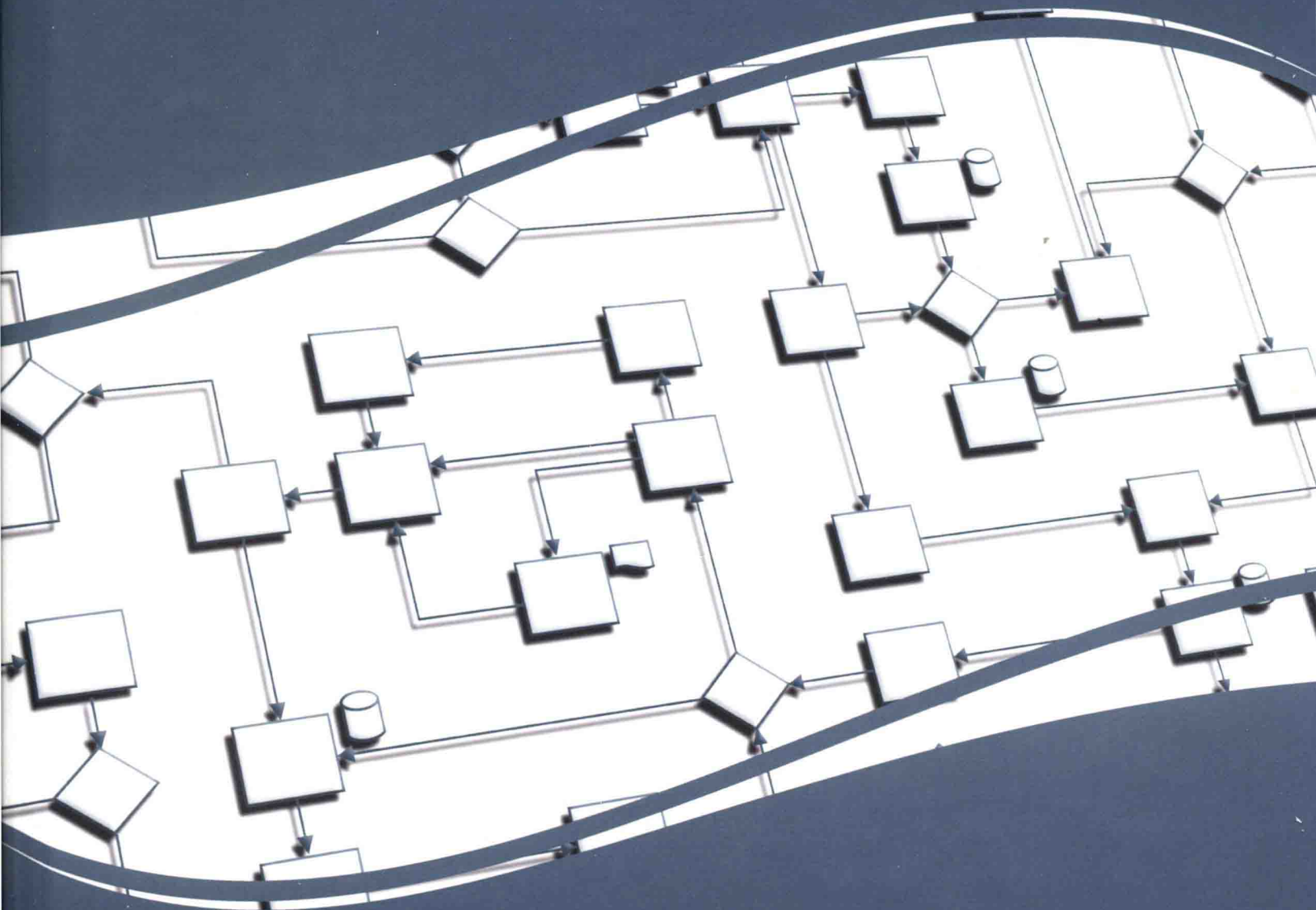


Web Data Mining and the Development of Knowledge-Based Decision Support Systems



G. Sreedhar

Web Data Mining and the Development of Knowledge-Based Decision Support Systems

G. Sreedhar

Rashtriya Sanskrit Vidyapeetha (Deemed University), India

A volume in the Advances in Data Mining and
Database Management (ADMDM) Book Series



www.igi-global.com

Published in the United States of America by

IGI Global
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: <http://www.igi-global.com>

