iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4179

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Analysis of Users Behaviour in Structured E-Commerce Websites

Nikhil Prasad[1], Priyansu Prasad[2], Manisha Shendre[3], Aarti Wakale[4], Prof. Dipashri Sonawale[5]
[1, 2, 3, 4, 5] *Dept. of Information Technology, M.G.M. College Of Engineering and Technology, Kamothe-410209, Navi Mumbai*

***Abstract*: Online shopping is becoming more common in our daily lives. Therefore, understanding user requirements is must. So that we should knowing user interests and behavior is essential for creating a more adaptive e-commerce website. Incorporating a view of user behaviour while a session can be of great interest to identify more complex behavioural patterns.**
With the help of user behavior, we can find most of user are purchasing or interest in which product So that we can provide more Product related to the category on e-commerce web site

I. INTRODUCTION

The way people shop has changed tremendously in last two decades. Instead of going traditional shopping people are buying more and more over the Internet. E-commerce provides users the opportunity of browsing many product category, being continuously informed, creating wish list, comparing prices, and enjoying a better service based on their individual interests. This ecommerce market is highly competitive, allowing users to easily migrate from one e-commerce when their necessities are not satisfied. As a consequence, e-commerce business analysts require to know and understand consumers' behavior when those navigate through the website, as well as trying to identify the reasons that motivated them to purchase, or to migrate to other ecommerce system.
In the field of e-commerce, most data mining techniques saves user navigation events. Nevertheless, these sequences are not directly mined; instead, each sequence is transformed into a session characterization. A characterization consists of in detailed details of what a particular user has done while navigating. The contents of these structures can be diverse. In the characterization contains the web browser used by the customer, visited pages, products added in cart, products added in Wishlist, time the customer spent on each page, or the keywords used in search engine etc. This solution tries to identify the user's interests from the contents of the visited pages. This will improve the efficiency of the system and will helps in overall system growth.

II. PROBLEM DEFINITION

As e-commerce growth rate is growing exponentially the data which is generated is tremendous. Project goal is to collect easy and every activity of user and to store in it proper format. Project not only covers the data which is generated by the user but also perform data mining operation on the collected data. The data mining operations includes classification clustering etc. The outcome will suggest some meaningful pattern or logic which will help in growth of organization.

III. PROPOSED SYSTEM

1) Currently, there are powerful commercial tools for analyzing logs of e-commerce websites, being Google Analytic one of the main ones. Google Analytic controls the network traffic, collects information about user sessions (first and last web page visited, pages visited, time spent on each page, etc.), and displays reports synthesizing users' behavior. These traffic-based data can also be combined with other users' personal and geographic information. Google Analytic is not able to import the web server logs of a website, but it works analyzing the information collected by means of page tagging techniques. These techniques have some disadvantages with respect to the log-based analysis, such as dependence on JavaScript and cookies, the necessity of adding page tags to every page, the complexity of tag-based implementations, the fact that, as a result, customers may experience a change in the download time of the website, or privacy concerns, for instance. Nevertheless, Google reports are rich in data that, in turn, require experts in the problem domain to exploit them. In any case, the conclusions of the analysis can be used to improve the website design, to design advertising and marketing campaigns, to analyze customers demographic information or to control real-time traffic. Similar commercial tools are Clicky, Piwik , Adobe Analytics, W3Counter Web Analytical tool .

2) The methodology and tool proposed in this work try to overcome some of the drawbacks of the previous approaches, providing with the possibility of getting a very accurate interpretation of users' behavior:

©IJRASET (UGC Approved Journal): All Rights are Reserved

3) In comparison to the clustering approaches and the commented commercial tools, the advantage of our mining technique is that this provide causal relations among events of a user trace, instead of providing with a global view of the whole session. Besides, it is the fact of avoiding the need of tagging the web pages.
4) With respect to those approaches whose main objective is predicting the coming possible events (as the case of Markov models, for instance), the approach allows having a global view of the sessions, making easier a global analysis of the user behavior, giving hints and facilitating the re-design of the website for a better adaptation to the user necessities.
5) An interesting feature of the approach followed in this paper is that it properly fits the open nature of the use of e-commerce websites, where there are very few constrains for the users to navigate among site web pages.
6) Another interesting feature of the followed mining approach is the fact of being able to analyse sequences of detailed events. The fact of considering the causal relations of events inside a user session, allowing to look for intra-session patterns (and not only patterns repeated in different sessions) can provide the analysts with a much more detailed perspective of a user behavior.

