

Classification of Customer to Upgrade Profits in Retail Market with Deep Learning Methodology

I. Gayathri Devi^{1*}, G. Surya Kala Eswari², G. Kumari³

^{1,2,3} CSE, Pragati Engineering College, JNTUK, Kakinada, India

*Corresponding Author: igayathridevi9@gmail.com

DOI: <https://doi.org/10.26438/ijcse/v7i5.188191> | Available online at: www.ijcseonline.org

Accepted: 15/May/2019, Published: 31/May/2019

Abstract— Capital investment in retail sector and competition in the market has changed the style of marketing. At the same time the enhancements in the field of information technology provided an upper hand to the marketer to know the exact need, preference and purchase trend of the customer. By knowing the actual need, preference and purchase trend of customers the marketer can make a future business plan to increase the sale and earn more profit. This paper provides a framework to the retail marketer to find the potential customer by analyzing the previous purchase history of the customer. This task can be accomplished by the use of data mining technique. In this paper we have used k-mean clustering algorithm and Naive Bayes' classifier for identifying potential customer for a particular section of products of the retailer.

Keywords—Naive Bayes, Cluster, Centroid, Foreign Direct Investment (FDI)

I. INTRODUCTION

According to the Global Retail Development Index 2012, India ranks fifth among the top 30 emerging markets for retail. The recent announcement by the Indian government with Foreign Direct Investment (FDI) ^[1] in retail, especially allowing 100% FDI in single brands and multi-brand FDI has created positive sentiments in the retail sector. Since revenue and the competition is increasing in the field of retail marketing therefore every marketer wishes is to increase profits through sales, but this can't be possible without managing customers.

Every business organization has a primary goal to increase sales and through which it earns profit. To increase sales they apply marketing and sales promotion strategies so that customers can know about their product and their promotion activities such as a discount on a particular item or an entire section. Generally for these activities organization apply mass marketing which causes decrease in intensity of effort. If they apply their effort into a particular direction then the intensity of effort will increase. The current marketing and sales promotion in retail field is almost dependent on the mass marketing. The marketer promotes the product to the mass of the customer without knowing their need of such products. Mass marketing is a market coverage strategy in which a firm decides to broadcast a message that will reach the largest number of people possible.

Traditional mass marketing has focused on radio, television and newspapers are the medium used to reach the broad audience. So there is a need to overcome this problem by using computer software methodologies like machine learning ^[10], Data Mining and several other fields are used.

II. RELATED WORK

The marketer may investigate the reasons a customer or a group has not purchased over a long period of time. On the basis of these three parameters the customers can be grouped into two categories, i.e. more profitable and less profitable category. For this purpose we have used K-mean clustering algorithm ^[11].

Basic k-means algorithm:

1. Select K points as initial centroids.
2. Repeat
3. Form k clusters by assigning each point to its closest centroid.
4. Recompute the centroid of each cluster.
5. until centroids do not change.

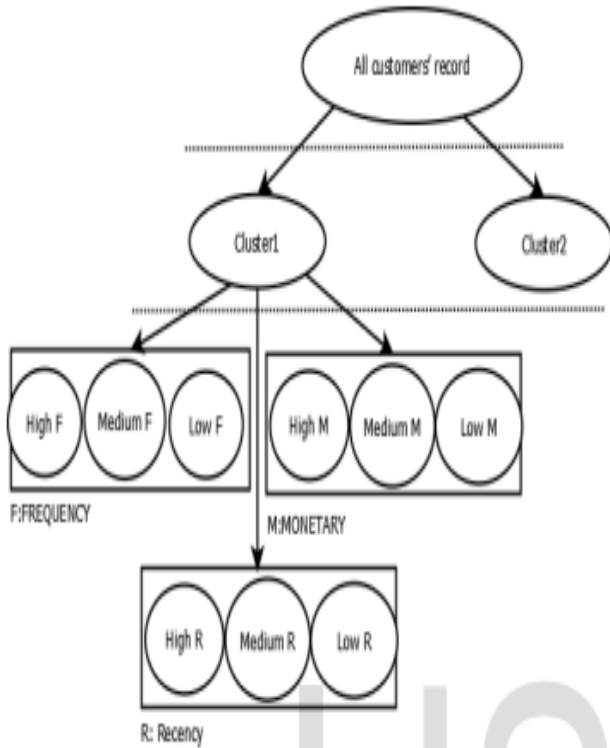
Since the k-means algorithm requires weightings point on the basis of that the transaction data can be clustered in the numbers of desired cluster.

