Class Maps for Visualizing Classification Results

Peter Rousseeuw

KU Leuven

Abstract

Classification is a major tool of statistics and machine learning. A classification method first processes a training set of objects with given classes (labels), with the goal of afterward assigning new objects to one of these classes. When running the resulting prediction method on the training data or on test data, it can happen that an object is predicted to lie in a class that differs from its given label. This is sometimes called label bias, and raises the question whether the object was mislabeled. The proposed class map reflects the probability that an object belongs to an alternative class, howfar it is fromthe other objects in its given class, and whether some objects lie far fromall classes. The goal is to visualize aspects of the classification results to obtain insight in the data. The display is constructed for discriminant analysis, the k-nearest neighbor classifier, support vector machines, logistic regression, and coupling pairwise classifications. It is illustrated on several benchmark datasets, including some about images and texts.

Keywords: Discriminant analysis; k-Nearest neighbors; Mislabeling; Pairwise coupling; Support vector machines