IV. LITERATURE SURVEY

P Devika, R C Jisha, G P Sajeev [1] Given approach of recommendation system in e-commerce environment.
Rana Alaa El-DeenAhmeda, Shereen Morsya, NermeenMekawiea [2] Studied different Classification Algorithms for Consumer Online Shopping Attitudes and Behavior.
Limei Zhang [3] Gives way of customer relationship managementwith help of Data Mining
HarshitaKotian, B. B. Meshram[4] Provide quality management framefork for ecommerce website.For evaluating the quality of the E-commerce application, an E-commerce Total Quality Management framework (E-TQMF) is proposed which takes into consideration the quality aspect both from the customer's and quality expert's perspective

Gu Hongjiu [5] describes the process, methods, and specific applications of data mining in e-commerce site.

Tanner Larsson[6]Clear how traditional e commerce has changed and Grow & Scale A Successful Ecommerce Business as per new technology
XiaofengXue [7] conceptualize and strategies customer knowledge management

Judah Phillips[8] Implement and unify ecommerce analytics related to product, transactions, customers, merchandising, and marketing
More effectively measure performance associated with customer acquisition, conversion, outcomes, and business impact
Use analytics to identify the tactics that will create the most value, and execute them more effectively.
Adrian Mejia[9] There is always an ever-growing list of requirements while designing an e-commerce application, which needs to be flexible enough for easy adaptation. The MEAN stack allows you to meet those requirements on time and build responsive applications using JavaScript.
Robert Kauffman[10] E-commerce is a multi-disciplinary area, which should be developed in co-operation with existing fields such as Information Systems and Technology; Marketing, Finance and Supply Chain Management; Business Strategy and Management; Public Policy; Computer Science and Telecommunications; and Legal Studies. We will solicit papers on current technologies from these areas, as well as publish papers on completely new topics.

V. SYSTEM ARCHITECTURE

A. *Architecture of E-commerce Management System*

In Below Fig. We Show how data will take and get processes for making a proper decision for e commerce business.
E-commerce management system take all information of user from website domain and website database then that data will forward for data cleaning and selection to get useful data Then it takes as a input for Data mining Engine to perform data mining methods on collected data. Data mining engine evaluate pattern for given data and store in Knowledge Base
On bases of knowledge based product suggestion mechanism apply on E-commerce management system.
Knowledge base is also show (accessed) to the admin to take particular action for increasing sells of the product

