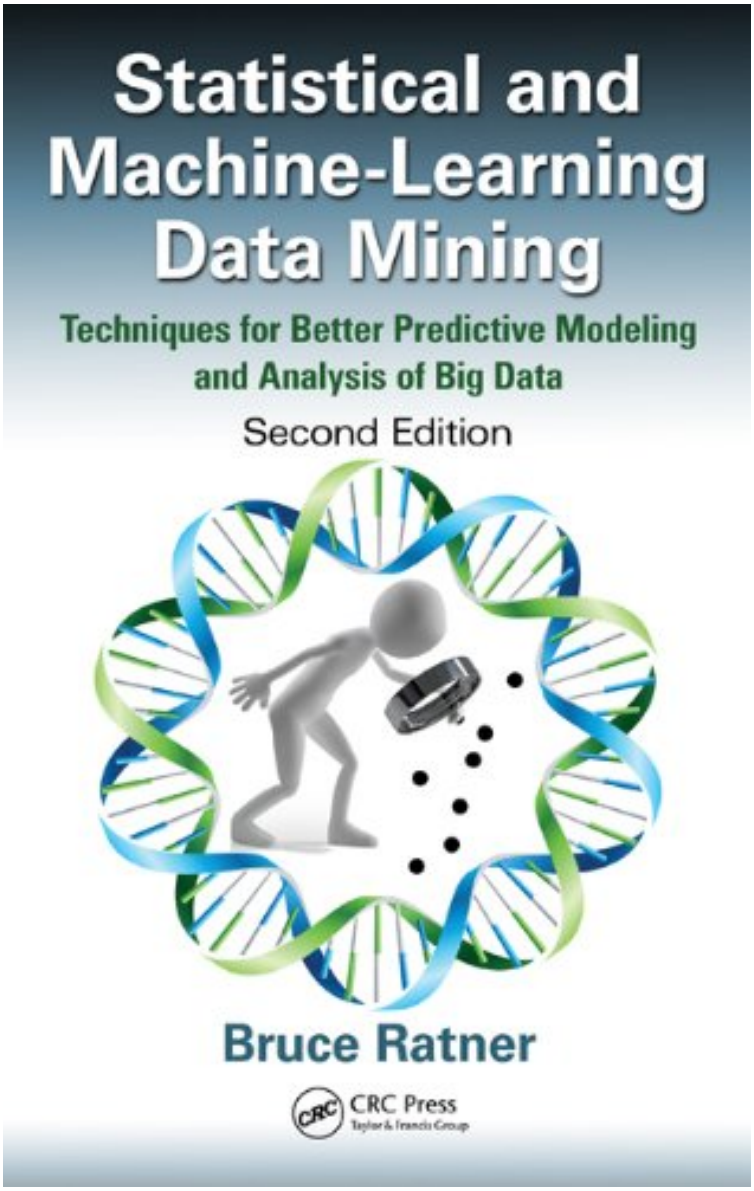
Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data, Second Edition

Bruce Ratner



#1520853 in eBooks 2012-02-10 2012-02-10File Name: B0077RR42Q | File size: 48.Mb

Bruce Ratner : Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data, Second Edition before purchasing it in order to gage whether or not it would be worth my

time, and all praised Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data, Second Edition:

