

## STATISTICAL ANALYSIS AND DATA MINING

Original Article

## Feature import vector machine: A general classifier with flexible feature selection

Samiran Ghosh , Yazhen Wang

First published: 26 January 2015

<https://doi.org/10.1002/sam.11259>

About



Access

als



are

## Abstract

The support vector machine (SVM) and other reproducing kernel Hilbert space (RKHS) based classifier systems are drawing much attention recently owing to its robustness and generalization capability. General theme here is to construct classifiers based on the training data in a high dimensional space by using all available dimensions. The SVM achieves huge data compression by selecting only few observations that lie close to the boundary of the classifier function. However when the number of observations is not very large (small  $n$ ) but the number of dimensions/features is large (large  $p$ ), then it is not necessary that all available features are of equal importance in the classification context. Possible selection of a useful fraction of the available features may result in huge data compression. In this paper, we propose an algorithmic approach by means of which such an *optimal* set of features could be selected. In short, we reverse the traditional sequential observation selection strategy of SVM to that of sequential feature selection. To achieve this we have modified the solution proposed by Zhu and Hastie in the context of import vector machine (IVM), to select an *optimal* sub - dimensional model to build the final classifier with sufficient accuracy.

About Wiley Online Library Help & Support Opportunities Connect with Wiley 