Copyright © 2017 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

CIP Data Pending

ISBN: 978-1-5225-1877-8

eISBN: 978-1-5225-1878-5

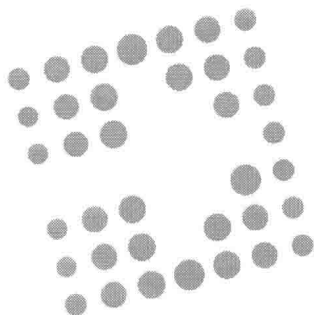
This book is published in the IGI Global book series Advances in Data Mining and Database Management (ADMMDM) (ISSN: 2327-1981; eISSN: 2327-199X)

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: eresources@igi-global.com.



Advances in Data Mining and Database Management (ADMDM) Book Series

David Taniar
Monash University, Australia

ISSN:2327-1981
EISSN:2327-199X

MISSION

With the large amounts of information available to organizations in today's digital world, there is a need for continual research surrounding emerging methods and tools for collecting, analyzing, and storing data.

The **Advances in Data Mining & Database Management (ADMDM)** series aims to bring together research in information retrieval, data analysis, data warehousing, and related areas in order to become an ideal resource for those working and studying in these fields. IT professionals, software engineers, academicians and upper-level students will find titles within the ADMDM book series particularly useful for staying up-to-date on emerging research, theories, and applications in the fields of data mining and database management.

COVERAGE

- Customer Analytics
- Profiling Practices
- Educational Data Mining
- Web-based information systems
- Data Warehousing
- Data Analysis
- Database Security
- Web mining
- Text Mining
- Data Mining

IGI Global is currently accepting manuscripts for publication within this series. To submit a proposal for a volume in this series, please contact our Acquisition Editors at Acquisitions@igi-global.com or visit: <http://www.igi-global.com/publish/>.

The Advances in Data Mining and Database Management (ADMDM) Book Series (ISSN 2327-1981) is published by IGI Global, 701 E. Chocolate Avenue, Hershey, PA 17033-1240, USA, www.igi-global.com. This series is composed of titles available for purchase individually; each title is edited to be contextually exclusive from any other title within the series. For pricing and ordering information please visit <http://www.igi-global.com/book-series/advances-data-mining-database-management/37146>. Postmaster: Send all address changes to above address. Copyright © 2017 IGI Global. All rights, including translation in other languages reserved by the publisher. No part of this series may be reproduced or used in any form or by any means – graphics, electronic, or mechanical, including photocopying, recording, taping, or information and retrieval systems – without written permission from the publisher, except for non commercial, educational use, including classroom teaching purposes. The views expressed in this series are those of the authors, but not necessarily of IGI Global.

Titles in this Series

For a list of additional titles in this series, please visit: www.igi-global.com

Emerging Trends in the Development and Application of Composite Indicators

Veljko Jeremic (University of Belgrade, Serbia) Zoran Radojicic (University of Belgrade, Serbia) and Marina Dobrota (University of Belgrade, Serbia)

Information Science Reference • copyright 2017 • 402pp • H/C (ISBN: 9781522507147) • US \$205.00 (our price)

Web Usage Mining Techniques and Applications Across Industries

A.V. Senthil Kumar (Hindusthan College of Arts and Science, India)

Information Science Reference • copyright 2017 • 424pp • H/C (ISBN: 9781522506133) • US \$200.00 (our price)

Social Media Data Extraction and Content Analysis

Shalin Hai-Jew (Kansas State University, USA)

Information Science Reference • copyright 2017 • 493pp • H/C (ISBN: 9781522506485) • US \$225.00 (our price)

Collaborative Filtering Using Data Mining and Analysis

Vishal Bhatnagar (Ambedkar Institute of Advanced Communication Technologies and Research, India)

Information Science Reference • copyright 2017 • 309pp • H/C (ISBN: 9781522504894) • US \$195.00 (our price)

Effective Big Data Management and Opportunities for Implementation

Manoj Kumar Singh (Adama Science and Technology University, Ethiopia) and Dileep Kumar G. (Adama Science and Technology University, Ethiopia)

Information Science Reference • copyright 2016 • 324pp • H/C (ISBN: 9781522501824) • US \$195.00 (our price)

Data Mining Trends and Applications in Criminal Science and Investigations

Omowunmi E. Isafiade (University of Cape Town, South Africa) and Antoine B. Bagula (University of the Western Cape, South Africa)

