Redução do conjunto de dados de treinamento para melhorar a eficiência do classificador SVM

Peterson Lima Sarmento, Guaraci José Erthal e Luciano Vieira Dutra

The Support Vector Machine classifier (SVM), a classifier that basically uses training data near the decision boundary is often used due to its good performance. A disadvantage of this method is the increase training time of the classifier as the training set size increases. The purpose of this paper is to apply techniques to reduce the training set size preserving the classification accuracy. The techniques used are editing, multiediting and condensing of data, already applied in data reduction for k-nearest neighbor method (k-NN), which is also a classifier that operates on the data near the border separating classes. It is expected that accuracy with reduced data sets and there is gain in time to estimate the classifier parameters.