

Abstract

The aim of this thesis is to develop and assess the performance of proximal classifiers, a class of kernel-based regularized mean square error type classifier that learns within the penalized modeling paradigm. The name proximal classifier indicates the fact of classification of a test pattern by its proximity either to a hyperplane or to a class centroid. A novel nonparallel plane classifier has been proposed in this thesis. The basic idea of the nonparallel plane classifier is to model each class of data by fitting separate hyperplane through it. The goal of the thesis is to overcome the limitation of the twin support vector machine (TWSVM) and to formulate an efficient nonparallel plane classifier. The proposed nonparallel plane proximal classifier (NPPC) is computationally efficient than TWSVM and SVM with comparable classification accuracy. Additionally, the binary NPPC is extended to multiclass NPPC using three decomposition techniques. The classification accuracy of multiclass NPPC is not better than that of the multiclass SVM. Moreover, it is not computationally efficient to handle large data set. In order to overcome the limitations of the decomposition techniques, an alternative approach of multiclass data classification by the proximity to a class centroid using discriminant analysis through vector-valued regularized kernel function approximation (VVRKFA) is proposed in this thesis. VVRKFA offers a vector-valued output at a single step. It is an extension of the fast regularized kernel function approximation (FRKFA) that has been proposed to overcome the limitations of computationally costly support vector regression (SVR) in fitting a nonlinear regression function to the large data set. Experimental results show that the proximity to a hyperplane may be effective for binary data. As opposed to this, the proximity to a class centroid approach is better for the classification of multiclass data. Finally, both NPPC and VVRKFA are applied on gene microarray data for cancer classification. A new NPPC ensemble method using proximity profile-based combination of multiple experts is proposed to achieve high classification accuracy on cancer diagnosis. On the other hand, VVRKFA is able to achieve high classification accuracy using a single classifier with a considerable gain in computational time.