

# Contents

## Part I Support Vector Machines: Theory and Algorithms

<b>1</b>	<b>Support Vector Machines for Classification Problems</b>	3
1.1	Method of Maximum Margin	3
1.2	Dual Problem	5
1.3	Soft Margin	6
1.4	$C$ -Support Vector Classification	8
1.5	$C$ -Support Vector Classification with Nominal Attributes	10
1.5.1	From Fixed Points to Flexible Points	10
1.5.2	$C$ -SVC with Nominal Attributes	11
1.5.3	Numerical Experiments	12
<b>2</b>	<b>LOO Bounds for Support Vector Machines</b>	15
2.1	Introduction	15
2.2	LOO Bounds for $\varepsilon$ -Support Vector Regression	16
2.2.1	Standard $\varepsilon$ -Support Vector Regression	16
2.2.2	The First LOO Bound	17
2.2.3	A Variation of $\varepsilon$ -Support Vector Regression	26
2.2.4	The Second LOO Bound	27
2.2.5	Numerical Experiments	30
2.3	LOO Bounds for Support Vector Ordinal Regression Machine	32
2.3.1	Support Vector Ordinal Regression Machine	33
2.3.2	The First LOO Bound	38
2.3.3	The Second LOO Bound	42
2.3.4	Numerical Experiments	44
<b>3</b>	<b>Support Vector Machines for Multi-class Classification Problems</b>	47
3.1	$K$ -Class Linear Programming Support Vector Classification Regression Machine ( $K$ -LPSVCR)	47
3.1.1	$K$ -LPSVCR	49
3.1.2	Numerical Experiments	50
3.1.3	$\nu$ - $K$ -LPSVCR	52

3.2	Support Vector Ordinal Regression Machine for Multi-class Problems . . . . .	54
3.2.1	Kernel Ordinal Regression for 3-Class Problems . . . . .	54
3.2.2	Multi-class Classification Algorithm . . . . .	56
3.2.3	Numerical Experiments . . . . .	57
<b>4</b>	<b>Unsupervised and Semi-supervised Support Vector Machines . . . . .</b>	<b>61</b>
4.1	Unsupervised and Semi-supervised $\nu$ -Support Vector Machine . . . . .	62
4.1.1	Bounded $\nu$ -Support Vector Machine . . . . .	62
4.1.2	$\nu$ -SDP for Unsupervised Classification Problems . . . . .	63
4.1.3	$\nu$ -SDP for Semi-supervised Classification Problems . . . . .	65
4.2	Numerical Experiments . . . . .	66
4.2.1	Numerical Experiments of Algorithm 4.2 . . . . .	66
4.2.2	Numerical Experiments of Algorithm 4.3 . . . . .	67
4.3	Unsupervised and Semi-supervised Lagrange Support Vector Machine . . . . .	69
4.4	Unconstrained Transductive Support Vector Machine . . . . .	72
4.4.1	Transductive Support Vector Machine . . . . .	73
4.4.2	Unconstrained Transductive Support Vector Machine . . . . .	74
4.4.3	Unconstrained Transductive Support Vector Machine with Kernels . . . . .	77
<b>5</b>	<b>Robust Support Vector Machines . . . . .</b>	<b>81</b>
5.1	Robust Support Vector Ordinal Regression Machine . . . . .	81
5.2	Robust Multi-class Algorithm . . . . .	93
5.3	Numerical Experiments . . . . .	94
5.3.1	Numerical Experiments of Algorithm 5.6 . . . . .	94
5.3.2	Numerical Experiments of Algorithm 5.7 . . . . .	95
5.4	Robust Unsupervised and Semi-supervised Bounded $C$ -Support Vector Machine . . . . .	96
5.4.1	Robust Linear Optimization . . . . .	97
5.4.2	Robust Algorithms with Polyhedron . . . . .	97
5.4.3	Robust Algorithm with Ellipsoid . . . . .	101
5.4.4	Numerical Results . . . . .	103
<b>6</b>	<b>Feature Selection via <math>l_p</math>-Norm Support Vector Machines . . . . .</b>	<b>107</b>
6.1	$l_p$ -Norm Support Vector Classification . . . . .	107
6.1.1	$l_p$ -SVC . . . . .	108
6.1.2	Lower Bound for Nonzero Entries in Solutions of $l_p$ -SVC . . . . .	109
6.1.3	Iteratively Reweighted $l_q$ -SVC for $l_p$ -SVC . . . . .	111
6.2	$l_p$ -Norm Proximal Support Vector Machine . . . . .	111
6.2.1	Lower Bounds for Nonzero Entries in Solutions of $l_p$ -PSVM . . . . .	113
6.2.2	Smoothing $l_p$ -PSVM Problem . . . . .	113
6.2.3	Numerical Experiments . . . . .	114

