1. Record Nr. UNINA9910437560803321

Autore	Cordeiro Robson L. F
Titolo	Data mining in large sets of complex data / / Robson L. F. Cordeiro, Christos Faloutsos, Caetano Traina Junior
Pubbl/distr/stampa	London ; ; New York, : Springer, c2013
ISBN	1-299-19712-4 1-4471-4890-8
Edizione	[1st ed. 2013.]
Descrizione fisica	1 online resource (xi, 116 pages) : illustrations (some color)
Collana	SpringerBriefs in computer science, , 2191-5768
Classificazione	ST 530
Altri autori (Persone)	FaloutsosChristos Traina JuniorCaetano
Disciplina	006.312
Soggetti	Data mining Database searching
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	"ISSN: 2191-5768."
Nota di bibliografia	Includes bibliographical references.
Nota di contenuto	Preface -- Introduction -- Related Work and Concepts -- Clustering Methods for Moderate-to-High Dimensionality Data -- Halite -- BoW -- QMAS -- Conclusion.
Sommario/riassunto	The amount and the complexity of the data gathered by current enterprises are increasing at an exponential rate. Consequently, the analysis of Big Data is nowadays a central challenge in Computer Science, especially for complex data. For example, given a satellite image database containing tens of Terabytes, how can we find regions aiming at identifying native rainforests, deforestation or reforestation? Can it be made automatically? Based on the work discussed in this book, the answers to both questions are a sound “yes”, and the results can be obtained in just minutes. In fact, results that used to require days or weeks of hard work from human specialists can now be obtained in minutes with high precision. Data Mining in Large Sets of Complex Data discusses new algorithms that take steps forward from traditional data mining (especially for clustering) by considering large, complex datasets. Usually, other works focus in one aspect, either data size or complexity. This work considers both: it enables mining complex data from high impact applications, such as breast cancer diagnosis, region classification in satellite images, assistance to climate

change forecast, recommendation systems for the Web and social networks; the data are large in the Terabyte-scale, not in Giga as usual; and very accurate results are found in just minutes. Thus, it provides a crucial and well timed contribution for allowing the creation of real time applications that deal with Big Data of high complexity in which mining on the fly can make an immeasurable difference, such as supporting cancer diagnosis or detecting deforestation.