Information Science Reference • copyright 2016 • 386pp • H/C (ISBN: 9781522504634) • US \$210.00 (our price)

Intelligent Techniques for Data Analysis in Diverse Settings

Numan Celebi (Sakarya University, Turkey)

Information Science Reference • copyright 2016 • 353pp • H/C (ISBN: 9781522500759) • US \$195.00 (our price)

Managing and Processing Big Data in Cloud Computing

Rajkumar Kannan (King Faisal University, Saudi Arabia) Raihan Ur Rasool (King Faisal University, Saudi Arabia) Hai Jin (Huazhong University of Science and Technology, China) and S.R. Balasundaram (National Institute of Technology, Tiruchirappalli, India)

Information Science Reference • copyright 2016 • 307pp • H/C (ISBN: 9781466697676) • US \$200.00 (our price)



www.igi-global.com

701 E. Chocolate Ave., Hershey, PA 17033

Order online at www.igi-global.com or call 717-533-8845 x100

To place a standing order for titles released in this series, contact: cust@igi-global.com
Mon-Fri 8:00 am - 5:00 pm (est) or fax 24 hours a day 717-533-8661

Acknowledgment

I am very much happy and thankful to IGI global for giving me the opportunity to produce my second *Web Data Mining and the Development of Knowledge-Based Decision Support Systems*, which is very much necessary in the present internet world. I express a deep sense of gratitude to the Jan Travers, David J Perreault, Eleana Wehr, Joshua Herring, Kayla Wolfe, and other members of IGI global who supported either directly or indirectly during book project development.

I am thankful to all authors who contributed their valuable efforts and ideas in the form of chapters in the book. I would like to express my sincere thanks to Prof. A. Anandaraja Chari for his continuous support and encouragement in my career.

I am thankful to vice chancellor, registrar and other faculty members of Rashtriya Sanskrit Vidyapeetha, Tirupati for their support and kindness in completing the book.

Finally, I am very much thankful to my parents and family members for their support and encouragement in achieving this target goal in my academic career.

Section 1

Developing Decision Support Systems

Editorial Advisory Board

A. Anandaraja Chari, *Rayalaseema University, India*
N. Kasiviswanath, *G. Pulla Reddy Engineering College, India*
A. V. Krishna Prasad, *K. L. University, India*
B. Eswara Reddy, *JNTUA College of Engineering, India*
R. J. Rama Sree, *Rashtriya Sanskrit Vidyapeetha (Deemed University), India*

List of Reviewers

Siva Prasad Akula, *Dr. V. S. Krishna Government Degree and PG College, India*
Madana Kumar Reddy C, *Annamacharya P. G. College of Computer Studies, India*
Suruchi Chawla, *Shaheed Rajguru College Delhi University, India*
Kijpokin Kasemsap, *Suan Sunandha Rajabhat University, Thailand*
Vudattu Kiran Kumar, *Dravidian University, India*
Varaprasad Rao M, *Anurag Group of Institutions, India*
Wing Shui Ng, *The Hong Kong Institute of Education, India*
Adiraju Prasantha Rao, *Anurag Group of Institutions, India*
P. Thendral, *VIT University, India*
Madhu Viswanatham Vankadara, *VIT University, India*
Carlos Ortiz Zezzatti, *Jaurez City University, Mexico*

Table of Contents

Preface	xvi
Acknowledgment	xvii

Section 1 **Developing Decision Support Systems**

Chapter 1

Development of Efficient Decision Support System Using Web Data Mining	1
<i>G. Sreedhar, Rashtriya Sanskrit Vidyapeetha (Deemed University), India</i>	
<i>A. Anandaraja Chari, Rayalaseema University, India</i>	

Chapter 2

Improving Decision-Making in a Business Simulator Using TOPSIS Methodology for the Establishment of Reactive Stratagem.....	12
<i>Carlos Alberto Ochoa Ortiz Zezzatti, Juarez City University, Mexico</i>	

Chapter 3

DSS for Web Mining Using Recommendation System	22
<i>Varaprasad Rao M, Anurag Group of Institutions, India</i>	
<i>Vishnu Murthy G, Anurag Group of Institutions, India</i>	

Chapter 4

Mastering Intelligent Decision Support Systems in Enterprise Information Management	35
<i>Kijpokin Kasemsap, Suan Sunandha Rajabhat University, Thailand</i>	

