



Contribution ID: 30

Type: **not specified**

Partitioning Metric Space Data

Wednesday, 13 September 2023 08:50 (20 minutes)

The partitioning of the data into clusters, carried out by the researcher in accordance with a certain criterion, is a necessary step in the study of a particular phenomenon. Subsequent research should confirm or refute the appropriateness of such a division, and in a positive case, evaluate the discriminating power of the criterion (or, in other words, the influencing power of the factor according to the level of which the data was divided). If the data comes from a metric space, this means that for any pair of data, a distance is defined that characterizes the dissimilarity between them. Speaking of data, we are not necessarily talking about numbers, it can be information of any kind about the objects under study (such as spectrograms, 3B forms, etc.) obtained as a result of measurement, observation, query, etc., however distance between data, expressing how far apart the objects of interest are represented by a scalar. The correct choice of the distance metric is a fundamental problem in quality control, pattern recognition, machine learning, cluster analysis, etc. We propose two universal discriminating statistics - SP (segregation power) based on the ratio and the difference of inter to intra clusters' correlated estimates of the distance between objects and discuss their specificity and sensitivity as well as their universalism and robustness in relation to the type of objects under study.

Keywords

data partitioning, segregation, clustering

Classification

Both methodology and application

Primary authors: MARMOR, Yariv (ORT Braude College of Engineering); Prof. BASHKANSKY, Emil (Braude College of Engineering)

Presenter: Prof. BASHKANSKY, Emil (Braude College of Engineering)

Session Classification: CONTRIBUTED Complex data and design