Abstract

Data mining provides a way for finding hidden and useful knowledge from the large amount of data. Usually, we find any information by finding normal trends or distribution of data. But sometimes, rare events or data objects may provide information which is very interesting to us. Outlier detection is one of the tasks of data mining. It finds abnormal data points or sequences hidden in the dataset. Data stream is an unbounded sequence of data with explicit or implicit temporal context. Data stream is uncertain and dynamic in nature. Traditional outlier detection techniques for static data, which require the whole dataset for modeling, are not suitable for data stream because the whole data stream cannot be stored. Network intrusion detection, web clickstream analysis, fraud detection, fault detection in machines, sensor data analysis are some of the applications of data stream outlier detection. In this paper, we have described several issues in data stream outlier detection and usual approaches or techniques for finding outlier in data stream.
1. Jiawei Han, Micheline Kamber and Jian Pei, “Data mining Concepts and Techniques”, Third Edition, Morgan Kaufmann Series in Data management Systems.
8. Manish Gupta, Jing Gao, Member, IEEE, Charu C. Aggarwal, Fellow, IEEE, and Jiawei Han, Fellow, IEEE, “Outlier Detection for Temporal Data: A Survey”, IEEE.

Index Terms

Computer Science

Information Sciences
Keywords

Data mining, Outliers, data stream mining.