©IJRASET (UGC Approved Journal): All Rights are Reserved

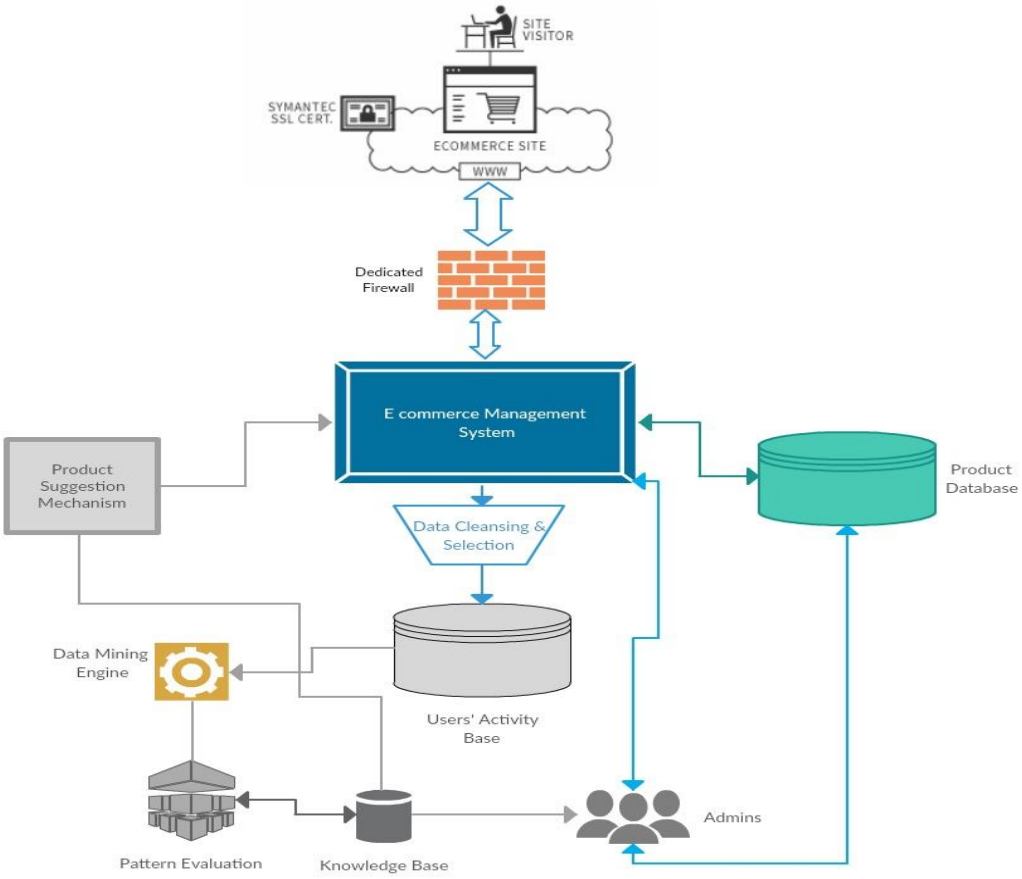


Fig. Architecture of E commerce management System

VI. ALGORITHMS

A. *Regression algorithms*

Regression algorithm is used for prediction. In an e-commerce website, regression algorithm is very much useful. It gives answer to similar types of questions:

1) Who will revisit the sites in next couple of days?

2) Who are likely to buy?

Regression algorithms implemented are as follows :

a) *Simple Linear Regression:* Simple linear regression is a statistical method that enables users to summarise and study relationships between two continuous (quantitative) variables. Linear regression is a linear model wherein a model that assumes a linear relationship between the input variables (x) and the single output variable (y). Here the y can be calculated from a linear combination of the input variables (x).

b) Multiple Linear Regression: Multiple linear regression is very similar to linear regression. The only difference is, there can be multiple input parZameter for a single output parameter.

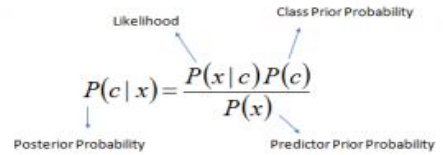
B. *Classification Algorithm*

In an e-commerce application, classification algorithm are implemented in Recommendation system. Following are the classification algorithms implemented:

1) *K-Nearest Neighbors K-NN:* In *k-NN classification*, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its *k* nearest neighbors (*k* is a positive integer, typically small). If *k* = 1, then the object is simply assigned to the class of that single nearest neighbor.

2) *Naive Bayes: Naive* Bayes Algorithm and is used in a Recommendation System that uses data mining techniques to filter unseen information and predict whether a user would like a given resource or not. It is based on Bayes Theorem.

©IJRASET (UGC Approved Journal): All Rights are Reserved



$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \cdots \times P(x_n|c) \times P(c)$$

Where,

a) $P(c|x)$ is the posterior probability of *class* (c, *target*) given *predictor* (x, *attributes*).
b) $P(c)$ is the prior probability of *class.*
c) $P(x|c)$ is the likelihood which is the probability of *predictor* given *class*.
d) $P(x)$ is the prior probability of *predictor*.

3) Decision Tree: In decision trees, for predicting a class label for a record we start from the **root** of the tree. We compare the values of the root attribute with record's attribute. On the basis of comparison, we follow the branch corresponding to that value and jump to the next node.

We continue comparing our record's attribute values with other **internal nodes** of the tree until we reach **a leaf node** with predicted class value.

C. Clustering Algorithm

1) K-Means Clustering: K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable *K*. The algorithm works iteratively to assign each data point to one of *K* groups based on the features that are provided. Data points are clustered based on feature similarity.

D. Association Rule Learning

1) *Apriority:* Apriori is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database: this has applications in domains such as market basket analysis.

