Table of Contents

Foreword ................................................................. xxxiii
Preface ............................................................................. xxxiv
Acknowledgment .............................................................. xxxv

Volume I

Section I
Document Preprocessing

Chapter I
On Document Representation and Term Weights in Text Classification ........................................ 1
Ying Liu, The Hong Kong Polytechnic University Hong Kong SAR, China

Chapter II
Deriving Document Keyphrases for Text Mining ................................................................. 23
Yi-fang Brook Wu, New Jersey Institute of Technology, USA
Quanzhi Li, Avaya, Inc., USA

Chapter III
Intelligent Text Mining: Putting Evolutionary Methods and Language Technologies Together .......... 37
John Atkinson, Universidad de Concepción, Chile

Section II
Classification and Clustering

Chapter IV
Automatic Syllabus Classification Using Support Vector Machines ............................................. 61
Xiaoyan Yu, Virginia Tech, USA
Manas Tungare, Virginia Tech, USA
Weigo Yuan, Virginia Tech, USA
Yubo Yuan, Virginia Tech, USA
Manuel Pérez-Quiñones, Virginia Tech, USA
Edward A. Fox, Virginia Tech, USA
Chapter V
Partially Supervised Text Categorization

Xiao-Li Li, Institute for Infocomm Research, Singapore

Chapter VI
Image Classification and Retrieval with Mining Technologies

Yu-Jin Zhang, Tsinghua University, Beijing, China

Chapter VII
Improving Techniques for Naïve Bayes Text Classifiers

Han-joon Kim, University of Seoul, Korea

Chapter VIII
Using the Text Categorization Framework for Protein Classification

Ricco Rakotomalala, University of Lyon 2, France
Faouzi Mhamdi, University of Jandouba, Tunisia

Chapter IX
Featureless Data Clustering

Wilson Wong, University of Western Australia, Australia
Wei Liu, University of Western Australia, Australia
Mohammed Bennamoun, University of Western Australia, Australia

Chapter X
Swarm Intelligence in Text Document Clustering

Xiaohui Cui, Oak Ridge National Laboratory, USA
Thomas E. Potok, Oak Ridge National Laboratory, USA

Chapter XI
Some Efficient and Fast Approaches to Document Clustering

P. Viswanth, Indian Institute of Technology Guwahati, India
Bidyut Kr. Patra, Indian Institute of Technology Guwahati, India
V. Suresh Babu, Indian Institute of Technology Guwahati, India

Chapter XII
SOM-Based Clustering of Textual Documents Using WordNet

Abdelmalek Amine, Djillali Liabes University, Alergia & Taher Moulay University Center, Algeria
Zakaria Elberrichi, Djillali Liabes University, Algeria
Michel Simonet, Joseph Fourier University, France
Ladjel Bellatreche, University of Poitiers, France
Mimoun Malki, Djillali Liabes University, Algeria
Chapter XIII
A Multi-Agent Neural Network System for Web Text Mining ............................................................. 201

Lean Yu, Chinese Academy of Sciences, China & City University of Hong Kong, China
Shouyang Wang, Chinese Academy of Sciences, China
Kin Keung Lai, City University of Hong Kong, China

Section III
Database, Ontology, and the Web

Chapter XIV
Frequent Mining on XML Documents ................................................................................................. 227

Sangeetha Kutty, Queensland University of Technology, Australia
Richi Nayak, Queensland University of Technology, Australia

Chapter XV
The Process and Application of XML Data Mining ................................................................................... 249

Richi Nayak, Queensland University of Technology, Australia

Chapter XVI
Approximate Range Querying over Sliding Windows ........................................................................ 273

Francesco Buccafurri, University “Mediterranea” of Reggio Calabria, Italy
Gianluca Caminiti, University “Mediterranea” of Reggio Calabria, Italy
Gianluca Lax, University “Mediterranea” of Reggio Calabria, Italy

Chapter XVII
Slicing and Dicing a Linguistic Data Cube ......................................................................................... 288

Jan H. Kroeze, University of Pretoria, South Africa
Theo J. D. Bothma, University of Pretoria, South Africa
Machdel C. Matthee, University of Pretoria, South Africa

Chapter XVIII
Discovering Personalized Novel Knowledge from Text ........................................................................ 301

Yi-fang Brook Wu, New Jersey Institute of Technology, USA
Xin Chen, Microsoft Corporation, USA

Chapter XIX
Untangling BioOntologies for Mining Biomedical Information .......................................................... 314

