

MASTER'S THESIS

Kernel based learning methods for pattern and feature analysis

Wu, Zhili

Date of Award:
2004

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

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

Declaration	i
Abstract	ii
Acknowledgements	iii
Contents	i
List of Tables	vi
List of Figures	viii
Chapter 1 Basic Concepts and Thesis Outline	1
1.1 Introduction to Basic Concepts	1
1.1.1 Pattern and Feature Analysis	1
1.1.2 Three Learning Modes	2
1.1.3 Kernel Methods	3
1.2 Organization of Thesis	4
1.2.1 Supervised Kernel Methods - Support Vector Machines	4

1.2.2	Unsupervised and Semi-Supervised Kernel-Based Learning . . .	6
Chapter 2 SVMs for Classification and Sequential Minimal Optimization		7
2.1	Support Vector Classifiers	7
2.1.1	Large Margin Formulation	8
2.1.2	Quadratic Programming	9
2.1.3	Imperfect Separation	11
2.1.4	Decision Function	13
2.1.5	Nonlinear SVMs Using Lernel Tricks	13
2.2	On Improving Sequential Minimal Optimization	15
2.2.1	Introduction	15
2.2.2	SMO Formulation	16
2.2.3	Working Pair Selection	18
2.2.4	Two Observations on SMO	20
2.2.5	Working Pair Selection for Larger Changes	21
2.3	Experiments	23
2.3.1	On Cache Policy and Working Pair Selection	23
2.3.2	On Selecting Pairs Maximizing the Objective Change	24
2.4	Chapter Summary	26
Chapter 3 SVMs for Multiclass Learning via Hadamard Codes		27
3.1	Introduction	27

3.2	ECOC and (Symmetric) Hamming Decoding	29
3.3	Hadamard Output Codes	33
3.4	Experimental Results	36
3.5	Chapter Summary and Discussion	38
Chapter 4 SVMs for Regression via a Missing Value Insensitive Kernel		40
4.1	Support Vector Regression	40
4.2	Missing Values	41
4.3	Summation of Exponentials Kernels	43
4.4	SoE vs. RBF Kernels	44
4.5	Related Works - ANOVA Decomposition Kernels	47
4.6	Experiment	48
4.6.1	Synthetic Data	50
4.6.2	Real Data	51
4.7	Chapter Summary	54
Chapter 5 Unsupervised Kernel Rival Penalized Competitive Learning		56
5.1	Introduction	57
5.2	Kernel RPCL Algorithm	60
5.2.1	Algorithm Formulation	60
5.2.2	Seeds Initialization by Kernel Spectrum	62
5.2.3	On Determining De-Learning Rates	64

5.2.4	Complexity Issues	65
5.3	Experimental Results	65
5.3.1	Synthetical Data	65
5.3.2	UCI Iris Data	69
5.3.3	Text Data Clustering	71
5.4	Chapter Summary	72
Chapter 6 Semi-supervised Feature Selection via Transductive SVMs		74
6.1	Introduction	75
6.2	SVMs and Transductive SVMs	76
6.2.1	Support Vector Machines for Feature Selection	76
6.2.2	Transductive Support Vector Machines	77
6.3	SVM-Related Feature Selection Methods	77
6.3.1	Correlation Scoring Measures (CORR-SVM)	78
6.3.2	Recursive Feature Elimination	78
6.3.3	Multiplicative Update	79
6.4	Transductive Feature Selection	80
6.5	Experimentation	82
6.5.1	Overall Performance on Competition Tasks	82
6.5.2	Correlation Scoring Measures	83
6.5.3	TSVMs for Classification and Feature Selection	85
6.5.4	Feature Saliency Check	87

6.6 Chapter Summary	89
Chapter 7 Summary and Conclusion	91
Bibliography	93
Curriculum Vitae	101