Section 2 **Developing Web Mining Systems**

Chapter 5

Web Data Mining in Education: Decision Support by Learning Analytics with Bloom's Taxonomy.....	58
<i>Wing Shui Ng, The Education University of Hong Kong, Hong Kong</i>	

Chapter 6

Discover Patterns from Web-Based Dataset	78
--	----

Raghvendra Kumar, LNCT College, India

Priyanka Pandey, LNCT College, India

Prasant Kumar Pattnaik, KIIT University, India

Chapter 7

Web Usage Mining: Improving the Performance of Web-Based Application through Web Mining.....	107
--	-----

Sathiyamoorthi V, Sona College of Technology, India

Chapter 8

Multi-Agent-Based Information Retrieval System Using Information Scent in Query Log Mining for Effective Web Search.....	131
--	-----

Suruchi Chawla, Shaheed Rajguru College Delhi University, India

Section 3

Developing Efficient Knowledge based Systems

Chapter 9

Knowledge Representation Technologies Using Semantic Web	158
--	-----

Vudattu Kiran Kumar, Dravidian University, India

Chapter 10

Discovering Knowledge Hidden in Big Data from Machine-Learning Techniques	167
---	-----

Adiraju Prashantha Rao, Anurag Group of Institutions, India

Chapter 11

Knowledge-Based Decision Support System for Analyzing the Relevancies of Various Attributes Based on Their Characteristics	184
--	-----

Madana Kumar Reddy C, Annamacharya PG College of Computer Studies, India

Chapter 12

Website Topology Modification with Hotlinks Using Mined Webusage Knowledge	194
--	-----

Thendral Puyalnithi, VIT University, India

Madhu Viswanatham V, VIT University, India

Section 4

Developing Social Media based Mining Systems

Chapter 13

Mining on Social Media.....	206
-----------------------------	-----

Ambati Venkata Krishna Prasad, KL University, India

Venkata Naresh Mandhala, KL University, India

Chapter 14

Some Other Applications in Community Graph under the Preview of Social Graph Using Graph-Mining Techniques..... 223

Bapuji Rao, iNurture Education Solutions Private Limited, India

Sasmita Mishra, IGIT, India

Saroja Nanda Mishra, IGIT, India

Chapter 15

Social Network Web Mining: Web Mining Techniques for Online Social Network Analysis 284

Balamurugan Balusamy, VIT University, India

Vegesna Tarun Sai Varma, VIT University, India

Sohil Sri Mani Yeshwanth Grandhi, VIT University, India

Section 5

Developing Data and Text Mining Systems

Chapter 16

Data Mining and Data Warehousing: Introduction to Data Mining and Data Warehousing 312

Sathiyamoorthi V, Sona College of Technology, India

Chapter 17

Text Mining: Current Trends and Applications 338

Kijpokin Kasemsap, Suan Sunandha Rajabhat University, Thailand

Compilation of References 359

About the Contributors 402

Index 407

Detailed Table of Contents

Preface.....	xvi
--------------	-----

Acknowledgment	xvii
----------------------	------

Section 1 Developing Decision Support Systems

Chapter 1

Development of Efficient Decision Support System Using Web Data Mining	1
<i>G. Sreedhar, Rashtriya Sanskrit Vidyapeetha (Deemed University), India</i>	
<i>A. Anandaraja Chari, Rayalaseema University, India</i>	

The management of web sites imposes a constant demand for new information and timely updates due to the increase of services and content that site owners wish to make available to their users, which in turn is motivated by the complexity and diversity of needs and behaviours of the users. Such constant labour intensive effort implies very high financial and personnel costs. The growth of World Wide Web and technologies has made business functions to be executed fast and easier. E-commerce has provided a cost efficient and effective way of doing business. Web mining is usually defined as the use of data mining techniques to automatically discover and extract information from web documents and services. Also, web data mining is commonly categorized into three areas: web content mining that describes the discovery of useful information from content, web structure mining that analyses the topology of web sites, and web usage mining that tries to make sense of the data generated by the navigation behaviour and user profile.

