

Mining the Web: Discovering Knowledge from Hypertext Data

By Soumen Chakrabarti

Download now

Read Online

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti

Mining the Web: Discovering Knowledge from Hypertext Data is the first book devoted entirely to techniques for producing knowledge from the vast body of unstructured Web data. Building on an initial survey of infrastructural issues?including Web crawling and indexing?Chakrabarti examines low-level machine learning techniques as they relate specifically to the challenges of Web mining. He then devotes the final part of the book to applications that unite infrastructure and analysis to bring machine learning to bear on systematically acquired and stored data. Here the focus is on results: the strengths and weaknesses of these applications, along with their potential as foundations for further progress. From Chakrabarti's work?painstaking, critical, and forward-looking?readers will gain the theoretical and practical understanding they need to contribute to the Web mining effort.

* A comprehensive, critical exploration of statistics-based attempts to make sense of Web Mining.
* Details the special challenges associated with analyzing unstructured and semi-structured data.
* Looks at how classical Information Retrieval techniques have been modified for use with Web data.
* Focuses on today's dominant learning methods: clustering and classification, hyperlink analysis, and supervised and semi-supervised learning.
* Analyzes current applications for resource discovery and social network analysis.
* An excellent way to introduce students to especially vital applications of data mining and machine learning technology.

Download Mining the Web: Discovering Knowledge from Hyperte ...pdf

Read Online Mining the Web: Discovering Knowledge from Hyper ...pdf

Mining the Web: Discovering Knowledge from Hypertext Data

By Soumen Chakrabarti

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti

Mining the Web: Discovering Knowledge from Hypertext Data is the first book devoted entirely to techniques for producing knowledge from the vast body of unstructured Web data. Building on an initial survey of infrastructural issues?including Web crawling and indexing?Chakrabarti examines low-level machine learning techniques as they relate specifically to the challenges of Web mining. He then devotes the final part of the book to applications that unite infrastructure and analysis to bring machine learning to bear on systematically acquired and stored data. Here the focus is on results: the strengths and weaknesses of these applications, along with their potential as foundations for further progress. From Chakrabarti's work?painstaking, critical, and forward-looking?readers will gain the theoretical and practical understanding they need to contribute to the Web mining effort.

* A comprehensive, critical exploration of statistics-based attempts to make sense of Web Mining.
* Details the special challenges associated with analyzing unstructured and semi-structured data.
* Looks at how classical Information Retrieval techniques have been modified for use with Web data.
* Focuses on today's dominant learning methods: clustering and classification, hyperlink analysis, and supervised and semi-supervised learning.
* Analyzes current applications for resource discovery and social network analysis.
* An excellent way to introduce students to especially vital applications of data mining and machine learning technology.

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti Bibliography

- Sales Rank: #612512 in Books
- Brand: Brand: Morgan Kaufmann
- Published on: 2002-10-23
- Original language: English
- Number of items: 1
- Dimensions: 9.50" h x 7.50" w x 1.00" l, 1.66 pounds
- Binding: Hardcover
- 344 pages

Download Mining the Web: Discovering Knowledge from Hyperte ...pdf

Read Online Mining the Web: Discovering Knowledge from Hyper ...pdf

Download and Read Free Online Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti

Editorial Review

Review
"...solid and beneficial to readers interested in Web data mining, especially those interested in the details of algorithmic implementation." = Bernard J. Jansen, *Information Processing & Management*

"The treatment is systematic, comprehensive and in-depth, yet very lucid and accessible to a wide range of Web technology developers. The author's insights and depth of knowledge as on of the pioneering researchers on hypertext information mining and retrieval are also evident in the extensive and useful bibliographic notes provided at the end of each chapter..." - Professor Joydeep Ghosh, University of Texas, Austin

"The author has done the community a great service by synthesizing all the important work in this field into an excellent book, which introduces fairly sophisticated material in an easy-to-read manner. This book for the first time, makes it possible to offer Web Mining as a real course." - Professor Jaideep Srivastava, University of Minnesota

" *Mining the Web: Discovering Knowledge from Hypertext from Hypertext Data*, by Soumen Chakrabarti, focuses extensively on building a better search engine crawler...Chakrabarti's book begins with a discussion of search engine crawlers in a chapter titled "Crawling the Web." The discussion in this chapter is technical and detailed. Readers learn about features such as the robots.txt file that can be written in a certain way to stop crawlers from visiting a page...The most interesting part of the book is perhaps Chapter 7, "Social Network Analysis." In this chapter, the author presents the most famous search engine algorithms (e.g., PageRank, HITS, SALSA)." - Journal of Marketing Research, Sandeep Krishnamurthy

