

Store Sales Forecasting using Artificial Neural Network

Arup Kumar Bhattacharjee, Manas Ghosh, Soumen Mukherjee and Sudipta Patra

RCC Institute of Information Technology, Bahelghata, Kolkata – 700015, West Bengal, India; arupk.b@gmail.com, soumou601@gmail.com, manasrcciit@gmail.com, sudipta.patra2@gmail.com

Abstract

Forecasting is the use of historic data to predict the future trends. Forecasting provides an important benchmark for organizations which have a long-term perspective of operations. In this paper Time series analysis is used to predict the future sales. Also various industrial areas where this technique can be used are explored. Application of the proposed forecasting technique is discussed in various industries.

Keywords: Artificial Neural Network, Forecasting, Neural Network

1. Introduction

Forecasting is important for almost every aspect of life. It's probability of accurate future results. It's very tough to predict sales because it depends on various real time factors not past statistics. Different competitors, Goodwill, Product Quality, make clear impact on sale¹.

The reasons behind forecasting are so important because of following

- Improved sells performance
- Marketing
- Business control
- Inventory management
- properly planning

To solve forecasting³ problems there are variety of forecasting techniques present. Each formula has separate scenario where it works perfectly. There are three basic types method present qualitative techniques, time series analysis and causal models.

The first uses qualitative information (expert opinions) and data regarding special events⁴ of the type already mentioned, and should or may not take the past into thought.

The second, on the opposite hand⁵, focuses entirely on patterns and pattern changes, and so depends entirely on historical information⁶.

The third uses extremely refined and specific data regarding relationships between system components, and is powerful enough to require special events into consideration.

Time series analysis techniques, the past is very important to causal models⁷.

*Author for correspondence

2. Related Works

Demand forecasting method presents the technique where the different products sales increase depends on product feature (colour, type etc). Author mainly focussed on analysis of forecasting approach using demand management of products session based¹. Frank et al. presents forecasting women's apparel sales using mathematical modelling. They used different models like ANN model, multivariate fuzzy logic etc. Typically, sales data have a high noise level. The problem is intensified by a number of outliers (with extremely high or low values). The authors worked on session based sales product demand difference¹⁶.

In¹⁷ author concentrate on festive centric sales. So for sales, stores mostly focus on holidays. And holidays are on specific week of the year like Christmas day (52nd week), thanks giving day (47th week) etc. Hence sales have direct relation with week number (holidays).

2.1 Database

Cultured datasets are-

Table 1: stores.csv

There are three fields

- i. store: the store number
- ii. Type: it's anonymous information about type of the store
- iii. Size: it's also anonymous information about size of the store

Table 2: train.csv

There are 5 fields

- i. Store: the store number
- ii. Dept: the department number
- iii. Date: the week
- iv. Weekly Sales: sales for the given department in the given store (may have negative data in weekly sales most likely due to product return by customers from the previous weeks purchase)¹²

- v. Is Holiday: whether the week is a special holiday week

Table 3: test.csv

Except the withheld of weekly sales this file is same as train.csv
There are 4 fields

- i. Store: the store number
- ii. Dept: the department number
- iii. Date: the week
- iv. IsHoliday: whether the week is a special holiday week

Table 4: features.csv

It contains additional data related to the store, date, department and regional activities like temperature etc. It contains the following fields:

- i. Store - the store number
- ii. Date - the week
- iii. Temperature - average temperature in the region
- iv. Fuel_Price - cost of fuel in the region
- v. Markdown1-5 -It anonymize data depends on various promotional offer. It's not available for all store every time and it's only available after month of November each year. Not available records denoted by NA
- vi. CPI - the consumer price index
- vii. Unemployment - rate of unemployment¹⁵
- viii. Holiday - whether the week is a special holiday week

3. Proposed Work

3.1 Time Series Analysis

Time series analysis is that when our data plotted across time. It's a set of method to analysis our data. It's mainly three components sessional components, trend components and irregular or error component.

Degree of prediction (Y_t) depends on identification of sessional and trend component (Equation No. 1).

$$Y_t = \text{trend} + \text{sessional} + \text{irregular (random)} \quad (1)$$

It's an example of number of birth from a simple birth ratio from January 1988 to 2002.

Figure 1 plot birth ratio vs time in time series analysis. Vertical axis plotted as time and horizontal axis plotted as birth ratio¹⁰.

Figure 2 shows time series analysis in decomposed form. Top plot shows original time series. Next (second from top) shows trends com-

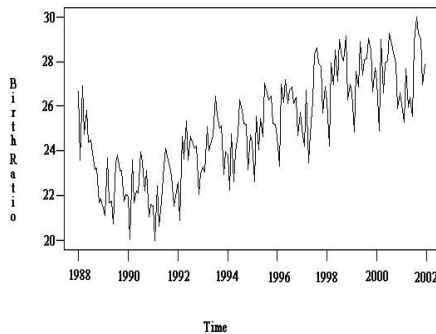


Figure 1. Birth Ratio vs Time Plot.

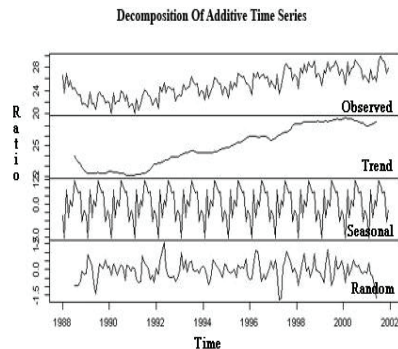


Figure 2. Time Series Analysis.

ponent and third plot shows seasonal component and the bottom plot shows random components.

