Abstract

Traditional data mining classifiers are used for mining the static data, in which incremental learning assumed data streams come under stationary distribution where data concepts remain unchanged. The concept of data can be changed at any time in real world application this refers to change in the class definitions over time. Classifier ensembles are rapidly gaining popularity in data mining Community, because they are comparatively more accurate, easy and react better to concept drift than single classifiers. They are general way of boosting classification accuracy. Their modularity provides natural path of absorbing changes by modifying ensemble member. The proposed approach uses ensemble classifiers to improve the accuracy of the classification in data streams. The performance of the classifiers tested with benchmark datasets from UCI machine learning repository. the experimental results prove that this approach great accuracy when comparing to the single classifier.

- Pedro Domingos, Geoff Hulten, "Mining High Speed Data Streams", KDD-00 in proceeding of sixth ACM SIGKDD international conference on knowledge discovery and data mining, USA, 2000, pp 71-80.

- Geoff Hulten, Laurie Spencer, Pedro Domingos, "Mining time changing data streams", ACM, USA, 2001, 97-106.

- Haixun Wang, Wei Fan, Philip S. Yu, Jiawei Han, "Mining Concept Drifting Data Streams Using Ensemble Classifiers", SIGKDD &apos;03, ACM, USA, 2003, pp 226-235.

- Mohammad M. Masud, Jing Gao, Latifur Khan, Jiawei Han, and Bhavani Thuraisingham, "A Multi-partition Multi-chunk Ensemble Technique to Classify Concept-Drifting Data Streams", Springer-Verlag, Berlin Heidelberg, 2009, pp 363–375.


**Index Terms**

Computer Science  
Data Mining
Keywords

Data stream  concept drift  boosting  static data.