

Enhancing Sales Forecasting by using Neuro Networks and the Popularity of Magazine Article Titles

Hani A. Omar, Duen-Ren Liu
 Institute of Information Management
 National Chiao Tung University, Hsinchu 300, Taiwan
 haniomar.iim97g@nctu.edu.tw; dliu@iim.nctu.edu.tw

Abstract—In this paper, we examine how the popularity information of magazines can be useful for sales forecasting. We propose a sales forecasting model based on Back Propagation Neural Network (BPNN) where the inputs are historical sales and the popularity indexes of magazine article titles. Our proposed model using the popularity of magazine article titles in the forecasting process can improve the accuracy of sales forecasting.

Keywords – Forecasting, Neural-Network, Popularity, Google Search engine.

I. INTRODUCTION

Sales forecasting is important for enterprises to make business plans and gain competitive advantages. The sales of magazines are usually affected by the contents of magazines. Popular contents can often boost the sales. It is interesting to estimate the effect of the magazines' popularity on sales forecasting by analyzing the contents of magazines and measuring their popularity using search engines. The objective of this research is to investigate how to utilize the popularity information of magazines derived from Google Search engine to improve sales forecasting.

Several contributions have been made in the field of forecasting [9]. Traditional time series methods are confined to the assumption of linearity, but some data are nonlinear. In order to overcome this limitation of traditional methods, many researchers use soft computing techniques such as fuzzy logic, neural network, fuzzy neural network, evolutionary algorithm etc [11]. If there are nonlinearities in the process being modeled, then a method which can account for these nonlinearities should produce superior forecast. Neural Networks are reported to be such a method [4]. Neural networks are also more noise tolerant, having the ability to learn complex systems with incomplete and corrupted data. In addition, they are more flexible, having the capability to learn dynamic systems through a retraining process using new data patterns [10].

In this paper, we proposed a sales forecasting model based on Back Propagation Neural Network (BPNN) where the inputs are historical sales and the popularity indexes of magazine article titles. We utilize the popularity of celebrity words appeared in article titles to predict the sales of magazines. Some tools and web sites provide functions to estimate the popularity of keywords based on search counts or web-page counts. If popularity count is tied directly to ad revenue (such as with ads shown with YouTube videos), revenue might fairly accurately be estimated ahead of time if all parties know how many views the video is likely to attract. Meanwhile, Digg

allows users to submit links to news, images, and videos that are interest to the site's general audience [14]. Moreover, search data have the potential to describe user interests in a variety of economic activities in real time [1]. While other research shows the observation that search counts are generally a prediction of consumer activities, such as purchasing music in the future [3].

These researches and the need to forecast are our motivation to use Google Search Engine for measuring the popularity of title words. The search engine results of title words of magazine articles reflect the readers' interests and are important indicators of the sales. Using non-linear historical data of sales and the popularity of title words in our proposed model can contribute to improve the forecasting performance. Our experiment evaluation shows that our proposed model using the popularity of magazine article titles outperforms traditional methods.

The rest of this paper is organized as follows. Section II presents a short survey of existing literatures on Neural Networks and Double Exponential Smoothing methods. Section III presents the proposed model for sales forecasting. Experimental evaluation is reported in section IV. The conclusions are summarized in section V.

II. THEORETICAL BACKGROUND

Several contributions have been made in the field of forecasting. Statistical forecasting methods outperform the simplistic models in terms of forecasting accuracy [6]. However, there are two major drawbacks of these methods. First, for each problem, an individual statistical model has to be chosen that makes some assumptions about underlying trends. Second, the power of deterministic data analysis needs to be exploited for single time series with some hidden regularity [15]. Artificial Neural Network has been adopted to replace traditional methods because of better performance, adaptive capability and so on [5].

A. Introduction to Artificial Neural Network

An artificial neural network (ANN), often called a "neural network" [12]. ANN is an information processing system that has been developed as generalization of mathematical models of human neural biology. ANN is composed of nodes or units connected by directed links. Each link has a numeric weight [8]. ANN adjusts the weights such that the predicted values and the real values are as close as possible [13]. There are more than one type of ANN depending on the way of learning and adjusting the weights for hidden layer(s). Back

Propagation Neural Network is one of the mostly used neural networks and has been applied to many fields [2]. The structure of BPNN is illustrated by Fig. 1.

Step 1: In the beginning, we need to construct the model and determine the number of inputs; initial weights; transfer function and number of hidden layers.

Step 2: After collecting and preprocessing the data, the data set is splitting into two sets, training set and testing set.

Step 3: The training set is used as an inputs to train and learn the weights.

Step 4: The testing set is used to evaluate the model.

