



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 10, Issue 6, June 2022**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.165**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

# Customer Behavior Analysis Using Web Mining

**Dhanashri Lambade, Vikrant Kalbhor, Ketan Pawar, Rameshwari Jadhav, Shilpa Kuchekar**

Department of Computer Science and Engineering, PCET Nutan College of Engineering and Research, Pune, India

**ABSTRACT:** Internet shopping is yet in its early stage in India yet developing at a high speed. To proceed with its development, it is important to get the client's inclinations. Investigation of customer's conduct regarding web-based shopping comprises itemized data regarding customers past buys as well as forecast of future buys. This developing requirement for refined data can't be met with straightforward information base programming. Information Mining is utilized for tracking down the concealed data from the pool of information. It has additionally been called as information

examination, information disclosure and insightful learning. The capacity to perceive and follow designs in information help organizations filter through layers of apparently irrelevant information for significant connections. Through this examination it becomes simple for the web-based retailers to decide the aspects that impact the take-up of web-based shopping and plan powerful advertising techniques. This paper assembles a guide for dissecting purchaser's internet purchasing conduct with the assistance of information mining. The point of this paper is to comprehend the job of information mining in development of on the web shopping. The main considerations that influence the purchaser's internet purchasing conduct are comfort, usability and seen benefits. Security is additionally a significant thought while selecting to direct shopping exercises on the web. This study will help in further examining the buyer web-based purchasing conduct towards Online shopping which will assist the retailers with planning proper showcasing systems for selling their items online which will further assistance being developed of the country.

**KEYWORDS:** Behavior Analysis, Market Basket Analysis, Customer Behavior Modeling, Classification, Data Mining, Regression, Data Selection .

## I. INTRODUCTION

Online business fundamentally represents electronic trade which connects with a site that sells items or benefits straightforwardly from the site with the assistance of a shopping basket or shopping bin framework and installments can be done through cards, e-banking and money down. It assists clients with purchasing anything structure a pen to an insurance contract from the solace of their home or office and gift it to somebody sitting miles separated just by click of the mouse. It offers different advantages to organizations for instance, simple reach to quickly developing internet based local area, giving limitless rack spot to items and administrations, consolidating worldwide business sectors at low working expenses. Ease to get to web is the main consideration in fast reception to E- business. For advancement of E-business in India the vitally fundamental elements are free from any danger installment modes. Despite the fact that there are different advantages in shopping on the web yet very much like each coin has different sides, there exist different purposes behind not shopping on the web for model absence of trust, security concerns, vulnerability about the item and administration quality, delay or non- conveyance of merchandise, and absence of touch-and-feel shopping experience. "Portable Commerce (M-business) is the subset of electronic-trade, which incorporates all web-based business exchanges completed utilizing a cell phone. Essentially, M- trade is the approach to carrying on with work in a condition of movement. "M- business relies upon the accessibility of portable network. M-business offers different benefits like omnipresence, personalization, adaptability, and conveyance, moment availability, promptness. There are numerous manners by which organizations, government and individuals benefit from m-trade like Selling an item or administration which is data based (conveyance straightforwardly to cell phones) or area based. Further developing efficiency by social affair time basic data (reports, photos) and SMS based up-to date data. The capacity to get to data on portable, at reasonable expense can transform people and vocations in country regions (Latest on the climate forecast or wellbeing administrations). It tends to be utilized as the medium to teach and make mindfulness among the country individuals. Utilizations of Internet on cell phones have led to data access conquering topographical boundaries and eliminated the preparation cost of innovation. Client conduct investigation depends on purchaser purchasing conduct, with the client assuming the parts of client, payer and purchaser. The worry of numerous associations is at this point not on the singular purchaser yet rather on aggregate or on the other hand authoritative



purchasing conduct which help in figuring out which clients merit creating and overseeing by setting up interesting procedures to draw in explicit clients. Through examination of customers’ conduct, precise profiles are being created by indicating needs and interest and permitting business to give clients what they need it, when they need, prompting a superior consumer loyalty along these lines keeping them to return for additional. Customer conduct incorporates the investigation of people, gatherings or associations about their course of choosing, getting, utilizing and arranging the items, administrations, encounters or thoughts to fulfill needs and the effect of these cycles on the customer and society. While enormous scope data innovation has been developing separate exchange and scientific frameworks, information mining gives the connection between the two. Information mining programming dissects connections and designs in put away exchange information in light of open-finished client questions. Information mining is the self-loader revelation of examples, affiliations, changes, abnormalities, and genuinely huge designs and occasions in information. Conventional information examination is suspicion driven as in a speculation is framed and approved against the information. Information mining, interestingly, is information driven as in designs are consequently removed from information. Different examinations on buyer buying conduct have been introduced and utilized in genuine issue. Information mining strategies are relied upon to be more powerful apparatuses for breaking down buyer conduct. Information mining has rapidly arisen as a profoundly advantageous device for utilizing current detailing abilities to reveal and comprehend stowed away designs in immense information base and these examples are then utilized in models that foresee individual conduct with high precision.

