Optimization overlap clustering based on the hybrid rough discernibility concept and rough K-Means

Subtitle:

Article type: Research Article

Authors: Setyohadi, Djoko Budyanto* | Bakar, Azuraliza Abu | Othman, Zulaila Ali

Affiliations: Data Mining and Optimization Research Group, Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, University Kebangsaan Malaysia, Bangi, Selangor Darul Ehsan, Malaysia

Correspondence: [*] Corresponding author: Djoko Budyanto Setyohadi, Data Mining and Optimization Research Group, Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, University Kebangsaan Malaysia, Bangi, Selangor Darul Ehsan 43000, Malaysia. E-mail: djokobdy@gmail.com

Abstract: Technically, the problem of overlap in a dataset is viewed as an uncertainty problem and is solved using a fuzzy set theoretical approach, specifically, fuzzy clustering. This approach is powerful but has some problems associated with it, of which the design of the membership function is the most serious. There are many different techniques for optimizing fuzzy clustering, including those based on similarity decomposition and centroids of clusters. Furthermore, the problem of overlap clustering is still being studied to improve its performance, especially with respect to the membership optimization. Rough set theory (RST) is the complement of fuzzy set theory and evidence theory, which use different techniques to address the uncertainty problem in overlap clustering. Considering the simplicity of the membership computation in RST, we propose an overlap clustering algorithm, which involves the use of the discernibility concept of RST to improve the overlap clusters as an existing variant of the overlap clustering algorithm. The experiment described here demonstrates that this new method improves the performance and increases the accuracy of clustering while avoiding the time-consuming optimization process.