

# SVM Inverse Problem and Its Applications to Decision Tree Induction

Prof. Xizhao Wang  
Department of Mathematics and Computer Science  
Hebei University  
Baoding, Hebei, CHINA  
[xizhaowang@ieee.org](mailto:xizhaowang@ieee.org)

## Abstract

Support Vector Machine (SVM) is a classification technique of machine learning based on statistical learning theory. When the instances are linearly separable in considering a classification problem with two classes, the goal of a SVM is to construct a hyper-plane with the maximum margin to classify them. The SVM inverse problem is to determine how to split a given dataset into two clusters such that the margin between the two clusters attains the maximum where we assume that examples of the dataset have no initial class labels.

It is difficult to exactly solve the SVM inverse problem since it is a NP-hard problem. In this presentation we propose an effective approximation method to solve the SVM inverse problem. It is based on a technique of clustering to reduce the computational complexity prior to the computation of the margin between the two clusters. The relationship between SVM margin and the distance among the clusters is investigated.

The application of the SVM inverse problem to decision tree induction is discussed where the expanded attribute at a node is selected based the maximum margin. The initial experiments showed a strong relationship between the increase of margin and the generalization capability of a classifier. The key point of this application is that maximum-margin can significantly improve the generalization capability of the classifier.