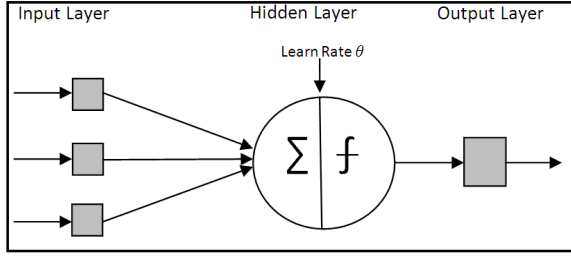


Figure 1. Paradigm of Neural Network

B. Double Exponential Smoothing Method

Double Exponential Smoothing method models a given time series using a simple linear regression that smoothes both the level and the trend components in the data; as such, it has the advantage of being a more realistic technique than the simple exponential smoothing which does not consider trend effect [7, 16]. In this research, we compare our proposed method with the Double Exponential Smoothing (DES) method. It has two parameters for data smoothing and trend smoothing. The prediction error could be reduced by tuning the smoothing parameter. So choosing the right smoothing parameter can produce a better prediction. Eq. (1) is used to estimate the predicted value.

$$S_{t+1} = \alpha x_t + (1 - \alpha)(S_t + b_t), \quad (1)$$

where S_{t+1} is the predicted sale value at time $t+1$; x_t is the sales value at time t ; t is a time and $t > 1$; α is a data smoothing factor, and $0 < \alpha < 1$; b is the trend effect. Eq. (2) is used to find the trend effect.

$$b_t = \beta(S_t - S_{t-1}) + (1 - \beta)b_{t-1}, \quad (2)$$

where β is trend smoothing factor and $0 < \beta < 1$.

III. PROPOSED METHODOLOGY

In this section, we describe our proposed sales forecasting model that utilizes Back Propagation Neural Networks (BPNN) and the popularity indexes of magazine article titles. Fig. 2 illustrates the proposed model that contains four parts described as follows.

A. Popularity analysis of magazine contents

Many factors may have influence on the sales such as content, social stratification, Age stratification, locations

of outlets, etc. We focus on the contents of magazine article titles. The buyers usually cannot get the whole content before buying magazines; they just can take a quick look for articles by reading article titles. So, publishers usually pick attractive and sometimes ambiguous headlines or article titles to increase the sales. Our model extracts the article titles for each issue of magazines. Then we tokenize the content into title words (e.g. celebrity words) for previous and new issues of magazines. We note that meaningful title words (e.g. celebrity names) are usually formed by a combination of several individual words.

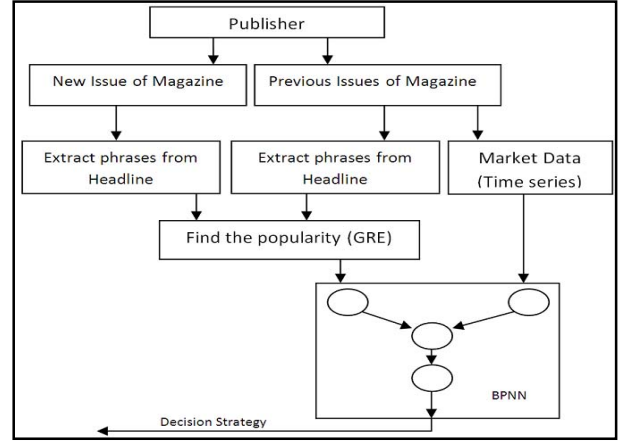


Figure 2. The framework of proposed forecasting model

Google Search engine is used to derive the popularity scores of the title words by getting the search result, i.e., the number of web-pages for each title word during one month search period before the publication date of each issue. We accumulate the popularity scores of all title words for each issue to represent the popularity index of the issue.

$$P_t = \sum_{j=1}^{N_t} R_{t,j}, \quad (3)$$

where P_t is the popularity index of issue t ; N_t is the number of title words in issue t ; and $R_{t,j}$ is the outcome of Google Search for title word j in issue t . The popularity indexes are normalized by dividing the maximum value over all issues.

B. Market Data

The most interesting part for the publishers is the sales. Our data has weekly sales because it's a weekly magazine. These sales values will be used as the input of the BPNN and Double exponential smoothing (DES) for sales forecasting. The sales values are normalized by dividing the maximum value over all issues.

