict The Final Selling Price Of An Apartment In Co-Operative Housing Society Sold In Stockholm City With Features Stemming From Open Data. Kth Royal Institute Of Technology Computer Science & Communication, Stockholm, Sweden.

White, H., (1989). Learning in artificial neural networks: A statistical perspective. Neural Computation 1, 425–464.

Witten, I.H. & Frank, E., (2005). Data Mining: Practical Machine Learning Tools and Techniques. Second Edition. United States of America: Morgan Kaufmann Publications.

Wong, F.S., (1991). Time series forecasting using backpropagation neural networks. Neurocomputing 2, 147–159.

Yu, C. C. & Liu, B.D., (2002). A backpropagation algorithm with adaptive learning rate and momentum coefficient. Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on. Honolulu, HI, USA, USA.

Zhang, G.P., (2003). “Time series forecasting using a hybrid ARIMA and neural network model”, Neurocomputing 50, pages: 159–175.

Zhang, G.P., (2007). “A neural network ensemble method with jittered training data for time series forecasting”, Information Sciences 177, pages: 5329–5346.