

A Survey Paper on Sales Analysis And Forecasting In Shopping Mart

Amit Kumar¹, Kartik Sharma², Anup Singh³, Dravid Kumar⁴

^{1,2,3,4} Dept of Computer Eng.

^{1,2,3,4} SKNSITS, Pune, INDIA

Abstract- The enormous number of Shopping Marts in the market today has lead to the need of having analysis tools which helps to determine whether the organization is meeting its desired sales goals. The goal is to analyze the database transactions of Shopping Marts using various data mining techniques and algorithms such as affinity analysis, logistic regression and linear regression. Idea is to develop a system which takes input, the database transactions of sold products, segments the data obtained, analyzes the graphs and extracts the market trends and product sales patterns. The system optimizes this data on the basis of market requirements thereby improving sales and merchandise planning in a way that to increase the overall productivity and profits of the organization.

Keywords- Transactions, Regression, Affinity analysis

I. INTRODUCTION

With the growing economy, the shopping trends have also been increased; therefore, the challenges for shop owners to perform better in market have also been increased. Therefore, shop owners and businessmen need to analyze their sales data and forecast future sales & ideas in order to raise their sales graph and profit percentages. In this way they can avoid going out of stock on most selling products, provide discounts on product at right time, avoid losses in investments. The regular and new customers can be handled much easily and profitably. With the increase in universal data volume, the technology of big data and its analytical processes are generally used to provide the description about massive datasets. Compared with other traditional datasets and its processes, big data includes semi structured and unstructured data that need more real time analysis. Big data also gets details about new prospects for determining new values, supports us to improve an in-depth understanding of the hidden values, and also incurs new challenges, for instance, how to exceptionally organize and manipulate such big datasets. The volume of information from various sources is growing large, it also provides about some challenging issues demanding rapid resolutions. Big data visualization process is another vital process which takes an important place in big data analytics

problems. Because through data visualization only the final report of data analytics will be visualized.

II. RELATED KNOWLEDGE ABOUT SALES FORECAST

Sales Forecast [1] is a technology which using the mathematical way to predict the sales of one or several varieties of products of a company in a specific period of time in the future. Based on various factors and combined with the company's sales performance, it presents us a feasible sales target through certain analyzing method. With the help of sales forecast, salesman can be greatly motivated to promote product sales as soon as possible to realize the products' value. Enterprise manager can also benefit a lot from it. They can rearrange the producing process accordingly so as to reduce operational risk and improve the company's competitiveness.

With the help of various information technology, enterprises have accumulated hundreds even thousands of gigabytes sales history data. Prediction usually involves massive data processing. However, in face of these massive data, traditional forecasting system cannot meet the new forecast requirements anymore, such as operating efficiency, computing performance, accuracy and storage space, large amounts of historical data is now in an offline state, turned into a kind of "data grave". Further more, the traditional database technology, which used to be used alot for forecasting, is weak in knowledge expressing and reasoning. For the two reasons mentioned above, a model that both have the capability of massive data processing and knowledge discovery can fix these sticky problems. Compared to traditional data processing tech, Data Ming is more specialized in massive data based knowledge discovery, which makes it a better solution than traditional database tech to be used in prediction model.

III. FORECASTING METHOD

A. Double Moving Average Model

Double Moving Average [3] is a average computing

method based on Single Moving Average model. Firstly, it uses the single moving average twice to get the one moving average value and one moving average based moving average value, here we call it twice moving average value. Then the algorithm uses the two kinds of value to calculate the target data, according to its compute model.

B. Exponential Smoothing Model

Exponential Smoothing [4] is a kind of moving average method developed on the basis of time series analysis and forecast, it is a most often used method in production prediction, especially in short-term forecast. It was built on the theory that the trend of time series has the characteristics of stability and regularity.

Common practice for the use of Exponential Smoothing is use it to get predicted value of the historical data, then use it again to predict recently from maximum recent demand and the predicted value mentioned above. The final step is using the recent trend factors to adjust the result. The model is listed down below:

Exponential smoothing is a very effective marketing budget, statistical methods. You can use Excel to predict which is time saving and effective. But there are some limitations. First, a more complete historical data is required before using this model; Second, if season factors influences business sales a lot, time series decomposition is more applicable than exponential smoothing. It would be wise to choose between exponential smoothing and qualitative forecasting according to the specific circumstance before making a final conclusion.

