Abstract

Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups. Boosting is the iterative process which aims to improve the predictive accuracy of the learning algorithms. Clustering with boosting improves quality of mining process. When supervised algorithms applied on training data for learning, there may be possibility of biased learning which affects the accuracy of prediction result. Boosting provides the solution for this. It generates subsequent classifiers by learning incorrect predicted examples by previous classifier. Boosting process possesses some limitations. Different approaches introduced to overcome the problems in boosting such as overfitting and troublesome area problem to improve performance and quality of the result. Cluster based boosting address limitations in boosting for supervised...
learning systems. In literature Cluster based boosting [6] is used to address limitations in
boosting for supervised learning systems. In paper [6], k-means is used as a clustering
algorithm. Encapsulation of another clustering method with CBB may result into increase in the
performance. In our proposed work we used fuzzy c means (FCM), Expectation Minimization
and Hierarchical algorithm with CBB and compared the results.

References

- L. Reyzin and R. Schapire, "How boosting the margin can also boost classifier
- D. Frossyniotis, A. Likas, and A. Stafylopatis, "A clustering method based on
- A. Vezhnevets and O. Barinova, "Avoiding boosting overfitting by removing
- Y. Sun, J. Li, and W. Hager, "Two new regularized adaboost algorithms," in Proc.
- A. Ganatra and Y. Kosta, "Comprehensive evolution and evaluation of
- Cluster-Based Boosting, L. Dee Miller and Leen-Kiat Soh, Member, IEEE, 2015
- W. Gao and Z-H. Zhou, "On the doubt about margin explanation of
- Yinghua Lu, Tinghuai Ma1, Changhong Yin2, Xiaoyu Xie2, Wei Tian and
ShuilMingZhong, "Implementation of the Fuzzy C-Means Clustering Algorithm in
Meteorological Data," 2013

Index Terms

Computer Science

Algorithms

Keywords

Clustering  Classification  Boosting  decision Tree