Abstract

The data mining techniques exploitation in the field of web is referred as web mining. The enormous data is present at the websites and this need to be tackled well with the help of different data mining techniques. Searching, pulling data together and analyzing the data are the main focus of web mining. The application of web mining is in the field of e-commerce and e-learning, web search, database, AI, information retrieval, system improvement etc. Information extraction from the web documents is a typical task and can be done efficiently after the thorough study of mining. This paper would facilitate to comprehend the concept of web mining by analyzing the facts retrieved from various sources. The paper presents the literature survey on web mining. It also explains the detailed view of three kinds of web mining techniques viz. web content mining, web structure mining and web usage mining. For the survey, different papers are analyzed and then presented as the study of web mining and its subtasks.

References

- Vel L., Royakkers L., "Ethical Issues in Web Mining", Ethics and

A Hand to Hand Taxonomical Survey on Web Mining


- Xiaoqiu T. and Min Y., "Mining Maximal Frequent Access Sequences Based on Improved WAPTree," Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06).


- Jalali M. and Mustapha N., "WebPUM: A Web-based recommendation system to
predict user future movements\textquotedbl;, international journal Expert Systems with Applications 37 (2010) 6201–6212
- Yong-gui W. and Zohn J. \textquotedbl;Research on Semantic Web Mining\textquotedbl;, 2010 International Conference On Computer Design And Applications (ICCDA 2010).
- Meo R. ,Lanzi L. P. ,Matera M. And Esposito R. \textquotedbl;Integrating Web Conceptual Modeling and Web Usage Mining\textquotedbl;, Proceedings of the sixth WEBKDD workshop: Webmining and Web Usage Analysis (WEBKDD\texttt{amp}aposs;04), in conjunction with the 10th ACM SIGKDD conference (KDD\texttt{amp}apos;04), Seattle, Washington, USA, 2004.
- Mobasher B. , Dai H. , Luo T. And Nakagawa M. \textquotedbl;Using Sequential and Non-Sequential Patterns in PredictiveWeb Usage Mining Tasks\textquotedbl;, Proceeding ICDM \texttt{amp}apos;02 Proceedings of the 2002 IEEE International Conference on Data Mining, IEEE Computer Society Washington, DC, USA \texttt{amp}apos;02.
- Esmin A. , Lima J. , Yano,Tiago E. T. And Carneiro G. S. \textquotedbl;ArchCollect - A Tool for WEB Usage Knowledge Acquisition from User\texttt{amp}apos;s Interactions\textquotedbl;, Proceedings of the Tenth International Conference on Enterprise Information Systems, Barcelona, Spain, pp. 375-380,2008.
- Shahabi C. , Faisal A. , Kashani F. B. and Faruque J. \textquotedbl;INSITE: A Tool for Real-Time Knowledge Discovery from Users Web Navigation\textquotedbl;, Proceedings of the 26th International Conference on Very Large Databases (VLDB), Cairo, Egypt, pp. 635-638,2000.
- Berendt B. \textquotedbl;Web usage mining, site semantics, and the support of navigation\textquotedbl;, KDD Workshop on Web Mining for ECommerce Challenges and Opportunities\textquotedbl; pp. 83–93,2000.
- Masseglia F. , Poncelet P. and Cicchetti R. \textquotedbl;WebTool: An Integrated Framework

- Cooley R., Tan P. N. and Srivastava J., ”WebSIFT: The Web Site Information Filter System”, Proceedings of Workshop on Web Usage Analysis and User Profiling WEBKDD in conjunction with ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 1999, San Diego, California, USA.


Index Terms

Computer Science
Information Sciences

Keywords

Web mining, web content mining, web structure mining, web usage mining, information retrieval, information extraction