Web Mining: Introduction, Need, Working applications and types

Simarjeet Kaur
Assistant Professor,
Department of Computer Science and Applications,
MindTrek College, Jalandhar, India
Email - jacksgirl64@gmail.com

Abstract: *In today's world, large storage capacity is used for backup or storage of the useful data. In this way database form from the large set of data. On the other hand side, it is also difficult to handle such type of databases that are big in the Size. Sometimes we need only a single piece or specific portion of the database. It is not easy to get that specific part from the huge database. So, a mechanism is needed to perform this task. For this, a process is used for mining the useful data from the database. Mining is the process of extracting the useful and required data or piece of data from the huge database. This is the basic concept of the data mining that makes it a useful and powerful way to explore the huge databases in easier ways. This research paper focuses on the all the important aspects of the web mining and data mining including introduction, need, working of the web mining, types and applications.*

Key Words: *Mining ,Mechanism ,Useful , Specific, Portion*

1. INTRODUCTION:

Database is collection of data that includes different types of data. It is easier to get the required part of the data when the size of the database is small.[1] On the other hand side, if the size of the database is big, then it is difficult to get the required or specific part of the data, that is needed.[2] For getting the required part from the database, mining is used. Mining is the process of extraction and when the data is extracted from the database, it is known as data mining. Various methods are available for doing the process of mining.[3] World wide web contains large amount of useful information and data. Traditional methods are used to extract the required and useful information from the world wide web. This process is known as web mining.[4] Web mining extract the useful information and data from the web by using data mining methods. After that it integrates all the extracted information that user got after using the data extraction methods from the world wide web.

2. NEED:

It is very difficult to find the required data from the database that is big in size. If we are unable to get the required data from the database, then database is useless. The process of finding the required data or specific data from the database is very time consuming, difficult and long process in case of huge database. So, a mechanism is needed for fulfilling that need of the database user. Web mining and data mining are helpful for accomplishing this task. By using these mechanisms, time can be saved and human effort can be reduce. Mining helps to extract the useful or required part of the data from the database and world wide web.

3. WORKING:

Various methods are used to get the useful information directly from the web. For this purpose, web mining uses the web pages, server logs including content on the web and content of web. Web mining includes various techniques. These includes :

- Web Content based mining,
- web usage mining
- web structure mining
- Clustering
- Classification
- Association

All these techniques and methods are used for web mining. These are basic techniques that are useful in mining the data from the web. Explanation of all these is given below-:

