River Publishers

Web Mining: A Synergic Approach Resorting to Classifications and Clustering

Authors:

V.S. Kumbhar, Shivaji University, Kolhapur, India

K.S. Oza, Shivaji University, Kolhapur, India

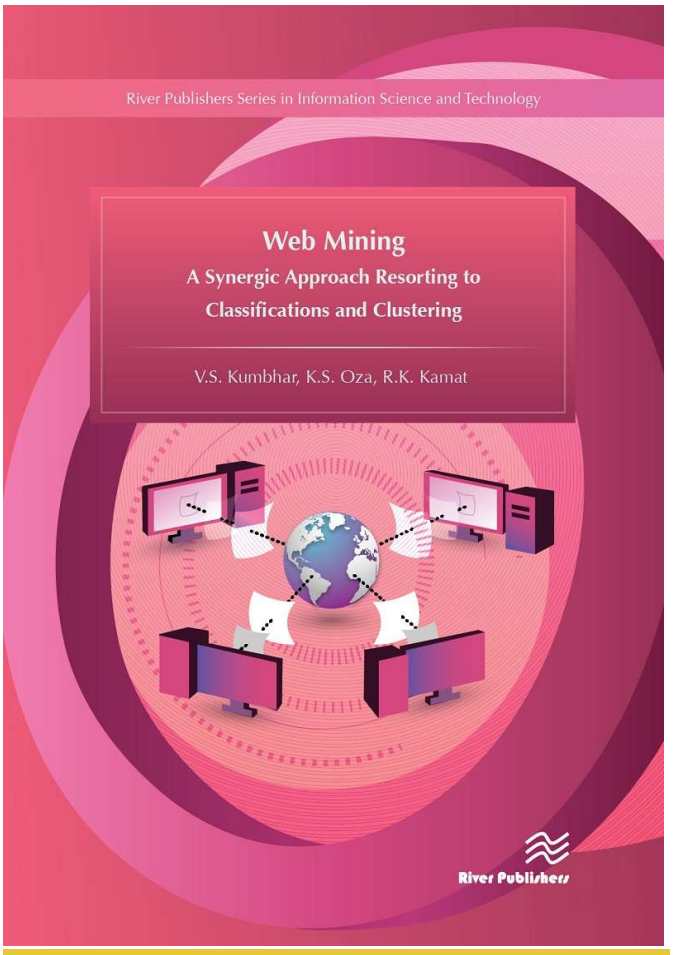
R.K. Kamat, Shivaji University, Kolhapur, India

Web mining is the application of data mining strategies to excerpt learning from web information, i.e. web content, web structure, and web usage data. With the emergence of the web as the predominant and converging platform for communication, business and scholastic information dissemination, especially in the last five years, there are ever increasing research groups working on different aspects of web mining mainly in three directions. These are: mining of web content, web structure and web usage. In this context there are good number of frameworks and benchmarks related to the metrics of the websites which is certainly weighty for B2B, B2C and in general in any e-commerce paradigm. Owing to the popularity of this topic there are few books in the market, dealing more on such performance metrics and other related issues. This book, however, omits all such routine topics and lays more emphasis on the classification and clustering aspects of the websites in order to come out with the true perception of the websites in light of its usability.

In nutshell, *Web Mining: A Synergic Approach Resorting to Classifications and Clustering* showcases an effective methodology for classification and clustering of web sites from their usability point of view. While the clustering and classification is accomplished by using an open source tool WEKA, the basic dataset for the selected websites has been emanated by using a free tool site-analyzer. As a case study, several commercial websites have been analyzed. The dataset preparation using site-analyzer and classification through WEKA by embedding different algorithms is one of the unique selling points of this book. This text projects a complete spectrum of web mining from its very inception through data mining and takes the reader up to the application level.

Salient features of the book include:

- Literature review of research work in the area of web mining
- Business websites domain researched, and data collected using site-analyzer tool
- Accessibility, design, text, multimedia, and networking are assessed
- Datasets are filtered further by selecting vital attributes which are Search Engine Optimized for processing using the Weka attributed tool
- Dataset with labels have been classified using J48, RBFNetwork, NaïveBayes, and SMO techniques using Weka
- A comparative analysis of all classifiers is reported
- Commercial applications for improving website performance based on SEO is given



River Publishers Series in Computing and Information Science and Technology

ISBN: 9788793379831
e-ISBN: 9788793379848
Available From: November 2016
Price: € 75.00

KEYWORDS:
Web Mining, Weka, Clustering, Classification, J48 algorithm, RBFNetwork, NaïveBayes, and SMO techniques

www.riverpublishers.com
marketing@riverpublishers.com