Dynamic data assigning assessment clustering of streaming data, Applied Soft Computing- Special Issue on Dynamic Data Mining

Abstract

Discovering interesting patterns or substructures in data streams is an important challenge in data mining. Clustering algorithm are very often applied to identify substructures, although they are designed to partition a data set. Another problem of clustering algorithms is that most of them are not designed for data streams. They assume that the data set to be analysed is already complete and will not be extended by new data. This paper discusses an extension of an algorithm that uses ideas from cluster analysis, but was designed to identify single clusters in large data sets without the necessity to partition the whole data set into clusters. The new extended version of this algorithm can applied to stream data and is able to identify new clusters in an incoming data stream. As a case study weather data are used.