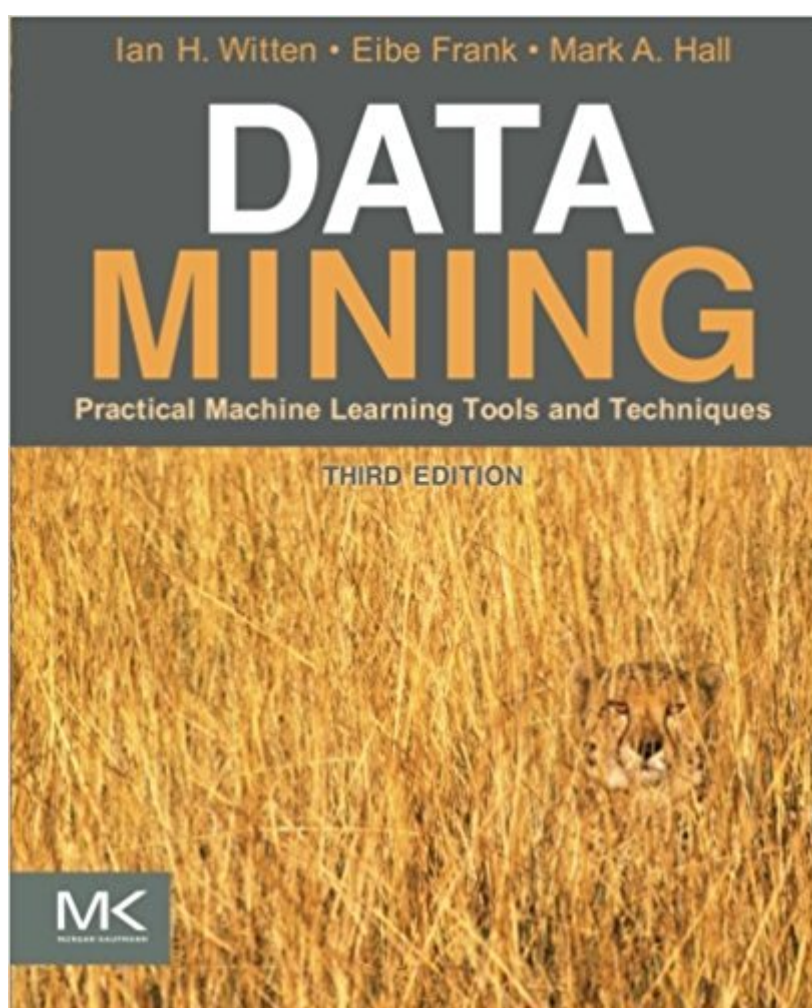The book was found

Data Mining: Practical Machine Learning Tools And Techniques, Third Edition (Morgan Kaufmann Series In Data Management Systems)

Ian H. Witten • Eibe Frank • Mark A. Hall

DATA MINING

Practical Machine Learning Tools and Techniques

THIRD EDITION

MK
MORGAN KAUFMANN

Synopsis

Data Mining: Practical Machine Learning Tools and Techniques, Third Edition, offers a thorough grounding in machine learning concepts as well as practical advice on applying machine learning tools and techniques in real-world data mining situations. This highly anticipated third edition of the most acclaimed work on data mining and machine learning will teach you everything you need to know about preparing inputs, interpreting outputs, evaluating results, and the algorithmic methods at the heart of successful data mining. Thorough updates reflect the technical changes and modernizations that have taken place in the field since the last edition, including new material on Data Transformations, Ensemble Learning, Massive Data Sets, Multi-instance Learning, plus a new version of the popular Weka machine learning software developed by the authors. Witten, Frank, and Hall include both tried-and-true techniques of today as well as methods at the leading edge of contemporary research. The book is targeted at information systems practitioners, programmers, consultants, developers, information technology managers, specification writers, data analysts, data modelers, database R&D professionals, data warehouse engineers, data mining professionals. The book will also be useful for professors and students of upper-level undergraduate and graduate-level data mining and machine learning courses who want to incorporate data mining as part of their data management knowledge base and expertise.Provides a thorough grounding in machine learning concepts as well as practical advice on applying the tools and techniques to your data mining projectsOffers concrete tips and techniques for performance improvement that work by transforming the input or output in machine learning methodsIncludes downloadable Weka software toolkit, a collection of machine learning algorithms for data mining tasks-in an updated, interactive interface. Algorithms in toolkit cover: data pre-processing, classification, regression, clustering, association rules, visualization

