

Designing Web Mining Technique for Customer Segmentation

¹Naresh Chandra, ²Dr Krishana Kumar

¹Naresh Chandra, Research Scholar, IFTM University, Moradabad, India.

²Dr. Krishna Kumar, Associate Professor, IFTM University, Moradabad, India.

ABSTRACT:

Different web mining techniques are available there. Our assignment is to analyze those web mining techniques for Customer Segmentation. We will analyze how efficiently we can keep track the likes, dislikes and feedback of the customers. Accordingly, we can segment by clustering and splitting customers into bunches that are not recently recognized. After identification of the segmented group we can emphasize on that group which is more profitable or we can develop effective market strategy for that group. Then we will propose a web mining technique that can be more beneficial to target the market and to predict the sale. We can build up the system for increasingly beneficial items by distinguishing the most and least gainful customers.

Based on Geographic, Benefit, Behavior, Demographic, and Psychographic there are five prime segmentation strategies available. Here are some valuable things we can perform for betterment of the customer and also for the business to grow after the analysis of the segmentation done:

- Can identify the foremost and least profitable customers of the market.
- Better focusing on marketing strategies.
- Improving customer service.
- Can build loyal relationships with customers.
- Price different products differently.
- Can develop better products.
- Can create Face values.
- Can provide various possible options.

Date of Submission: 17-01-2020

Date Of Acceptance: 05-02-2020

I. INTRODUCTION

A. Web Mining:

Web mining is a process to finding out the new patterns or extracting the new knowledge from the **World Wide Web**. Name also proposes, that info gathered by mining the web. In context to client relationship management (CRM), Web mining is the mix up of the different ancient data mining procedures and techniques available which we can apply on the info accumulated over the World Wide Web.

B. Types of Web Mining:

As shown in the figure given below, there are three different types of web mining:



i. Web Content Mining: Whenever we mine or extract the useful info from the content of the web page, we refer it as Web Content Mining. We can apply many of the text mining techniques over the content mining because most of the content of the web pages are in the text form. Since most of the content of the web page are semi structured or unstructured in nature, therefore, there is difference between Web content mining and data mining as data mining deals with structured data only. Web content mining needs more imaginative applications of knowledge mining as well as text mining methods and additionally its own distinct methods. Within the past few years, there was fasten activities generations within the area of Web content mining. This is often not stunning as a result of the exceptional growth of the Web contents and vital economic good thing about such mining. However, because of the heterogeneousness and the Web data structure's lack ness, self-discovery of directed or surprised knowledge info still presents several exciting investigation issues.

ii. Web Structure Mining: Web structure mining will classify the Web pages on the basis of

likeness and relationship between different Internet websites and generate the info. Web structure mining tries to determine helpful info from the hyperlinks configuration.

iii. Web Usage Mining: In **Web Usage Mining** we have a tendency to acknowledge the user contact arrangements from usage of Web logs using the solicitation of data mining methods to know and higher serve the needs of Web-based applications. Usage knowledge arrests the distinctiveness or derivation of Web users in conjunction with their surfing patterns of an internet website.

C. Web Mining Techniques:

Different web mining techniques are there for various form of web mining.

i. Some of the techniques used in Web Content Mining are:

- Unstructured Data Mining
- Structured Data Mining
- Semi structured Data Mining
- Multimedia Data Mining

ii. Some of the techniques used in Web Structure Mining are:

- Cardinality wise
- Type wise
- Strength wise
- According to Classification
- According to Cluster Analysis

iii. Some of the techniques used in Web Usage Mining are:

- Data Pre processing
 - Cleansing of Data
 - Identification of Client and
 - Identification of Session
- Pattern Discovery
 - Statistical Investigation
 - Association Regulation
 - Clustering
 - Classification
 - Sequential Patterns
- Pattern Analysis
 - Knowledge Query Mechanism
 - OLAP
 - Intelligent Agent

D. Customer Segmentation:

The way toward partitioning potential markets or purchasers into explicit groups based on the similarities of that group is **Segmentation**. Five prime segmentation approaches are there like:

- a. *Based on Geographic,*
- b. *Benefit,*
- c. *Behavior,*
- d. *Demographic, and*

e. *Psychographic*

E. Advantages of Customer Segmentation:

Here are some valuable things we can perform for betterment of the customer and also for the business to grow after the analysis of the segmentation done:

- Can identify the foremost and least profitable customers of the market.
- Better focusing on marketing strategies.
- Improving customer service.
- Can build loyal relationships with customers.
- Price different products differently.
- Can develop better products.
- Can create Face values.
- Can provide various possible options.