**II. PROBLEM STATEMENT**

Customer behavior refers to the purchasing behavior of final customer or individual or household who buys goods and services for personal use. Customer behavior is very important as it supports product positioning, development of effective marketing strategy and enhancement of long-term customer relationship. To make the system more flexible for the customers who are using online system for buying things. The main aim of our Project is to support and promote online businesses by providing a proper system using some methodology and techniques like Web mining, Information mining, Data mining and Data analysis. These techniques provide accurate results about the customer behavior which in return is helpful for the retail as well as for wholesale businesses.

**III. BACKGROUND STUDY AND TECHNOLOGY GAPS IDENTIFIED**

Sr. No	Paper Title	Year	Description
1	A Data Mining Based Approach to Customer Behavior in an Electronic Settings.	May-2019	The understanding of customer incidents and behavior is crucial to the success of any organization. Evidence from literature shows a prediction pattern of products to customer. These studies predicted product characteristics leaving out the customers characteristics. To address this gap, this study aims to design data-mining system and implement it on an electronic commerce organization website. The customer information and history (clickstreams) from the electronic commerce website was used to predict the customers’ behavior.
2	Do Retailers Benefit from Deploying Customer Analytics?	June-2020	This research focuses on the retailing industry, an industry characterized by tight margins that lead to careful scrutiny of all business investments. Using survey data from 418 top managers based in the Americas, Europe Middle East and Africa (EMEA) and Asia, we show that of the eight industries in the study, firms in the retail industry have the most to gain from deploying customer analytics. However, we also find that not only do many retailers not perceive this potential gain, they do not invest in customer analytics at an economically appropriate level. Thus we identify a gap between perception and reality concerning the potential for customer analytics in the retail industry that has both theoretical and practical implications.
3	Consumer Behavior Research: A Synthesis of the Recent Literature	June-2016	This article analyzes 12 years of recent scholarly research on consumer behavior published in the five leading international journals in this field. The methodology used here involves the classification of topics to evaluate key trends in consumer behavior literature. It includes a ranking of topics published, typology of the published articles, the research classification in terms of methodologies, and analysis techniques. The most cited articles in the field and within each journal are also examined. The comprehensive literature review of consumer behavior research undertaken in this article could advance the discipline of consumer behavior research



			by elucidating the evolution of consumer behavior literature in the studied period.
4	Mining the E-commerce Data to Analyze the Target Customer Behavior	July-2008	This research used the principles of data mining to cluster customer segments by using K-Means algorithm and data from web log of various e-commerce websites. Consequently, the results showed that there was a clear distinction between the segments in terms of Customer Behavior.
5	Customer Behavioral Model Using Data Mining.	September-2014	This report shows the problem of customer relationship management (CRM) and how data mining tools are used to support the decision making. We describe the methods towards predicting customer's behavior, such as collection and preparation of data, segmentation and profiling modeling. This report covers discussion about web mining which can be considered as a separate section due to its current popularity e-commerce. Data mining technologies extract hidden information and knowledge from large data stored in databases or data warehouses, thereby supporting the decision making process of firms. Data mining helps marketing professionals improve their understanding of customer behavior. In turn, this better understanding allows them to target marketing campaigns more accurately and to align campaigns more closely with the needs, wants and attitudes of customers and prospects.
6	Data Mining Techniques: A Source For Consumer Behavior Analysis.	October-2018	The author aim of this paper is to analyze about various studies on consumer purchasing behaviors that have been presented and used in real problems. Data mining techniques are expected to be a more effective tool for analyzing consumer behaviors. However, the data mining method has disadvantages as well as advantages. Therefore, it is important to select appropriate techniques to mine databases. The objective of this paper is to know consumer behavior, his psychological condition at the time of purchase and how suitable data mining method apply to improve conventional method. Moreover, in an experiment, association rule is employed to mine rules for trusted customers using sales data in a super market industry.
7	Mining the customer behavior using web usage mining in e-commerce .	July-2012	The main purpose of this paper is to study the customer's behavior using the Web mining techniques and its application in e-commerce to mine customer behavior. The concept of Web mining describing the process of Web data mining in detail: source data collection, data preprocessing, pattern discovery, pattern analysis and cluster analysis. With the advanced information technologies, server are now able to collect and store mountains of data, describing their numerous contributions and different customer profiles, from which they seek to derive information about their customer's requirements. Conventional methods are no longer appropriate for these business situations to find the customer behavior. The principle of data mining is to cluster customer segments by using K-Means algorithm in which input data comes from web log of various e-commerce websites. Hence, determine the relationship between Web data mining and ecommerce and also to apply Web mining technology in ecommerce.
8	Customer Behaviour Analysis: Identifying risky customers based on their purchased product on e-commerce.	September - 2020	Today's world many people are using e-commerce to buy their day to day products. E-commerce plays a vital role in purchasing products as the e-commerce industry grows many problems related to this industry start growing. The main problem the e-commerce industry is facing is to identify risky customers based on their purchase of products. Many criminals are today using e-commerce sites to buy antisocial or harmful or unusual or explosive/chemicals products to attempt crime. Many criminals and terror organizations are using e-commerce sites to buy chemicals, acids, electronics components. This research paper focuses on identifying risky customers based on purchased products using machine learning methods.

