

1. Record Nr.	UNINA9910483961103321
Titolo	Clustering High--Dimensional Data : First International Workshop, CHDD 2012, Naples, Italy, May 15, 2012, Revised Selected Papers // edited by Francesco Masulli, Alfredo Petrosino, Stefano Rovetta
Pubbl/distr/stampa	Berlin, Heidelberg : , : Springer Berlin Heidelberg : , : Imprint : Springer, , 2015
ISBN	3-662-48577-X
Edizione	[1st ed. 2015.]
Descrizione fisica	1 online resource (IX, 149 p. 41 illus. in color.)
Collana	Information Systems and Applications, incl. Internet/Web, and HCI ; ; 7627
Disciplina	005.74
Soggetti	Database management Application software Artificial intelligence Information storage and retrieval Data mining Algorithms Database Management Information Systems Applications (incl. Internet) Artificial Intelligence Information Storage and Retrieval Data Mining and Knowledge Discovery Algorithm Analysis and Problem Complexity
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Intro -- Preface -- Organization -- Contents -- Clustering High-Dimensional Data -- 1 Introduction -- 2 Defining Clustering -- 3 The Century of Big Data -- 4 Approaches to High Dimensional Data Clustering -- 4.1 Subspace Clustering -- 4.2 Projected Clustering -- 4.3 Biclustering -- 4.4 Hierarchical Clustering -- 5 Conclusions -- References -- What are Clusters in High Dimensions and are they Difficult to Find? -- 1 Introduction -- 2 Properties of High-Dimensional Data -- 3 Cluster Analysis -- 4 What are Clusters, Especially in Higher Dimensions? -- 5 Consequences for Clustering Algorithms -- 6

Conclusions -- References -- Efficient Density-Based Subspace Clustering in High Dimensions -- 1 Introduction -- 2 Density-Based Subspace Clustering -- 3 Dimensionality Unbiased Density -- 4 Redundancy-Removal -- 5 Pruning Subspace Clusters -- 6 Indexing Subspace Clustering -- 7 Approximate Jump Clustering -- 8 Conclusion -- References -- Comparing Fuzzy Clusterings in High Dimensionality -- 1 Introduction -- 2 Fuzzy Clustering -- 2.1 Some Notations and Definitions -- 2.2 Fuzzy Clustering -- 2.3 Methods for Fuzzy Clustering -- 2.4 Possibilistic Clustering Models -- 2.5 Graded Possibilistic Models -- 3 Comparing Fuzzy Clusterings -- 3.1 Approaches to the Comparison of Clusterings -- 3.2 Notation -- 3.3 Co-association -- 3.4 Fuzzy Coassociation -- 3.5 Comparing Two Partitions -- 4 Partition Similarity Indexes -- 4.1 The Rand and Jaccard Indexes -- 4.2 The Fuzzy Jaccard Index -- 4.3 The Fuzzy Rand Index -- 4.4 The Probabilistic Rand Index -- 4.5 The Probabilistic Jaccard Index -- 5 Applications of Fuzzy Similarity Indexes -- 5.1 Visual Stability Analysis Based on Comparing Fuzzy Clusterings -- 5.2 Tracking Deterministic Annealing -- 6 Conclusion -- References -- Time Series Clustering from High Dimensional Data -- 1 Introduction. 2 Financial High Dimensional Data Characteristics -- 3 Beanplot Time Series -- 4 Parameterizing Beanplot Time Series Data -- 5 Time Series Factor Analysis on Beanplot Time Series -- 6 From Time Series Factor Analysis to the Feature Clustering Approach -- 7 Using the Self Organizing Maps -- 8 Simulation Study -- 9 Application on Real Data -- 10 Conclusions -- References -- Data Dimensionality Estimation: Achievements and Challenges -- 1 Introduction -- 2 Global Methods -- 2.1 Projection Techniques -- 2.2 Fractal-Based Methods -- 2.3 Multidimensional Scaling and Other Methods -- 3 Local Methods -- 3.1 Fukunaga-Olsen's Algorithm -- 3.2 TRN-Based and Local MDS Methods -- 4 Mixed Methods -- 4.1 Levina-Bickel Algorithm -- 5 ID Estimation Methods Benchmarking -- 6 Conclusions -- References -- A Novel Intrinsic Dimensionality Estimator Based on Rank-Order Statistics -- 1 Introduction -- 2 Related Works -- 3 Theoretical Results -- 4 The Algorithm -- 5 Algorithm Evaluation -- 5.1 Dataset Description -- 5.2 Experimental Setting -- 5.3 Experimental Results -- 6 Conclusions and Future Works -- A Algorithm Implementation -- References -- Dimensionality Reduction in Boolean Data: Comparison of Four BMF Methods -- 1 Matrix Decompositions, Dimensionality Reduction, and Boolean Data -- 2 Boolean Matrix Factorization -- 3 The Four Methods Being Compared -- 4 Experimental Comparison -- 4.1 Method of Comparison -- 4.2 Datasets Used -- 4.3 Results -- 5 Conclusions and Further Issues -- References -- A Rough Fuzzy Perspective to Dimensionality Reduction -- 1 Introduction -- 2 Related Works -- 3 Rough-Fuzzy Sets -- 4 Rough-Fuzzy Product Feature Selection -- 4.1 Feature Granularization -- 4.2 Feature Selection -- 5 Experimental Results -- 6 Conclusions -- References -- Author Index.

Sommario/riassunto

This book constitutes the proceedings of the International Workshop on Clustering High-Dimensional Data, CHDD 2012, held in Naples, Italy, in May 2012. The 9 papers presented in this volume were carefully reviewed and selected from 15 submissions. They deal with the general subject and issues of high-dimensional data clustering; present examples of techniques used to find and investigate clusters in high dimensionality; and the most common approach to tackle dimensionality problems, namely, dimensionality reduction and its application in clustering.