

## Fuzzy Optimization Multi-Objective Clustering Ensemble Model for Multi-Source Data Analysis

**Authors :** C. B. Le, V. N. Pham

**Abstract :** In modern data analysis, multi-source data appears more and more in real applications. Multi-source data clustering has emerged as a important issue in the data mining and machine learning community. Different data sources provide information about different data. Therefore, multi-source data linking is essential to improve clustering performance. However, in practice multi-source data is often heterogeneous, uncertain, and large. This issue is considered a major challenge from multi-source data. Ensemble is a versatile machine learning model in which learning techniques can work in parallel, with big data. Clustering ensemble has been shown to outperform any standard clustering algorithm in terms of accuracy and robustness. However, most of the traditional clustering ensemble approaches are based on single-objective function and single-source data. This paper proposes a new clustering ensemble method for multi-source data analysis. The fuzzy optimized multi-objective clustering ensemble method is called FOMOCE. Firstly, a clustering ensemble mathematical model based on the structure of multi-objective clustering function, multi-source data, and dark knowledge is introduced. Then, rules for extracting dark knowledge from the input data, clustering algorithms, and base clusterings are designed and applied. Finally, a clustering ensemble algorithm is proposed for multi-source data analysis. The experiments were performed on the standard sample data set. The experimental results demonstrate the superior performance of the FOMOCE method compared to the existing clustering ensemble methods and multi-source clustering methods.

**Keywords :** clustering ensemble, multi-source, multi-objective, fuzzy clustering

**Conference Title :** ICMLC 2021 : International Conference on Machine Learning and Cybernetics

**Conference Location :** London, United Kingdom

**Conference Dates :** June 28-29, 2021