#### IV. ARCHITECTURE

##### Architecture Of System Working Principle :

Data mining is mainly extracting information and knowledge which is not known by people and potentially useful from a large number of incomplete and vague random data of practical application. Web mining is the application of data

mining technology, which is to extract interesting and potentially useful patterns and hidden information from web documents and web activities. Web Mining is broadly categorized into Web content mining (WCM), Web structure mining (WSM), and Web usage mining (WUM). Web content mining is related to the uncovering of useful information from web contents, including text, image, audio, video, etc. Research in web content mining encompasses resource discovery from the web, document categorization and clustering, and information extraction from web pages. Web structure mining studies the web's hyperlink structure. It usually involves analysis of the in-links and out-links of a web page. Web usage mining focuses on analyzing search logs or other activity logs to find interesting patterns.

## WEB MINING

The process of Web mining is divided into four stages:

- source data collection
- data pre-processing
- pattern discovery
- pattern analysis.

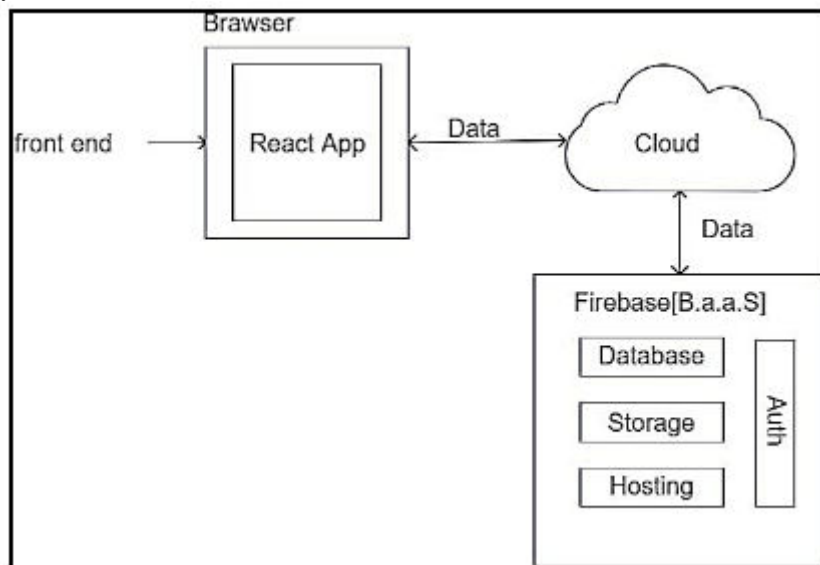


Fig 1. System Architecture

### 1. SOURCE DATA COLLECTION

In mining of Web data, Web log files on the Web server are the main source of data. Web log files contain the history of the visitor's browsing behavior. Web log files include the server log, agent log and client log.

### 2. DATA PRE-PROCESSING

The actual data collected have certain features such as redundancy, ambiguity and incomplete. In order to mine the knowledge more effectively, pre-processing the data collected is essential. Pre-processing can provide accurate, concise data for data mining. Data pre-processing, includes data cleaning, user identification, user session identification, access path supplement and transaction identification.

