Abstract

Naïve Bayes Classifier (NBC) is one of classification method which is simple and efficient, so that it employed widely to dealing with various problems of classification in real world. In its classification process, NBC requires probability of each attribute and its class label pair. The classification problems in real world involve numeric attributes and the ways to estimate its probability are different from categorical attributes. Discretization is a popular approach to map numeric attribute into categorical attribute which can reduce the number of classification errors on NBC.

A value in simple dicretization method must belong to one interval, whereas that value may have a relation with another interval. Discretization with Fuzzy approach makes a value belong to one or two interval when it located in a boundary of adjacent interval. However, to determine number of interval, boundary of interval, and cut point of adjacent interval is not easy. Therefore, Genetic Algorithm can used to optimize these parameters, so that Genetic Algorithm Fuzzy Discretization (GAFD) is a discretization method that will be used to improve the accuracy of NBC. Experimental result showed NBC with GAFD generated higher accuracy and smaller classification error than NBC without discretizaiton.

Key Words: Naïve Bayes Classifier, Fuzzy Discretization, Genetic Algorithm, Discretization.