Book Information

Series: Morgan Kaufmann Series in Data Management Systems

Paperback: 664 pages

Publisher: Morgan Kaufmann; 3 edition (January 20, 2011)

Language: English

ISBN-10: 0123748569

ISBN-13: 978-0123748560

Product Dimensions: 7.5 x 1.5 x 9.2 inches

Shipping Weight: 3.1 pounds (View shipping rates and policies)

Average Customer Review: 4.2 out of 5 stars   79 customer reviews

Best Sellers Rank: #54,858 in Books (See Top 100 in Books)   #23 inÃ Â Books > Textbooks > Computer Science > Artificial Intelligence   #56 inÃ Â Books > Computers & Technology > Computer Science > AI & Machine Learning > Intelligence & Semantics   #75 inÃ Â Books > Textbooks > Computer Science > Database Storage & Design

Customer Reviews

"...offers a thorough grounding in machine learning concepts as well as practical advice on applying machine learning tools and techniques in real-world data mining situations." "Co-author Witten is the author of other well-known books on data mining, and he and his co-authors of this book excel in statistics, computer science, and mathematics. Their in- depth backgrounds and insights are the strengths that have permitted them to avoid heavy mathematical derivations in explaining machine learning algorithms so they can help readers from different fields understand algorithms. I strongly recommend this book to all newcomers to data mining, especially to those who wish to understand the fundamentals of machine learning algorithms."--INFORMS Journal of Computing "The third edition of this practical guide to machine learning and data mining is fully updated to account for technological advances since its previous printing in 2005 and is now even more closely aligned with the use of the Weka open source machine learning, data mining and data modeling application. Beginning with an introduction to data mining, the volume explores basic inputs, outputs and algorithms, the implementation of machine learning schemes and in-depth exploration of the many uses of the Weka data analysis software. Numerous illustration, tables and equations are included throughout and additional resources are available through a companion website. Witten, Frank and Hall are academics with the department of computer science at the University of Waikato, New Zealand, the home of the Weka software project."--Book News, Reference & Research "I would recommend this book to anyone who is getting started in either data mining or machine learning and wants to learn how the fundamental algorithms work. I liked that the book slowly teaches you the different algorithms piece by piece and that there are also a lot of examples. I plan on taking a machine learning course this upcoming fall semester and feel that the book gave me great insight that the course will be based on mathematics more than I had originally expected. My favorite part of the book was the last chapter where it explains how you can solve different practical data mining scenarios using the different algorithms. If there were more chapters like the last one, the book would have been perfect. This book might not be that useful if you do not plan on using the Weka software or if you are already familiar with the various machine learning algorithms. Overall, Data

Mining: Practical Machine Learning Tools and Techniques is a great book to learn about the core concepts of data mining and the Weka software suite."--ACM SIGSOFT Software Engineering Notes "This book is a must-read for every aspiring data mining analyst. Its many examples and the technical background it imparts would be a unique and welcome addition to the bookshelf of any graduate or advanced undergraduate student. The book is written for both academic and application-oriented readers, and I strongly recommend it to any reader working in the area of machine learning and data mining."--Computing Reviews.com

Data Mining: Practical Machine Learning Tools and Techniques offers a thorough grounding in machine learning concepts as well as practical advice on applying machine learning tools and techniques in real-world data mining situations. This highly anticipated third edition of the most acclaimed work on data mining and machine learning will teach you everything you need to know about preparing inputs, interpreting outputs, evaluating results, and the algorithmic methods at the heart of successful data mining. Thorough updates reflect the technical changes and modernizations that have taken place in the field since the last edition, including new material on Data Transformations, Ensemble Learning, Massive Data Sets, Multi-instance Learning, plus a new version of the popular Weka machine learning software developed by the authors. Witten, Frank, and Hall include both tried-and-true techniques of today as well as methods at the leading edge of contemporary research.

