

# Contents

<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>xiii</b>
<b>List of Symbols and Operations</b>	<b>xvii</b>
<b>List of Abbreviations</b>	<b>xxi</b>
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Supervised Binary Classification Problem . . . . .	3
1.3 Support Vector Machine . . . . .	3
1.3.1 Kernel Trick and Nonlinear SVM . . . . .	5
1.3.2 Limitations of SVM . . . . .	6
1.4 Variants of SVM . . . . .	7
1.4.1 Least Square Support Vector Machine (LS-SVM) . . . . .	8
1.4.2 Proximal Support Vector Machine (PSVM) . . . . .	8
1.4.3 Generalized Eigen Value Proximal Support Vector Machine (GEPSVM) . . . . .	10
1.4.4 Twin Support Vector Machine (TWSVM) . . . . .	11
1.5 Motivation and Goal . . . . .	13
1.6 Brief Overview of Contribution . . . . .	14
1.7 Organization of the Thesis . . . . .	14
1.8 Data Sets . . . . .	15
1.9 Machine Specification and Software . . . . .	22

<b>2</b>	<b>Unity Norm Twin Support Vector Machine Classifier</b>	<b>25</b>
2.1	Introduction . . . . .	25
2.2	Limitation of TWSVM . . . . .	25
2.2.1	Illustration with 2-dimensional Example . . . . .	26
2.3	Unity Norm TWSVM . . . . .	27
2.3.1	Solution of the Objective Functions . . . . .	28
2.3.2	Training Algorithm of UNTWSVM . . . . .	32
2.3.3	Extension to Nonlinear UNTWSVM Classifier . . . . .	33
2.4	Experimental Results . . . . .	34
2.4.1	Implementation . . . . .	34
2.4.2	Selection of the Penalty Parameter . . . . .	35
2.4.3	Visual Illustration with 2D Data Sets . . . . .	35
2.4.4	Performance on Benchmark Data Sets . . . . .	37
2.4.5	Performance of Nonlinear Classifier . . . . .	40
2.5	Conclusion . . . . .	41
<b>3</b>	<b>Nonparallel Plane Proximal Classifier</b>	<b>43</b>
3.1	Introduction . . . . .	43
3.2	Formulation of NPPC . . . . .	44
3.2.1	Solution Methodology of the Objective Functions . . . . .	45
3.2.2	Training Algorithm . . . . .	47
3.3	Nonlinear Kernel Nonparallel Plane Proximal Classifier (NKNPPC) Formulation . . . . .	48
3.4	Efficient Implementation of NPPC and its Time Complexity . . . . .	50
3.5	Experimental Results and Comparison . . . . .	51
3.5.1	Experimental Setup . . . . .	52
3.5.2	Limitation of NPPC . . . . .	52
3.5.3	Performance on Cross Planes Data Set . . . . .	54
3.5.4	Performance Comparison of Linear Classifiers . . . . .	54
3.5.5	Performance Comparison of Nonlinear Kernel Classifiers . . . . .	56
3.5.6	Effectiveness of Generating Nonlinear Surfaces . . . . .	58
3.5.7	Comparison of Computational Efficiency on Large Data Set . . . . .	59
3.6	Conclusion . . . . .	61

---

<b>4</b>	<b>Multiclass Nonparallel Plane Proximal Classifier</b>	<b>63</b>
4.1	Introduction . . . . .	63
4.2	Decomposition of Multiclass Problems . . . . .	64
4.2.1	OVR Method . . . . .	64
4.2.2	OVO Method . . . . .	65
4.2.3	DAG Method . . . . .	66
4.3	Data Unbalance Problem . . . . .	66
4.4	Linear Multiclass NPPC . . . . .	67
4.4.1	Linear OVR NPPC . . . . .	67
4.4.2	Linear OVO and DAG NPPC . . . . .	68
4.5	Nonlinear Multiclass NPPC . . . . .	68
4.5.1	OVR Nonlinear Kernel NPPC . . . . .	69
4.5.2	OVO and DAG Nonlinear Kernel NPPC . . . . .	69
4.6	Experimental Results . . . . .	70
4.6.1	Tuning of Multiclass NPPC . . . . .	70
4.6.2	Experimental Setup . . . . .	72
4.6.3	Visual Illustration with 2-dimensional Data Sets . . . . .	72
4.6.4	Effectiveness of Nonlinear Multiclass NPPC . . . . .	73
4.6.5	Performance of Multiclass NPPC on Benchmark Data Set . . . . .	73
4.6.5.1	Performance of Linear Multiclass NPPC . . . . .	74
4.6.5.2	Performance of Nonlinear Multiclass NPPC . . . . .	79
4.7	Conclusion . . . . .	83
<b>5</b>	<b>Discriminant Analysis for Fast Multiclass Data Classification Through Regularized Kernel Function Approximation</b>	<b>85</b>
5.1	Introduction . . . . .	85
5.2	Function Approximation or Regression . . . . .	87
5.3	Support Vector Regression (SVR) . . . . .	88
5.4	FRKFA Formulation . . . . .	89
5.4.1	Solution of FRKFA . . . . .	90
5.4.2	Performance of FRKFA . . . . .	91
5.4.2.1	Experimental Setup . . . . .	92
5.4.2.2	Selection Methodology of Reduced Set Size . . . . .	92
5.4.2.3	Performance on Synthetic Data Sets . . . . .	93
5.4.2.4	Performance on Benchmark Data Sets . . . . .	94

