Different Clustering Techniques – Means for Improved Knowledge Discovery

OLIVERA GRLJEVIC *(University of Novi Sad, Faculty of Economics, Serbia, oliverag@eccf.su.ac.yu)*
SASA BOSNJAK *(University of Novi Sad, Faculty of Economics, Serbia, bsale@eccf.su.ac.yu)*
ZITA BOSNJAK *(University of Novi Sad, Faculty of Economics, Serbia, bzita@eccf.su.ac.yu)*

Each clustering technique, sometimes even the same algorithm applied several times on initial data set, can result in different source data set partitions putting an accent on a specific aspect of resulting clusters. Apart from diverse outputs, clustering algorithms use different visualization techniques that enable better insight into the structure of derived clusters and grouping relationships of similar entities. This variety of visualization techniques, that different clustering techniques offer, makes possible for a much broader and even non-technical audience, to understand the discovered relationships among data. Furthermore, each clustering technique comes with some specifics; one can denote clustering centers, other will emphasize typical representatives of resulting clusters or least typical cluster instances, etc. Each additional information attained from different tools is very useful for analysts, given the fact that they have a great responsibility to successfully carry out an interpretation of the results gained through some of the available tools and to give meaning to what makes a qualitative set of clusters.
In this article we presented the clustering results of small and medium sized enterprises' (SME) data in the province of Vojvodina, attained in DataEngine, Weka and iDataAnalyzer tools for intelligent analysis. Each tool supports a different clustering algorithm. We proposed the composite approach that implies diversity of tools and hence we obtained results that we believe significantly simplifies a work of analysts in knowledge discovery, facilitates results interpretation and derivation of detailed and clear conclusions.

Keywords: DataEngine, Weka, iDataAnalyser, clustering techniques