C. Grey-Markov Model

To begin with, $x(0) \square \square \{x(0) (1), x(0) (2), \square, x(0) (n)\}$ is assumed as the original time sequence, Grey-Markov prediction model can be divided into two parts.

1. GM(1,1) Model

Research shows that there is no positive connection between the first element of original time sequence and the GM(1,1) model. So, a kind of new measure is taken to build an optimal GM(1,1) forecast Model. Detailed information about constructing this model has beyond the scope of this article, and the details can be found in reference.

2. Markov model

This step is based on the result of GM(1,1) model. According to the difference range between actual value and predicted value by GM(1,1), the series is divided into different groups. Each group is called a state. Here we assume the states are E_1, E_2, \dots, E_n .

IV. EXISTING SYSTEM

Existing systems helps in classifying the various systems which are already in the market. The classification of existing systems is given as follows:

Mindtree:

- It has a reporting capability which provides a complete view of customers.
- It consists of a manager module which is used for planning, preparing and tracking customer interactions.

Micro Strategy:

- To create superior data visualizations it uses a powerful visual data exploration interface.
- The data from multiple sources is combined.
- Feasible advanced analytics are used for trend analysis and financial analysis.

Tableau:

- It has an excellent user interface: It has highest number of customers as it provides convenient, straightforward and manageable user interface.
- Integration: It integrates well with big data platforms; including Hadoop it also offers support for Google Big Query API – a boon for organizations that want highly detailed analytics.

Drawbacks of Existing Systems:

- Tools which use big data sets are imprecise.
- Data breach can occur due to big data analytics..
- Existing systems like Micro Strategy have a very complex development environment. Despite of the fact that it uses Schema & SQL Engine it has a very high development speed.

V. OUR APPROACH

As shown in (Fig.1) The Sales Analytics tool is one which takes input the sales transactions data by using data mining techniques [4]. The raw data is then segmented as products that are sold, the combination of products sold together, product ratings and product review [3]. These are then profiled into graphs to obtain the market trends and patterns from which we obtain the fast selling and slow selling products. Then the tool decides which product’s price needs to be increased and which products should be discounted or discarded. Merchandize planning is updated accordingly to balance the stocks so that there is no shortage of products. The prices of products are revised according to their demand to increase the profit of organizations.

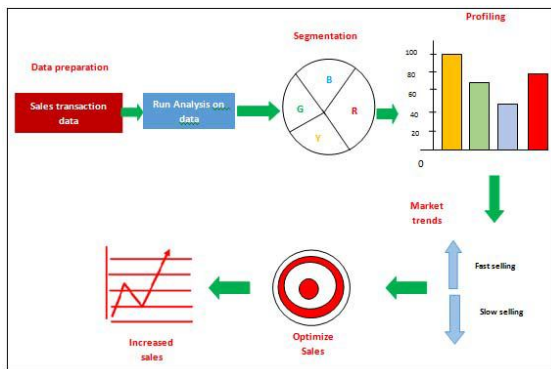
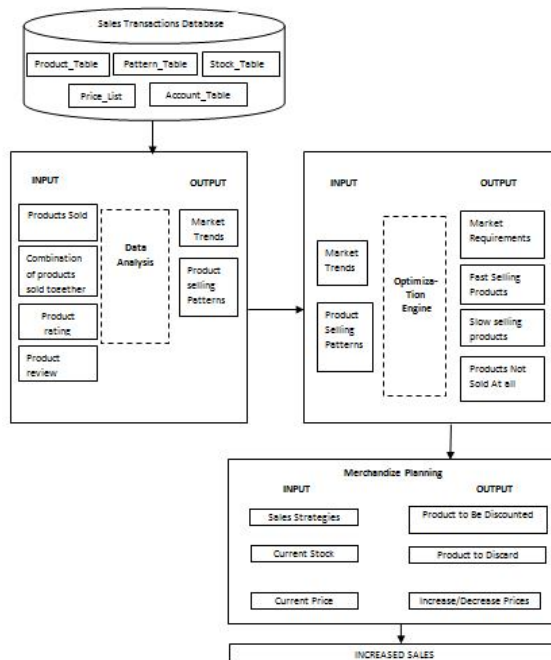