II. LITERATURE REVIEW

[1] **Tak-Lam Wong** et al (2010) uses the Bayesian learning of new attribute. First objective of their method is to learned from a source Web site to a new hidden site. Another aim of their approach is to handle with new attribute which was not learned prior. Semantic labels i.e. part of the text on the Web page representing the attribute name found here can be discovered.

[2] **Bharati M. Ramageri** (2010) study different algo and techniques of data mining. Whenever voluminous data is gathered there are very broad possibilities of data mining to find the patterns or new knowledge from the data, which can be very helpful for growing of the business fast. Therefore, data mining is very bright future and very important development of info technology.

[3] **Dr. Sankar Rajagopal** (2011) uses the clustering for discovery of high-profit and low risk client. For business to grow fast and taking the right decision, it is very important for clustering and for finding the pattern from the data of customers. First he performed data cleaning and found the patterns with demographic clustering algorithm using IBM I-Miner. Then after profiling the data, identification of new clusters and identification of the high-value low-risk clients can be done. 80% of the revenue from 10-20% of the customers achieved from this cluster.

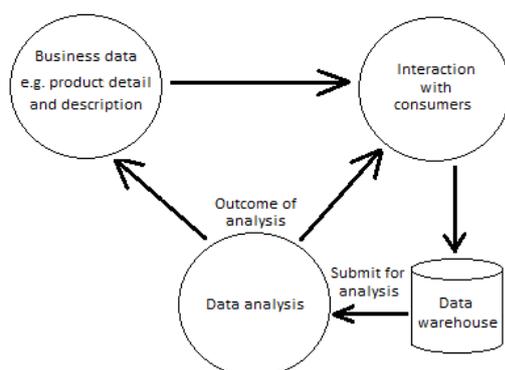
[4] **Hao Ma** et al (2012) generates the graph of data source and then mine that graph for recommendation. A node of the graph represents the similarities of the attributes. They also suggested generalising various recommendations problem through graph. Their methods of graph mining have the bright future for large scale data source of web.

[5] **Rupam Some** (2012) provides the different activities and updating done in all the three form of mining i.e. web content mining, web usage mining and the web structure mining. A concurrent evaluation is done for past present and future of the

web mining. Authors show the importance of the knowledge extraction through web mining.

[6] **Pradnyesh Bhisikar** et al (2013) article provides the review and analysis of different three classes of web mining i.e. web content mining, web usage mining and the web structure mining. Different tasks of web usage mining are told by authors like gathering, preparation, pattern analysis, pattern visualisation and application. At the time of website designing, the mining of server log under web usage mining can enhance the usability of website. Web customisation can be achieved by web usage mining techniques.

[7] **Ahmad Tasnim Siddiqui** et al (2013) proposed system model to integrate web mining with e-commerce applications shown in the following figure :



Business data collection regarding the consumer's choice, clicks on web pages and links and the nature if consumers can be stored in data warehouse. The accuracy and low recall page are very low. For example, if we search anything via any search engine, we get inappropriate and superfluous info. Predictive tools of SQL and Oracle's data mining tools can be used for mining the data stored in warehouse.

[8] **Kishana R** et al (2013) states that for the prediction of sale of market according to season, an online system can be developed using k-means clustering method and SPSS tools. Author states the importance of the clustering methods for segmentation of customers for finding out the trends in market to research further for betterment of customers. In the model developed by the authors, the data of the end of the day is compared with the data used for prediction. Their model gives the good and accurate in the form of high sale.

