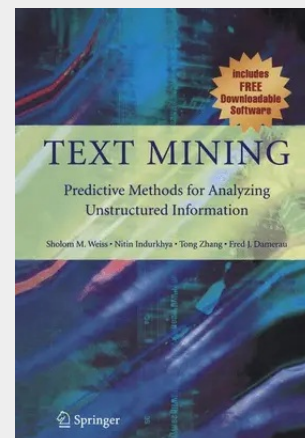


Text Mining

Predictive Methods for Analyzing Unstructured Information

Text mining searches for regularities, patterns or trends in natural language text. Inspired by data mining, which discovers major patterns from highly structured databases, text mining aims to extract useful knowledge from unstructured text. This book focuses on the concepts and methods needed to expand horizons beyond structured, numeric data to automated mining of text samples. This authoritative and highly accessible text/reference, written by a team of authorities on text mining, develops the foundation concepts, principles, and methods needed to expand beyond structured, numeric data to automated mining of text samples. Researchers, computer scientists, and advanced undergraduates and graduates with work and interests in data mining, machine learning, databases, and computational linguistics will find the work an essential resource.

Data mining is a mature technology. The prediction problem, looking for predictive patterns in data, has been widely studied. Strong methods are available to the practitioner. These methods process structured numerical information, where uniform measurements are taken over a sample of data. Text is often described as unstructured information. So, it would seem, text and numerical data are different, requiring different methods. Or are they? In our view, a prediction problem can be solved by the same methods, whether the data are structured - numerical measurements or unstructured text. Text and documents can be transformed into measured values, such as the presence or absence of words, and the same methods that have proven successful for predictive data mining can be applied to text. Yet, there are key differences. Evaluation techniques must be adapted to the chronological order of publication and to alternative measures of error. Because the data are documents, more specialized analytical methods may be preferred for text. Moreover, the methods must be modified to accommodate very high dimensions: tens of thousands of words and documents. Still, the central themes are similar.



160,49 €

149,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9781441929969

Medium: Buch

ISBN: 978-1-4419-2996-9

Verlag: Springer

Erscheinungstermin: 19.11.2010

Sprache(n): Englisch

Auflage: 1. Auflage. Softcover version of original hardcover Auflage 2005

Produktform: Kartoniert

Gewicht: 388 g

Seiten: 237

Format (B x H): 155 x 235 mm

