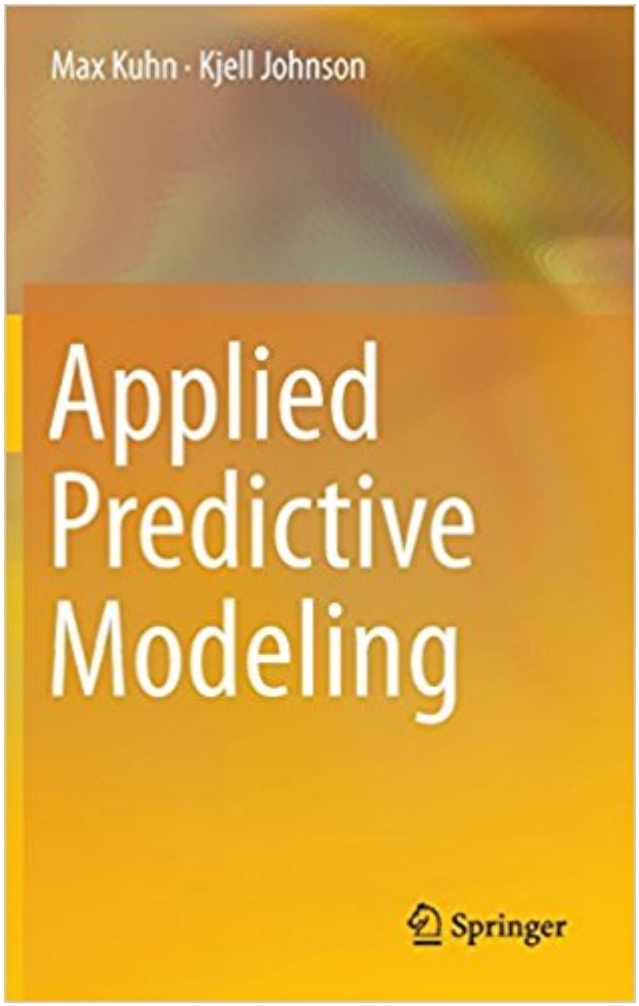
Applied Predictive Modeling



Synopsis

Winner of the 2014 Technometrics Ziegel Prize for Outstanding BookApplied Predictive Modeling covers the overall predictive modeling process, beginning with the crucial steps of data preprocessing, data splitting and foundations of model tuning.Â The text then provides intuitive explanations of numerous common and modern regression and classification techniques, always with an emphasis on illustrating and solving real data problems.Â Addressing practical concerns extends beyond model fitting to topics such as handling class imbalance, selecting predictors, and pinpointing causes of poor model performanceâ •all of which are problems that occur frequently in practice.Â The text illustrates all parts of the modeling process through many hands-on, real-life examples.Â And every chapter contains extensive R code for each step of the process.Â The data sets and corresponding code are available in the book's companion AppliedPredictiveModeling R package, which is freely available on the CRAN archive.Â This multi-purpose text can be used as an introduction to predictive models and the overall modeling process, a practitioner's reference handbook, or as a text for advanced undergraduate or graduate level predictive modeling courses.Â To that end, each chapter contains problem sets to help solidify the covered concepts and uses data available in the book's R package.Â Readers and students interested in implementing the methods should have some basic knowledge of R.Â And a handful of the more advanced topics require some mathematical knowledge.

Book Information

Hardcover: 600 pages

Publisher: Springer; 2013 edition (September 15, 2013)

Language: English

ISBN-10: 1461468485

ISBN-13: 978-1461468486

Product Dimensions: 6.1 x 1.3 x 9.2 inches

Shipping Weight: 2.3 pounds (View shipping rates and policies)

Average Customer Review: 4.8 out of 5 starsÂ Â See all reviewsÂ (52 customer reviews)

Best Sellers Rank: #17,677 in Books (See Top 100 in Books) #1 inÂ Books > Textbooks > Medicine & Health Sciences > Research > Biostatistics #2 inÂ Books > Medical Books > Basic Sciences > Biostatistics #11 inÂ Books > Computers & Technology > Software > Mathematical & Statistical

