WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on Customer Behavior Analysis using Data Mining

Prof. Rahenaaz Pathan[1], Prof. Kinnari Mishra[2]

[1]Lecturer, Parul University,Gujarat,India
[2]Lecturer,Parul University,Gujarat,India
[1]rahenaazbanu.pathan17716@paruluniversity.ac.in
[2]kinnari.mishra12751@paruluniversity.ac.in

ABSTRACT

This paper attempts to determine Using web data sources, uncovering disguised information and identifying user behavior on the web. Predicting user behavior may be done quite simple with the use of this information. On the website of the particular organization, the user pattern is assessed using the free tool similar web and page score. This tool attempts to mine the web in a domain-independent manner. The findings of the experiment make it easier to traverse the website and improve its design architecture. This project examines the unique outcomes of a website in the context of a specific education domain application. The primary goal of doing research on university websites is to identify patterns of behavioral expressions and user characteristics so that the organization may improve the structure of the website and properly upload additional material.

***Keywords*:**Datamining,ConsumerBehavior, Pattern, Analyzing, Statistics

I.INTRODUCTION

The Internet's popularity continues to rise, and it now offers a wide range of services in virtually every industry. The World Wide Web, or Www, is one of the Internet's most important services, having become a standard tool for practically every activity needing a knowledge base. It not only meets the needs of various communities, whether a student wants to learn more about a specific topic, any university for admission, projects, or a want to buy something, but it also assists Internet-based business web sites that are facing competition due to the rapid growth of E commerce and the large amounts of data generated by these web-portals. This data consists of numerous logs and transaction details, and using this data, marketanalysesand predictionscan bedoneveryfairlyaccurately.

Web-portals are an important tool for sharing and accessing information, as well as for forming networks and collaborations, in many

organizations and enterprises. The majority of the time, the online transaction process encourages transactions. These are now the organizations' and corporate offices' primary strategies. Even the top education organizations have started offering online courses and teaching methods have also been modified with webinars etc. In fact, these online opportunities have revolutionized the education system. This is also true for distance education system and given the multitasking requirement in the current situation, online opportunities meet these requirements with efficiency. Because online processes make it relatively simple to store and retrieve information, data mining opportunities abound, and it's even better when logs can be accessed for research purposes. Various research attempts have been made to uncover the quantum of information and how it might be leveraged to expand e-commerce potential.

II. DATAMINING PROCESSES

The need to create data mining techniques to extract information from large web-log files is constantly there. These are frequently automated and methodical processes for analysing new data and user response patterns. Given the large amount of data, delving into all of the specifics of website users is a difficult undertaking.

Data Collection: Log files are typically accessible via a variety of sources, including portal-specific servers, open-source portals, client personal sites, proxy servers, and so on. The most readily available source of information about websites and user interfaces is a web log server. Users can access the source files, which are saved in plain text format.

Preprocessing: This is mostly for the purpose of removing click stream data and categorising transactions based on visiting facts. Because not all

servers have the correct or required format, preprocessing is essential.

Data Cleaning: Data cleaning is a time-consuming step in the data mining from web sources process. Because there is so much information and so many opportunities, it is critical to filter out the noise, misinformation, and redundant data. This usually entails the deletion of files like jpeg, text, gif, sound, and animation. Following that, IP addresses are utilised to identify users, and log files are used to understand their habits, such as frequency of visits and duration of website access.

PatternDiscovery: Studying particular patterns of information storage and retrieval, user behaviours, and specific transaction data is what pattern discovery is all about. Summary statistics, pattern recognition, and machine learning techniques such as artificial neural networks and clustering algorithms are all part of this process. On preprocessed log data, developing pattern discovery approaches include path identification, Frequent Pattern based Association rule, grouping, and classification. For finding resemblance patterns, a clustering approach is utilized.

Pattern Analysis: The knowledge discovery technique is used to separate the predictive emergent patterns from web log files. The final stage of the web usage mining process is pattern analysis. It incorporates the translation of mined pattern data to web log data.

III. TOOLS AND DATA COLLECTION

There are a variety of programmers for analyzing web log files. Free web analyzer programmers are used to collect data. A comparative study is made between the Top most universities of Haryana by knowing the customer interaction between the period Dec 2016 to Feb 2017 on the web sites of the universities mentioned in table 1 by using similar web tool and page score tool. There are a variety of experiments that may be used to analyse these web server access logs as input. It generates the reports for frequent access and identifies the user access patterns.

Table:1

Code Name	University Name
A	Maharashtra university
B	Madhurai kamrazUniversity
C	Ch. Deviray University
D	Jambeshwar University
E	M.D.University,Rohtak
F	Sigma Institute, vadodara
G	B.P.Singh Vishvawalyadaya, Khanpur
H	YMCA University of Scienceand Technology

Table 1 Shows the Universities name and its code which is temporary assigned for experiment purpose Table: 1.1

	Ranking	
CodeName	GlobalRank	CountryRank
A	657947	27617
B	89270	3552
C	458553	18602
D	349861	16700
E	102399	4097
F	875610	36476

G	2539932	103937
H	690777	32654

Table 1.1 gives the comparative Global Ranking and county Rank of the Universities Analysis Report from similar web and page score.

Table: 1.2

Code.	Loading Time(in Sec)
A	4.46
B	23.81
C	27.52
D	80.29
E	17.01
G	73.12
H	4.24

This table shows the how many visitor visit the websites ofthe Universities on Mobile phone and Desktops to find the information.