Chapter 2

Improving Decision-Making in a Business Simulator Using TOPSIS Methodology for the Establishment of Reactive Stratagem.....	12
<i>Carlos Alberto Ochoa Ortiz Zezzatti, Juarez City University, Mexico</i>	

This study combines Fuzzy Logic and multicriteria TOPSIS method for the selection, from three different alternatives, which machines of high productivity is more convenient to a construction company. The evaluation of each alternative is made through group decision making which identifies the most important criteria according to the requirements presented by the company. To assess the selected criteria in the TOPSIS method is weighted by a group of experts who, based on their experience and knowledge of this type of machinery, assess the relevance of these in the operation and functioning of the hydraulic excavator. Both qualitative and quantitative studies are used in this work, however the experts evaluate,

through surveys based on Likert scale all the criteria in which they want to measure the perception. Data provided from the surveys is used for the construction and association of the groups of expert's opinion through the use of fuzzy sets to avoid ambiguity problems of the linguistic variables.

Chapter 3

DSS for Web Mining Using Recommendation System	22
<i>Varaprasad Rao M, Anurag Group of Institutions, India</i>	
<i>Vishnu Murthy G, Anurag Group of Institutions, India</i>	

Decision Supports Systems (DSS) are computer-based information systems designed to help managers to select one of the many alternative solutions to a problem. A DSS is an interactive computer based information system with an organized collection of models, people, procedures, software, databases, telecommunication, and devices, which helps decision makers to solve unstructured or semi-structured business problems. Web mining is the application of data mining techniques to discover patterns from the World Wide Web. Web mining can be divided into three different types – Web usage mining, Web content mining and Web structure mining. Recommender systems (RS) aim to capture the user behavior by suggesting/recommending users with relevant items or services that they find interesting in. Recommender systems have gained prominence in the field of information technology, e-commerce, etc., by inferring personalized recommendations by effectively pruning from a universal set of choices that directed users to identify content of interest.

Chapter 4

Mastering Intelligent Decision Support Systems in Enterprise Information Management	35
<i>Kijpokin Kasemsap, Suan Sunandha Rajabhat University, Thailand</i>	

This chapter explains the overview of Intelligent Decision Support Systems (IDSSs); the overview of Enterprise Information Management (EIM); the IDSS techniques for EIM in terms of Expert System (ES), Multi-Agent System (MAS), Fuzzy Logic (FL), Artificial Neural Network (ANN), Evolutionary Computation (EC), and Hybrid System (HS); and the multifaceted applications of IDSSs in EIM. IDSS techniques are rapidly emerging as the modern tools in information management systems and include various techniques, such as ES, MAS, FL, ANN, EC, and HS. IDSS techniques can increase the sensitiveness, flexibility, and accuracy of information management systems. IDSS techniques should be implemented in modern enterprise in order to gain the benefits of using the decision-making process concerning EIM. The chapter argues that utilizing IDSS techniques for EIM has the potential to increase organizational performance and reach strategic goals in global operations.

Section 2 **Developing Web Mining Systems**

Chapter 5

Web Data Mining in Education: Decision Support by Learning Analytics with Bloom's Taxonomy	58
<i>Wing Shui Ng, The Education University of Hong Kong, Hong Kong</i>	

Web data mining for extracting meaningful information from large amount of web data has been explored over a decade. The concepts and techniques have been borrowed into the education sector and the new research discipline of learning analytics has emerged. With the development of web technologies, it has

been a common practice to design online collaborative learning activities to enhance learning. To apply learning analytics techniques to monitor the online collaborative process enables a lecturer to make instant and informed pedagogical decisions. However, it is still a challenge to build strong connection between learning analytics and learning science for understanding cognitive progression in learning. In this connection, this chapter reports a study to apply learning analytics techniques in the aspect of web usage mining and clustering analysis with underpinning Bloom's taxonomy to analyze students' performance in the online collaborative learning process. The impacts of intermediate interventions are also elaborated.

Chapter 6

Discover Patterns from Web-Based Dataset 78

Raghvendra Kumar, LNCT College, India

Priyanka Pandey, LNCT College, India

Prasant Kumar Pattnaik, KIIT University, India

The Web can be defined as a depot of varied range of information present in the form of millions of websites dispersed around us. Often users find it difficult to locate the appropriate information fulfilling their needs with the abundant number of websites in the Web. Hence multiple research work has been conducted in the field of Web Mining so as to present any information matching the user's needs. The application of data mining techniques on web usage, web content or web structure data to find out useful data like users' way in patterns and website utility statistics on a whole can be defined as Web mining. The main cause behind development of such websites was to personalize the substance of a website on user's preference. New methods are developed to deal with a Web site using a link hierarchy and a conceptual link hierarchy respectively on the basis of how users have used the Web site link structure.