**Part II Multiple Criteria Programming: Theory and Algorithms**

**7 Multiple Criteria Linear Programming . . . . . 119**

7.1 Comparison of Support Vector Machine and Multiple Criteria Programming . . . . . 119

7.2 Multiple Criteria Linear Programming . . . . . 120

7.3 Multiple Criteria Linear Programming for Multiple Classes . . . . 123

7.4 Penalized Multiple Criteria Linear Programming . . . . . 129

7.5 Regularized Multiple Criteria Linear Programs for Classification . . . . . 129

**8 MCLP Extensions . . . . . 133**

8.1 Fuzzy MCLP . . . . . 133

8.2 FMCLP with Soft Constraints . . . . . 136

8.3 FMCLP by Tolerances . . . . . 140

8.4 Kernel-Based MCLP . . . . . 141

8.5 Knowledge-Based MCLP . . . . . 143

8.5.1 Linear Knowledge-Based MCLP . . . . . 143

8.5.2 Nonlinear Knowledge and Kernel-Based MCLP . . . . . 147

8.6 Rough Set-Based MCLP . . . . . 150

8.6.1 Rough Set-Based Feature Selection Method . . . . . 150

8.6.2 A Rough Set-Based MCLP Approach for Classification . . . . . 152

8.7 Regression by MCLP . . . . . 155

**9 Multiple Criteria Quadratic Programming . . . . . 157**

9.1 A General Multiple Mathematical Programming . . . . . 157

9.2 Multi-criteria Convex Quadratic Programming Model . . . . . 161

9.3 Kernel Based MCQP . . . . . 167

**10 Non-additive MCLP . . . . . 171**

10.1 Non-additive Measures and Integrals . . . . . 171

10.2 Non-additive Classification Models . . . . . 172

10.3 Non-additive MCP . . . . . 178

10.4 Reducing the Time Complexity . . . . . 179

10.4.1 Hierarchical Choquet Integral . . . . . 179

10.4.2 Choquet Integral with Respect to  $k$ -Additive Measure . . . 180

**11 MC2LP . . . . . 183**

11.1 MC2LP Classification . . . . . 183

11.1.1 Multiple Criteria Linear Programming . . . . . 183

11.1.2 Different Versions of MC2 . . . . . 186

11.1.3 Heuristic Classification Algorithm . . . . . 189

11.2 Minimal Error and Maximal Between-Class Variance Model . . . . 191

### Part III Applications in Various Fields

<b>12 Firm Financial Analysis</b>	195
12.1 Finance and Banking	195
12.2 General Classification Process	196
12.3 Firm Bankruptcy Prediction	199
<b>13 Personal Credit Management</b>	203
13.1 Credit Card Accounts Classification	203
13.2 Two-Class Analysis	207
13.2.1 Six Different Methods	207
13.2.2 Implication of Business Intelligence and Decision Making	211
13.2.3 FMCLP Analysis	213
13.3 Three-Class Analysis	219
13.3.1 Three-Class Formulation	219
13.3.2 Small Sample Testing	222
13.3.3 Real-Life Data Analysis	227
13.4 Four-Class Analysis	228
13.4.1 Four-Class Formulation	228
13.4.2 Empirical Study and Managerial Significance of Four-Class Models	230
<b>14 Health Insurance Fraud Detection</b>	233
14.1 Problem Identification	233
14.2 A Real-Life Data Mining Study	233
<b>15 Network Intrusion Detection</b>	237
15.1 Problem and Two Datasets	237
15.2 Classify NeWT Lab Data by MCMP, MCMP with Kernel and See5	239
15.3 Classify KDDCUP-99 Data by Nine Different Methods	240
<b>16 Internet Service Analysis</b>	243
16.1 VIP Mail Dataset	243
16.2 Empirical Study of Cross-Validation	244
16.3 Comparison of Multiple-Criteria Programming Models and SVM	247
<b>17 HIV-1 Informatics</b>	249
17.1 HIV-1 Mediated Neuronal Dendritic and Synaptic Damage	249
17.2 Materials and Methods	251
17.2.1 Neuronal Culture and Treatments	251
17.2.2 Image Analysis	252
17.2.3 Preliminary Analysis of Neuronal Damage Induced by HIV MCM Treated Neurons	252
17.2.4 Database	253
17.3 Designs of Classifications	254

- 17.4 Analytic Results . . . . . 256
  - 17.4.1 Empirical Classification . . . . . 256
- 18 Anti-gen and Anti-body Informatics . . . . . 259**
  - 18.1 Problem Background . . . . . 259
  - 18.2 MCQP, LDA and DT Analyses . . . . . 260
  - 18.3 Kernel-Based MCQP and SVM Analyses . . . . . 266
- 19 Geochemical Analyses . . . . . 269**
  - 19.1 Problem Description . . . . . 269
  - 19.2 Multiple-Class Analyses . . . . . 270
    - 19.2.1 Two-Class Classification . . . . . 270
    - 19.2.2 Three-Class Classification . . . . . 271
    - 19.2.3 Four-Class Classification . . . . . 271
  - 19.3 More Advanced Analyses . . . . . 272
- 20 Intelligent Knowledge Management . . . . . 277**
  - 20.1 Purposes of the Study . . . . . 277
  - 20.2 Definitions and Theoretical Framework of Intelligent Knowledge . . . . . 280
    - 20.2.1 Key Concepts and Definitions . . . . . 280
    - 20.2.2 4T Process and Major Steps of Intelligent Knowledge Management . . . . . 288
  - 20.3 Some Research Directions . . . . . 290
    - 20.3.1 The Systematic Theoretical Framework of Data Technology and Intelligent Knowledge Management . . . . . 291
    - 20.3.2 Measurements of Intelligent Knowledge . . . . . 292
    - 20.3.3 Intelligent Knowledge Management System Research . . . . . 293
- Bibliography . . . . . 295**
- Subject Index . . . . . 307**
- Author Index . . . . . 311**



<http://www.springer.com/978-0-85729-503-3>

Optimization Based Data Mining: Theory and Applications

Shi, Y.; Tian, Y.; Kou, G.; Peng, Y.; Li, J.

2011, XVI, 316 p., Hardcover

ISBN: 978-0-85729-503-3