

# Contents

<b>Acknowledgements</b>	<b>ix</b>
<b>Abstract</b>	<b>xi</b>
<b>I FUNDAMENTAL CONCEPTS</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Supervised Machine Learning . . . . .	3
1.2 Active Learning . . . . .	5
1.3 Overview of Thesis . . . . .	8
<b>2 Kernel Machines</b>	<b>11</b>
2.1 Linear Classifiers . . . . .	11
2.2 Kernel Feature Map and Feature Space . . . . .	12
2.3 Version Space Model . . . . .	20
2.4 Support Vector Machines . . . . .	21
2.5 Bayes Point Machines . . . . .	27
<b>II ACTIVE LEARNING</b>	<b>33</b>
<b>3 Active Learning Models</b>	<b>35</b>
3.1 General Models in Active Learning . . . . .	35
3.2 Pool-based Active Learning . . . . .	38
<b>4 Binary Classification</b>	<b>41</b>
4.1 Volume-based Selection Strategies . . . . .	41
4.2 Kernel Billiard Approach to Active Learning . . . . .	46
4.3 Experiments . . . . .	52
4.3.1 Artificial Data . . . . .	52
4.3.2 Real-World Data . . . . .	53
4.4 Discussion . . . . .	54

<b>5</b>	<b>Selection of Multiple Examples</b>	<b>59</b>
5.1	Introduction . . . . .	59
5.2	Generalization of the SIMPLE Selection Strategy . . . . .	60
5.3	Incorporating Diversity into Batch Selection . . . . .	62
5.4	Computational Complexity . . . . .	66
5.5	Experiments . . . . .	67
5.5.1	Experimental Setting . . . . .	67
5.5.2	Experimental Results . . . . .	69
5.6	Discussion . . . . .	70
<b>6</b>	<b>Multiclass Classification</b>	<b>73</b>
6.1	Introduction . . . . .	73
6.2	Reducing Multiclass to Binary Classification Learning . . . . .	75
6.2.1	One-versus-all Decomposition . . . . .	76
6.2.2	One-versus-one Decomposition . . . . .	76
6.2.3	DAG Multiclass Classification . . . . .	77
6.3	Active Multiclass Learning . . . . .	78
6.4	Experiments . . . . .	81
6.4.1	Experimental Setting . . . . .	81
6.4.2	Experimental Results . . . . .	82
6.5	Discussion . . . . .	84
<b>7</b>	<b>Label Ranking Learning</b>	<b>85</b>
7.1	Introduction . . . . .	85
7.2	Ranking Problems . . . . .	87
7.2.1	Constraint Classification . . . . .	88
7.2.2	Implementation of Kernel Constraint Classification . . . . .	90
7.2.3	Pairwise Ranking . . . . .	93
7.3	Computational Complexity . . . . .	95
7.4	Active Ranking Learning . . . . .	95
7.5	Experiments . . . . .	98
7.5.1	Experimental Setting . . . . .	98
7.5.2	Experimental Results . . . . .	101
7.6	Related Work . . . . .	104
7.7	Discussion . . . . .	104
<b>8</b>	<b>Convergence Analysis</b>	<b>107</b>
8.1	Convergence Analysis . . . . .	107
<b>9</b>	<b>Conclusion</b>	<b>117</b>

<b>III</b>	<b>APPENDIX</b>	<b>121</b>
<b>A</b>	<b>Bound for Sampling Method</b>	<b>125</b>
<b>B</b>	<b>Experimental Results: Multiclass Learning</b>	<b>129</b>
<b>C</b>	<b>Experimental Results: Label Ranking Learning</b>	<b>139</b>
	<b>List of Figures</b>	<b>149</b>
	<b>List of Tables</b>	<b>151</b>
	<b>Notation</b>	<b>152</b>
	<b>References</b>	<b>155</b>