Customer Reviews

I read "Applied predictive modeling" (which I will shorten to APM) shortly after I read "Introduction to statistical learning" (ISL) by James, Witten, Hastie and Tibshirani, and find that book both closest to APM, and helpful in highlighting APM's strengths.The two books cover the same broad subject. If you google "kuhn caret", you will find Max Kuhn's (very informative) presentation of his "caret" R package, and its first slide will tell you that he uses "predictive modeling" as a synonym of "machine learning" - what Hastie and Tibshirani call "statistical learning". Adopting H&T's terminology choice, I will say that both books combine theory of "statistical learning" with hands-on illustrations and exercises implemented in R; the get-your-hands-dirty, try-it-out element is, in fact, ISL's key difference from the earlier, venerable "Elements of statistical learning".Both books, inevitably, go over a catalog of statistical-learning techniques. The shorter ISL, in my opinion, is superior at explaining the concepts and communicating the principles, while APM takes the more straightforward approach of "beefing up" the catalog, by spending more pages on each item and including more items. While ISL is by design very accessible, APM can be more technical - the detail will surely be appreciated by any practitioner - and, as it talks about the various methods, it can and does discuss recent extensions, offering an extensive and "fresh" bibliography. R-wise, APM's advantage is not decisive (if you look at content, not line count) but big; the book naturally favors "caret" - which has a useful role, "wrapping" a plethora of third-party R packages, and providing a common interface, plus helpful utilities - but both references and uses the specialist packages as well.

There are many fine math-oriented predictive modeling books, such as Hastie (The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)). Kuhn et al consider them "sister texts" and begin immediately to differentiate-- their approach is hands on and practical, for the express purpose of demonstrating HOW to sort, structure and predict via Python or R, for the purpose of accuracy and understanding of the DATA and trends, NOT learning the underlying math.For a couple of pharmaceutical guys, (who BTW use R extensively, I've been an analyst in that industry), you'd think the examples would be new chemical or biological entities. Not so! The cases are fun and exciting, ranging from the nontrivial compression strength of concrete (want that bridge to hold when you cross?) to fuel economy, credit scoring, success in grant applications (boy their colleagues will love that one!), and cognitive impairment. I evaluate technology for patents at payroy dot com, and we have a log likelihood model using Bayesian and Monte Carlo that their grant section helped translate seamlessly to R! We're NOT talking pie in the sky pseudo code here, but real life, real results recipes.The authors talk about

the "scholarly veil" -- meaning we general workers and researchers don’t always "deserve" to see the underlying process, software and data (and, other than open source, often can’t afford it). Wow, do they pop that myth!

tl;dr: A brilliant book covering Predictive modelling in R. With a strong practical bent it walks the reader through the application of modern classification and regression techniques to a broad number of varied and interesting data sets. It uses existing packages where possible so you can jump straight in (great for Kagglers) but there is a lot here to master. It is especially strong on preprocessing (both unsupervised and supervised), model tuning and model assessment. Should not be your first book on R or data analytics but the best balance of Practical application without foregoing theory that I have seen. It is wonderful to see how professional data analysts approach predictive modelling tasks. The data sets are not toy models to highlight approaches but interesting and complex problems from a wide variety of disciplines.(Note that this book does not cover Time Series, Generalised Additive Models and Ensemble’s of different models).Review:Data science has become very popular due to the increase in computing power (including things like AWS), the amount of data that is accessible on the internet and a number of open-source tools (R and Python for example) that allow even relative beginners to complete quite sophisticated models. Coursera allows for one to complete courses on Machine Learning for free and sites like Kaggle have even turned it into something of a sport where people compete to create predictive models for money or even job interviews. Part of the excitement is that Predictive models can be applied to almost any field you can think of.

Groundwater and Geothermal Systems: Fundamentals of Mass, Energy and Solute Transport in Poroelastic Rocks (Multiphysics Modeling) Geochemical Modeling of Groundwater, Vadose and Geothermal Systems (Multiphysics Modeling) Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences (Modeling and Simulation in Science, Engineering and Technology) Student Solutions Manual for Differential Equations: Computing and Modeling and Differential Equations and Boundary Value Problems: Computing and Modeling Applied Groundwater Modeling, Second Edition: Simulation of Flow and Advective Transport Modeling and Simulation in Medicine and the Life Sciences (Texts in Applied Mathematics) Applied Cryptography: Protocols, Algorithms, and Source Code in C [APPLIED CRYPTOGRAPHY: PROTOCOLS, ALGORITHMS, AND SOURCE CODE IN C BY Schneier, Bruce (Author) Nov-01-1995 Elena Bablenis Haveles BS Pharm Pharm D's Applied Pharmacology 6th (Sixth) edition(Applied Pharmacology for the Dental Hygienist [Paperback])(2010) Applied Therapeutics: The Clinical Use of Drugs (APPLIED THERAPEUTICS (KODA-KIMBLE))