[9] **Hasan Ziafat** et al (2014) outlined that whenever the knowledge of business and the knowledge of data mining techniques occurs simultaneously then the result of mining can be very surprised. In the framework of Customer Relation

Management there are some of the data mining techniques are very useful such like:

- Customer segmentation:
 - Segmentation according to value:
 - Behavioral segmentation:
 - Value-at-risk segmentation:
- Targeted marketing campaigns:
 - Voluntary churn modeling and estimation of the customer's likelihood/propensity to churn.
 - Estimation of the chance/tendency to take up an add-on product, to switch to a more gainful produce, or to upsurge practice of an current produce.
 - Estimation of the lifetime value (LTV) of customers.

The different steps in segmentation can be as:

Step 1: Database can be analysed for customer segments

Step 2: Evaluation and fixing the position the segments

Step3: Cost-benefit analysis to be done for various actions per segment

Step 4: Preparation and delivery of different mechanism for different segment.

[10] **Y.Raju** et al (2015) have offered in this paper the importance of web mining methods related to web personalisation. Investigation of the targets and business requirement development entailed by Personalization. Personalisation can be categorised in three different types:

1. Based on particular group or profile
2. Antiquates based
3. Association based

[11] **Pranit B** et al (2015) described some characteristics of big data which are threats and opportunities also for mining and discovery of useful info or knowledge from the web. These are:

- (a) Info present on web is heterogeneous in nature. Same info on different web pages may be presented in differently by words or other contents.
- (b) Info on the web is linked. Internet sites have the hyperlink whether in same page or different site.
- (c) Most of the info on web is noisy.
- (d) Most of the web pages are dynamic in nature and are updating continuously.
- (e) The web pages are cybernetic society. Map Reduce, Apache Hadoop and visualization-based method called as Visual Web Mining (VWM) are studied by authors to mining the big data.

[12] **V. Anitha** et al (2016) prescribed system to discover user behaviour through the web server log file. All the attributes related to links are kept with client. Then cleansing of the data is done to eliminate the noise and unwanted data. Then on the basis of links stored in the database the behaviour of the customer is predicted.

[13] **Er. Rupampreet Kaur** et al (2017) states that the main target of segmentation was to isolate the

objects that are homogeneous and heterogeneous with the external market. The result of segmentation depends mainly on the knowledge variables. Clustering and subgroup discovery are two types of customer segmentation. To determine groups and structures in the data, Clustering methods are used. Discovery of subgroup can be identified by the dependency of the variables on another variables. Different methodologies used are compared on the base of usage and validation.

[14] **Keerti. S.** et al (2017) says that customer satisfaction is now a day is very important thing for sustainable growth of market in the today's competition. Therefore an intelligence online system for well-organized marketing is required.

[15] **Kiril Griazev** et al (2018) presented a new web mining arrangement for web mining. This arrangement will serve as classification method. They categorised the existing web mining techniques on the basis of different criteria.

III. PROBLEM DESCRIPTION

For doing the analysis of different existing web mining techniques we have to take huge data set from online shopping portal such as Amazon, Flipcart and then applying those web mining techniques on the data for customer segmentation. After applying those web mining techniques on the data taken, we are able to segment the groups of customers. We will analyze how efficiently we can keep track the likes, dislikes and feedback of the customers. Accordingly, we can segment by clustering and splitting customers into bunches that are not recently recognized. After identification of the group we can emphasize on that group which is more profitable or we can develop effective market strategy for that group. Then we will propose a web mining method that can be more beneficial to capture the more market and to predict the sale. We can build up the system for increasingly beneficial items by distinguishing the most and least gainful customers.

IV. FINDINGS/RESULT

	Types	Explanation
Web Mining	Web Content Mining Web Usage Mining Web Structure Mining	All the literature available categorizes mainly the web mining into these three types.
	Methods	
Customer Segmentation	Classification Clustering Regression Neural network Association Rule Decision Tree Nearest Neighbor	Different segmentation approaches in web mining are available. Customer segmentation is more important in terms of profit.

V. CONCLUSION

Different web mining techniques are available there. Our task is to analyze different web mining techniques for Customer Segmentation. We can segment by clustering and splitting customers into bunches that are not recently recognized on the foundation of different characteristics of the clients. After identification of the segmented group we can emphasize on that group which is more profitable or we can develop effective market strategy for that group. Then we will propose a web mining method that can be more beneficial for capturing the more market and to predict the sale.