Calculation of weighted score:

For the analysis purpose the transaction records of the customers are chosen between the specific periods of time. The time period can vary from research to research or it may depend on the condition and requirement of the marketer.

Here, we are taking time period of 12 months for analyzing customer’s transaction database. During these 12 months all transaction records from the database will be analyzed and weighted value is assigned to each customer that can be distinguished by their mobile number (Mb). This will act as the primary key in the database.

The total weighted score ^[12] is calculated on the basis of three individual factors that is frequency, monetary, Recency.

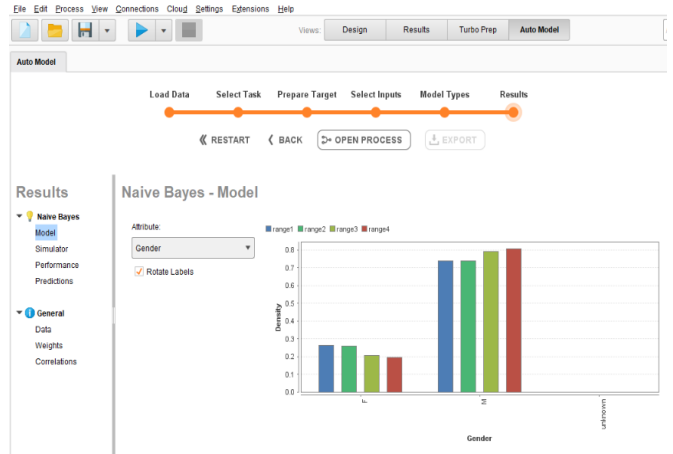


III. METHODOLOGY

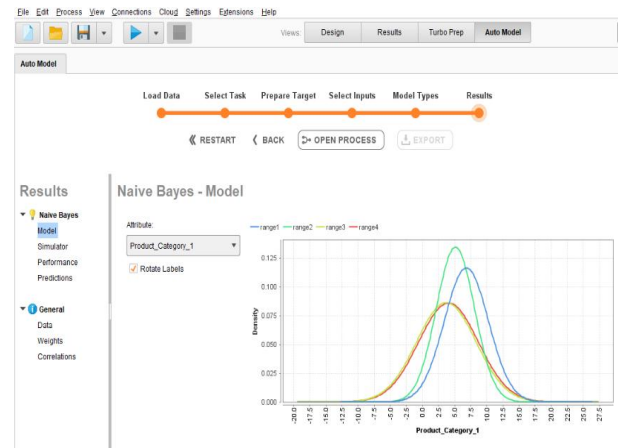
Identifying right customer and providing right service at right time and treating different types of customers differently is the key to success in business. So, a classifier model will be used to classify customers ^[18] into different categories based on their purchase history. Then the model can be used to predict category of future coming customer. Here, we will use Deep learning techniques with Keras for classification of the customer.

