Intrinsic Dimensionality and Small Sample Properties of Classifiers.

Sarūnas Raudys

Abstract: Small learning-set properties of the Euclidean distance, the Parzen window, the minimum empirical error and the nonlinear single layer perceptron classifiers depend on an "intrinsic dimensionality" of the data, however the Fisher linear discriminant function is sensitive to all dimensions. There is no unique definition of the "intrinsic dimensionality". The dimensionality of the subspace where the data points are situated is not a sufficient definition of the "intrinsic dimensionality". An exact definition depends both, on a true distribution of the pattern classes, and on the type of the classifier used.

Keywords: