MASTER'S THESIS

Kernel based learning methods for pattern and feature analysis

Wu, Zhili

Date of Award:
2004

Kernel Based Learning Methods for Pattern and Feature Analysis

WU Zhili

A thesis submitted in partial fulfillment of the requirements

for the degree of

Master of Philosophy

Principal Supervisor: Dr. LI Chun-hung

Hong Kong Baptist University

September 2004

Abstract

Kernel-based learning methods (kernel methods) have significant influences on recent development of machine learning research. This thesis is on devising and improving kernel methods, and on applying them to pattern and feature analysis.

Part of our research focuses on improving Support Vector Machines (SVMs). In solving SVMs, we find that some proposed cache policies for sequential minimal optimization (SMO) result in low efficiency. A better strategy is to cache gradients for all vectors frequently checked. Moreover, we propose a strategy that utilizes the nearest neighboring vectors to speed up the convergence of SMO. We also suggest the use of Hadamard codes for multiclass label prediction by SVMs. We prove that the Hadamard codes are optimal in correcting the wrong labels predicted by base classifiers. Furthermore, we design a new *summation of exponential* (SoE) kernel for solving regression tasks with missing values. We show SoE kernels are admissible to kernel conditions and insensitive to missing values.

This thesis also deals with unsupervised and semi-supervised kernel methods. Specifically, we transform the Rival Penalized Competitive Learning (RPCL) from data space to feature space for automatic clustering. In addition, we use spectral analysis of kernel matrices to address the seeds initialization problem associated with RPCL. We also improve the SVM-based feature selection in a semi-supervised manner by utilizing both labeled and unlabeled data. The new feature selection method exhibits good performance in solving feature selection benchmark problems.

Contents