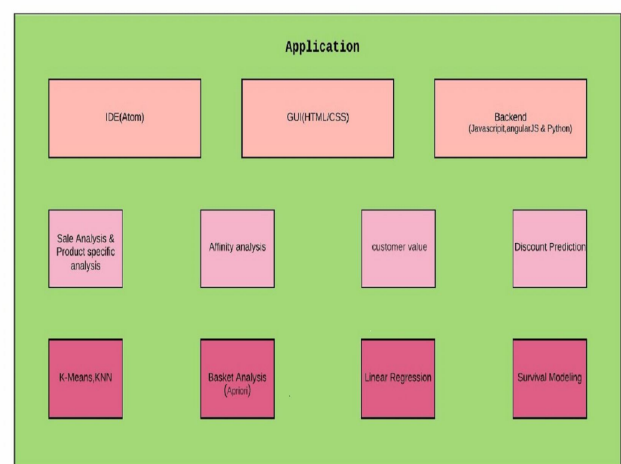
Fig 1. Overview of the system.

The block diagram (Fig.2) shows the working of every unit of the Sales Analytics Tool.



- Database: The database contains all details of the products sold, product selling patterns, stocks of products, prices of products, account details.
- Product Sold: Here we analyze the products which are sold and those which are not sold.
- Product Rating and Reviews: The reviews from customers who have made a purchase from the organization helps to understand which products are liked by them.
- Market Trends: Sales patterns of the product sold indicate market trends, here we do the changes in customer demands e.g. increasing or decreasing product price also add new product or any service launch for growing sales.
- Product Selling Patterns: According to the sales of products, generate a pattern of sales which shows fast-selling and less-selling products.
- Sales Strategies: It will perform various sales strategies to optimize online sales like Recommend Products and Up sells, special offers on products, Increase Urgency, Add reviews and ratings.
- Current Stock: It will Check the product stock if product is out of stock then balance the stock and also make offer discount on less-selling products.
- Current Prices: It will Evaluate and change the product prices (increase/decrease) according to sales.
- Increased Sales: The overall aim of this system is to increase productivity of products and increase the overall sales profits.

VI. SYSTEM ARCHITECTURE



VII. CONCLUSION

Due to tremendous growth of shopping mart it has become increasingly necessary for organizations to utilize automated tools to find the desired information resources, and to track and analyze the sales transaction. Patterns of the type of products that are sold frequently can be found out using data mining of sales data. Thus by observing the vast increase in shopping mart there was a need to make use of various data mining algorithms to optimize the sales of our ecommerce website. By performing analysis on database of the products we will update the stocks of our inventory by using algorithms like Market Basket Analysis. This algorithm helps to produce patterns product sales. Then we will categorize the products as fast-selling and less-selling products using algorithms like logistic and linear regression. Accordingly we will update stocks of fast-selling products.

REFERENCES

- [1] Dr. Zainab Pirani, Anuja Marewar, Zainab Bhavnagarwala, Madhuri Kamble."Analysis and Optimization of online sale of products" international Conference on innovation in information in information ,Embedded and Communication system(ICIIIECS) 2017
- [2] Xiaoyan Wang "Application Research on Marketing Data Analysis using Data Mining technology" 2016 International Conference on Smart Grid and Electrical Automation
- [3] Hyung-il Ahn, W. Scott Spangler" Sales Prediction with Social Media Analysis" 2014 SRII Global Conference
- [4] Clemens Schwenke, Johannes Ziegenbalg, Klaus Kabitzsch, Volodymyr Vasyutynskyy" Simulation based Forecast of Supermarket Sales" 2012 IEEE
- [5] M.D. Anto Praveena, Dr. B. Bharathi "A Survey Paper on Big Data Analytics" INTERNATIONAL CONFERENCE ON INFORMATION, COMMUNICATION & EMBEDDED SYSTEMS (ICICES 2017)
- [6] Yu-shui Geng, Xin-wu Du " The Research Of Data Mining Based Sales Forecast" 2010 IEEE