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 through 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


- Dzitac I. and Moisil I. "Advanced AI Techniques for Web Mining; Proceeding of the 10th WSEAS international conference on Mathematical methods, computational techniques and intelligent systems, 2008.
- Chang C., Lui S., Wu Y. "Applying Pattern Mining to Web Information Extraction; Advances in Knowledge Discovery and Data, 2001 – Springer.
- Yusifov F. F. "Web Traffic Mining using Neural Networks; World Academy of Science, Engineering and Technology 21 2008
- Eirinaki M., Vazirgiannis M. "Web Mining for Web Personalization; ACM Transactions on Internet Technology (TOIT), Volume 3 Issue 1, February 2003.
- Fu Y., Shih M., Creado M., Ju1 C. "Reorganizing Web Sites Based on User Access Patterns; Proceeding of the tenth international conference on Information and knowledge management, Pages 583 – 585, 2001.
- Jain D., Sinhal A., Gupta N., Narwariya P., Sarawat D. and Pandey A. "Hiding Sensitive Association Rules without Altering the Support of Sensitive Item(s);
A Hand to Hand Taxonomical Survey on Web Mining


- Fayyad U., "From data mining to knowledge discovery: An overview", Advances in Knowledge Discovery and Data Mining, pp. 1-34, AAAI Press, 1996.
Search Results Based on Web Content Mining Techniques," wi, pp. 301-304, 2006
IEEE/WIC/ACM International Conference on Web Intelligence (WI’06), 2006.
- Hosseini M., Hassani A. H., "Mining Search Engine Query Log for Evaluating
Content and Structure of a Web Site"; in Proceedings of the 2007 IEEE/WIC/ACM
International Conference on Web Intelligence.
Content-Based Image Retrieval"; IEEE Transactions on Pattern Analysis and Machine
Intelligence, vol. 30, no. 11, pp. 1902-1912, Nov. 2008
- Poonkuzhali G., Thiagarajan K., "Signed Approach for Mining Web Content
Outliers"; World Academy of Science, Engineering and Technology 56, 2009
- Madria K. S., Bhowmick S. S., Ng K. W., and Lim E., "Research Issues in Web
Data Mining"; Proceedings of the First International Conference on Data Warehousing and
Knowledge Discovery, p. 303-312, September 01, 1999
- Kleinberg J. K., "Authoritative sources in a hyperlinked environment"; Journal
- Furnkranz J., "Web structure mining — Exploiting the graph structure of the
worldwide web"; OGAI-J. 21(2) (2002) 17–26
- Smith A. K. and Ng A., "Web page clustering using a self-organizing map of user
- Fang X. and Sheng O., "LinkSelector: A web mining approach to hyperlink
- Hay B., Wets G., and Vanhoof K., "Mining navigation patterns using a sequence
alignment method"; Knowledge Inform. Syst. 6(2) (2004) 150–163
- Guan S. and McMullen P., "Organizing information on the next generation web
—design and implementation of a new bookmark structure"; International Journal Inform
- Song Q. and Shepperd M., "Mining web browsing patterns for e-commerce";
- Chikhi F. N., Rothenburger B. and Aussenac-Gilles N., "A Comparison of
Dimensionality Reduction Techniques for Web Structure Mining"; Proceedings of the
- Moussiades L. and Nakali A., "Mining the Community Structure of a Web
Site"; bci, pp. 239-244, 2009 Fourth Balkan Conference in Informatics, 2009
- Zhu J. and Hong J., "Using Markov Models for Web Site Link Prediction";
College Park, Maryland, USA ACM June 11-15, 2002
- Borges and Levene M., "A dynamic clustering-based markov model for web usage
Mining"; cs. IR/0406032, 2004
- 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)
- Tao H. Y., Hong P. T., and Su M. Y., "Web usage mining with intentional browsing
- Jalali M. and Mustapha N., "A Recommender System Approach for Classifying
User Navigation Patterns Using Longest Common Subsequence Algorithm"; American
- Jalali M. and Mustapha N., "WebPUM: A Web-based recommendation system to
- Cooley R. , Tan P. N. and Srivastava J. ,&quot;WebSIFT: The Web Site Information Filter System&quot;, 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

6 / 7