Chapter 7

Web Usage Mining: Improving the Performance of Web-Based Application through Web Mining..... 107

Sathiyamoorthi V, Sona College of Technology, India

In recent days, Internet technology has provided a lot of services for sharing and distributing information across the world. Among all the services, World Wide Web (WWW) plays a significant role. The slow retrieval of Web pages may lessen the interest of users from accessing them. To deal with this problem, Web caching and Web pre-fetching are the two techniques used. Web proxy caching plays a key role in improving Web performance by keeping Web objects that are likely to be used in the near future in the proxy server which is closer to the end user. It helps in reducing user perceived latency, network bandwidth utilization, and alleviating loads on the Web servers. Thus, it improves the efficiency and scalability of Web based system. This chapter gives an overview of Web usage mining and its application on Web and discusses various approaches for improving the performance of Web.

Chapter 8

Multi-Agent-Based Information Retrieval System Using Information Scent in Query Log Mining for Effective Web Search..... 131

Suruchi Chawla, Shaheed Rajguru College Delhi University, India

This chapter explains the multi-agent system for effective information retrieval using information scent in query log mining. The precision of search results is low due to difficult to infer the information need

of the small size search query and therefore information need of the user is not satisfied effectively. Information Scent is used for modeling the information need of user web search session and clustering is performed to identify the similar information need sessions. Hyper Link-Induced Topic Search (HITS) is executed on clusters to generate the Hubs and authorities for web page recommendations to users who search with similar intents. This multi-agent system based on clustered query sessions uses query operations like expansion and recommendation to infer the information need of user search queries and recommends Hubs and authorities for effective web search.

Section 3 Developing Efficient Knowledge based Systems

Chapter 9

Knowledge Representation Technologies Using Semantic Web	158
<i>Vudattu Kiran Kumar, Dravidian University, India</i>	

The World Wide Web (WWW) is global information medium, where users can read and write using computers over internet. Web is one of the services available on internet. The Web was created in 1989 by Sir Tim Berners-Lee. Since then a great refinement has done in the web usage and development of its applications. Semantic Web Technologies enable machines to interpret data published in a machine-interpretable form on the web. Semantic web is not a separate web it is an extension to the current web with additional semantics. Semantic technologies play a crucial role to provide data understandable to machines. To achieve machine understandable, we should add semantics to existing websites. With additional semantics, we can achieve next level web where knowledge repositories are available for better understanding of web data. This facilitates better search, accurate filtering and intelligent retrieval of data. This paper discusses about the Semantic Web and languages involved in describing documents in machine understandable format.

Chapter 10

Discovering Knowledge Hidden in Big Data from Machine-Learning Techniques	167
<i>Adiraju Prashantha Rao, Anurag Group of Institutions, India</i>	

As the speed of information growth exceeds in this new century, excessive data is making great troubles to human beings. However, there are so much potential and highly useful values hidden in the huge volume of data. Big Data has drawn huge attention from researchers in information sciences, policy and decision makers in governments and enterprises. Data analytic is the science of examining raw data with the purpose of drawing conclusions about that information. Data analytics is about discovering knowledge from large volumes data and applying it to the business. Machine learning is ideal for exploiting the opportunities hidden in big data. This chapter able to discover and display the patterns buried in the data using machine learning.

Chapter 11

Knowledge-Based Decision Support System for Analyzing the Relevancies of Various Attributes Based on Their Characteristics	184
<i>Madana Kumar Reddy C, Annamacharya PG College of Computer Studies, India</i>	

Data mining extracts novel and useful knowledge from large repositories of data and has become an effective analysis and decision means in any organization. The resource of the World Wide Web is almost infinite. The growing importance of electronic media for storing and disseminating text documents has

created an urgent need for tools and techniques that assist users in finding and extracting relevant and previously unknown information from massive collection of documents available in the web. Thus the development of techniques for mining unstructured, semi-structured, and fully structured textual data has become quite important in both academia and industry. Information management of well organized databases has been a focus of the Data mining research. When to specify too many attributes, system will slow down thus exclude irrelevant or weakly relevant attributes. The general idea behind attribute relevance analysis is to compute some measure that is used to quantify the relevance of an attribute with respect to a given class or concept.