In previous presentation author worked on simple linear regression and median model for prediction.

4. Results

Figure 3 shows prediction for a specific department of a store. Here vertical axis define week and horizontal axis define total sales. Around data of 3 years (148 weeks) have been used. There are 4 time most sales pick up every year in this image which was very closely affected for 4 festivals every year. The black line plotted our cultured data set and the red line shows our prediction result. Black line ended after 148 weeks (cultured data set). The difference between black and red lines represents error.

5. Applications in Industry

Sales forecasting is a real world application. Various types of prediction can be done based on data

1. Short-Term Forecasting -The general trend of sales is less important for short-term fluctuation.
2. Medium-Term Forecasting- It is most important in the area of business budgeting. For periods up to 1 year ahead.
3. Long-Term Forecasting-Needed mainly by financial accounts for long term resource implications. For period of 3 years and upwards. (our projects closely follows this)

6. The Evolution of Topics over Time

The prediction is evaluated based on the weighted mean absolute error (WMAE). WMAE (Equation No. 2) gives information about how the error obtained but it did not guide which is good prediction.

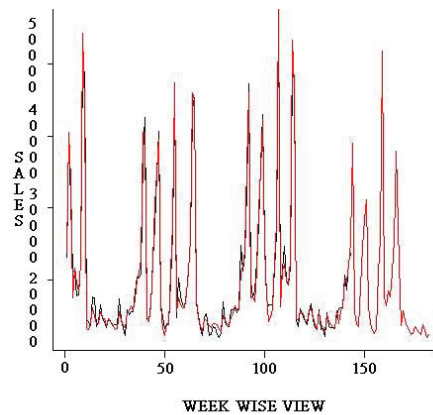


Figure 3. Week wise Sales Plot.

$$WMAE = \frac{1}{\sum w_i} \sum_{i=1}^n w_i |y_i - \hat{y}_i| \quad (2)$$

where

n denotes number of rows

\hat{y} denotes predicted sales

y_i denotes actual sales

w_i are weights. $w = 5$ if the week is a holiday week, 1 otherwise

7. Conclusion and Future scope

The main barrier in this forecasting system is lack of data. Sales world are very noisy. It's very tough for datasets to fit in a specific model easily. Different data models perform works best in different scenario. Here work is done with 3 years' dataset but sometimes sales closely changed with environment or political status of state. Even several companies functioned on different region (like amazon, eBay, Alibaba) with different geographical location people habituated with different type of

life style and products. And these factors make effect on sales. This paper did not work on different product, brand or feature whereas the prediction system is based on total sale. Neural network may also be used extensively for prediction.

8. References

1. Nenni ME, Giustiniano L, Pirolò L. Demand Forecasting in the Fashion Industry: A Review. 2012.
2. Matias AJ. Budgeting and Forecasting - The Quick Reference Handbook. Publisher: Matias Interactive Learning; 1st Edition. 2012 Feb 27. ISBN: B007FPK904.
3. Gilliland M. The Business Forecasting Deal: Exposing Myths, Eliminating Bad Practices, Providing Practical Solutions, Publisher: John Wiley & Sons. ISBN-10: 0470574437, ISBN-13: 978-0470574430. 2010 Jun 25.
4. Krueger RA. Business Forecasting: A Practical, Comprehensive Resource for Managers and Practitioners. Publisher: Book Surge Publishing. ISBN-10: 1419677799, ISBN-13: 978-1419677793. 2008 Dec 2.
5. Dougherty J. Sales and Operations Planning - Best Practices, Publisher: Trafford Publishing. ISBN-10: 1412082102, ISBN-13: 978-1412082105. 2006 May 2.
6. Kabacoff R. R in Action. Publisher: Manning Publications; 1 edition. ISBN-10: 1935182390, ISBN-13: 978-1935182399. 2011 Aug 27.
7. Sales Forecasting & Market Potential: Best Practices in India by Abhishek Gupta, Publication Date: 2013 Sep 21. Available from: <http://management.cloud-journals.com/index.php/IJAMR/article/view/Mgmt-25>
8. Fildes R, Beard C. Forecasting systems for production and inventory control. International Journal of Operations and Production Management. 1992; 12(5):4-27.
9. Makridakis S. Forecasting: its role and value in planning and strategy. International Journal of Forecasting. 1996; 12:513-37.
10. Plossl G. Getting the most from forecasts. Production and Inventory Management Journal, First Quarter. 1973; 1-15.

11. Wacker JG, Sprague L. The impact of institutional factors on forecast accuracy: a manufacturing executive perspective. *International Journal of Production Research*. 1995; 33(11):2945–58.
12. Wacker JG. A definition of theory: implications for different theory-building research methods in operations management. *Journal of Operations Management*. 1998; 16:361–85.
13. Available from: <http://www.entrepreneur.com/article/204878>
14. Available from: <http://www.forbes.com/sites/scottedinger/2013/06/03/four-principles-for-great-sales-forecasts/>
15. Cleger-Tamayo S, Fernandez-Luna JM, Huete JF. On the use of Weighted Mean Absolute Error in Recommender Systems. 2012.
16. Frank C, Garg A, Raheja A. Forecasting women's apparel sales using mathematical modeling. 2003.
17. Bhattacharjee AK, Patra S, Barman RD, Bera B, Mukherjee S, Ghosh M. Store Sales Forecasting: A Case Study. IV International Conference on Dynamism and Contemporary Practices of Business Management. 2015 Dec 05.