Catia Pesquita, University of Lisbon, Portugal
Daniel Faria, University of Lisbon, Portugal
Tiago Grego, University of Lisbon, Portugal
Francisco M. Couto, University of Lisbon, Portugal
Mário J. Silva, University of Lisbon, Portugal
Chapter XX
Thesaurus-Based Automatic Indexing ................................................................. 331
   Luis M. de Campos, University of Granada, Spain
   Juan M. Fernández-Luna, University of Granada, Spain
   Juan F. Huete, University of Granada, Spain
   Alfonso E. Romero, University of Granada, Spain

Chapter XXI
Concept-Based Text Mining ................................................................................. 346
   Stanley Loh, Lutheran University of Brazil, Brazil
   Leandro Krug Wives, Federal University of Rio Grande do Sul, Brazil
   Daniel Lichtnow, Catholic University of Pelotas, Brazil
   José Palazzo M. de Oliveira, Federal University of Rio Grande do Sul, Brazil

Chapter XXII
Statistical Methods for User Profiling in Web Usage Mining ............................ 359
   Marcello Pecoraro, University of Naples Federico II, Italy
   Roberta Siciliano, University of Naples Federico II, Italy

Chapter XXIII
Web Mining to Identify People of Similar Background ...................................... 369
   Quanzhi Li, Avaya, Inc., USA
   Yi-fang Brook Wu, New Jersey Institute of Technology, USA

Chapter XXIV
Hyperlink Structure Inspired by Web Usage ..................................................... 386
   Pawan Lingras, Saint Mary’s University, Canada
   Rucha Lingras, Saint Mary’s University, Canada

Chapter XXV
Designing and Mining Web Applications: A Conceptual Modeling Approach ... 401
   Rosa Meo, Università di Torino, Italy
   Maristella Matera, Politecnico di Milano, Italy

Volume II

Chapter XXVI
Web Usage Mining for Ontology Management ............................................... 418
   Brigitte Trousse, INRIA Sophia Antipolois, France
   Marie-Aude Aufaure, INRIA Sophia and Supélec, France
   Bénédicte Le Grand, Laboratoire d’Informatique de Paris 6, France
   Yves Lechevallier, INRIA Rocquencourt, France
   Florent Masseglia, INRIA Sophia Antipolois, France
Chapter XXVII
A Lattice-Based Framework for Interactively and Incrementally Mining Web Traversal Patterns ................................................................. 448
Yue-Shi Lee, Ming Chuan University, Taiwan, ROC
Show-Jane Yen, Ming Chuan University, Taiwan, ROC

Chapter XXVIII
Privacy-Preserving Data Mining on the Web: Foundations and Techniques .................................................. 468
Stanley R. M. Oliveira, Embrapa Informática Agropecuária, Brazil
Osmar R. Zaïane, University of Alberta, Edmonton, Canada

Section IV
Information Retrieval and Extraction

Chapter XXIX
Automatic Reference Tracking ................................................................. 483
G.S. Mahalakshmi, Anna University, Chennai, India
S. Sendhilkumar, Anna University, Chennai, India

Chapter XXX
Determination of Unithood and Termhood for Term Recognition ................................................................. 500
Wilson Wong, University of Western Australia, Australia
Wei Liu, University of Western Australia, Australia
Mohammed Bennamoun, University of Western Australia, Australia

Chapter XXXI
Retrieving Non-Latin Information in a Latin Web: The Case of Greek .................................................. 530
Fotis Lazarinis, University of Sunderland, UK

Chapter XXXII
Latent Semantic Analysis and Beyond ................................................................. 546
Anne Kao, The Boeing Phantom Works, USA
Steve Poteet, The Boeing Phantom Works, USA
Jason Wu, The Boeing Phantom Works, USA
William Ferng, The Boeing Phantom Works, USA
Rod Tjoelker, The Boeing Phantom Works, USA
Lesley Quach, The Boeing Phantom Works, USA

Chapter XXXIII
Question Answering Using Word Associations ................................................................. 571
Ganesh Ramakrishnan, IBM India Research Labs, India
Pushpak Bhattacharyya, IIT Bombay, India
Chapter XXXIV
The Scent of a Newsgroup: Providing Personalized Access to Usenet Sites through Web Mining ................................................................. 604
   Giuseppe Manco, Italian National Research Council, Italy
   Riccardo Ortale, University of Calabria, Italy
   Andrea Tagarelli, University of Calabria, Italy

