Abstract

Data stream mining is an emerging area for extracting useful information from continuous arriving data. Web click stream, weather monitoring, network traffic, shopping history, web log are some key resources of generating data stream. Clustering is one of the most useful technique for analyzing stream data, as it does not require any predefined class labeling. Data stream mining is challenging as the data is massive and arriving continuously. The traditional clustering algorithms cannot be directly applied on the data streams. Data stream mining needs one scan algorithms to extract rich data in the form of data streams. In this paper we discuss various data stream clustering algorithms with their limitations and required data structures. This paper also provides a comparative study of these algorithms. Real world applications of data streams, data resources and publicly available softwares are also discussed.

References

1. Gaber, Mohamed Medhat, Arkady Zaslavsky, and Shonali Krishnaswamy. "Mining data
A Survey on Clustering Algorithms for Data Streams


Index Terms

Computer Science

Algorithms
Keywords

Data Mining, Data Stream, Clustering