- a. The main task of data cleaning is to remove the Web log redundant data which is not associated with the useful data, narrowing the scope of data objects .
- b. Determining the single user must be done after data cleaning. Users are identified by their unique ID. Unique ID is given to each user. The purpose of user identification is to identify the user's uniqueness. User

identification can be complete by means of cookie technology, user registration techniques and heuristic rules.

- c. User session identification should be done on the basis of the user identification. The purpose is to divide each user's access information into several separate session processes. The simplest way is to use time-out estimation approach, that is, when the time interval between the page requests exceeds the given value, namely, that the user has started a new session.
- d. Because the widespread use of the page caching technology and the proxy servers, the access path recorded by the Web server access logs may not be the complete access path of users. Incomplete access log does not accurately reflect the user's access patterns, so it is necessary to add access path. Path supplement can be achieved using the Web site topology to make the page analysis.
- e. The transaction identification is based on the user's session recognition, and its purpose is to divide or combine transactions according to the demand of data mining tasks in order to make it appropriate for demand of data mining analysis.

#### **V. PATTERN DISCOVERIES**

A pattern discovery includes classification analysis, association rule discovery, sequential pattern discovery, clustering analysis and dependency modelling.

- a. In classification analysis, data is classified according to the pre-defined categories which are mainly attributing the user profile to established categories of users. After classification, activities according to the characteristics of type of clients, personalized information services are started.
- b. The association rule discovery is used for determining the relevant rules from the Web log database access information. It also determines the inter-relationship which is unidentified from data by analysing the potential linkages between users access to web pages.
- c. Sequential pattern discovery is to tap the model which has time series relations. On the site server logs, the user's sequence of behavior is uneven due to odd visit of the user that generates a discontinuous time series. The focus of sequential pattern mining is analysis of relations between data. Sequential pattern was found to aid e-commerce organizers to predict customer's access model and to carry out targeted services.
- d. The clustering analysis is used to classify user or data items with similar personality and is used to collect user or data items with similar personality together. It can help with marketing decisions.
- e. The dependency modelling is used to develop a model which can state the dependence between the various variables in Web field. It can provide a theoretical framework for the analysis of user behavior, and has the potential of predicting the Web resource utilization.

#### **VI. PATTERN ANALYSIS**

A pattern set is generated using a pattern discovery algorithm. A pattern is analysed on the basis of the search event of a particular user. The data we get from the users activity on the website is used for pattern analysis

VII. IMPLEMENTATION DETAILS

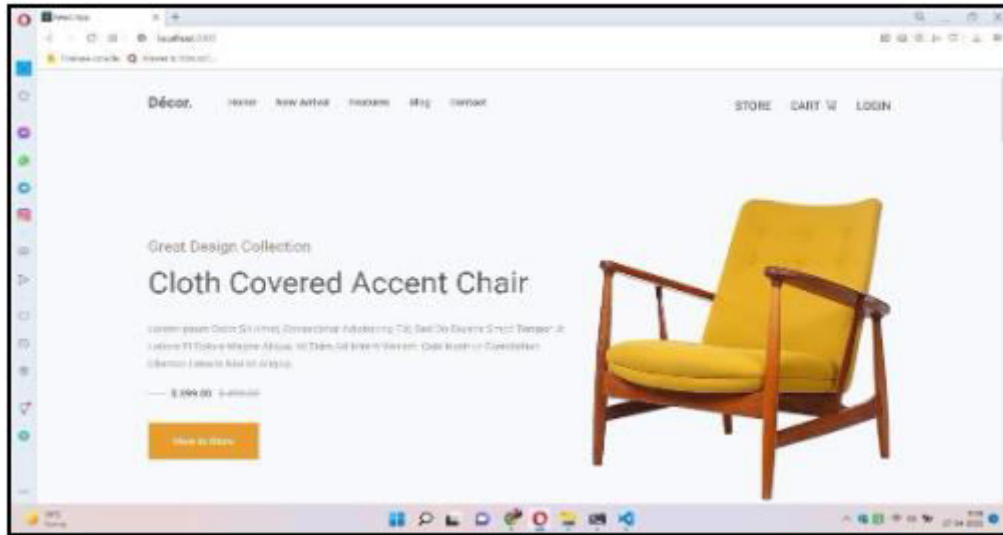


Fig 2. Home Page.

