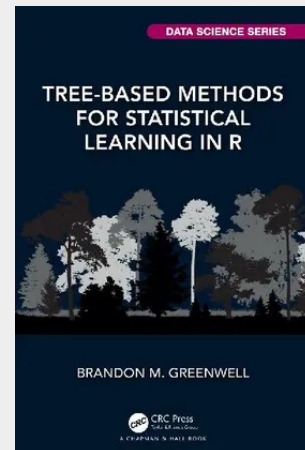


Tree-Based Methods for Statistical Learning in R

Tree-based Methods for Statistical Learning in R provides a thorough introduction to both individual decision tree algorithms (Part I) and ensembles thereof (Part II). Part I of the book brings several different tree algorithms into focus, both conventional and contemporary. Building a strong foundation for how individual decision trees work will help readers better understand tree-based ensembles at a deeper level, which lie at the cutting edge of modern statistical and machine learning methodology. The book follows up most ideas and mathematical concepts with code-based examples in the R statistical language; with an emphasis on using as few external packages as possible. For example, users will be exposed to writing their own random forest and gradient tree boosting functions using simple for loops and basic tree fitting software (like rpart and party/partykit), and more. The core chapters also end with a detailed section on relevant software in both R and other opensource alternatives (e.g., Python, Spark, and Julia), and example usage on real data sets. While the book mostly uses R, it is meant to be equally accessible and useful to non-R programmers. Consumers of this book will have gained a solid foundation (and appreciation) for tree-based methods and how they can be used to solve practical problems and challenges data scientists often face in applied work. Features: - Thorough coverage, from the ground up, of tree-based methods (e.g., CART, conditional inference trees, bagging, boosting, and random forests). - A companion website containing additional supplementary material and the code to reproduce every example and figure in the book. - A companion R package, called treemisc, which contains several data sets and functions used throughout the book (e.g., there's an implementation of gradient tree boosting with LAD loss that shows how to perform the line search step by updating the terminal node estimates of a fitted rpart tree). - Interesting examples that are of practical use; for example, how to construct partial dependence plots from a fitted model in Spark MLlib (using only Spark operations), or post-processing tree ensembles via the LASSO to reduce the number of trees while maintaining, or even improving performance.



163,20 €

163,20 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9780367532468

Medium: Buch

ISBN: 978-0-367-53246-8

Verlag: Chapman and Hall/CRC

Erscheinungstermin: 23.06.2022

Sprache(n): Englisch

Auflage: 1. Auflage 2022

Serie: Chapman & Hall/CRC Data Science Series

Produktform: Gebunden

Gewicht: 772 g

Seiten: 404

Format (B x H): 161 x 240 mm