C. Back Propagation Neural Network

This is the main part of our model. We use the simple structure of BPNN, one layer for the input, hidden layer and output layer, respectively. For the input layer, we have three inputs that represent the popularity index of the predicted issue $t+1$ and the most recent two sales amounts

of issues t and $t-1$. In order to avoid computational problems, to meet the BPNN requirement and to enhance the learning process, the input data of neural network must be normalized. In our model, we use the simple normalization by dividing all the values with the largest value presented in the time series data. The synaptic weights and the threshold for the hidden layer are initialized, as shown in Eq. (4).

$$\sum_{i=1}^m x_i \times w_i - \theta, \quad (4)$$

where m is the number of inputs; x_i is the input value after normalization; w_i is the synaptic weight and θ is the threshold for summation. The result of previous formula is the output of the input layer and also becomes an input for hidden layer where the Sigmoid Function is used as the activation function. Sigmoid Function produces values in $[0, 1]$. This value could be fluctuating depending on the value of inputs. But the input layer has different initialized weights, so tuning the synaptic weights, the threshold of summation, and the learning rate, can help to make the forecasting converges to the real value. Fig. 3 illustrates the process of BPNN.

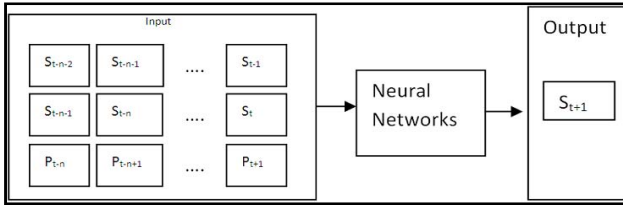


Figure 3. The process of BPNN

BPNN has two stages of process: training weight process and the testing process. BPNN needs weight training to train the synaptic weights as well as the threshold to enhance the result by comparing the predicted sales value and the real sales value.

D. Making Decision

After the testing process, the predicted value needs to be evaluated by considering the prediction error between the predicted value and the real value. Then, the publisher can take decision depending on the predicted value and other factors. These factors could be related to age stratification of the outlet location or the other constraints related to publishing process. Also, the popularity of title words and sales predictions can help the editors to pick the title words which could increase the sales of magazine.

IV. EXPERIMENTAL RESULT AND ANALYSIS

A. Experiment Design

Our data set is the sales data of a Chinese publication magazine which is categorized as weekly magazine. The data set contains article titles of 133 weekly issues, sold through 6205 outlets from June 2009 to December 2011. We pick the top 10 outlets with the highest sales from

those outlets which have positive correlations of sales and the popularity indexes of magazine issues. For the sales data of those top 10 outlets, we use the first 80 issues of sales for each of the top ten outlets as the training set and the remaining 53 issues of sales as the testing set. After deriving the prediction values, the measurement of the accuracy is required to compare with other methods. We use the Root Mean Square Error, as defined in Eq. (5), to evaluate the performance of our proposed method.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (sales_t - prediction_t)^2}{n}} \quad (5)$$

where n is the number of weeks in the testing set for each outlet; $Sales_t$ is the real sales and $prediction_t$ is the predicted sale value of a sales forecasting method.

In our experiments, 10 outlets are chosen to predict the sales by using BPNN. For these outlets, we normalize the sales and the popularity indexes. Training weights need initializations. So, we try to train our data by taking different weights and threshold to get the best result. Three methods have been implemented to evaluate the effect of using the popularity of magazine contents in the prediction process. The first one is BPNN by using the popularity and sales as inputs while the second is BPNN by taking the sales values as inputs; the last method is Double Exponential Smoothing for sales forecasting.

BPNN with popularity and sales: According to the methodology section, BPNN depends on the normalized inputs for historical sales data and the popularity of title words. After using the training set in BPNN, the testing set is used to get the predicted values.

BPNN with sales only: For evaluating the effect of using the popularity of magazine contents (title words), in this experiment, we used just the sales of two previous issues as inputs for BPNN without considering the popularity of title words.

DES with sales only: This experiment uses statistical method for predicting sales. As mentioned in section II, this method depends on data and trend smoothing factor as shown in Eqs (1) and (2). The time series of Sales data are the inputs for DES. We take different data and trend smoothing parameters to find the best result of DES.

B. Comparison and Analysis

Fig. 4 shows the comparison of BPNN (with popularity and without) and DES using RMSE. As we can see, using the popularity of title words in BPNN for forecasting sales of magazines outperforms both of the BPNN without popularity and the DES method. Fig. 4 shows that the popularity of the title words of articles in the celebrity section of the magazine has positive effect on increasing the prediction accuracy of sales forecasting for individual outlets and their average. Accordingly, the prediction by using the popularity indexes performs better than the other

two methods which just use the sales data in the forecasting process.