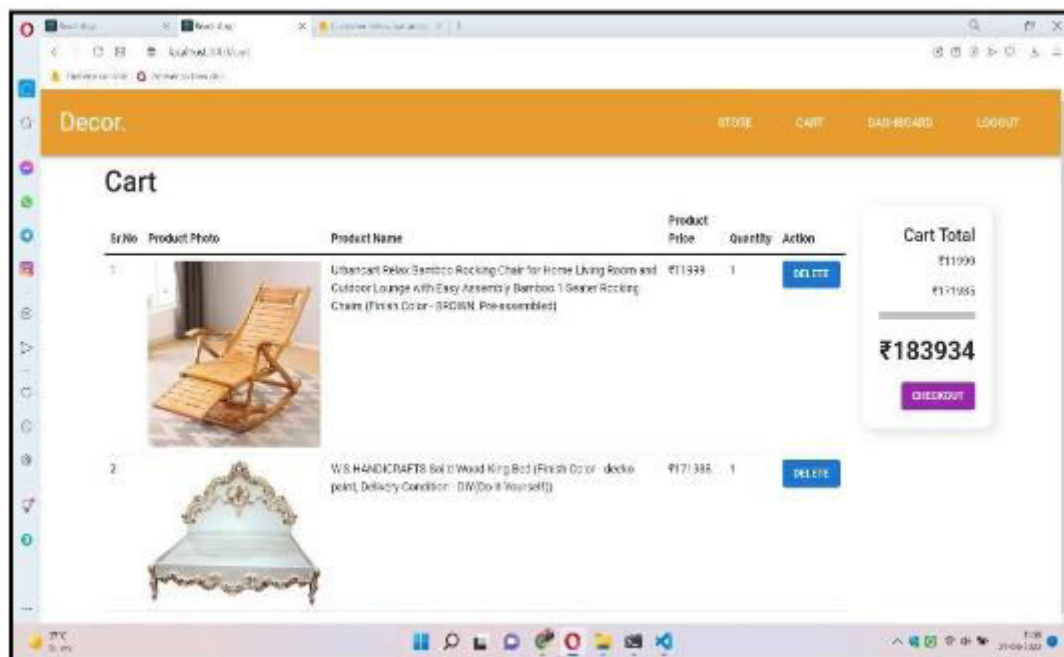


Fig 3. Checkout Window.

