

Feature Selection for Rule-Based Classification of Fish Types Using Information Gain

Yadanar Wai, Aye Aye
Computer University (Pyay)
yatanarwai456@gmail.com

Abstract

Classification is a process of finding the common properties among various features and classifying them into classes. The feature selection process is very important which selects a subset of relevant features, so that the selected subset is sufficient to perform the classification task. Classification of the real world problem can be made easier and less time consuming by removing the irrelevant features. In this paper, an efficient feature selection method based on information gain is presented for finding and selecting relevant features which maximized the accuracy of classifier. Features that are relevant for rule-based classifier are selected by using information gain. Information gain calculates the gain value of each feature in the dataset to select the best(relevant) feature subset. Then, rule-based classifier generates the rules by using selected(best) feature subset to classify the class label of an unknown sample. The feature selection based on information gain yields the significant improvement of accuracy by using cross-validation for classification of fish dataset from fish profile.com repository when using rule-based classifier.