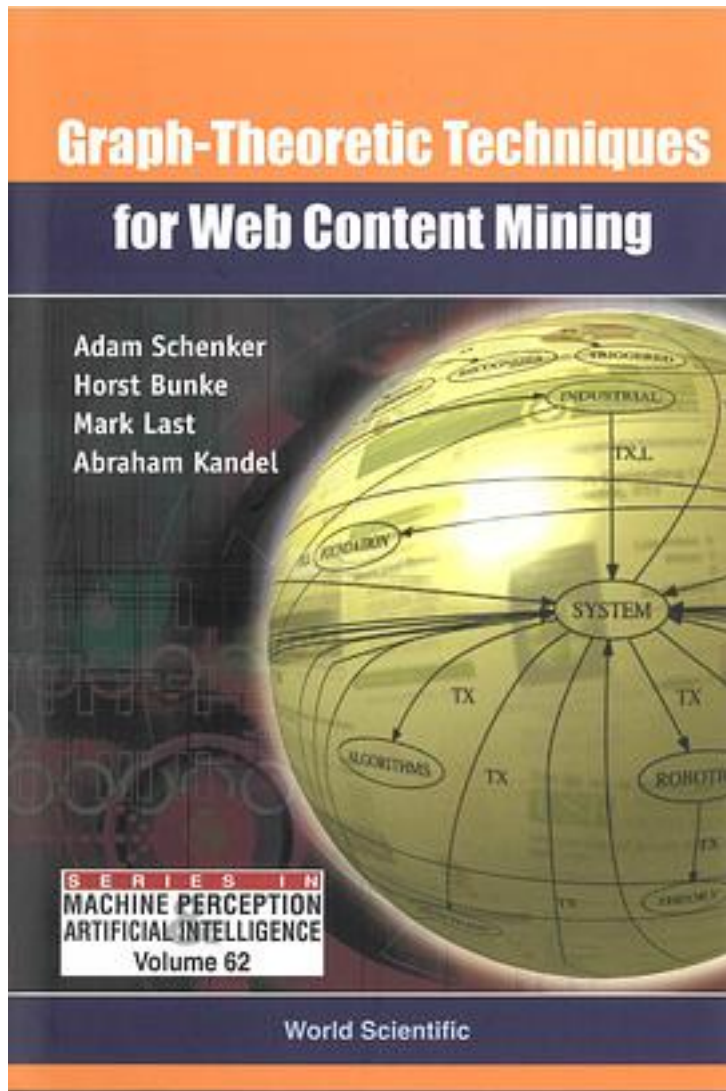


Graph-theoretic Techniques For Web Content Mining



[Graph-theoretic Techniques For Web Content Mining_ 下载链接1_](#)

著者:Abraham Kandel

出版者:World Scientific Publishing Co Pte Ltd

出版时间:2005-5-31

装帧:Hardcover

isbn:9789812563392

This book describes exciting new opportunities for utilizing robust graph representations of data with common machine learning algorithms. Graphs can model additional information which is often not present in commonly used data representations, such as vectors. Through the use of graph distance - a relatively new approach for determining graph similarity - the authors show how well-known algorithms, such as k-means clustering and k-nearest neighbors classification, can be easily extended to work with graphs instead of vectors. This allows for the utilization of additional information found in graph representations, while at the same time employing well-known, proven algorithms. To demonstrate and investigate these novel techniques, the authors have selected the domain of web content mining, which involves the clustering and classification of web documents based on their textual substance. Several methods of representing web document content by graphs are introduced; an interesting feature of these representations is that they allow for a polynomial time distance computation, something which is typically an NP-complete problem when using graphs. Experimental results are reported for both clustering and classification in three web document collections, using a variety of graph representations, distance measures, and algorithm parameters. In addition, this book describes several other related topics, many of which provide excellent starting points for researchers and students interested in exploring this new area of machine learning further. These topics include creating graph-based multiple classifier ensembles through random node selection and visualization of graph-based data using multidimensional scaling.

作者介绍:

目录:

[Graph-theoretic Techniques For Web Content Mining_下载链接1_](#)

标签

评论

[Graph-theoretic Techniques For Web Content Mining_下载链接1_](#)

书评

[Graph-theoretic Techniques For Web Content Mining 下载链接1](#)