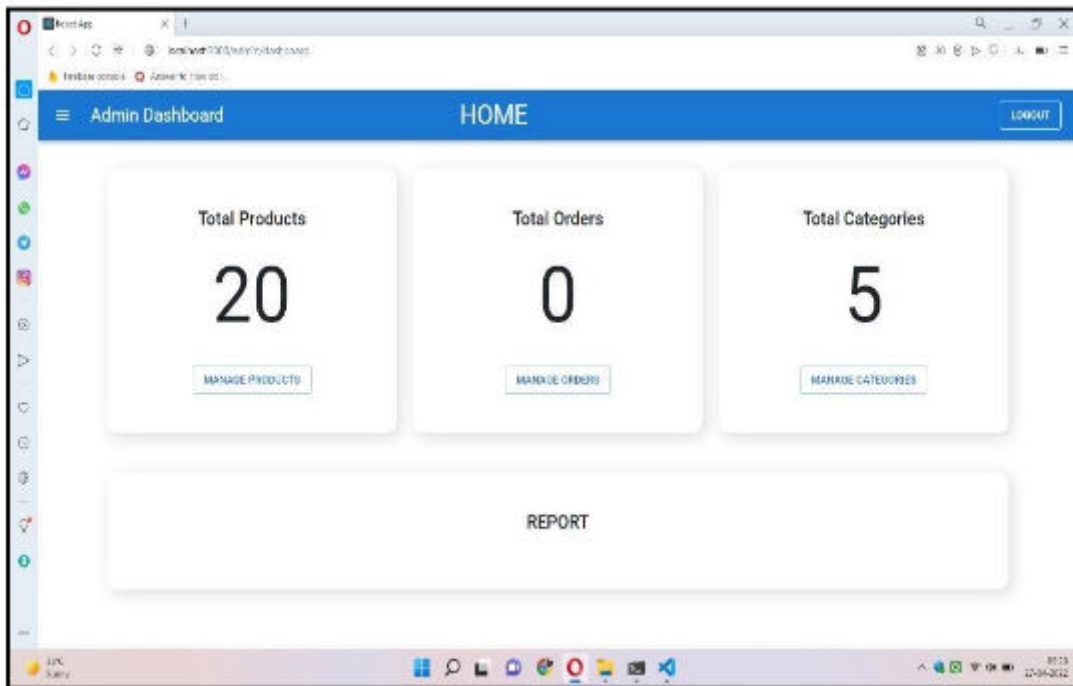


Fig 4. Admin Dashboard

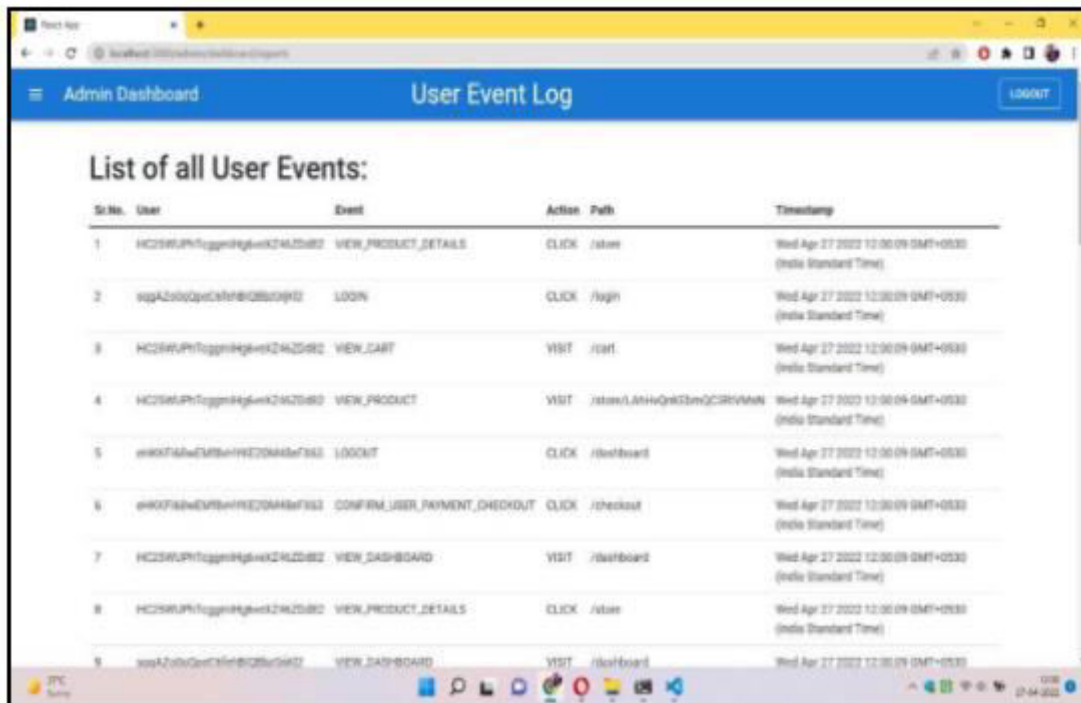


Fig 5 . Log File

VIII. CONCLUSION

Data mining is primarily used today by companies with a strong customer focus-retail, financial, communication and marketing organizations. Data mining has a lot of importance because of its huge applicability. It





is being used increasingly in business applications for understanding and then predicting valuable data, like customer buying actions and buying tendency, profiles of customers, industry analysis, etc. Data Mining is used in several applications like market research, customer behavior, direct marketing, bioinformatics, genetics, e-commerce, customer relationship management and financial services. Studying consumer Online buying behavior with the help of data mining as the analysis tool will give a better insight about the perception of the users when they shop online. This will further help the online retailers to plan effective marketing strategies, which will lead to growth in the online market that will generate profits and further will lead to growth of the Indian economy.

#### **REFERENCES**

- [1] A Data Mining Based Approach to Customer Behavior in an Electronic Settings, May-2019.
- [2] Do Retailers Benefit from Deploying Customer Analytics? June- 2020.
- [3] Consumer Behavior Research: A Synthesis of Recent Literature. June-2016.
- [4] Mining the E-commerce Data to Analyze the Target Customer Behavior. July-2008.
- [5] Customer Behavioral Model Using Data Mining, September-2014.
- [6] Data Mining Techniques: A Source For Consumer Behavior Analysis. October -2018.
- [7] Mining the customer behavior using web usage mining in e- commerce, July-2012.
- [8] Customer Behavior Analysis: Identifying risky customers based on their purchased product on e-commerce, September-2020.



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
Impact Factor: 8.165



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details