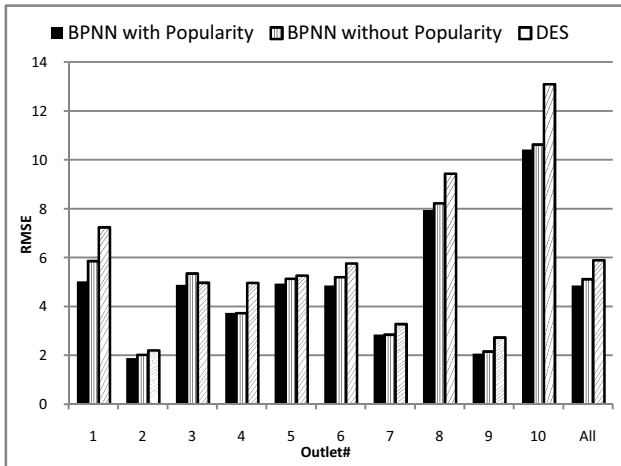


Figure 4. RMSE for BPNN (with/without popularity) and DES

From our experiments, the results clearly show that the proposed BPNN with popularity adds value to the forecasting process comparing to the BPNN without popularity or to the DES method.

V. CONCLUSION

In this paper, a novel approach is developed to improve the forecasting process for sales predictions. This method combines the historical sales data and the popularity of title words of magazine articles to enhance the accuracy of predictions. The effect of popularity has been tested in three different methods. The model has been tested with popularity and without considering the popularity. Using the popularity as input in BPNN has added value to the forecasting process and increased the prediction accuracy. Also, our proposed prediction model outperforms the DES method. This result also suggests the magazine editors to get helpful way to pick most effective titles and headlines for their articles. In this paper, we analyze the article titles of the celebrity section. In our future work, we will investigate the article titles of other sections for enhancing the forecasting results.

VI. ACKNOWLEDGEMENTS

This research was supported in part by the National Science Council of Taiwan, under grant NSC 99-2410-H-009-034-MY3.

VII. REFERENCES

- [1] Z.H.I. Da, J. Engelberg, and P. Gao: 'In Search of Attention', *The Journal of Finance*, 2011, 66, (5), pp. 1461-1499
- [2] M. Gao, Y.-c. Guo, Z.-x. Liu, Y.-c. GUO, and Y.-p. ZHANG: 'Feed-forward Neural Network Blind Equalization Algorithm Based on Super-Exponential Iterative', *International Conference on Intelligent Human-Machine Systems and Cybernetics*, 2009

- [3] S. Goel, J.M. Hofman, S. Lahaie, D.M. Pennock, and D.J. Watts: 'Predicting consumer behavior with Web search', *Proceedings of the National Academy of Sciences*, 2010, 107, (41), pp. 17486-17490
- [4] L. Hamm, and B.W. Brorsen: 'Forecasting Hog Prices with a Neural Network', *Journal of Agribusiness*, 1997, 15
- [5] M. Hayati, and Y. Shirvany: 'Artificial Neural Network Approach for Short Term Load Forecasting for Illam Region', *IJECSE*, 2007, 1, (2)
- [6] E.J. Marien: 'Demand Planning and Sales Forecasting: A Supply Chain Essential', *Supply Chain Management Review*, 1999, 2, (4), pp. 76-86
- [7] J. Mentzer: 'Forecasting with adaptive extended exponential smoothing', *Journal of the Academy of Marketing Science*, 1988, 16, (3), pp. 62-70
- [8] C.A. Mitrea, C.K.M. Lee, and Z. Wu: 'A Comparison between Neural Networks and Traditional Forecasting Methods: A Case Study', *International Journal of Engineering Business Management*, 2009, 1, (2), pp. 19-24
- [9] H.-N. Nguyen, Q. Ni, and M.D. Rossetti: 'Exploring the Cost of Forecast Error in Inventory Systems'. *Proc. Proceedings of the 2010 Industrial Engineering Research Conference 2010*
- [10] A.A. Philip, A.A. Taofiki, and A.A. Bidemi: 'Artificial Neural Network Model for Forecasting Foreign Exchange Rate', *World of Computer Science and Information Technology Journal (WCSIT)*, 2011, 1, (3), pp. 110-118
- [11] M. Shah: 'Fuzzy based trend mapping and forecasting for time series data', *Expert Systems with Applications*, 2012, 39, (7), pp. 6351-6358
- [12] Y. Singh, and A.S. Chauhan: 'NEURAL NETWORKS IN DATA MINING', *Journal of Theoretical and Applied Information Technology*, 2009, 5, (1)
- [13] D. Svozil, V. Kvasnicka, and J.i. Pospichal: 'Introduction to multi-layer feed-forward neural networks', *Chemometrics and Intelligent Laboratory Systems*, 1997, 39, (1), pp. 43-62
- [14] G. Szabo, and B.A. Huberman: 'Predicting the popularity of online content', *Commun. ACM*, 2010, 53, (8), pp. 80-88
- [15] F.M. Thiesing, and O. Vornberger: 'Sales forecasting using neural networks'. *Proc. Neural Networks Computational Intelligence 1997*
- [16] M. Xie: 'A study of exponential smoothing technique in software reliability growth prediction', *Quality and reliability engineering international*, 1997, 13, (6), pp. 347