INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume 4, Issue 2)

Available online at: www.ijariit.com

Analysis and prediction of E-customers behavior by mining clickstream data using Naive Bayes

Namrata Pawar
pawarnamrata1604@gmail.com
Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra

Monali Gaikwad
monaligaikwad25.MG@gmail.com
Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra

Sarika Kalyani
sarikakalyani200@gmail.com
Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra

Margi Savla
margie.js96@gmail.com
Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra

ABSTRACT

Nowadays, online shopping has become a trend. In online shopping, it is very difficult to analyze and do prediction of the customer whether he will buy the product or not. So to predict that Naïve Bayes is used. Data mining extracts the information from a large amount of data which stores in multiple heterogeneous data bases. This model extracts information and makes predictions about customers shopping behavior and helps to analyze click streams of e-customers on a digital market place. After collection of the dataset from the database, data mining is applied to the dataset collected and online customer behavior is predicted. Naïve Bayes is applied to the dataset which will predict the customer behavior and also predict about the customer's interest about the item.

Keywords: *Data mining, Clickstream, e- customer, Customer behavior, Digital market.*

1. INTRODUCTION

Nowadays, the online shopping is very much growing business. As the need of increase sale and business point of view the satisfaction of the customer is very much important, so to overcome the customer trouble while getting a suggestion or while getting purchasing product to give him accurate output that he wants that is main motivation behind this application. As we are seeing that trend of Online Shopping is Growing Day by day, Customer is not getting proper suggestions and sufficient benefit on buying product So to overcome these difficulties of customers we are using Naive Bayes approach. The websites nowadays are using different Algorithm for suggestions as they are not sufficient enough to provide better services to the customer. We are developing our application by mining click stream data in order to give better suggestions to customers.

2. OBJECTIVES

- To collect the history data
- Training the dataset To Add and Manage products
- Extraction of behavior feature
- Applying mining Algorithms
- Find out users interest
- Give some specific offers based on user's interest

3. DATASET USED IN THE STUDY

a. Buy Count: This involves a count of the customer purchasing the product

b. Click Count: It is the count of the user clicks for the particular product

c. Search Count: It is the count of customer search for the particular product

d. Time of day: The time of day such as morning, evening, etc. of customer

e. Add to cart: It takes into account the product added to cart by customer

f. Special Day: If is one week or earlier than official or

A religious day such as Christmas day, Independence Day, etc

g. Day: represents the day of the week: Sunday, Monday, and Tuesday etc.

h. Period of Day: We have four periods for this variable; morning, afternoon, evening and after midnight

i. Time Spent on site: This variable includes total time spent on the site.

j. Category of Search: Products have been categorized into 4 i.e. Clothing, Ornaments, Cosmetic, Stationary.

k. No. Of Items in basket: It shows the number of different items in the online shopping basket.

l. Discounted items in basket: The e-commerce company makes promotion campaigns or applies discounts.

m. Product Category of the item in the basket: There are five categories in this field: Female, male, unisex, child (girl), child (boy). Item

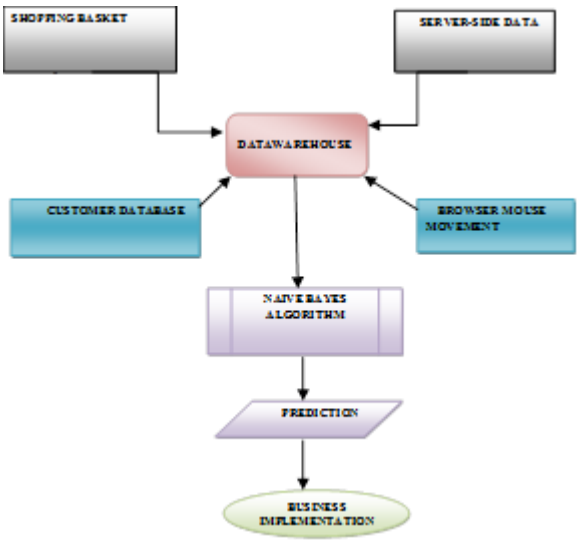
n. Item adds time: It shows the time (in seconds) of the first item added to basket. If the basket is empty it takes 0 values.

