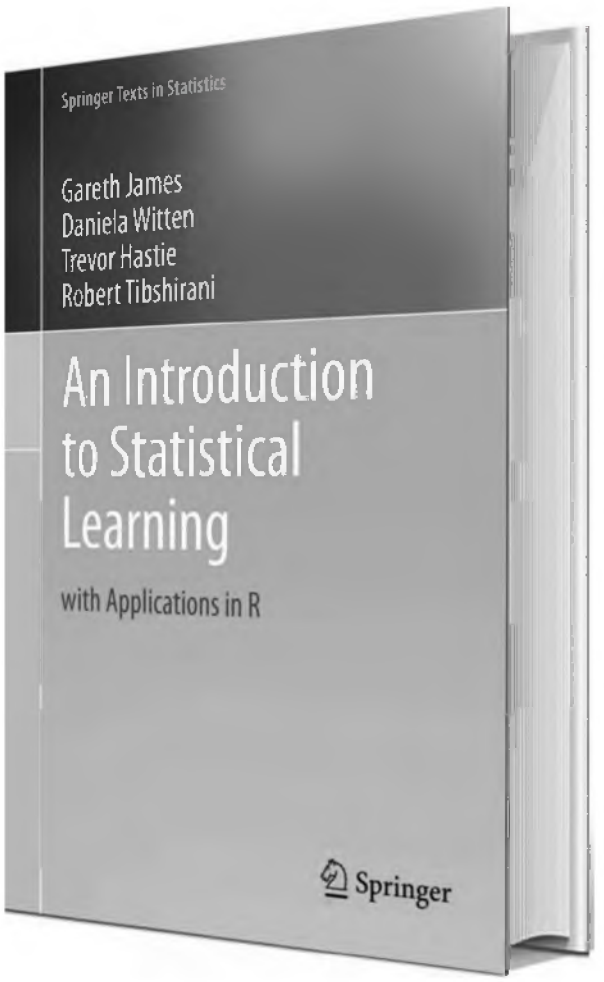
Book Review

An Introduction to Statistical Learning with Applications in R

by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani

Published by: Springer
Published in: 2022
Pages: 607
Language: English
ISBN: 9781071614204



About the Author

"An Introduction to Statistical Learning" is a widely acclaimed textbook written by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. The book provides a comprehensive introduction to statistical learning methods and their application in data analysis and prediction.

Gareth James: Gareth James is a professor of data sciences and operations at the University of Southern California. He has made significant contributions to statistical learning and its applications, particularly in the fields of bioinformatics and sports analytics. James has published numerous research papers and co-authored several influential books, including "Introduction to Statistical Learning" and "Applied Predictive Modeling."

Daniela Witten: Daniela Witten is a professor of biostatistics and statistics at the University of Washington. Her research focuses on developing statistical methods for high-dimensional data analysis, with applications in genomics and neuroscience. Witten has received several prestigious awards for her contributions to statistical methodology and has been recognized as one of the leading young statisticians in the field.

Trevor Hastie: Trevor Hastie is a professor of statistics and data science at Stanford University. He has made significant contributions to the fields of statistical learning and machine learning, particularly in the development of tree-based methods, regularization techniques, and statistical modeling. Hastie has authored or co-authored numerous influential books, including "The Elements of Statistical Learning" and "Statistical Learning with Sparsity."

Robert Tibshirani: Robert Tibshirani is a professor of health research and policy and statistics at Stanford University. He is widely recognized for his contributions to statistical learning and the development of regularization methods, including the Lasso and Elastic Net. Tibshirani has received numerous awards, including the prestigious COPSS Presidents' Award, for his influential work in statistics and its applications.

The authors' collective contributions have had a significant impact on the fields of statistics, machine learning, and data science. Their research and publications have not only advanced the theoretical foundations of statistical learning but also provided practical tools and techniques for analyzing complex datasets in various domains. Their work has greatly influenced the way statistical modeling and prediction are conducted, making statistical learning accessible to a broader audience.

Appreciation and Critique

"Introduction to Statistical Learning" is a widely acclaimed textbook written by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. As a reviewer, I have thoroughly examined the book and its contributions to the literature in the domain of statistical learning. This review aims to provide an appreciation of the book's strengths while also

highlighting certain areas that could benefit from improvement.

Expected Contributions to the Literature

"Introduction to Statistical Learning" is a valuable addition to the literature on statistical learning. Its strength lies in its accessibility and practical approach to teaching complex concepts. By bridging the gap between theory and application, the book empowers readers to apply statistical learning methods effectively. It serves as an excellent introductory resource for students, researchers, and practitioners in the fields of statistics, machine learning, and data science.

However, to further contribute to the literature, future editions of the book could expand on advanced topics and provide more detailed discussions on cutting-edge techniques. Additionally, addressing the mathematical rigor to enhance accessibility would widen the book's appeal and make it accessible to a broader audience.

"Introduction to Statistical Learning" is an invaluable resource that offers a comprehensive introduction to statistical learning methods. It excels in providing clear explanations, practical examples, and a solid foundation in statistical learning. While there is room for improvement in terms of mathematical rigor and coverage of advanced topics, the book makes a significant contribution to the literature by empowering readers with practical skills for data analysis and prediction. Overall, it is a commendable resource for anyone interested in the field of statistical learning.

difficult to interpret or can easily be misinterpreted. The first issue the author discussed is that of aspect ratio i.e., relative dimensions of the graph. Giving clear examples, the author describes how choosing an incorrect aspect ratio can visually convey wrong information if the person reading the graph is not careful. Proceeding further, the author discusses pitfalls related to change of scale with a graph, problems with inverted representation, etc. The author also discussed about charts like stacked area bar charts which are difficult to visually process accurately.

In the fifth chapter, the author explains some basic guidelines for creating attractive and informative graphs. Intuitive yet important concepts like the usage of grid lines, how much number precision should be displayed, coloring schemes for displaying multiple lines/bars, etc. are commented upon by the author. Then the author shows the proper usage of pie charts and bar charts.

The sixth chapter draws from concepts of psychology and information processing and suggests some key attributes like form, motion, and spatial position for creating compelling graphs. The importance of text content while displaying graphs is also discussed at length and suggestion regarding how to improve storytelling using texts is discussed at length.

Overall, the book covers all major types of graphs that are used in research and academia. The book is easy to read and understand and the usage of colorful images makes it easy to understand. The author has taken several examples from different domains (like that of finance, marketing, and sales, etc.) which makes this book ideal for students in any specialization. Illustrative examples are often based on data from Indian companies which makes it easier for students to understand the story that is being told using the graphs. Psychological aspects of the visual processing of data are also discussed which will help students decide on the techniques that can be used to create greater pre-processing impact in the mind of the readers. Several practice problems and their solution are also presented so that readers can practice to better understand the concepts through actual implementation.

Although the book focuses on the conceptual side, it can be greatly enhanced by providing little coding in terms of how to implement the charts in Excel, etc. Further, a brief explanation of the calculations associated with the charts could help students understand the logic of the graph better. Finally, some more sample datasets can be provided to the students for more practice.

The book will be very useful for students who want to learn about basic data visualization techniques irrespective of their domain. This book can also be used in conjunction with other textbooks in introductory statistical courses as well.

Mr. Mithun
Co-ordinator, Business Analytics Programme
Rashtreeya Vidyalaya Institute of Management (RVIM)
#CA-17, 36 th Cross, 26 th Main
4 th "T" Block, Jayanagar
Bengaluru – 560041, India.
E-mail: mithun.rvim@rvei.edu.in