Chapter 12

Website Topology Modification with Hotlinks Using Mined Webusage Knowledge 194

Thendral Puyalnithi, VIT University, India

Madhu Viswanatham V, VIT University, India

The hotlinks are the special links introduced in the website to reduce the time to access certain webpages in a webpage that is present in the deeper levels of the topology. Hotlinks selection mechanism plays a vital role in quick access of webpages. The problem is to decide which webpage should be having hotlinks and where the hotlinks should be placed in the website tree topology. We have proposed a methodology which starts by finding the frequent webpage access pattern of visitors of the website. The frequent pattern is found using Associative mining, Apriori algorithm or Frequent Pattern Tree algorithm. Then the frequent patterns are passed through page ranking mechanism. We find the pattern which is having the highest priority. Then the hotlinks are created for the members (webpages hyperlinks) of the pattern. Thus, the work is about assigning hotlinks for a set of pages which are frequently visited. Thus, by updating the topology by introducing hotlinks we can reduce the time to access the web pages.

Section 4

Developing Social Media based Mining Systems

Chapter 13

Mining on Social Media..... 206

Ambati Venkata Krishna Prasad, KL University, India

Venkata Naresh Mandhala, KL University, India

Social media mining is the process of representing, analyzing, and extracting actionable patterns and trends from raw social media data. Social media is favored by many users since it is available to individuals without any limitations to share their opinions, educational learning experiences and concerns via their status. Twitter API, twitter4j, is processed for searching the tweets based on the geo location. Student's posts on social network offers us a stronger concern to take decisions concerning the particular education system's learning method of the system. Evaluating knowledge in social media is sort of a difficult method. Bayes classifier are enforced on deep-mined knowledge for analysis purpose to urge the deeper understanding of the information. It uses multi label classification technique as every label falls into completely different classes. Label based measures are mostly taken to research the results and comparing them with the prevailing sentiment analysis technique.

Chapter 14

Some Other Applications in Community Graph under the Preview of Social Graph Using Graph-Mining Techniques..... 223

Bapuji Rao, iNurture Education Solutions Private Limited, India

Sasmita Mishra, IGIT, India

Saroja Nanda Mishra, IGIT, India

The retrieval of sub-graph from a large graph in structured data mining is one of the fundamental tasks for analyze. Visualization and analyze large community graph are challenging day by day. Since a large community graph is very difficult to visualize, so compression is essential. To study a large community graph, compression technique may be used for compression of community graph. There should not be any loss of information or knowledge while compressing the community graph. Similarly to extract desired knowledge of a particular sub-graph from a large community graph, then the large community graph needs to be partitioned into smaller sub-community graphs. The partition aims at the edges among the community members of dissimilar communities in a community graph. Sometimes it is essential to compare two community graphs for similarity which makes easier for mining the reliable knowledge from a large community graph. Once the similarity is done then the necessary mining of knowledge can be extracted from only one community graph rather than from both which leads saving of time.

Chapter 15

Social Network Web Mining: Web Mining Techniques for Online Social Network Analysis 284

Balamurugan Balusamy, VIT University, India

Vegetna Tarun Sai Varma, VIT University, India

Sohil Sri Mani Yeshwanth Grandhi, VIT University, India

Today, social networks are major part of everyone's lives. They provide means to communicate with people across the globe with ease. As of July 2016, there are over 1.71 billion monthly active Facebook users. They generate significant amount of data, which if analysed well will provide us with valuable information. This can be done by analysing the log data collected at the respective social networking service. This chapter focuses on extraction and analysis of Facebook data since it is presently the most used social network. The result of analysis can be used in building decision support systems for an organization to help with the decision making process.

Section 5

Developing Data and Text Mining Systems

Chapter 16

Data Mining and Data Warehousing: Introduction to Data Mining and Data Warehousing 312

Sathiyamoorthi V, Sona College of Technology, India

It is generally observed throughout the world that in the last two decades, while the average speed of computers has almost doubled in a span of around eighteen months, the average speed of the network has doubled merely in a span of just eight months! In order to improve the performance, more and more researchers are focusing their research in the field of computers and its related technologies. Data Mining is one such research area. It extracts useful information the huge amount of data present in the database. The discovered knowledge can be applied in various application areas such as marketing, fraud detections and customer retention. It discovers implicit, previously unknown and potentially useful information out of datasets. Recent trend in data mining include web mining where it discover knowledge from web based information to improve the page layout, structure and its content.