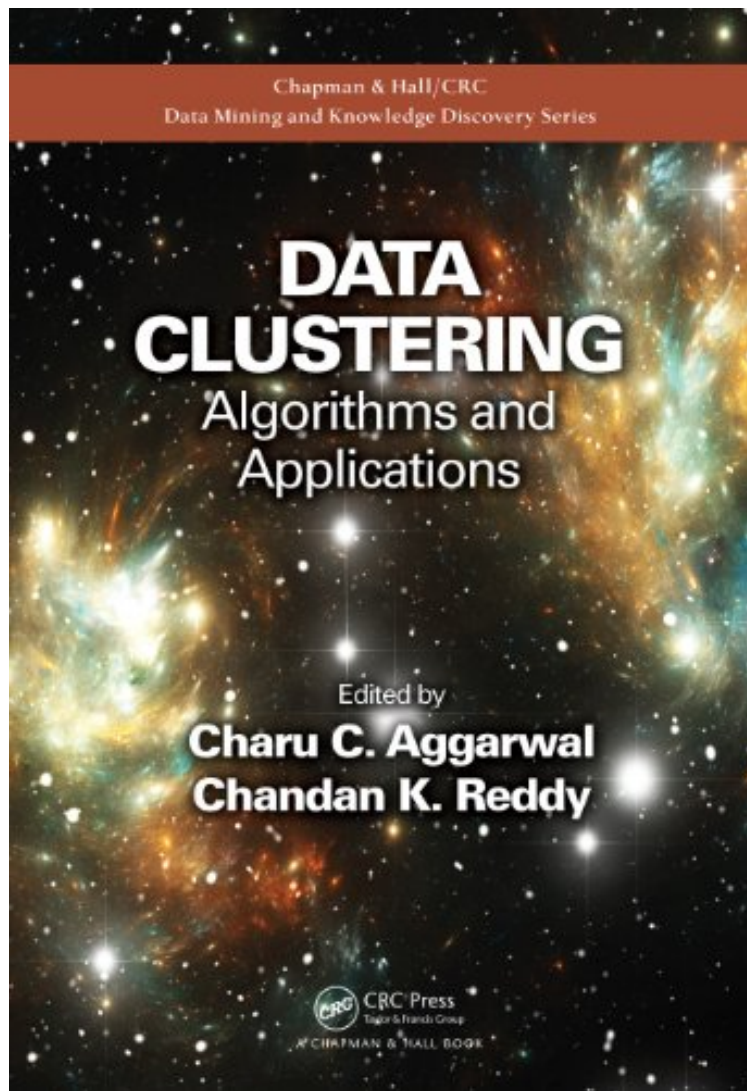


(Pdf free) Data Clustering: Algorithms and Applications (Chapman Hall/CRC Data Mining and Knowledge Discovery Series)

Data Clustering: Algorithms and Applications (Chapman Hall/CRC Data Mining and Knowledge Discovery Series)

*From Chapman and Hall/CRC
ePub | *DOC | audiobook | ebooks | Download PDF*



[Download](#)

[Read Online](#)

#377429 in eBooks 2013-08-21 2013-08-21 File Name: B00EYROAQU | File size: 69.Mb

From Chapman and Hall/CRC : Data Clustering: Algorithms and Applications (Chapman Hall/CRC Data Mining and Knowledge Discovery Series) before purchasing it in order to gage whether or not it would be worth my time, and all praised Data Clustering: Algorithms and Applications (Chapman Hall/CRC Data Mining and Knowledge Discovery Series):

4 of 4 people found the following review helpful. Must buy for Data Scientists who want to really understand clustering..By Robert J. LakeI have read this book twice. Can not put it down. As a practicing Data Scientist it has

really helped me to tune my clustering algorithms. Well worth the money.0 of 0 people found the following review helpful. Very comprehensive and up-to-dateBy Kevin S. GrayVery comprehensive and up-to-date. My main issue was that some chapters could have used better editing.7 of 8 people found the following review helpful. Worth the readBy JonquilleGood collection of articles - freely accessible online.Check out Taming Text, Programming Collective Intelligence, and SciKit-Learnfor less mathematical, more road-ready, how-to with code.For more accessible reading/understanding, check out Stanford's Text Mining book (free online).

