Abstract

Feature usability index is introduced here as a measure for evaluating classification efficacy of features. It is defined using measures of homogeneity, class specificity, and error in decision making. Homogeneity measures the extent of outlying observations, class specificity assesses the separation between distributions of different labeled classes, and error in decision making is computed using overlap in posteriori decision boundary. This is followed by feature ranking and optimal feature subset selection through ordering of features based on feature usability index and involves a complexity of O(DlogD) for D features. The results validating classifier independent feature ranking and optimal feature subset selection are also presented along with a comparative analysis using χ2 statistics for feature selection.

**Index Terms**

Computer Science

Algorithms

**Key words**

Feature ranking

feature selection

knowledge discovery

knowledge engineering

pattern recognition