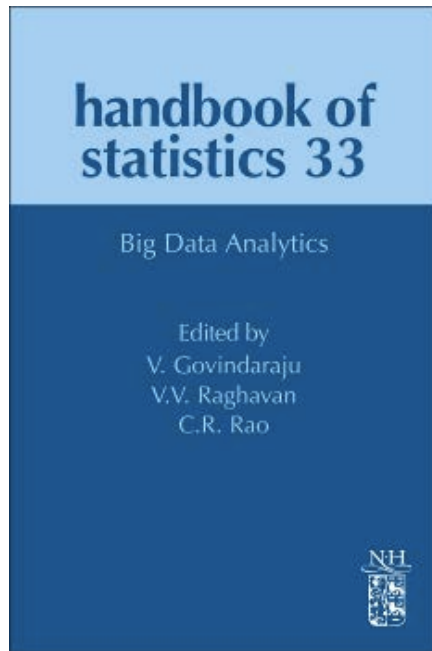# Big Data Analytics

BY V. GOVINDRAJU, V.V. RAGHAVAN AND C.R. RAO (EDITORS) - ISBN: 9780444634924

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.

REVIEWED BY
SUGATA SANYAL

TRACKING THE EVOLUTION OF BIG DATA, FOCUSING ON TIMELY TOPICS SUCH AS DATA MINING AND ANALYTICS

Big Data Analytics has become an often repeated name. I was curious as always about any new issues. But to learn a lot about a Big Subject, one needs to study a lot. When I got an invitation from Elsevier to review a, now famous, book titled, Big Data Analytics, I was thrilled for few techno-academic reasons:

This book is a compendium of various chapters where it deals with theory of Big Data and its applications in real life issues.

Editors are all world famous academician and the standard of the book is very high.

I read the book thoroughly (for writing a critical review you need to) and learned a lot. Big Data Analytics is changing the way we handle various issues and it is yielding results which were not possible a few years back.

This subject is not a narrow one; it is a science of science. And it is ever expanding and some new applications are becoming tractable due to Big Data Analytics application.

The book was edited by Profs. Venu Givindaraju and C. R. Rao and is now available with Elsevier. A brief description below:

While the term Big Data is open to varying interpretation, it is quite clear that the Volume, Velocity, and Variety (3Vs) of data have impacted every aspect of computational science and its applications. The volume of data is increasing at a phenomenal rate and a majority of it is unstructured. With big data, the volume is so large that processing it using traditional database and software techniques is difficult, if not impossible. The drivers are the ubiquitous sensors, devices, social

networks and the all-pervasive web. Scientists are increasingly looking to derive insights from the massive quantity of data to create new knowledge. In common usage, Big Data has come to refer simply to the use of predictive analytics or other certain advanced methods to extract value from data, without any required magnitude thereon. Challenges include analysis, capture, curation, search, sharing, storage, transfer, visualization, and information privacy. While there are challenges, there are huge opportunities emerging in the fields of Machine Learning, Data Mining, Statistics, Human-Computer Interfaces and Distributed Systems to address ways to analyze and reason with this data. The edited volume focuses on the challenges and opportunities posed by "Big Data" in a variety of domains and how statistical techniques and innovative algorithms can help glean insights and accelerate discovery. Big data has the potential to help companies improve operations and make faster, more intelligent decisions.

My detailed review is available at: http://goo.gl/pzIBoH
I wish and hope that you enjoy Big Data Analytics and its applications as much as I did.

THE BOOK:

V. GOVINDRAJU, V.V. RAGHAVAN, C.R. RAO (EDS) (2015), BIG DATA ANALYTICS, 390 P.
ELSEVIER
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.
Contact him at: pawan@cs.smu.ca

SUGATA SANYAL
Member, School of Computing and Informatics "Brain Trust", University of Louisiana at Lafayette, USA.
Honorary Professor, IIT, Guwahati.
Member, Senate, IIIT, Guwahati.
Contact him at: sanyals@gmail.com