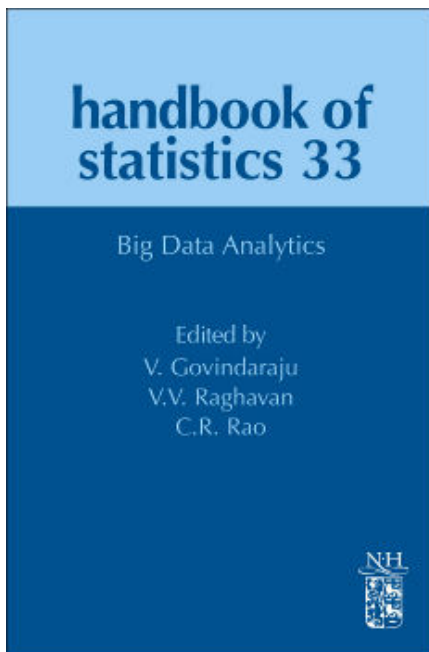


Big data analytics

BY V. GOVINDRAJU, V.V. RAGHAVAN, C.R. RAO,
PRINT BOOK ISBN : 9780444634924 EBOOK ISBN : 9780444634979



REVIEWED BY PAWAN LINGRAS

A BOOK THAT BALANCES THE NUMERIC,
TEXT, AND CATEGORICAL DATA MINING
WITH A TRUE BIG DATA PERSPECTIVE

The book is edited by leaders in both text mining/information retrieval and numeric data. It is a handbook meant for researchers and practitioners that are familiar with the basic concepts and techniques of data mining and statistics. Most people will go through the book perfunctorily to have a good understanding of broad range of topics covered in the book, and revisit the detailed treatment as needed. However, serious research students who want to have a comprehensive understanding of the world of Big Data may find it useful to spend a month going through most of the chapters in detail and also follow a long list of citations that provide specifics.

No one person can have a mastery of the topics covered in the book. The editors have sought experts in various aspects.

Despite the large array of authors, the book has managed a consistent and smooth flowing writing style.

The general areas covered in the book include text mining, web and social network analytics, images, biometrics, and health/epidemiology and customer relationship management. It is unlikely that a single book can contain all the areas which can use big data analysis. However, the breadth of techniques and application domains means that a person looking at a new area may find similarity with one of the chapters discussed in the book.

As mentioned before, the basic data mining and statistical techniques are not part of this book. However, some of these techniques are revisited from the big data perspective. There is also more emphasis on graph theory, which has not received as much attention in the earlier data mining research. I also liked the chapter on simulation to generate additional data, since not all the possible future conditions will be captured by existing real-world datasets. For example, rare events are under-represented in a dataset. The big data analysis will also result in the prescription of new operating conditions that were not previously seen and hence not part of the existing datasets. Simulation helps enhance datasets overcome such deficiencies.

While the book is divided into techniques and applications sections, even the techniques are presented with a firm sight of applicability in mind. Implementation aspects, including MapReduce and Hadoop, emphasize the big data aspect of data mining.

THE BOOK

GOVINDRAJU, V., RAGHAVAN, V.V.,
RAO, C.R. (EDS) (2015), BIG DATA
ANALYTICS, 390 P.
SPRINGER
PRINT BOOK ISBN : 9780444634924
EBOOK ISBN : 9780444634979

ABOUT THE REVIEWER

PAWAN LINGRAS
PROFESSOR AND DIRECTOR
COMPUTING AND DATA ANALYTICS
SAINT MARY'S UNIVERSITY
HALIFAX, NOVA SCOTIA, B3H3C3
CANADA

E-mail: pawan@cs.smu.ca