Clustering by two methods simultaneously

Nánásiová Oľga[1], Sabo Miroslav[2]

SvF STU Radlinského 11, Bratislava, Slovak Republic
[1]nanasiova@stuba.sk, [2]sabo@math.sk

Clustering is a method of unsupervised learning, that finds clusters (objects with similar properties) in data. Since fifties, there were proposed many different clustering algorithms. Our method combine two arbitrary clustering algorithms to obtain more accurate results. Method is based on concept of similarity coefficients. Also practical application on hydrological data in R language will be shown.

Key words: Similarity coefficients, symmetric diference, cluster analysis

References:
[1] Dalirsefat, S.B., Meyer.A, & Mirhoseini, S.Z. 2009: Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silkworm, Bombyx mori. Journal of Insect Science , Vol.9, No.71; 681-689.
[2] Haldiki, M., Batistakis,Y., & Vazirgiannis, M. 2001: On Clustering Validation Techniques.Journal of Intelligent Information Systems,Vol.17, No 2/3; 107-145
[3] Jackson, A.D., Somers K.M., & Harvey, H.H. 1989: Similarity coefficients: Measures of co-occurrence and association or simply measures of occurrence? The American Naturalist Vol.133, No.3; 436-453.
[4] Murgula, M.& Villasenor, J.L. 2003: Estimating the effect of similarity coefficient and the cluster algorithm on biogeographic classifications. Ann. Bot. Fennici 40; 415-421