Section V
Application and Survey

Chapter XXXV
Text Mining in Program Code ............................................................... 626
   Alexander Dreweke, Friedrich-Alexander University Erlangen-Nuremberg, Germany
   Ingrid Fischer, University of Konstanz, Germany
   Tobias Werth, Friedrich-Alexander University Erlangen-Nuremberg, Germany
   Marc Wörlein, Friedrich-Alexander University Erlangen-Nuremberg, Germany

Chapter XXXVI
A Study of Friendship Networks and Blogosphere ..................................... 646
   Nitin Agarwal, Arizona State University, USA
   Huan Liu, Arizona State University, USA
   Jianping Zhang, MITRE Corporation, USA

Chapter XXXVII
An HL7-Aware Decision Support System for E-Health .................................. 670
   Pasquale De Meo, Università degli Studi Mediterranea di Reggio Calabria, Italy
   Giovanni Quattrone, Università degli Studi Mediterranea di Reggio Calabria, Italy
   Domenico Ursino, Università degli Studi Mediterranea di Reggio Calabria, Italy

Chapter XXXVIII
Multitarget Classifiers for Mining in Bioinformatics ..................................... 684
   Diego Liberati, Istituto di Elettronica e Ingegneria dell’Informazione e delle Telecomunicazioni Consiglio Nazionale delle Ricerche Politecnico di Milano, Italy

Chapter XXXIX
Current Issues and Future Analysis in Text Mining for Information Security Applications ............... 694
   Shuting Xu, Virginia State University, USA
   Xin Luo, Virginia State University, USA

Chapter XL
Collaborative Filtering Based Recommendation Systems .................................. 708
   E. Thirumaran, Indian Institute of Science, India
   M. Narasimha Murty, Indian Institute of Science, India
Chapter XLII
Performance Evaluation Measures for Text Mining .......................................................... 724
Hanna Suominen, Turku Centre for Computer Science (TUCS), Finland & University of Turku, Finland
Sampo Pyysalo, Turku Centre for Computer Science (TUCS), Finland & University of Turku, Finland
Marketta Hiissa, Turku Centre for Computer Science (TUCS), Finland & Åbo Akademi University, Finland
Filip Ginter, Turku Centre for Computer Science (TUCS), Finland & University of Turku, Finland
Shuhua Liu, Academy of Finland, Finland & Åbo Akademi University, Finland
Dorina Marghescu, Turku Centre for Computer Science (TUCS), Finland & Åbo Akademi University, Finland
Tapio Pahikkala, Turku Centre for Computer Science (TUCS), Finland & University of Turku, Finland
Barbro Back, Turku Centre for Computer Science (TUCS), Finland & Åbo Akademi University, Finland
Helena Karsten, Turku Centre for Computer Science (TUCS), Finland & Åbo Akademi University, Finland
Tapio Salakoski, Turku Centre for Computer Science (TUCS), Finland & University of Turku, Finland

Chapter XLII
Text Mining in Bioinformatics: Research and Application .................................................. 748
Yanliang Qi, New Jersey Institute of Technology, USA

Chapter XLIII
Literature Review in Computational Linguistics Issues in the Developing Field of Consumer Informatics: Finding the Right Information for Consumer’s Health Information Need .............. 758
Ki Jung Lee, Drexel University, USA

Chapter XLIV
A Survey of Selected Software Technologies for Text Mining ............................................ 766
Richard S. Segall, Arkansas State University, USA
Qingyu Zhang, Arkansas State University, USA

Chapter XLV
Application of Text Mining Methodologies to Health Insurance Schedules ....................... 785
Ah Chung Tsoi, Monash University, Australia
Phuong Kim To, Tedis P/L, Australia
Markus Hagenbuchner, University of Wollongong, Australia
Chapter XLVI
Web Mining System for Mobile-Phone Marketing ................................................................. 807
    Miao-Ling Wang, Minghsin University of Science & Technology, Taiwan, ROC
    Hsiao-Fan Wang, National Tsing Hua University, Taiwan, ROC

Chapter XLVII
Web Service Architectures for Text Mining: An Exploration of the Issues via an E-Science
Demonstrator ............................................................................................................................ 822
    Neil Davis, The University of Sheffield, UK
    George Demetriou, The University of Sheffield, UK
    Robert Gaizauskas, The University of Sheffield, UK
    Yikun Guo, The University of Sheffield, UK
    Ian Roberts, The University of Sheffield, UK