

1. Record Nr.	UNINA9910494563103321
Autore	Scitovski Rudolf
Titolo	Cluster Analysis and Applications / / by Rudolf Scitovski, Kristian Sabo, Francisco Martínez-Álvarez, Šime Ungar
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2021
ISBN	3-030-74552-X
Edizione	[1st ed. 2021.]
Descrizione fisica	1 online resource (277 pages)
Collana	Computer Science Series
Disciplina	519.53
Soggetti	Artificial intelligence - Data processing Computer science Pattern recognition systems Artificial intelligence Algorithms Machine learning Data Science Theory and Algorithms for Application Domains Automated Pattern Recognition Artificial Intelligence Design and Analysis of Algorithms Machine Learning
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	1 -- Introduction. 2 Representatives -- 2.1 Representative of data sets with one feature. 2.1.1. Best LS-representative -- 2.1.2 Best ℓ_1 -representative -- 2.1.3 Best representative of weighted data -- 2.1.4 Bregman divergences -- 2.2 Representative of data sets with two features -- 2.2.1 Fermat–Torricelli–Weber problem -- 2.2.2 Centroid of a set in the plane -- 2.2.3 Median of a set in the plane -- 2.2.4 Geometric median of a set in the plane -- 2.3 Representative of data sets with several features -- 2.3.1 Representative of weighted data -- 2.4 Representative of periodic data -- 2.4.1 Representative of data on the unit circle -- 2.4.2 Burn diagram -- 3 Data clustering -- 3.1

Optimal k-partition -- 3.1.1 Minimal distance principle and Voronoi diagram -- 3.1.2 k-means algorithm -- 3.2 Clustering data with one feature -- 3.2.1 Application of the LS-distance-like function -- 3.2.2 The dual problem -- 3.2.3 Least absolute deviation principle -- 3.2.4 Clustering weighted data -- 3.3 Clustering data with two or several features -- 3.3.1 Least squares principle -- 3.3.2 The dual problem -- 3.3.3 Least absolute deviation principle -- 3.4 Objective function $F(c_1, \dots, c_k) = \sum_{i=1}^n \min_{j=1, \dots, k} d(c_j, a_i)$ -- 4 Searching for an optimal partition -- 4.1 Solving the global optimization problem directly -- 4.2 k-means algorithm II -- 4.2.1 Objective function F using the membership matrix -- 4.2.2 Coordinate Descent Algorithms -- 4.2.3 Standard k-means algorithm -- 4.2.4 k-means algorithm with multiple activations -- 4.3 Incremental algorithm -- 4.4 Hierarchical algorithms -- 4.4.1 Introduction and motivation -- 4.4.2 Applying the Least Squares Principle. 4.5 DBSCAN method -- 4.5.1 Parameters MinPts and 97 4.5.2 DBSCAN algorithm -- 4.5.3 Numerical examples -- 5 Indexes -- 5.1 Choosing a partition with the most appropriate number of clusters -- 5.1.1 Calinski–Harabasz index -- 5.1.2 Davies–Bouldin index -- 5.1.3 Silhouette Width Criterion -- 5.1.4 Dunn index -- 5.2 Comparing two partitions -- 5.2.1 Rand index of two partitions -- 5.2.2 Application of the Hausdorff distance -- 6 Mahalanobis data clustering -- 6.1 Total least squares line in the plane. 6.2 Mahalanobis distance-like function in the plane -- 6.3 Mahalanobis distance induced by a set in the plane -- 6.3.1 Mahalanobis distance induced by a set of points in R^n -- 6.4 Methods to search for optimal partition with ellipsoidal clusters -- 6.4.1 Mahalanobis k-means algorithm 139 CONTENTS v -- 6.4.2 Mahalanobis incremental algorithm -- 6.4.3 Expectation Maximization algorithm for Gaussian mixtures -- 6.4.4 Expectation Maximization algorithm for normalized Gaussian mixtures and Mahalanobis k-means algorithm -- 6.5 Choosing partition with the most appropriate number of ellipsoidal clusters -- 7 Fuzzy clustering problem -- 7.1 Determining membership functions and centers -- 7.1.1 Membership functions. 7.1.2 Centers -- 7.2 Searching for an optimal fuzzy partition with spherical clusters -- 7.2.1 Fuzzy c-means algorithm -- 7.2.2 Fuzzy incremental clustering algorithm (FInc) -- 7.2.3 Choosing the most appropriate number of clusters -- 7.3 Methods to search for an optimal fuzzy partition with ellipsoidal clusters -- 7.3.1 Gustafson–Kessel c-means algorithm -- 7.3.2 Mahalanobis fuzzy incremental algorithm (MFInc) -- 7.3.3 Choosing the most appropriate number of clusters -- 7.4 Fuzzy variant of the Rand index -- 7.4.1 Applications -- 8 Applications -- 8.1 Multiple geometric objects detection problem and applications -- 8.1.1 Multiple circles detection problem -- 8.1.2 Multiple ellipses detection problem -- 8.1.3 Multiple generalized circles detection problem -- 8.1.4 Multiple lines detection problem -- 8.1.5 Solving MGOD-problem by using the RANSAC method -- 8.2 Determining seismic zones in an area -- 8.2.1 Searching for seismic zones -- 8.2.2 The absolute time of an event -- 8.2.3 The analysis of earthquakes in one zone -- 8.2.4 The wider area of the Iberian Peninsula -- 8.2.5 The wider area of the Republic of Croatia -- 8.3 Temperature fluctuations -- 8.3.1 Identifying temperature seasons -- 8.4 Mathematics and politics: How to determine optimal constituencies? -- 8.4.1 Mathematical model and the algorithm -- 8.4.2 Defining constituencies in the Republic of Croatia -- 8.4.3 Optimizing the number of constituencies -- 8.5 Iris -- 8.6 Reproduction of Escherichia coli. 9 Modules and the data sets -- 9.1 Functions -- 9.2 Algorithms -- 9.3 Data generating -- 9.4 Test examples -- 9.5 Data sets -- Bibliography -- Index.

amount of data and wide availability of tools for processing these data, the biggest limitation is the lack of trained experts who are qualified to process and interpret the results. This textbook is intended for graduate students and experts using methods of cluster analysis and applications in various fields. With clear explanations of ideas and precise definitions of notions, accompanied by numerous examples and exercises together with Mathematica programs and modules, Cluster Analysis and Applications is meant for students and researchers in various disciplines, working in data analysis or data science.
