

1. Record Nr.	UNINA9910483571603321
Autore	Genuer Robin
Titolo	Random forests with R // Robin Genuer, Jean-Michel Poggi
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2020] ©2020
ISBN	3-030-56485-1
Edizione	[1st ed. 2020.]
Descrizione fisica	1 online resource (X, 98 p. 49 illus., 5 illus. in color.)
Collana	Use R!, , 2197-5736
Disciplina	519.5
Soggetti	Mathematical statistics R (Computer program language)
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Introduction -- CART trees -- Random forests -- Variable importance -- Variable selection -- References.
Sommario/riassunto	This book offers an application-oriented guide to random forests: a statistical learning method extensively used in many fields of application, thanks to its excellent predictive performance, but also to its flexibility, which places few restrictions on the nature of the data used. Indeed, random forests can be adapted to both supervised classification problems and regression problems. In addition, they allow us to consider qualitative and quantitative explanatory variables together, without pre-processing. Moreover, they can be used to process standard data for which the number of observations is higher than the number of variables, while also performing very well in the high dimensional case, where the number of variables is quite large in comparison to the number of observations. Consequently, they are now among the preferred methods in the toolbox of statisticians and data scientists. The book is primarily intended for students in academic fields such as statistical education, but also for practitioners in statistics and machine learning. A scientific undergraduate degree is quite sufficient to take full advantage of the concepts, methods, and tools discussed. In terms of computer science skills, little background knowledge is required, though an introduction to the R language is recommended. Random forests are part of the family of tree-based

methods; accordingly, after an introductory chapter, Chapter 2 presents CART trees. The next three chapters are devoted to random forests. They focus on their presentation (Chapter 3), on the variable importance tool (Chapter 4), and on the variable selection problem (Chapter 5), respectively. After discussing the concepts and methods, we illustrate their implementation on a running example. Then, various complements are provided before examining additional examples. Throughout the book, each result is given together with the code (in R) that can be used to reproduce it. Thus, the book offers readers essential information and concepts, together with examples and the software tools needed to analyse data using random forests. .