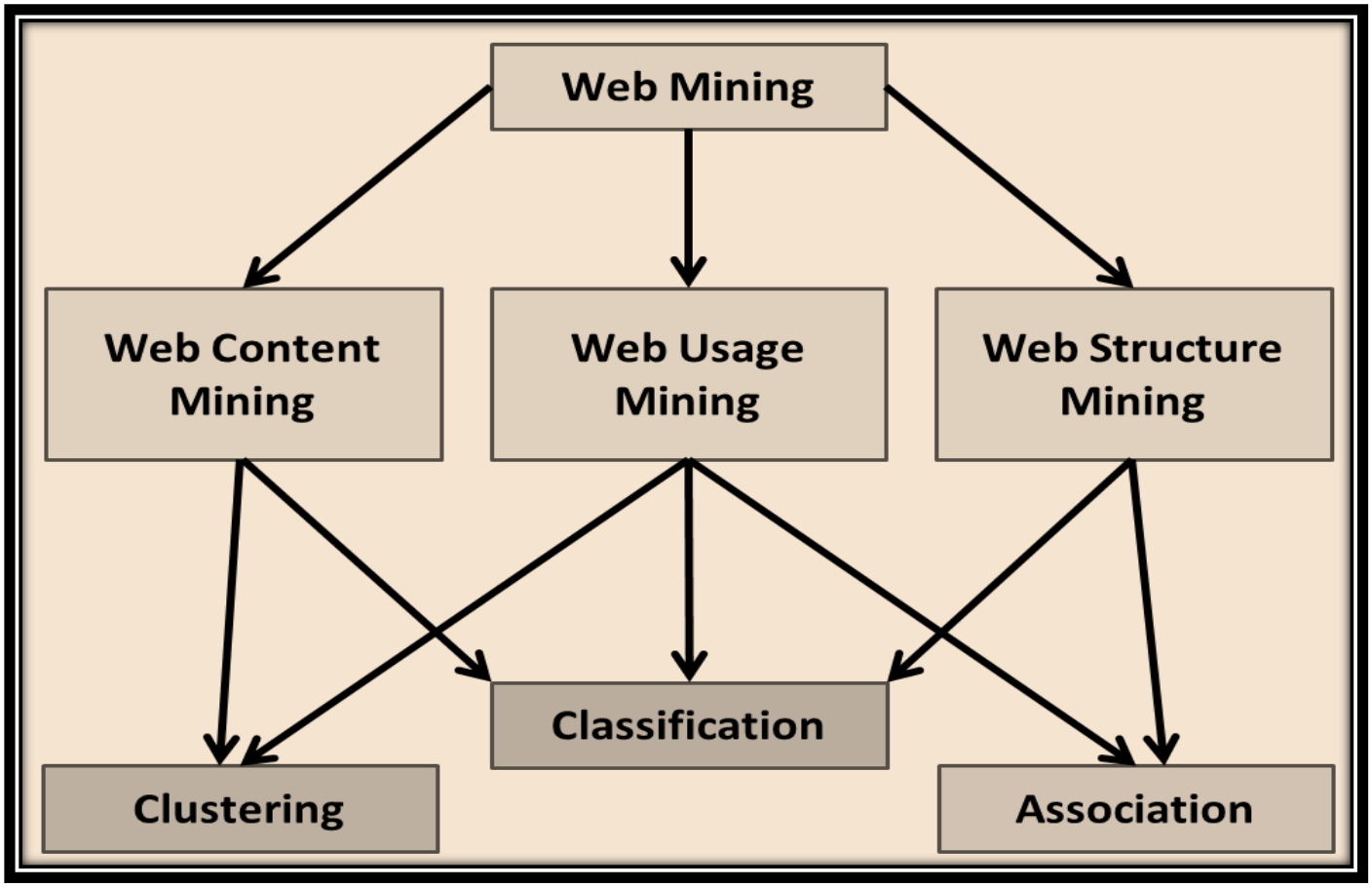

Fig. 1: WORKING OF WEB MINING

3.1. Web Content Mining: In this technique, Useful information from web is extracted from the web. For this purpose, contents of the web pages are used. Content is considered in this technique, that's why it is known as web content mining.

3.2. Web Usage Mining: In this technique, the patterns of the data is discovered by checking the client server transactions. Most frequent data patterns that are used frequently by client and server are selected for this mining.

3.3. Web Structure Mining: In this type of the mining, structure of the web is discovered , which includes the various web pages that act as node of graph and hyperlinks that act as edges of the graph. Structure is represented by the graph.

4.4. Clustering: Clustering technique uses various clusters. Clusters are subsets of the data. Set of clusters are used in this technique. Clustering includes hierarchical and partitional clustering.

- **Partitional Clustering:** In this type of clustering, data is divided into various parts but considered as a single subset of clustering only.
- **Hierarchical Clustering:** This technique uses the nested clusters. This technique uses the hierarchical structure of the clusters.

3.5. Classification: In this technique, various sets of the data are used to build data models. For this purpose, various classes are assigned to the data sets depending upon the categories of the data. Process of Categorization is used for the purpose of the classification.

3.6. Association: Association is used for web mining and data mining. Various Association rules are used for finding and generating various data sets from the mining process. Various rules are generated for finding the associations between the various data sets. But this process of association is very costly and not efficient .

4. APPLICATIONS: There are many applications of the web mining. Some of them are given below-:

- Web mining is used for finding the various web documents related to any type of particular subject or topic b which these web documents are presented.
- These techniques are also helpful to divide the web documents into various categories. In this way, it become easier to find the particular topic.
- Web mining also helps to find out the web documents or web pages that are related with each other. For this, various techniques are used for finding the related web documents from multiple servers.
- Web mining is also helpful in the field of education. Web mining is helpful in the area of online education and e-learning, where web based environment is needed.
- Web mining is also used in the field of online services that are provided b the government to the public via online mode. Thus, web mining is helpful in the field of e-governance.

5. CONCLUSION:

The use of modern user either computer or mobile user is increased after the introduction of the Modern technology. That's why the storage need increased and database that stores the important data of the user, is also increased in size. For getting the required and useful information from the database, web mining is very helpful and effective tool for the user to save the time and human effort both. So, we can't deny the need of web mining and data mining and their tools. In upcoming days, there will more changes and advancements in the field of web mining and data mining.

6. ACKNOWLEDMENT:

While making this research paper, I have put all of my best and possible efforts for making this research paper. I hope this research paper will be useful and helpful for the authors who want to do further research Related to web mining and working of web mining.

References:

1. "Zdravko Markov, Daniel T. Larose "Data Mining the Web: Uncovering Patterns in Web Content, Structure, and Usage", Wiley, 2007
2. Bing Liu, "Web Data Mining: Exploring Hyperlinks, Contents and Usage Data", Springer, 2007Bing Liu, "Web Data Mining: Exploring Hyperlinks, Contents and Usage Data", Springer, 2007.
3. Web Mining and Web Usage Analysis 2004 - revised papers from 6 th workshop on Knowledge Discovery on the Web, Bamshad Mobasher, Olfa Nasraoui, Bing Liu, Brij Masand, Eds., Springer Lecture Notes in Artificial Intelligence, 2006
4. Cooley, R., Mobasher, B. and Srivastava, J. "Data Preparation for Mining World Wide Web Browsing Patterns", Journal of Knowledge and Information System, Vol.1, Issue. 1, pp. 5–32, 1999Cooley, R., Mobasher, B. and Srivastava, J. "Data Preparation for Mining World Wide Web Browsing Patterns", Journal of Knowledge and Information System, Vol.1, Issue. 1, pp. 5–32, 1999
5. Kohavi, R., Mason, L. and Zheng, Z. (2004) "Lessons and Challenges from Mining Retail E-commerce Data" Machine Learning, Vol 57, pp. 83–113.
6. Eirinaki, M., Vazirgiannis, M. (2003) "Web Mining for Web Personalization", ACM Transactions on Internet Technology, Vol.3, No.1, February 2003.

Web References:

- http://webmining.spd.louisville.edu
- https://www.researchgate.net