IV. RESULTS AND DISCUSSION

As we considered, the data set regarding ‘Black Friday Market’

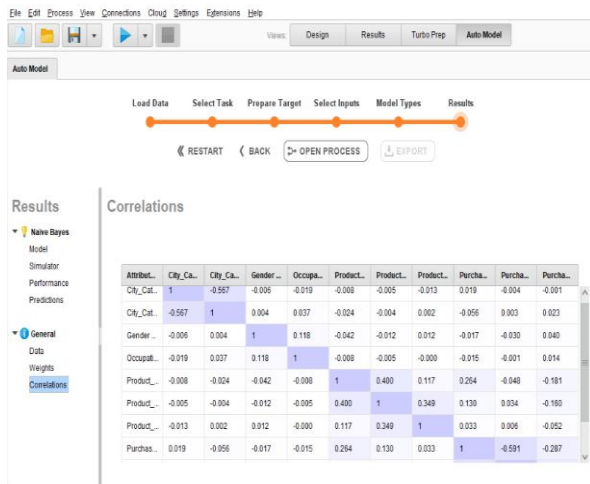
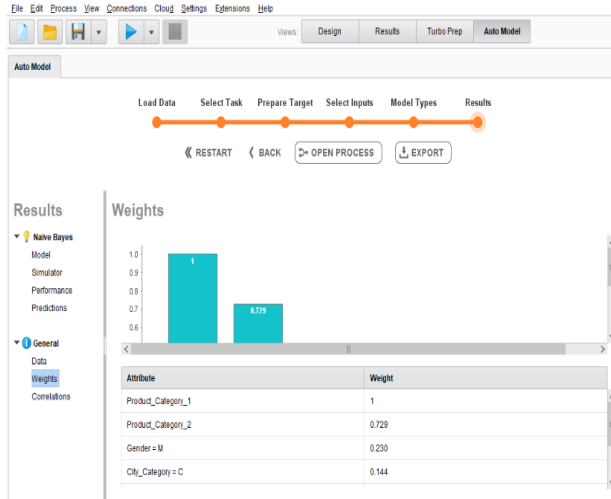


Gender Classification of customers



We have to classify the customers based on their historical purchase data and allocate them best servers for better priority ^[19].

Row No.	Purchase	prediction(P...	confidence...	confidence...	confidence...	confidence...	City_Category	Gender
5	range3	range3	0.010	0.229	0.384	0.378	B	M
6	range2	range2	0.406	0.594	0.000	0.000	A	M
7	range1	range2	0.403	0.597	0.000	0.000	A	M
8	range1	range2	0.332	0.668	0.000	0.000	A	M
9	range3	range3	0.000	0.007	0.827	0.166	A	M
10	range2	range2	0.470	0.530	0.000	0.000	A	F
11	range2	range3	0.009	0.373	0.453	0.165	B	M
12	range4	range3	0.014	0.432	0.481	0.072	C	M
13	range2	range2	0.234	0.766	0.000	0.000	C	M
14	range3	range3	0.009	0.166	0.687	0.139	C	M



V. CONCLUSION AND FUTURE SCOPE

In the proposed model we have provided a framework for classification of profitable customer for the retail marketer. The framework analyzes the customer's previous purchase history to identify his/her purchase behavior. We have used k-mean clustering technique and Naive Bayes's classifier for this purpose to identify the class of customer that is potential buyer of a particular product in the retail store. In this, marketer can perform one to one marketing instead of mass marketing. Since as per our assumption if the effort is applied in a particular direction, the sale will increase, and this will result in increased profit which is the ultimate goal of the marketer.

In the future, data mining will include more complex data types. In addition, for any model that has been designed, further refinement is possible by examining other variables and their relationships. Research in data mining will result in new methods to determine the most interesting characteristics

in the data. As models are developed and implemented, they can be used as a tool in enrollment management.

REFERENCES

- [1] Revenue Management Through Dynamic Cross Selling in E-Commerce Retailing Operations Research 54(5 – Serguei Netessine, See Profile, Available from Serguei Netessine, Serguei Netessine -1997
- [2] Insurance Databases for Effective Prediction of Customer Preferences over Life Insurance Products – S. Balaji, S. K. Srivatsa, PhD Senior Professor - A LITERATURE REVIEW – Olof Wahlberg, Christer Strandberg, Karl W. Sandberg
- [3] Unsupervised Learning in Large Datasets for Intelligent Decision Making – S. Balaji, Dr. S. K. Srivatsa.
- [4] important influences on classification accuracy – Mohammad Saad Al-Ahmadi, Peter A. Rosen, Rick L. Wilson
- [5] Decision Support System for A Customer Relationship Management Case Study – Ozge Kart, Alp Kut, Vladimir Radevski
- [6] Proceedings of the 40th Hawaii International Conference on System Sciences- 2007 Synergies of Data Mining and Operations Research –Stephan Meisel, Dirk Chr Mattfeld
- [7] Stakeholder Perceptions Regarding eCRM: A Franchise Case Study –unknown authors
- [8] Xin Zhao, Yi Wang, and Hongwang Cha. "A New Prediction Model of Customer Churn Based on PCA Analysis". In: Information Science and Engineering (ICISE), 2009 1st International Conference on. IEEE. 2009, pp. 4657–4661.
- [9] Qing-an Cui et al. "Using PCA and ANN to identify significant factors and modeling customer satisfaction for the complex service processes". In: Industrial Engineering and Engineering Management (IE&EM), 2011 IEEE 18th International Conference on. IEEE. 2011, pp. 1800–1804.
- [10] Volodymyr Kuleshov. "Fast algorithms for sparse principal component analysis based on Rayleigh quotient iteration". In: International Conference on Machine Learning. 2013, pp. 1418–1425.
- [11] Zhexue Huang. "Extensions to the k-means algorithm for clustering large data sets with categorical values". In: Data mining and knowledge discovery 2.3 (1998), pp. 283–304.
- [12] Charu C. Aggarwal, Alexander Hinneburg, and Daniel A. Keim. "On the Surprising Behavior of Distance Metrics in High Dimensional Space". In: Lecture Notes in Computer Science. Springer, 2001, pp. 420–434.
- [13] James Bennett, Stan Lanning, et al. "The netflix prize". In: Proceedings of KDD cup and workshop. Vol. 2007. New York, NY, USA. 2007, p. 35.
- [14] James Bennett. "The Cinematch system: Operation, scale coverage, accuracy impact". In: Summer School on the Present and Future of Recommender Systems (2006).
- [15] Yehuda Koren. "The bellkor solution to the netflix grand prize". In: Netflix prize documentation 81 (2009), pp. 1–10.
- [16] Andreas Töscher, Michael Jahrer, and Robert M Bell. "The bigchaos solution to the netflix grand prize". In: Netflix prize documentation (2009), pp. 1–52.
- [17] Martin Piotte and Martin Chabbert. "The pragmatic theory solution to the netflix grand prize". In: Netflix prize documentation (2009).
- [18] A Literature Review on Text Mining Techniques and Methods, IJCSE, Vol.06, Issue.02, pp.96-99, Mar-2018.
- [19] Metric for evaluating availability of an information system: A Quantitative approach based on component dependency, IJNSA, Vol-9, number-2, Mar-2017.

Authors Profile

Ms.I.GAYATHRI DEVI: is working as an Assistant Professor in department of Computer Science and Engineering, Pragati Engineering College. She acquired her Bachelor of Technology, Master's in computer science and engineering. She has 8 years of teaching experience. Her areas of interest include Network Security, Wireless Sensor Networks and Cloud Computing, Deep Learning.



Mrs.G.Kumari: is working as an Assistant Professor in department of Computer Science and Engineering, Pragati Engineering College. She acquired her Bachelor of Technology, Master's in computer science and engineering. She has 13 years of teaching experience. Her areas of interest include Network Security, Data mining, and Mobile Networks.



Mrs.G.Surya kala Eswari is working as Assistant Professor in department of Computer science and engineering, pragati engineering college. She acquired bachelor of technology and masters in CSE. She has 6 years of teaching experience. Her areas of interest include Networks, Cloud computing, Data Mining, Mobile networks.