REFERENCES

- [1]. Tak-Lam Wong and Wai Lam, **Learning to Adapt Web Info Extraction Knowledge and Discovering New Attributes via a Bayesian Approach** IEEE April, 2010
- [2]. Bharati M. Ramageri, **Data Mining techniques and applications**, Indian Journal of Computer Science and Engineering Vol. 1 No. 4 301-305, December, 2010
- [3]. Dr. Sankar Rajagopal, **Customer data clustering using data mining technique**, International Journal of Database Management Systems (IJDMS) Vol.3, No.4, November 2011
- [4]. Hao Ma, Irwin King, Senior Member, IEEE, and Michael Rung-Tsong Lyu, Fellow, **Mining Web Graphs for Recommendations**, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 6, JUNE 2012
- [5]. Rupam Some, **Performance Analysis of Web Mining Techniques and Future Directions**, IJCST Vol. 3, Issue 2, April - June 2012

- [6]. Pradnyesh Bhisikar, Prof. Amit Sahu, **Overview on Web Mining and Different Technique for Web Personalisation**, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 Vol. 3, Issue 2, March -April 2013
- [7]. Ahmad Tasnim Siddiqui, Sultan Aljahdali, **Web Mining Techniques in E-Commerce Applications**, International Journal of Computer Applications (0975 – 8887) Volume 69– No.8, May 2013
- [8]. Kishana R. Kashwan, Member, IACSIT, and C. M. Velu, **Customer Segmentation Using Clustering and Data Mining Techniques**, International Journal of Computer Theory and Engineering, Vol. 5, No. 6, December 2013
- [9]. Hasan Ziafat , MajidShakeri, **Using Data Mining Techniques in Customer Segmentation**, International Journal of Engineering Research and Applications ISSN : 2248-9622, Vol. 4, Issue 9(Version 3), September 2014, pp.70-79
- [10]. Y.Raju1, Dr. D. Suresh Babu, **A Novel Approaches is web mining techniques in case of web personalization**, International Journal of Research in computer application and robotics ISSN 2320-7345, February, 2015
- [11]. Pranit B. Mohata, Prof. Sheetal Dhande, **Web Data Mining Techniques and Implementation for Handling Big Data**, ISSN 2320-088XIJCSMC, Vol. 4, Issue. 4, April 2015
- [12]. V.Anitha, Dr.P.Isakki, **A Survey on Predicting User Behavior Based on Web Server Log Files in a Web Usage Mining**, 2016
- [13]. Er. Rupampreet Kaur, Er.Kiranbir Kaur, **Data Mining on Customer Segmentation: A Review**, International Journal of Advanced Research in Computer Science, Volume 8, No. 5, May-June 2017
- [14]. Keerti. S. Mahajan, S.S.Jamsandekar, Dr. A M. Gurav, **Machine Learning Approach for Marketing Intelligence: Managerial Application**, July 2017, International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 6 Issue 7
- [15]. Kiril Griazev, Simona Ramanauskaitė, **Web Mining Taxonomy**, 978-1-5386-6737-8/18/\$31.00 ©2018 IEEE
<https://www.researchgate.net>
- [17]. <https://www.ieee.org>
- [18]. <https://www.elsevier.com>

AUTHORS PROFILE



¹**Naresh Chandra** having the Master Degree of Computer Application (MCA-1998) from Gurukula Kangri Vishwavidyalaya, Haridwar, Uttarakhand, India. He is working with Planning Department of Uttarakhand State Government. At present he is pursuing his Ph.D. also in Computer Science and Application from IFTM University, Moradabad, India.



²**Dr Krishna Kumar** having the Ph.D. in Computer Science and Application (2014-2018) from IFTM University, Moradabad, India. At present, he is working as Associate Professor in IFTM University, Moradabad, India. Many of his manuscripts are published in various journals.

Naresh Chandra. “Designing Web Mining Technique for Customer Segmentation.” *International Journal of Engineering Research and Applications (IJERA)*, vol.10 (02), 2020, pp 01-05.