Intelligent Web Data with Semantic Web Mining

S. Surya[1], E. Premalatha[2], S. Muthulakshmi[3]

[1]MCA. B.Ed., M.Phil., Assistant professor (Department of computer science), Sri Akilandeswari Women's College, Wandiwash, Tamil Nadu, India
[2]PG Student (Department of Computer Science), Sri Akilandeswari Women's College, Wandiwash, Tamil Nadu, India
[3]PG Student (Department of Computer Science), Sri Akilandeswari Women's College, Wandiwash, Tamil Nadu, India

Abstract: *Semantic web mining is the outcome of two new and fast developing domains the semantic is web and data mining. The semantic web is implementation of the present web in which the information is well formed data mining is the non-trivial process of identifying valid, novel potentially useful and ultimately understandable patterns in data. Semantic web mining considered to application of data mining method and extract information for www. The web data considered for many kind of knowledge communication including web log data user profile data and web structure data. Semantic web mining relevant current place general way of two or more different areas, interaction research from knowledge back to special expert from business studies fields and computer science. Web data mining consists of web content mining web usage structure and web structure mining Data mining is used in various techniques such as classification, clustering, prediction, sequential discovery, decision tree, combination of long-term processing.*

Keywords: Data mining, Web, Web mining, Semantic

1. Introduction

The WWW drastically changed the data representation and its accessing the huge amount of data electronically accessed by the machine with in no time. The computer are developed for numeric calculation but now a day they have been totally used for information processing, text processing using data bases and playing games[1]. There is difference between the ways of information presented to user than actually it is present in the repositories.

The internet contains more than 10 billion static pages of information to be used by more than 1000 million users spread over the world. It is difficult to access and maintain the enormous amount of data using natural languages. The difficult to gap between available information and techniques used for accessing it. Along with required pages of irrelevant or merely relevant pages web retrieved in thousands [1].

The search engines are keyword sensitive, if string of keyword is not found it separates the words and presents an irrelevant matter. It is required to collect the information from several web pages if spread over various documents. The output result for query is relevant pages or meaningful pages on the internet.

2. Semantic web mining Taxonomy

There are three types of broad categories of web mining

1. Web content mining
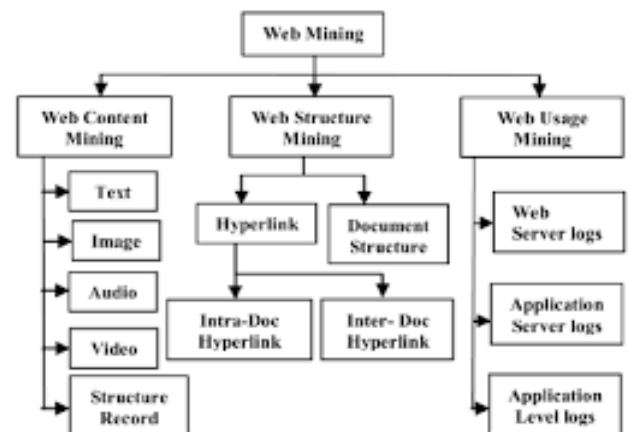2. Web structure mining
3. Web usage mining

Figure 1.1: Web Mining

2.1 Web Content Mining:

Web content mining is the process of extracting useful information from the contents of web documents content data is the collection of facts a web page is designed to contain. It consists of text, images, audio, video or structured records such as lists and tables [5].

Web content mining is also different from text mining because semi structure nature of the web, while text mining focuses on unstructured text. It is requires create an application of data mining and text mining techniques and also its unique approaches.

Figure 1.2: Web Content Mining

2.2 Web structure mining:

The structure of a typical web graph consists of web pages as nodes and hyperlinks as edges connecting related pages. Web structure mining is the process of discovering structure information from the web [2].

