Linear Discriminant Classifier (LDC) for Streaming Data with Concept Drift

Ludmila I. Kuncheva and Indre Zliobaite

Abstract. Simple classifiers, including LDC, have often been praised for their robustness and accuracy. Here we consider an online version of LDC applied to streaming data with concept drift. The classifier is trained on a moving window containing the latest N observations. Current approaches to determining the window size are mostly heuristical. The talk presents a framework within which theoretical relationship can be sought between the window size and the classification error. The framework is based upon two ideas. First, past literature offers formal relationships between the size of the training set and the asymptotic accuracy for several classifier models. Such a relationship can be used as a guide to determining the moving window size. Second, after a sudden change, the "old" classifier may be better than an undertrained "new" classifier that uses only the data coming after the change. We state the optimal window size for the case of LDC applied to two Gaussian classes and a sudden change in the form of rotation and translation of the feature space. Simulation results are included in order to investigate the sensitivity of the theoretical results to violation of the underlying assumptions.