

Behavioral Prediction for E-commerce and Recommendations By ClickStream Data

Sapna Malik¹

¹Department of Computer Science and Engineering ,MSIT,Delhi

Abstract - In a regular retail shop, it is easier to analyse the behavior of customers with the help of a shop assistant and understand what the customer wants. But it becomes difficult to understand and analyze customer behavior when it comes to online shopping. In this case, clickstreams or the mouse movements of e-customers proves to be very efficient in providing information about their buying behaviour. In this study, we provide a very efficient model to analyse e-customers behaviour based on which predictions about their shopping behavior on a digital platform. After collecting data, we extracted online customers' behavior patterns about buying or not. The model presented in research paper predicts whether customers will or will not buy their items added to shopping baskets on an e-commerce site and also recommends items based on the analysis. For the analysis we are mining clickstream data and using Naive Bayes approach to give better suggestions.

Key Words: Behavioral analysis, Clickstream data, e-Customer , Naive Bayes classifier

1.INTRODUCTION

Today, the people's way of shopping has changed. Now, people buy more and more over the Internet instead of going traditional shopping. E-commerce provides customers with the opportunity to browse various products , compare prices, create wishlists and enjoy a better service based on their individual interests.

As the need to increase sales and business, customer satisfaction is very much important. It is very important to know the consumer behavior for any company before the introduction of any product. It is the consumer behavior, on the basis of which companies make the decisions regarding what to produce, and for whom to produce and in what quantity. Hence, the companies are making great effort to understand customers and are trying to identify the reasons that motivated them to purchase, or not, a product [1]–[3]. With this knowledge, the e-commerce websites can deliver more personalized services and d increasing benefits [4].

However, understanding customer behavior is a very complex task [1]. E-commerce websites provide customers with various navigational options and actions which users can freely move through. Usually, these user activities are recorded in the web server logs

[1], [5]. The very valuable patterns found in these logs are analysed [6].

Data mining techniques[23] have been very useful for discovering patterns in log files (when applied to the analysis of web server logs the term web usage mining [7] is used). It helps in discovering patterns trying to explain the users' interests. Various techniques are used in the field of e-commerce, such as classification techniques, clustering, association rules or, sequential patterns [8], [9].

In many application domains these techniques are used in conjunction with process mining techniques. Such techniques are part of the business intelligence domain and apply specific algorithms to discover hidden patterns and relationships in large data sets [10].

2. Literature Review

Many researches have been done on how to analyze customer behavior on online platforms.

In papers [11], [12], experiments provide information on how the online decision- making processes of the users affects the complexity of the online process behaviour. It emphasizes that collecting clickstream logs at the application layer will help to integrate Web usage data with other customer-related data sources. Here Clickstream data is used analyzing the customer behavior to actively influence the customer's decision making.

In paper [13], research uses principles of data mining to cluster customer segments. It uses k-means algorithms for this purpose and the research uses data from web-logs of various e-commerce websites.

Paper [14] surveys and classifies DMT in knowledge, analysis and architecture types with their applications in practical and research domains.

In this paper[15], a collaborative filtering-based Web service recommender system to help users select services with optimal Quality-of-Service (QoS) performance. It employs the location information and QoS values to cluster users and services, and makes personalized service recommendations for users based on the clustering results. It also focuses on achieving considerable improvement on the recommendation accuracy.

In [16] survey, reviews developments made recently in recommender systems based on clickstream data . It discusses major challenges faced. It compares and evaluates various algorithms used. It discusses techniques devised for e-commerce websites with their drawbacks and a relative comparison of their performance.

On paper [17] a Bayesian model is used that not only provides recommendations and can make predictions. Collaborative filtering recommends items to users based on their historical rating information. The model is based on both user-based and item-based collaborative filtering approaches, which recommends items by using similar users' and items' information, respectively.

3. Proposed Work

The purpose of this project is to provide the Data Analysts some hints about the user buying behavior. The model presented analyse the clickstream data data collected for individual users and extracts the relevant information required to make predictions about their online shopping behavior. It also predicts whether a customer will or will not buy the items added in their shopping baskets. It also makes recommendations based on the data collected.