4. PROPOSED ARCHITECTURE

Fig .1. System Architecture

Naive Bayes classifiers rely on Bayes' decision rule with the assumption that the probability of each attribute value is independent of the values of other attributes. This assumption is not valid in general, but provides the advantage of computational simplicity and often leads to very good results. Inspired by this approach we used the posterior distributions of the class given each feature and combined these using the sum rule.

5. FLOWCHART



6. PSEUDOCODE

```
TRAINMULTINOMIALNB(C, D)
 1  V ← EXTRACTVOCABULARY(D)
 2  N ← COUNTDOCS(D)
 3  for each c ∈ C
 4  do N_c ← COUNTDOCSINCLASS(D, c)
 5     prior[c] ← N_c / N
 6     text_c ← CONCATENATETEXTOFALLDOCSINCLASS(D, c)
 7     for each t ∈ V
 8     do T_ct ← COUNTTOKENSOFTERM(text_c, t)
 9     for each t ∈ V
10     do condprob[t][c] ← (T_ct + 1) / Σ_t'(T_ct' + 1)
11  return V, prior, condprob

APPLYMULTINOMIALNB(C, V, prior, condprob, d)
1  W ← EXTRACTTOKENSFROMDOC(V, d)
2  for each c ∈ C
3  do score[c] ← log prior[c]
4     for each t ∈ W
5     do score[c] += log condprob[t][c]
6  return arg max_c∈C score[c]
```

7. CONCLUSION

As we are seeing that trend of Online Shopping is Growing Day by day, customer is not getting proper suggestions and sufficient benefit on buying product, So to overcome these difficulties of customers we are using Naive Bayes approach.

Naïve Bayes gives accurate output with a minimal error rate as compared to all data mining algorithms.The Naïve Bayes will help to improve and Increase sales. It will also find out users interest in the product by using both static and dynamic information of the customer. Based on his search, time logs, and various features, appropriate suggestions and discounts will be provided to customers which will make customer satisfy by will surfing on online shopping website which indirectly will make the customer to buy the product they want to this will increase sales as well as will make a profit.

8. ACKNOWLEDGMENT

We take this opportunity to thank our project guide Prof. A.V.Kanade and Head of the Department Prof.Dr.D.A.Godse and our Principal Prof.Dr.H.V.Vankudre for their valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this project report. We are also thankful to all the staff members of the Department of Information Technology of Bharati Vidyapeeth College of Engineering For Women, Pune for their valuable time, support, comments, suggestions, and persuasion. We would also like to thank the Institute for providing the required facilities, Internet access, and important books.

9. REFERENCES

[1] Misha Kakkar, Divya Upadhyay"., Web Browsing Behaviour Based Age Detection", International Journal Soft Computing and Engineering(IJSCE) ISSN:2331-2307, volume-3, Issue-1, March 2013.
[2] Edin Osmanbegovic, Mirza Suljic.," Data Mining Approach For Predicting Student Performance", Economic Review- Journal of Economic And Business, Vol. X, Issue 1, May 2012.
[3] Qasem A. Al-Radaieh, Eman Al Nagi., "Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance", International Journal Of Advanced Computer Science And Application Soft Computing and Engineering (IJACSA), Vol. 3, No. 2, 2012.
[4] 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. [2] Aikaterini C. et.al. 2013.
[5] Online and mobile customer behavior: a critical evaluation of Grounded Theory studies, Behaviour & Information Technology, 32(7), 655–667. [3]
[6] Han J., & Kamber, M., 2011. Data Mining Concepts and Techniques, 3rd ed., Morgan Kaufman Publishers Academic Press.
[7] Wu HE, 2013. Improving user experience with case-based reasoning systems using text mining and Web 2.0, Expert Systems with Applications, 40(2), 500–507. [7] Al-Zaidy, Rabeah; Fung, Benjamin C. M.; Yusuf, Amr M.; et al., 2012.
[8] Senekal, S., Kalczynski, P. J., & Nantel, J., 2005. Consumers Decision making process and their online shopping behavior: A click stream analysis. Journal of Business Research, 58, 1599 – 1608