5.5	Multiclass Data Classification Through Regression . . . . .	95
5.5.1	Vector-Valued Regularized Kernel Function Approximation (VVRKFA) . . . . .	96
5.5.2	Solution Methodology . . . . .	97
5.5.3	Decoding of the Label Vectors to the Class Label . . . . .	98
5.5.4	Training Algorithm . . . . .	99
5.6	Experimental Study . . . . .	100
5.6.1	Experimental Setup . . . . .	100
5.6.2	Results on Synthetic Data Sets . . . . .	101
5.6.3	Results on Benchmark Data Sets . . . . .	102
5.6.4	Analysis of Results . . . . .	106
5.6.5	Data Visualization . . . . .	109
5.7	Conclusion . . . . .	109
<b>6</b>	<b>Application of NPPC and VVRKFA on Cancer Classification from Gene Expression Data</b>	<b>111</b>
6.1	Introduction . . . . .	111
6.2	Microarray Technique . . . . .	111
6.3	Microarray Data Classification Techniques . . . . .	112
6.3.1	Methods of Feature Selection . . . . .	112
 <b>PART-I: Cancer Classification from Gene Expression Data by NPPC</b>		
<b>Ensemble</b>		<b>117</b>
6.4	Introduction . . . . .	117
6.5	NPPC Ensemble . . . . .	118
6.5.1	Gene Selection by Filter Method . . . . .	119
6.5.1.1	MRMR Method of Feature Selection . . . . .	119
6.5.1.2	Gene Selection by MRMR Method for NPPC Ensemble	120
6.5.2	Formation of NPPC Experts by Genetic Algorithm . . . . .	121
6.5.3	Selection of Expert Models in the NPPC Ensemble . . . . .	123
6.5.4	Decision Combination of the Selected NPPC Experts in the Ensemble . . . . .	123
6.5.4.1	Proximity Profile-based Decision Combination Scheme	124

---

6.5.4.2	Proximity Profile-based Decision Combination Scheme for OVO NPPC Ensemble . . . . .	125
6.6	Experimental Results . . . . .	127
6.6.1	Microarray Data Sets . . . . .	127
6.6.2	Experiment Design . . . . .	129
6.6.3	Results on Binary Data sets . . . . .	130
6.6.3.1	The Effect of GA-based Model Selection . . . . .	130
6.6.3.2	Effectiveness of NPPC Ensemble . . . . .	132
6.6.4	Results on Multiclass Cancer Classification . . . . .	134
6.6.5	Performance of Nonlinear Classifier Ensemble . . . . .	135
6.7	Analysis of Result and Discussion . . . . .	138
6.7.1	The Effect of GA-based Simultaneous Feature and Parameter Selection on Class Separability . . . . .	138
6.7.2	The Effect of Classifier Ensemble on Classification Accuracy .	140
6.7.3	Biological Relevance of the Selected Genes . . . . .	140
<b>PART-II Performance of VVRKFA on Multiclass Cancer Classification</b>		<b>145</b>
6.8	Introduction . . . . .	145
6.9	Experimental Results . . . . .	145
6.10	Conclusion . . . . .	147
<b>7</b>	<b>Conclusion and Future Scope</b>	<b>149</b>
7.1	Conclusion . . . . .	149
7.2	Future Scope . . . . .	153
<b>A</b>	<b>Derivation of the expression (2.7)</b>	<b>155</b>
<b>B</b>	<b>Derivation of the Gradient and Hessian of the Expression (3.3)</b>	<b>159</b>
<b>C</b>	<b>List of Selected Genes used for Experiment</b>	<b>161</b>
<b>References</b>		<b>165</b>
<b>Author's Publications</b>		<b>183</b>
<b>Author's Biography</b>		<b>185</b>