"All in all this is an excellent book. I enjoyed the book and highly recommend it as a textbook for web data mining classes at graduate or senior undergraduate levels. Chakrabarti has a rich vocabulary and is a gifted writer. I bet he will write new, good books in the future, and he should. I look forward to them." - Fazli Can - Miami University

From the Back Cover
Mining the Web: Discovering Knowledge from Hypertext Data is the first book devoted entirely to techniques for extracting and producing knowledge from the vast body of unstructured Web data. Building on an initial survey of infrastructural issues?including Web crawling and indexing?Chakrabarti examines machine learning techniques as they relate specifically to the challenges of Web mining and provides applications of machine learning to sytematically acquire, store, and analyze data. Here the focus is on results: the strengths and weaknesses of these applications, along with their potential as foundations for further progress toward a Web that is more aware of content semantics. This thorough and forward-looking book gives the theoretical and practical foundations you need to build innovative applications for mining the Web.

Features

- A comprehensive, critical exploration of statistics-based attempts to make sense of Web data.
- Details the special challenges associated with analyzing unstructured and semi-structured data.
- Looks at how classical Information Retrieval techniques have been modified for use with Web data.

- Focuses on today's dominant learning methods: clustering and classification, hyperlink analysis, and supervised and semi-supervised learning.
- Analyzes current applications for resource discovery and social network analysis.
- An excellent way to introduce students to especially vital applications of data mining and machine learning technology.

About the Author
Soumen Chakrabarti is assistant Professor in Computer Science and Engineering at the Indian Institute of Technology, Bombay. Prior to joining IIT, he worked on hypertext databases and data mining at IBM Almaden Research Center. He has developed three systems and holds five patents in this area. Chakrabarti has served as a vice-chair and program committee member for many conferences, including WWW, SIGIR, ICDE, and KDD, and as a guest editor of the IEEE TKDE special issue on mining and searching the Web. His work on focused crawling received the Best Paper award at the 8th International World Wide Web Conference (1999). He holds a Ph.D. from the University of California, Berkeley.

Users Review

From reader reviews:

Jonathan Nelson:

Your reading 6th sense will not betray anyone, why because this Mining the Web: Discovering Knowledge from Hypertext Data e-book written by well-known writer we are excited for well how to make book that can be understand by anyone who also read the book. Written in good manner for you, leaking every ideas and creating skill only for eliminate your personal hunger then you still skepticism Mining the Web: Discovering Knowledge from Hypertext Data as good book not just by the cover but also with the content. This is one reserve that can break don't ascertain book by its handle, so do you still needing yet another sixth sense to pick this!? Oh come on your examining sixth sense already said so why you have to listening to yet another sixth sense.

Michael Vu:

Do you like reading a guide? Confuse to looking for your favorite book? Or your book ended up being rare? Why so many query for the book? But almost any people feel that they enjoy regarding reading. Some people likes examining, not only science book and also novel and Mining the Web: Discovering Knowledge from Hypertext Data or others sources were given knowledge for you. After you know how the truly amazing a book, you feel need to read more and more. Science e-book was created for teacher or perhaps students especially. Those publications are helping them to add their knowledge. In other case, beside science reserve, any other book likes Mining the Web: Discovering Knowledge from Hypertext Data to make your spare time more colorful. Many types of book like here.

Teresa Riggs:

A lot of guide has printed but it takes a different approach. You can get it by web on social media. You can choose the top book for you, science, comedian, novel, or whatever through searching from it. It is called of book Mining the Web: Discovering Knowledge from Hypertext Data. You can add your knowledge by it.

Without leaving behind the printed book, it can add your knowledge and make a person happier to read. It is most crucial that, you must aware about e-book. It can bring you from one spot to other place.

Jeremy Quick:

Publication is one of source of understanding. We can add our expertise from it. Not only for students but also native or citizen need book to know the revise information of year to help year. As we know those guides have many advantages. Beside we add our knowledge, may also bring us to around the world. With the book Mining the Web: Discovering Knowledge from Hypertext Data we can get more advantage. Don't one to be creative people? To get creative person must like to read a book. Merely choose the best book that ideal with your aim. Don't be doubt to change your life at this book Mining the Web: Discovering Knowledge from Hypertext Data. You can more appealing than now.

Download and Read Online Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti #WMG1NZJ9KD3

Read Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti for online ebook

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti books to read online.

Online Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti ebook PDF download

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti Doc

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti Mobipocket

Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti EPub

WMG1NZJ9KD3: Mining the Web: Discovering Knowledge from Hypertext Data By Soumen Chakrabarti