Anti-learning in real and synthetic data

dr inż. Adam Kowalczyk, Politechnika Warszawska, the University of Melbourne

We shall discuss a binary classification problem in which standard supervised learning algorithms such as linear and kernel SVM, naive Bayes, ridge regression, k-nearest neighbours, shrunken centroid, multilayer perceptron and decision trees perform in an unusual way. On certain data sets they classify a randomly sampled training subset nearly perfectly, but systematically perform worse than random guessing on cases unseen in training.

We demonstrate this phenomenon in classification of natural data sets, in particular, a response to oesophageal cancer to chemo-radio-therapy treatment. We shall also discuss a range of synthetic datasets for which we analyse this phenomenon in the i.i.d. setting. Furthermore, we propose and evaluate a remedy that yields promising results for classifying such data as well as normal datasets. We simply transform the classifier scores by an additional 1-dimensional linear transformation developed, for instance, to maximize classification accuracy of the outputs of an internal cross-validation on the training set.

Finally, also discuss the relevance of anti-learning to a range of machine learning topics, such as theory of learning, boosting, regularization, sample bias and application of kernels.