93 of 94 people found the following review helpful. Low qualityBy Dimitri ShvorobMr. Ratner is one, ahem, passionate marketer -a. I first learnt of him and his GenIQ software from the e-mails he was posting to a LinkedIn group on a daily basis.b. Watch "Statistical modeling and analysis for database marketing" become "Statistical and machine-learning data mining" in its second edition - the first time I see a book go into (n+1) edition with a completely different title - adding three "hot" search terms: "data mining", "machine learning" and "big data". And it works (on some people) - the book has zilch to say about big data, yet here it is in the "Big data" section of a ("prestigious", according to Mr. Ratner) list published by IBM.c. He posts a five-star rating of own book - anonymously, as "The Significant Statistician". Two more reviews are from fans, who earlier gave five stars to the first edition, and both decided to upgrade; in total, three of the seven five-star reviews (excluding Mr. Ratner's own) come from his home state, NY, which is an interesting geographical spike. With one exception, the positive reviews have 2-3 positive votes each; coincidentally, my own review got two negative votes on the day of posting; literally putting two and two (multiplied by seven? or eight?) together, I can make a guess about where the votes are coming from.Mr. Ratner is also a bona fide statistics PhD, but it seems that in the decades after leaving grad school, he has not invested time in keeping his knowledge of the field up to date - or in writing his books, which are just another channel of GenIQ promotion. I have reviewed the first edition - take a look at the comments under that 2009 review - and am disappointed to see the second one just as poorly written (a half-page passage shows up, unchanged, three times - on pp. 18, 48, 90 - labor of love, you say?), poorly typeset and visually ugly, and, well, shallow.If this is your first book on statistical and non-statistical methods of data analysis, you may well be impressed, but at $80, the wisdom is a tad overpriced, and why not get a proper book by a recognized author? (I recommend "Introduction to statistical learning" by James, Witten, Hastie and Tibshirani; "Doing data science" by Schutt and O'Neil is another, very different option). America has a proud tradition of garage inventors, but this one needs to spend more time in the garage.13 of 14 people found the following review helpful. Not worth the money.By Brian SpieringThe best part of the book is the easy going writing style and story telling.However:It is very repetitive. It would ~20% shorter if verbatim passages were removed. Many ideas are restated throughout.It is a relatively expensive book. I expected more technical and complete coverage for the money.It does not adequately cover algorithms, mathematical concepts, or how to completely run the analysis described.It barely covers SAS implementation.It is not well organized. It comes across as a series of lectures or blog posts, in contrast to being a coherent book.9 of 23 people found the following review helpful. Finally, machine learning and data mining made accessibleBy Brad BurrisDr Ratner's overview of these topics hits some sweetspots with me. I'm always looking to understand trends in analytic tools and techniques that can provide my company with a competitive advantage. The probem I find in up and coming tools/techniques, is you need to be the esoteric statistician or mathematacian in the particular field to grasp the capabilities, let alone, how to use them. This book is much more positively utlitarian in design. It provides plenty of examples and allows someone with a basic understanding of scientific or engineering statistics access to machine learning and data mining techniques. This is a great text for: 1)those in business or NGO's exploring "Big Data" analysis; 2) manager's determining buget allocations for tools (which I think this book shows you don't necessarily need enterprise sofware packages); or 3) a survey course on these two topics. And some added bonuses: the author writes in an easy to read style and is more than willing to answer questions or have discussions on the topics.

The second edition of a bestseller, Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data is still the only book, to date, to distinguish between statistical data mining and machine-learning data mining. The first edition, titled Statistical Modeling and Analysis for Database Marketing: Effective Techniques for Mining Big Data, contained 17 chapters of innovative and practical statistical data mining techniques. In this second edition, renamed to reflect the increased coverage of machine-learning data mining techniques, the author has completely revised, reorganized, and repositioned the original chapters and produced 14 new chapters of creative and useful machine-learning data mining techniques. In sum, the 31 chapters of simple yet insightful quantitative techniques make this book unique in the field of data mining literature.The statistical data mining methods effectively consider big data for identifying structures (variables) with the appropriate predictive power in order to yield reliable and robust large-scale statistical models and analyses. In contrast, the author's own GenIQ Model provides machine-learning solutions to common and virtually unapproachable statistical problems. GenIQ makes this possible mdash; its utilitarian data mining features start where statistical data mining stops.This book contains essays offering detailed background, discussion, and illustration of specific methods for solving the most commonly experienced problems in predictive modeling and analysis of big data. They address each methodology and assign its application to a specific type of problem. To better ground readers, the book provides an in-depth discussion of the basic methodologies of predictive modeling and analysis. While this type of overview has been attempted before, this approach offers a truly nitty-gritty, step-by-step method that both tyros and experts in the field

can enjoy playing with.

About the AuthorBruce Ratner, DM STAT-l Consulting