Abstract

Feature subset selection is of immense importance in the field of data mining. The increased dimensionality of data makes testing and training of general classification method difficult. Mining on the reduced set of attributes reduces computation time and also helps to make the patterns easier to understand. In this paper a wrapper approach for feature selection is proposed. As a part of feature selection step we used wrapper approach with Genetic algorithm as random search technique for subset generation ,wrapped with different classifiers/ induction algorithm namely decision tree C4.5, NaïveBayes, Bayes networks and Radial basis function as subset evaluating mechanism on four standard datasets namely Pima Indians Diabetes Dataset, Breast Cancer, Heart Stat log and Wisconsin Breast Cancer. Further the relevant attributes identified by proposed wrapper are validated using classifiers. Experimental results illustrate, employing feature subset selection using proposed wrapper approach has enhanced classification accuracy.

Reference

Feature Subset Selection in Medical Data Mining. International Conference on IEEE International Advance Computing Conference (IACC’09), Thapar University, Patiala, Punjab India.


Index Terms

Computer Science  Data Mining
Key words

Feature Selection filters
wrappers