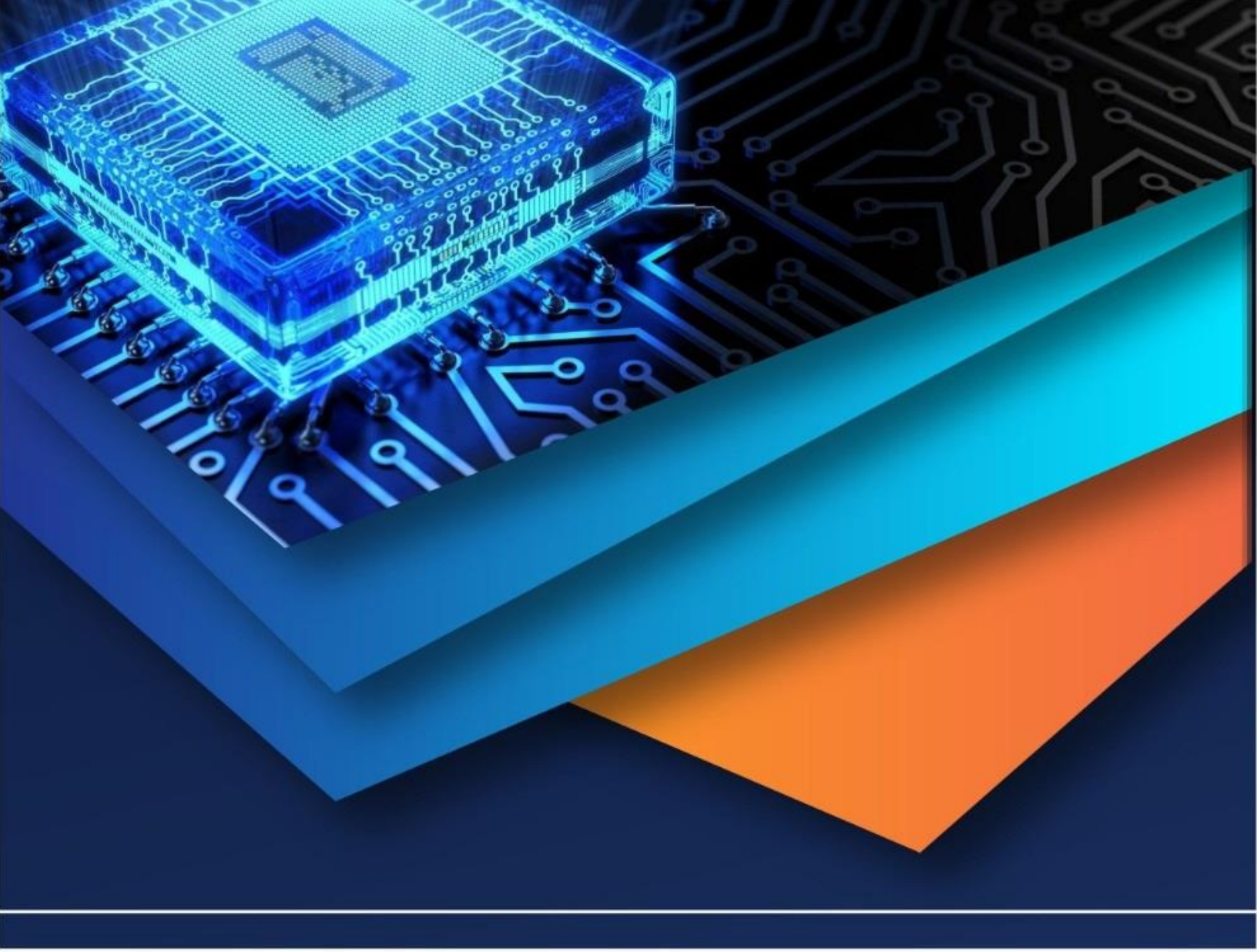
VII. CONCLUSIONS

It is mainly designed to record user's activity and detect and monitor the user's behaviour by this website we get to know the Product Ranking e.g. Which Product Got Highest Sale This Week(Trending), how many users are adding product in cart etc with the help of data mining technology with the help of different parameter like (age, gender, business). By doing this we can help an organization to grow with exponential rate.

REFERENCES

[1] P Devika, R C Jisha, G P Sajeev.Approach for book recommendation systems.
[2] Rana Alaa El-DeenAhmeda, , Shereen Morsya, NermeenMekawiea.Study of Classification Algorithms for Consumer Online Shopping Attitudes and Behavior Using Data Mining.
[3] Limei Zhang .Data mining application in customer relationship management.
[4] HarshitaKotian, B. B. MeshramA framework for quality management of e-commerce website.
[5] Gu HongjiuData Mining in the Application of E-Commerce Website.
[6] Tanner Larsson. Ecommerce Evolved: The Essential Playbook To Build, Grow & Scale A Successful Ecommerce Business
[7] XiaofengXue. Business management based on data mining under E-commerce.
[8] Judah Phillips, "Ecommerce Analytics: Analyze and Improve the Impact of Your Digital Strategy (FT Press Analytics)".
[9] Adrian Mejia, "Building an E-Commerce Application with MEAN ".
[10] Robert Kauffman, "Electronic Commerce Research and Applications" journal
[11] Swarupa Rani, A. and Jyothi, S. "AN AGRICULTURE SURVEY OF BIG DATA MINING APPLICATIONSSri PadmavathiMahilaVisvavidyalayam, Tirupati, AP, India, June2017.

©IJRASET (UGC Approved Journal): All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)