A. Behavior Analysis

To understand what an user wants, we have used clickstream data. Clickstream is a record of a user's activity on the internet; Mouse clicks that a user makes tell us a lot about the behavior of the user if it is analyzed in an appropriate way [18]. These movements are sort of the behavior of online customers [19]. It tells us about what interests the user and what not. It is very useful to predict where the customers are more likely to click next, given their previously followed patterns.

This type of analysis is known as web mining or web farming approach used to discover patterns in the navigation of websites and web contents [20], [21]. By analyzing these patterns and their relation with web content one can redesign a website, portal or e-business along with the behavior of the online users [22].

As the user surfs and clicks on the products, the data is collected by the model. All the records related to a user are maintained and are used for predictions.

B. Recommendation System

To predict a potential customer,, we have used Naive Bayes algorithm. It is a classification technique based on Bayes theorem with an assumption of independence among predictors. It assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

It also helps in finding out what interests users by using both static and dynamic information of the customer. Based on the search, time logs, and various features, appropriate suggestions are made to customers which will make customers satisfied and also surfing on online shopping websites will indirectly make the customer to buy the product they want and this will increase sales as well as make profits.

Our model uses it on the clickstream data collected to find whether the user will buy the product or not[24]. Also based on the predictions, the model also recommends similar items.

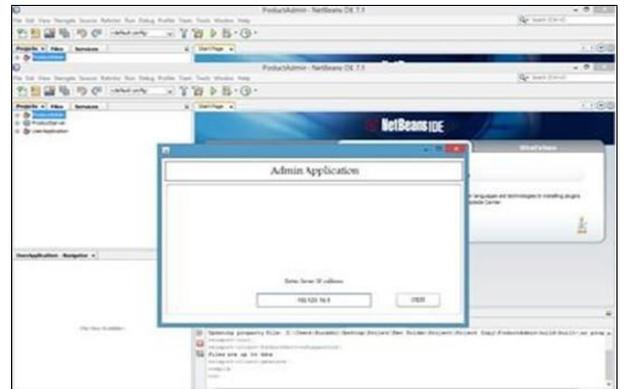
Naïve Bayes is used for this purpose as it gives accurate output with a minimal error rate as compared to all data mining algorithms.It is generally used for larger datasets.

4. Result & Discussion

A. Admin's Login : It is only used by the administrators of the websites. Each admin has a unique Id and password.

Figure 1: Admin's Login Page

B. Admin's Menu and Product management form : From here the admin can easily move for either



product management or user and product analysis. Product management form contains all the information about a particular product. Each product has a unique Id. The admin can easily add or remove any number of products.

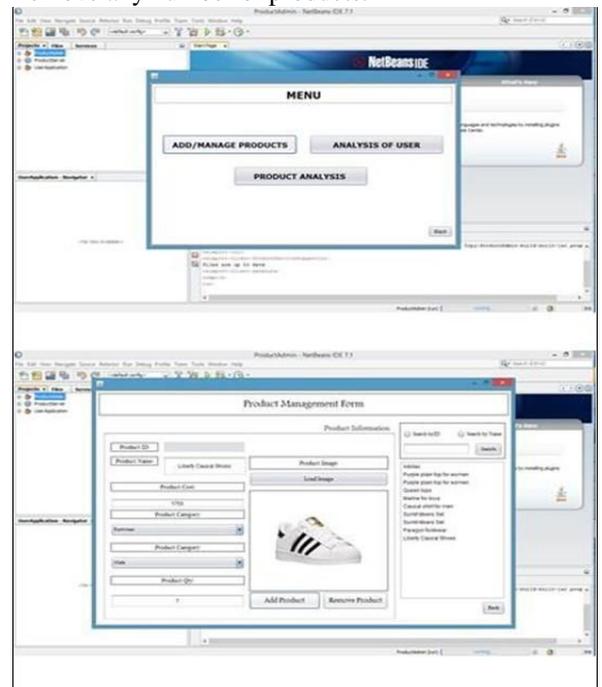


Figure 2: Admin's Menu and Product management form

C. Online shopping website and user analysis : the user can shop online from different varieties and different products . The heading of this page is online shopping website or the particular name of the website

