

# Table of Contents

Mining for Gems of Information .....	1
<i>Spiros Sirmakessis</i>	
From Text to Information: Document Processing and Visualization, a Text Mining Approach .....	7
<i>Martin Rajman, and Martin Vesely</i>	
Web Mining: The Past, the Present, and Future .....	25
<i>Konstantinos Markellos, Penelope Markellou, Maria Rigou, and Spiros Sirmakessis</i>	
Applications, Sectors and Strategies of Text Mining, a First Overall Picture.....	37
<i>Sergio Bolasco, Francesco Baiocchi, Alessio Canzonetti, Francesca Della Ratta, and Ana Feldman</i>	
Text Classification of News Articles with Support Vector Machines .....	53
<i>Gerhard Paaß, Joerg Kindermann, and Edda Leopold</i>	
A Review of Web Document Clustering Approaches .....	65
<i>N. Oikonomakou and M. Vazirgiannis</i>	
Supervised Term Weighting for Automated Text Categorization .....	81
<i>Franca Debole and Fabrizio Sebastiani</i>	
Machine Learning for Information Extraction in Genomics – State of the Art and Perspectives .....	99
<i>Claire Nédellec</i>	
Processing Multilingual Collection for Text Mining Applications .....	119
<i>Eric Gaussier</i>	
Text Mining Tools: Evaluation Methods and Criteria.....	131
<i>Antonis Spinakis and Paraskevi Peristera</i>	
Knowledge Advantage through online Text Mining. Research Trends in Competitive Intelligence and Virtual Communities Applications.....	151
<i>Alessandro Zanasi</i>	
Real Time Customer Opinion Monitoring.....	159
<i>Luca Dini and Giampaolo Mazzini</i>	

Validation Techniques in Text Mining (with Application to the Processing of Open-ended Questions).....	169
<i>Ludovic Lebart</i>	
Clickstream Analysis, Semiotic Interpretation and Semantic Text Mining for a Distance Measurement on the Hypertextual Map of an Internet-portal.....	179
<i>Furio Camillo</i>	
Text Mining in Official Statistic.....	189
<i>Mónica Becue, Bert Fridlund, Alf Fyhrlund, Albert Prat, and Bo Sundgren</i>	

Text Mining and its Applications  
Results of the NEMIS Launch Conference  
Sirmakessis, S. (Ed.)  
2004, X, 204 p. 24 illus., Hardcover  
ISBN: 978-3-540-20238-7