Table1.3

Code.	Direct	Referrals	Search	Social	mail
A	15.85%	6.36%	77.64%	0.15%	0.00%
B	31.50%	8.08%	59.57%	0.68%	0.17%
C	20.28%	4.51%	71.98%	0.00%	3.23%
D	30.38%	9.05%	60.35%	0.22%	0.00%
E	35.94%	5.72%	56.93%	0.52%	89%
F	14.64%	8.09%	76.35%	0.92%	0.00%
G	23.81%	13.89%	61.65%	0.66%	0.00%
H	21.15%	2.40%	75.65%	0.81%	0.00%

This tables shows that when a customer browse the website ofa particular universities then how much loading time it takes

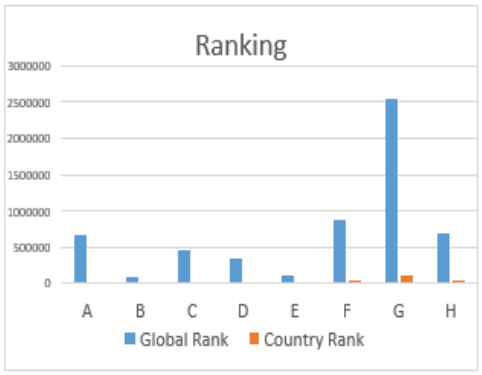
Table1.4

CodeName	Total Visit (On desktop and mobile Phone)
A	34.40K
B	444.60k
C	53.10k
D	104.70k
E	449.30k
F	29K
G	5.70k
H	38.70k

This tables shows that when a customer browse the website of a particular universities then how much loading time it takes.

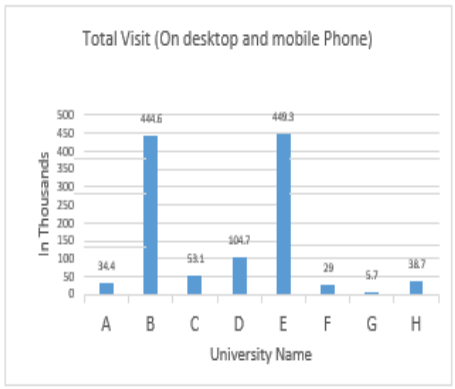
IV. TOOLSANALYSISREPORT (RESULT AND FINDNGS)

From December 7, 2016 to January 31, 2017, data was collected from the websites of Haryana's top universities. Different free tools, such as Page Score and Similar web, are used to examine the collected data. On the basis of web log data from educational universities' websites, a comprehensive experimental study was carried out, and a comparison analysis report was generated. The scope of such work is limited, and it takes time to develop and execute. A comparison chart of university educational websites is available



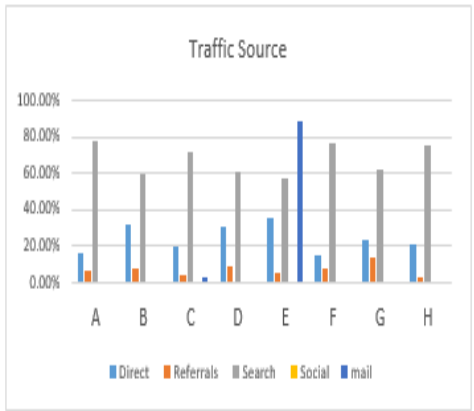
.

Graph1.



Graph2.

Graph3.



Graph 4.

Findings: (7 Dec 2016 to 31 Jan 2017)

1. Fig 4 clearly shows that University code H i.e. Bhagat Phool Singh University has good Global and Country ranking approaching that maximum customer visit this university website as compared toother universities.
2. Website loading time of university A i.e.MMU Mullana is less as compared to other universities
3. 449.3K customers visit the MDU Rohtak websiteduring Dec 2016 to Jan 2017.

V. CONCLUSION

As the usage of websites for online purchase and everything else that a client desires at the touch of a mouse grows, so does the use of websites for everything else. Web usage mining is an important study topic in which data surfing and navigation may be predicted with ease. For online usage mining, several academics have presented a variety of methodologies and technologies.This research examined web user pattern detection and pattern analysis in the education domain utilizing similar web and page scores. The findings should be extremely beneficial to website analysts, maintainers, and designers

REFERENCES

Li Yong-hong; Liu Xiao-liang, "Research of data mining based on e-commerce," Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on , vol.4, no., pp.719,722, 9-11 July 2010

Chaoyang Xiang; Shenghui He; Lei Chen, "A Studying System Based on Web Mining," Intelligent Ubiquitous Computing and Education, 2009, International Symposium pp.433,435, 15-16 May 2009

Li Mei; Feng Cheng, "Overview of Web mining technology and its application in e-commerce," Computer Engineering and Technology (ICCET), 2010, 2nd International Conference on , vol.7, no., pp.V7-277,V7-280,16-18 April 2010

Yadav, M.P.; Feeroz, M.; Yadav, V.K., "Mining the customer behavior using web usage mining in e-commerce," Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on , vol., no., pp.1,5, 26-28 July 2012 .

Zhiwu Liu; Li Wang, "Study of Data Mining Technology Used for E-commerce," Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on , vol., no., pp.509,512, 1-3 Nov. 2010

S.Yadav, K.Ahmad, J.Shekhar, ”Analysis of Web Mining Applications and Beneficial Areas”, International Journal of IUM, vol.12, Issue 2, 2011.

Shen Zihao; Wang Hui, "Research on E-Commerce

Application Based on Web Mining," Intelligent Computing and Cognitive Informatics (ICICCI), 2010 International Conference on , vol., no., pp.337,340, 22-23 June 2010

Tang, Hewen; Yan, Honglin; Zengfang Yang; Yu Ma; Chunping Li, "Application of Data Mining in Electronic Commerce," Information Engineering, 2009. ICIE '09. WASE International Conference on, vol.1, no., pp.631,634, 10-11 July 2009