Incremental Reduced Support Vector Machine

李育杰

台灣科技大學資訊工程系

Abstract

We will briefly introduce Support Vector Machines (SVMs) which are established as one of the most powerful tools for machine learning as well as data mining. We then address the computational difficulties, long CPU time and huge memory usage, in generating a nonlinear support vector machine classifier for a massive dataset. We propose a new algorithm, Incremental Reduced Support Vector Machine (IRSVM). We begin with an extremely small subset of the entire dataset which is called the reduced set. We then expand the reduced set incrementally according to an information criterion from the feature space viewpoint. This can be achieved by solving a series of small least squares problems. Once we have the reduced set, we utilize the reduced kernel technique to generate the nonlinear SVM classifier. We also apply IRSVM to the feature selection task. Finally, we test our approach on many publicly available datasets to demonstrate the accuracy and speed of IRSVM.