First of all, I would advise to think of this as a 400-page book with a WEKA appendix. Its price is about right for a 400-page machine learning textbook, and you don't even need to know that WEKA exists for the first 400 pages. I never read any of the WEKA stuff and got tons out of the textbook part.There's very little actual math or theory in this book. The average explanation amounts to "There's a technique called X, where you do this... it has a couple problems, but you could try fixing them in these ways." It's great for getting a lot of machine learning and data mining ideas in your head without having to get confused by learning the math behind them.Problems mostly come from the lack of organization. Most of these are in Chapter 6, which is by far the most important chapter. For instance, this chapter begins with two or three pages describing what's going on in Figure 1.3 from two-hundred pages earlier. Each section of the chapter references its corresponding section in Chapter 4 a lot. The authors also assume that you memorized, in intimate detail, their examples in the first five pages because they keep referencing them in detail throughout the book. Finally, the explanations of a couple algorithms -- decision trees, in particular -- can get disorganized and

confusing; however, these are exceptions to the rule.But, this is a good book. I got a lot of new ideas out of it for how to improve some the algorithms I work on, or for new things to try. It's great to have explanations of these machine learning algorithms and concepts that give you an intuition for what their goal/purpose is without going into too much detail about why they work -- there are ten other books for that.

The book is really good to start learning machine learning and data mining.Pros- It doesn't jump into algorithms with mathematical details. It starts with what is it all about, what input and output look like in typical machine learning problems.- One point that I really liked is that the book gives algorithms in two chapters (chapter 4 and 6). The first chapter is about basics and latter one gives detail about these algorithms.- It also covers well that I think it is mostly ignored by other books/tutorials: practical issues. How to normalize data, what happens your data have both categorical and numerical features, discretizing numerical features and so on.- If you consider using Weka, you should have this book. Authors are from the team who built Weka. For each algorithm described in the book, corresponding names of implementations in Weka are given too. With the book it is easier to understand parameters of Weka implementations of algorithms. Also last part of the book is like extensive Weka tutorial.Cons- In a few points, the book contains unnecessary details, although it is not the case for overall of the book. One of such things that I remember is chapter 4.7. The book spends 5 whole pages to how to find nearest neighbor efficiently (not-easy stuff), which I think it is really implementation detail. Instead of it, it could explain what nearest neighbor is, or something else.- The part about Weka has several figures, mostly Weka screen shots. It was difficult to follow these figures, because of black-white screen shots. I think these figures should be in color in the next edition, which will make much easier to follow.

To get the most out of this book, you need to either be a statistician, AI professional, or be willing to invest some time. But: if you commit yourself, then this book goes a long way to substitute for a graduate-level course on data mining. Don't get me wrong - it is not written with an academic audience in mind; as a matter of fact, it is unusually rich with application examples. But there is a lot to digest conceptually and many of the examples are quite involved. As such, it addresses the opposite end the O'Reilly series of how-to books. This one gets you up to speed with one of if not the best software package for data mining in all its many facets. With Weka and 'R', you have the tools to tackle many of the World's problems, and this book is the best introduction to one part of the duo.

Data Mining: Practical Machine Learning Tools and Techniques, Third Edition (Morgan Kaufmann Series in Data Management Systems) Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems) Data Mining: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems) Big Data For Business: Your Comprehensive Guide to Understand Data Science, Data Analytics and Data Mining to Boost More Growth and Improve Business - Data Analytics Book, Series 2 Discovering Knowledge in Data: An Introduction to Data Mining (Wiley Series on Methods and Applications in Data Mining) Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data (Princeton Series in Modern Observational Astronomy) Computer Networks, Fifth Edition: A Systems Approach (The Morgan Kaufmann Series in Networking) Learning Processing, Second Edition: A Beginner's Guide to Programming Images, Animation, and Interaction (The Morgan Kaufmann Series in Computer Graphics) VLSI Test Principles and Architectures: Design for Testability (The Morgan Kaufmann Series in Systems on Silicon) Computer Networks: A Systems Approach (The Morgan Kaufmann Series in Networking) Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Computer Organization and Design MIPS Edition, Fifth Edition: The Hardware/Software Interface (The Morgan Kaufmann Series in Computer Architecture and Design) Computer Organization and Design, Fourth Edition: The Hardware/Software Interface (The Morgan Kaufmann Series in Computer Architecture and Design) See MIPS Run, Second Edition (The Morgan Kaufmann Series

in Computer Architecture and Design) Foundations of Analog and Digital Electronic Circuits (The Morgan Kaufmann Series in Computer Architecture and Design) Self-Checking and Fault-Tolerant Digital Design (The Morgan Kaufmann Series in Computer Architecture and Design) Logical Effort: Designing Fast CMOS Circuits (The Morgan Kaufmann Series in Computer Architecture and Design) Skew-Tolerant Circuit Design (The Morgan Kaufmann Series in Computer Architecture and Design)