Abstract

Incremental K-means and DBSCAN are two very important and popular clustering techniques for today’s large dynamic databases (Data warehouses, WWW and so on) where data are changed at random fashion. The performance of the incremental K-means and the incremental DBSCAN are different with each other based on their time analysis characteristics. Both
algorithms are efficient compared to their existing algorithms with respect to time, cost and effort. In this paper, the performance evaluation of incremental DBSCAN clustering algorithm is implemented and most importantly it is compared with the performance of incremental K-means clustering algorithm and it also explains the characteristics of these two algorithms based on the changes of the data in the database. This paper also explains some logical differences between these two most popular clustering algorithms. This paper uses an air pollution database as original database on which the experiment is performed.

Reference

- Xiang Li, Rahul Ramachandran, Sunil Movva and Sara Graves, “Storm Clustering for Data-driven Weather Forecasting”, International conference in University of Alabama in Huntsville.
- Ester, M., Kriegel, H., Sander, J, Wimmer, M., Xiaowei Xu, “Incremental
clustering for mining in a data ware Housing”, University of Munich Oettingenstr.67, D-80538 München, Germany.

Index Terms

Computer Science

Data Mining

Key words

Clustering

DBSCAN

Incremental

K-means

Threshold