Theoretical Guarantees for the Construction of Informative Projection Ensembles Using k-NN Classifiers

Extracting compact and communicative models is fundamental to decision support systems. Specifically, the use of Informative Projections has been shown to facilitate the decision process, making automatic classification transparent and providing domain experts useful views of the data. Our research provides guarantees for ensembles of classifiers trained on low-dimensional Informative Projections. We analyze the theoretical properties of such ensembles, presenting three methods of training them. The guarantees hold under the assumption that the sample-specific information dictates the use of the classifiers in the ensembles. Our experiments demonstrate that high classification accuracy can be obtained using the low-dimensional models extracted by our methods.

Predictive systems designed to keep human operators in the loop rely on the existence of comprehensible classification models, which present an interpretable view of the data. Often, the domain experts require that the test data be represented using a small number of the original features, which serves to validate the classification outcome by highlighting the similarities to relevant training data. Informative Projection Ensembles (IPEs), alternatively use one of several compact submodels that ensure compliance with the stringent requirement on model size, while also attaining high performance through the use of specialized models. The decision is presented to the user together with a representation of how the classification label was assigned in the projected subspace.

The ensemble construction can be reduced to a combinatorial problem by optimizing over a matrix of loss estimators computed for every data point and each low-dimensional projection. We introduce three ways of solving this combinatorial problem. First, we formulate an integer linear program that computes the optimal point-to-projection allocation for the training sample given a limited number of projections. An alternative is a two-step procedure, similar to the adaptive lasso, which replaces the constraint on the number of projections with an L1 penalty with adaptive weights. Finally, we consider greedy projection selection, which is a promising option in this case because of the super-modularity of the loss.

We provide a guarantee for the risk consistency of k-NN ensembles in low-dimensional spaces. In addition, our experiments show that the methods we introduce can discover and leverage low-dimensional structure in data, if it exists, yielding accurate and compact models. The variety of of techniques we present for the selection of Informative Projections, allows the precise recovery of the patterns. Moreover, it is now possible to scale IPE model learning to larger datasets and multi-class problems, as shown in our experiments on public datasets. The IPE classifier has fast convergence rates, as the reduced number of features used by the $k-NN$ classifier makes it possible for models to achieve better performance with fewer training samples compared with kNN classifiers combined with other feature selection methods. Our algorithms are particularly useful in applications involving multivariate numeric data in which expert assessment of the results is of the essence.