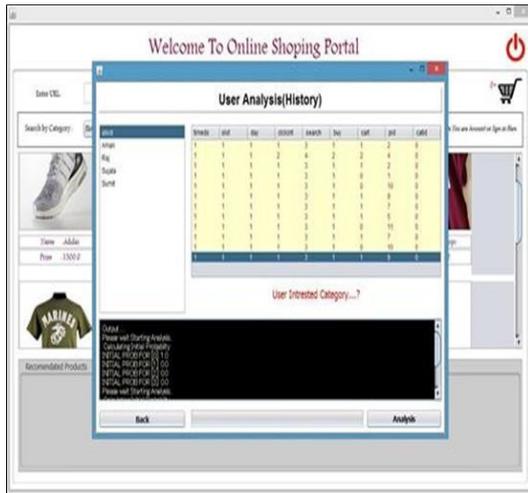


Figure 3: Online shopping website and user analysis

In the section of user analysis , the user can analyse the product and also user can compare the products and website can tell that which product will be best for the user

D. Product analysis : In this section the analysis of products can be done easily . The total sales from particular year can be seen easily



Figure 4:Product Analysis

Also the average sale of all the products can be seen in this section. Also it provides the administrators with knowledge about which product is being widely purchased.

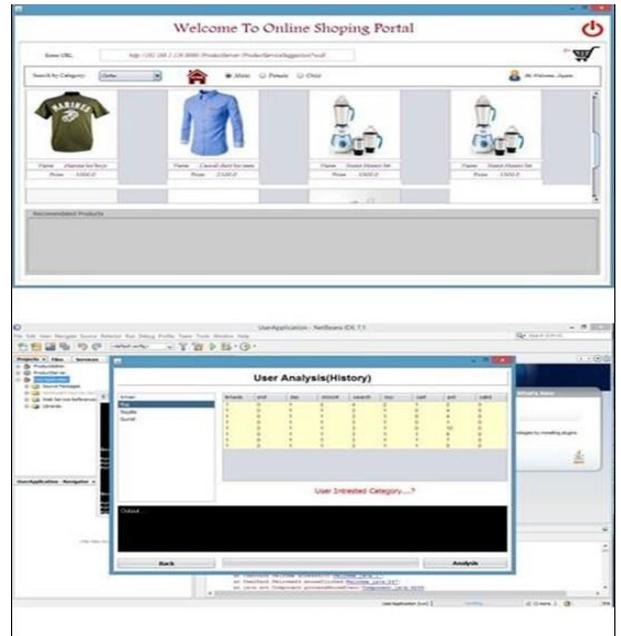


Figure 5:User Analysis



3. CONCLUSIONS

With the advent of technology and fast migration of business to digital platforms, becoming familiar with the customer to meet their demands has become the need of the hour. A wide range of literature survey is done to study and learn the technologies used till date to predict the behaviour of the customer over the digital market. This paper addresses the above mentioned aspects and makes the use of Naive Bayes classifier to predict the customer behaviour and recommend different products as per his/her visit to the e-commerce website. There is a remarkable shift from the traditional purchase of products as people now prefer to purchase the products online rather than visiting the shops and purchasing. The proposed model can be extended to another dimension of machine learning techniques, where purchase patterns of different customers will be analysed and accordingly the predictions will be made. Based on these predictions, recommendations to different customers purchasing the similar product will be made. This will eventually pave way to automation and also the customers will feel comfortable while purchasing the products from e-commerce platforms.

