Abstract

In this paper a hybrid feature selection method is proposed which takes advantages of wrapper subset evaluation with a lower cost and improves the performance of a group of classifiers. The method uses combination of sample domain filtering and resampling to refine the sample domain and two feature subset evaluation methods to select reliable features. This method utilizes both feature space and sample domain in two phases. The first phase filters and resamples the sample domain and the second phase adopts a hybrid procedure by information gain, wrapper subset evaluation and genetic search to find the optimal feature space. Experiments carried out on different types of datasets from UCI Repository of Machine Learning databases and the results show a rise in the average performance of five classifiers (Naive Bayes, Logistic, Multilayer Perceptron, Best First Decision Tree and JRIP) simultaneously and the classification error for these classifiers decreases considerably. The experiments also show that this method outperforms other feature selection methods with a lower cost.
A Hybrid Feature Selection Method to Improve Performance of a Group of Classification Algorithms

- Novakovic, J., 2010. The Impact of Feature Selection on the Accuracy of Naïve Bayes Classifier, 18th Telecommunications Forum TELFOR, page(s) 1114-1116, November 23-25, Serbia, Belgrade.
Intrusion Detection Dataset for Selection of Relevance Features, Proceedings of the World Congress on Engineering and Computer Science, October 20-22, San Francisco, USA.


Index Terms

Computer Science Data Mining

Keywords

Feature Selection Resampling Information Gain Wrapper Subset Evaluation