Research on the problem of clustering tends to be fragmented across the pattern recognition, database, data mining, and machine learning communities. Addressing this problem in a unified way, *Data Clustering: Algorithms and Applications* provides complete coverage of the entire area of clustering, from basic methods to more refined and complex data clustering approaches. It pays special attention to recent issues in graphs, social networks, and other domains. The book focuses on three primary aspects of data clustering: **Methods**, describing key techniques commonly used for clustering, such as feature selection, agglomerative clustering, partitional clustering, density-based clustering, probabilistic clustering, grid-based clustering, spectral clustering, and nonnegative matrix factorization; **Domains**, covering methods used for different domains of data, such as categorical data, text data, multimedia data, graph data, biological data, stream data, uncertain data, time series clustering, high-dimensional clustering, and big data; and **Variations and Insights**, discussing important variations of the clustering process, such as semisupervised clustering, interactive clustering, multiview clustering, cluster ensembles, and cluster validation. In this book, top researchers from around the world explore the characteristics of clustering problems in a variety of application areas. They also explain how to glean detailed insight from the clustering process, including how to verify the quality of the underlying clusters through supervision, human intervention, or the automated generation of alternative clusters.

About the Author Charu C. Aggarwal is a Research Scientist at the IBM T. J. Watson Research Center in Yorktown Heights, New York. He completed his B.S. from IIT Kanpur in 1993 and his Ph.D. from Massachusetts Institute of Technology in 1996. His research interest during his Ph.D. years was in combinatorial optimization (network flow algorithms), and his thesis advisor was Professor James B. Orlin. He has since worked in the field of performance analysis, databases, and data mining. He has published over 200 papers in refereed conferences and journals, and has applied for or been granted over 80 patents. He is author or editor of nine books, including this one. Because of the commercial value of the above-mentioned patents, he has received several invention achievement awards and has thrice been designated a Master Inventor at IBM. He is a recipient of an IBM Corporate Award (2003) for his work on bio-terrorist threat detection in data streams, a recipient of the IBM Outstanding Innovation Award (2008) for his scientific contributions to privacy technology, and a recipient of an IBM Research Division Award (2008) for his scientific contributions to data stream research. He has served on the program committees of most major database/data mining conferences, and served as program vice-chairs of the SIAM Conference on Data Mining, 2007, the IEEE ICDM Conference, 2007, the WWW Conference 2009, and the IEEE ICDM Conference, 2009. He served as an associate editor of the IEEE Transactions on Knowledge and Data Engineering Journal from 2004 to 2008. He is an associate editor of the ACM TKDD Journal, an action editor of the Data Mining and Knowledge Discovery Journal, an associate editor of the ACM SIGKDD Explorations, and an associate editor of the Knowledge and Information Systems Journal. He is a fellow of the IEEE for "contributions to knowledge discovery and data mining techniques", and a life-member of the ACM. Chandan K. Reddy is an Assistant Professor in the Department of Computer Science at Wayne State University. He received his PhD from Cornell University and MS from Michigan State University. His primary research interests are in the areas of data mining and machine learning with applications to healthcare, bioinformatics, and social network analysis. His research is funded by the National Science Foundation, Department of Transportation, and the Susan G. Komen for the Cure Foundation. He has published over 40 peer-reviewed articles in leading conferences and journals. He received the Best Application Paper Award at the ACM SIGKDD conference in 2010 and was a finalist of the INFORMS Franz Edelman Award Competition in 2011. He is a member of IEEE, ACM, and SIAM.