Abstract

Current research in data mining concentrates on the development of new techniques for mining high-speed data streams. The fundamental data generation mechanism changes over the time, this is common in most real-world data streams, which introduces concept drift into the data. Mobile devices, streaming, remote sensing applications which are networked digital information systems, encounter the issue of the size of data and the capacity to be adaptive to changes in concept in real-time. In this paper the main issue of concept drift is addressed with real and synthetic data streams and the comparison of ensemble classifiers has been made in view of concept drift for the assessment of the performance. Various classifiers were applied on data stream with and without concept drift for analysis. This has resulted in better performance of the classifiers on the type of data whether it is categorical, numeric or alphanumeric.

References

Analyzing Performance of Classification Algorithms on Concept Drifted Data Streams


Index Terms

Computer Science  Algorithms
Keywords

Data mining, Data Stream, Concept Drift, Classification