REFERENCES

1. R. Kohavi, "Mining e-commerce data: The good, the bad, and the ugly," in Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2001, pp. 8–13.
2. W. W. Moe and P. S. Fader, "Dynamic conversion behavior at e-commerce sites," *Manage. Sci.*, vol. 50, no. 3, pp. 326–335, 2004.
3. G. Liu et al., "Repeat buyer prediction for e-commerce," in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, New York, NY, USA, 2016, pp. 155–164.
4. Y. S. Kim and B.-J. Yum, "Recommender system based on click stream data using association rule mining," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13320–13327, 2011.
5. R. Kosala and H. Blockeel, "Web mining research: A survey," *SIGKDD Explor. Newsl.*, vol. 2, no. 1, pp. 1–15, Jun. 2000.
6. F. M. Facca and P. L. Lanzi, "Mining interesting knowledge from Weblogs: A survey," *Data Knowl. Eng.*, vol. 53, no. 3, pp. 225–241, 2005.
7. J. Srivastava, R. Cooley, M. Deshpande, and P.-N. Tan, "Web usage mining: Discovery and applications of usage patterns from Web data," *SIGKDD Explor. Newsl.*, vol. 1, no. 2, pp. 12–23, Jan. 2000.
8. Q. Zhang and R. S. Segall, "Web mining: A survey of current research, techniques, and software," *Int. J. Inf. Technol. Decision Making*, vol. 7, no. 4, pp. 683–720, 2008.
9. B. Singh and H. K. Singh, "Web data mining research: A survey," in Proc. IEEE Int. Conf. Comput. Int. Comput. Res. (ICIC), Sep. 2010, pp. 1–10.
10. W. M. P. van der Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, 1st ed. Heidelberg, Germany: Springer-Verlag, 2011.
11. . Senecal, S., Kalczynski, P. J., Nantel, J., 2005. Consumers Decision making process and their online shopping behaviour: A clickstream analysis. *Journal of Business Research*, 58, 1599–1608.
12. Hu, Jia, Zhong, Ning, 2008. Web farming with clickstream *International Journal Of Information Technology Decision Making*, 7(2), 291308.
13. Y. Jiang, S. YU, 2008. Mining the e-commerce data to analyse the target customer behaviour, *IEEE*, DOI 10.1109.
14. Y. Liu, J. Mao, M. Zhang, S. Ma, L. Ru, Y. Liu, 2011. *Expert Systems with Applications*, Elsevier, 0957-4174/:
15. Xi. Chen. Z. Zheng, 2014. Web Service Recommendation via Exploitation location and QoS Information, *IEEE*, DOI 1939-1374/413/31:0:
16. Shalini Gupta, Mayank Rawat, 2016, Recommendations through click stream: Tracking the need, current work and future directions, *IEEE*
17. Priscila Valdiviezo, Fernando Ortega, Eudrado Cobos, Raul Lara-Cabrera, 2019, A Collaborative Filtering Approach Based on Naive Bayes Classifier, *IEEE*, 2169-3536.
18. Premchaiswadi, Wichian; Romsaiyud, Walisa, 2012
19. .Extracting WebLog of Siam University for Learning User Behavior on MapReduce, *Proceedings of 4th International Conference on Intelligent and Advanced Systems (ICIAS) and A Conference of World Engineering, Science and Technology Congress.*
20. Senecal, S., Kalczynski, P. J., & Nantel, J., 2005. Consumers Decision making process and their online shopping behaviour: A clickstream analysis. *Journal of Business Research*, 58, 1599 – 1608.
21. Hu, Jia, Zhong, Ning, 2008. Web farming with clickstream *International Journal Of Information Technology & Decision Making*, 7(2), 291-308.
22. Ting, I-Hsien; Kimble, Chris; Kudenko, 2009. Daniel, Finding Unexpected Navigation Behaviour In Clickstream Data For Website Design Improvement, *Journal Of Web Engineering*, 8(1), 71-92.
23. Detlor B., 2000. The corporate portal as information structure: Towards a framework for portal design. *International Journal of Information Technology*
24. Gkhan S. , Hale D. , 2015. Analysis and Prediction of E-Customers Behaviour by Mining Clickstream Data , *IEEE International Conference on Big Data (Big Data)* , 978-1-4799-9926-2.
25. Thompson S., H. Teo, 2006. To buy or not to buy online: adopters and non-adopters of online shopping in Singapore. *Behaviour Information Technology*, 25 (6), 497–509.
26. Aikaterini C. et.al., 2013. Online and mobile customer behaviour: a critical evaluation of Grounded Theory studies, *Behaviour Information Technology*, 32(7), 655667.
27. Han J., Kamber, M., 2011. *Data Mining Concepts and Techniques*, 3rd ed., Morgan Kaufman Publishers Academic Press.
28. Wu HE, 2013. Improving user experience with case-based reasoning systems using text mining and Web 2.0, *Expert Systems with Applications*, 40(2), 500507.