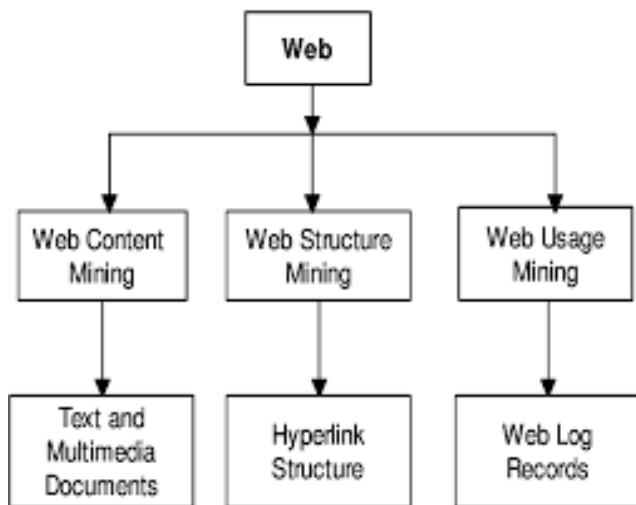Two kinds of structure information that are,

*Hyperlinks
* Document structure

Figure 1.3: Web Structure Mining

2.3 Web usage mining:

Web usage mining is the application of data mining techniques to extract knowledge from web data including web documents, hyperlinks between documents, usage log of web sites, etc.A panel organized at ICTAI1997.There are two different approaches were taken in initially defining web mining first was a "process-centric view" and second was a "data-centric view"[7].
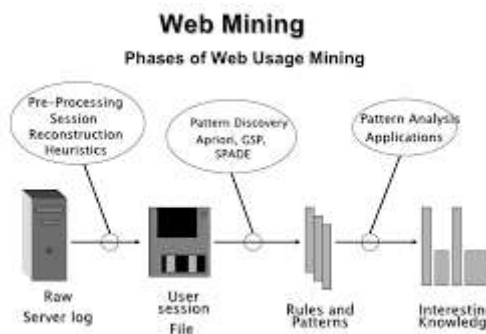

Figure 1.3: Web Usage Mining

3. Visualization of www

World Wide Web is a Collection of information it as been central development of information. This Web may contain images, video, audio and software components. The WWW is a way of exchanging information between computer and internet. It is an open Web platform for application development. The languages of C, C++, JAVA, Ruby on Rains [6].

4. Data Mining Techniques

4.1 Clustering

Clustering is similar to classification expect that groups are not in predefined by the data alone. It is alternatively referred to as unsupervised learning segmentation. Segmenting data into groups that might or might not be disjoined. The most similar data are group into clusters. Clustering is one of the descriptive modeling of data mining [3, 4].

4.2 Classification:

Classification maps data into predefined groups are classes it referred to as supervised learning because the classes are determined examining the data. Pattern recognition is type of classification where input pattern is classified into one or several classes.

4.3 Prediction:

Many real-world data mining application. Predicting future data states based on past and current data. Prediction can be viewed as type of classification prediction application is flooding, speech recognition, machine learning and pattern recognition.

4.4 Sequential discovery:

Sequential discovery is used to determined sequential pattern in data. These patterns are based on the time sequence of action it is similar to associations data. Sequential is unlike the market basket analysis.

4.5 Decision tree:

Decision tree is a predictive modeling technique used in classification, clustering and prediction tasks. Decision tree use a "Divide and conquer" technique to split the problem search space into subset.

5. Conclusion

World Wide Web is provides the collection of information in which whole world through domain organization. This paper suggested to how to invoke the web services and usage web mining visualization. Its scope to retrieved extract data from huge collection database inefficient manner. This paper has also discussed about the data mining and web mining and also used in new techniques.

References

[1] Berners-Lee, T., Hendler, J., and Lassila, o.(2001).The Semantic Web. Scientific American, 284(5):34-43.

[2] Desikan, P., Srinivastava, J., Kumar, V., and Tan, P.N.(2002).Hyperlink Analysis: Techniques and Applications, Technical Report(TR 2002-0152), Army High Performance Computing Center.

[3] Yao, Y.Y., Hamilton, H.J., and Wang, X.(2002).Page Prompter: An Intelligent Agent for Web Navigation Created Using Data Mining Techniques, Lecture Notes In Computer Science, 2475:506-513.

[4] Junkie Chen and Wei Liu, "Research for Web Usage Mining Model", International Conference on Computational Intelligence for Modeling Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce(CIMCA-IAWTIC'06), IEE, vol-2, pp.8, 2006.

[5] Arijit Abraham, "Business Intelligence from Web Content mining", 2003.

[6] Jaideep Srivastava et al."World Wide Web concepts, Usage, and Research Directions".

[7] Jaideep Srivastava et al., "Web usage mining: Discovery and applications of usage patterns from web data" published in ACM SIGKDD Explorations, Volume 1, Issue 2, Jan 2000, pp.12-23