

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	General Background	1
1.2	Historical Background of Multi-source Learning and Data Fusion	4
1.2.1	Canonical Correlation and Its Probabilistic Interpretation	4
1.2.2	Inductive Logic Programming and the Multi-source Learning Search Space	5
1.2.3	Additive Models	6
1.2.4	Bayesian Networks for Data Fusion	7
1.2.5	Kernel-based Data Fusion	9
1.3	Topics of This Book	18
1.4	Chapter by Chapter Overview	21
	References	22
<b>2</b>	<b>Rayleigh Quotient-Type Problems in Machine Learning</b>	<b>27</b>
2.1	Optimization of Rayleigh Quotient	27
2.1.1	Rayleigh Quotient and Its Optimization	27
2.1.2	Generalized Rayleigh Quotient	28
2.1.3	Trace Optimization of Generalized Rayleigh Quotient-Type Problems	28
2.2	Rayleigh Quotient-Type Problems in Machine Learning	30
2.2.1	Principal Component Analysis	30
2.2.2	Canonical Correlation Analysis	30
2.2.3	Fisher Discriminant Analysis	31
2.2.4	$k$ -means Clustering	32
2.2.5	Spectral Clustering	33
2.2.6	Kernel-Laplacian Clustering	33

2.2.7	One Class Support Vector Machine .....	34
2.3	Summary .....	35
	References .....	37
<b>3</b>	<b><math>L_n</math>-norm Multiple Kernel Learning and Least Squares Support Vector Machines .....</b>	<b>39</b>
3.1	Background .....	39
3.2	Acronyms .....	40
3.3	The Norms of Multiple Kernel Learning .....	42
3.3.1	$L_\infty$ -norm MKL .....	42
3.3.2	$L_2$ -norm MKL .....	43
3.3.3	$L_n$ -norm MKL .....	44
3.4	One Class SVM MKL .....	46
3.5	Support Vector Machine MKL for Classification .....	48
3.5.1	The Conic Formulation .....	48
3.5.2	The Semi Infinite Programming Formulation .....	50
3.6	Least Squares Support Vector Machines MKL for Classification .....	53
3.6.1	The Conic Formulation .....	53
3.6.2	The Semi Infinite Programming Formulation .....	54
3.7	Weighted SVM MKL and Weighted LSSVM MKL .....	56
3.7.1	Weighted SVM .....	56
3.7.2	Weighted SVM MKL .....	56
3.7.3	Weighted LSSVM .....	57
3.7.4	Weighted LSSVM MKL .....	58
3.8	Summary of Algorithms .....	58
3.9	Numerical Experiments .....	59
3.9.1	Overview of the Convexity and Complexity .....	59
3.9.2	QP Formulation Is More Efficient than SOCP .....	59
3.9.3	SIP Formulation Is More Efficient than QCQP .....	60
3.10	MKL Applied to Real Applications .....	63
3.10.1	Experimental Setup and Data Sets .....	63
3.10.2	Results .....	67
3.11	Discussions .....	83
3.12	Summary .....	84
	References .....	84
<b>4</b>	<b>Optimized Data Fusion for Kernel <math>k</math>-means Clustering .....</b>	<b>89</b>
4.1	Introduction .....	89
4.2	Objective of $k$ -means Clustering .....	90
4.3	Optimizing Multiple Kernels for $k$ -means .....	92
4.4	Bi-level Optimization of $k$ -means on Multiple Kernels .....	94
4.4.1	The Role of Cluster Assignment .....	94
4.4.2	Optimizing the Kernel Coefficients as KFD .....	94

4.4.3	Solving KFD as LSSVM Using Multiple Kernels . . . . .	96
4.4.4	Optimized Data Fusion for Kernel $k$ -means Clustering (OKKC) . . . . .	98
4.4.5	Computational Complexity . . . . .	98
4.5	Experimental Results . . . . .	99
4.5.1	Data Sets and Experimental Settings . . . . .	99
4.5.2	Results . . . . .	101
4.6	Summary . . . . .	103
	References . . . . .	105
<b>5</b>	<b>Multi-view Text Mining for Disease Gene Prioritization and Clustering</b> . . . . .	<b>109</b>
5.1	Introduction . . . . .	109
5.2	Background: Computational Gene Prioritization . . . . .	110
5.3	Background: Clustering by Heterogeneous Data Sources . . . . .	111
5.4	Single View Gene Prioritization: A Fragile Model with Respect to the Uncertainty . . . . .	112
5.5	Data Fusion for Gene Prioritization: Distribution Free Method . . . . .	112
5.6	Multi-view Text Mining for Gene Prioritization . . . . .	116
5.6.1	Construction of Controlled Vocabularies from Multiple Bio-ontologies . . . . .	116
5.6.2	Vocabularies Selected from Subsets of Ontologies . . . . .	119
5.6.3	Merging and Mapping of Controlled Vocabularies . . . . .	119
5.6.4	Text Mining . . . . .	122
5.6.5	Dimensionality Reduction of Gene-By-Term Data by Latent Semantic Indexing . . . . .	122
5.6.6	Algorithms and Evaluation of Gene Prioritization Task . . . . .	123
5.6.7	Benchmark Data Set of Disease Genes . . . . .	124
5.7	Results of Multi-view Prioritization . . . . .	124
5.7.1	Multi-view Performs Better than Single View . . . . .	124
5.7.2	Effectiveness of Multi-view Demonstrated on Various Number of Views . . . . .	126
5.7.3	Effectiveness of Multi-view Demonstrated on Disease Examples . . . . .	127
5.8	Multi-view Text Mining for Gene Clustering . . . . .	130
5.8.1	Algorithms and Evaluation of Gene Clustering Task . . . . .	130
5.8.2	Benchmark Data Set of Disease Genes . . . . .	132
5.9	Results of Multi-view Clustering . . . . .	133
5.9.1	Multi-view Performs Better than Single View . . . . .	133
5.9.2	Dimensionality Reduction of Gene-By-Term Profiles for Clustering . . . . .	135

5.9.3	Multi-view Approach Is Better than Merging Vocabularies .....	137
5.9.4	Effectiveness of Multi-view Demonstrated on Various Numbers of Views .....	137
5.9.5	Effectiveness of Multi-view Demonstrated on Disease Examples .....	137
5.10	Discussions .....	139
5.11	Summary .....	140
	References .....	141
<b>6</b>	<b>Optimized Data Fusion for <math>k</math>-means Laplacian Clustering .....</b>	<b>145</b>
6.1	Introduction .....	145
6.2	Acronyms .....	146
6.3	Combine Kernel and Laplacian for Clustering .....	149
6.3.1	Combine Kernel and Laplacian as Generalized Rayleigh Quotient for Clustering .....	149
6.3.2	Combine Kernel and Laplacian as Additive Models for Clustering .....	150
6.4	Clustering by Multiple Kernels and Laplacians .....	151
6.4.1	Optimize $A$ with Given $\theta$ .....	153
6.4.2	Optimize $\theta$ with Given $A$ .....	153
6.4.3	Algorithm: Optimized Kernel Laplacian Clustering .....	155
6.5	Data Sets and Experimental Setup .....	156
6.6	Results .....	158
6.7	Summary .....	170
	References .....	171
<b>7</b>	<b>Weighted Multiple Kernel Canonical Correlation .....</b>	<b>173</b>
7.1	Introduction .....	173
7.2	Acronyms .....	174
7.3	Weighted Multiple Kernel Canonical Correlation .....	175
7.3.1	Linear CCA on Multiple Data Sets .....	175
7.3.2	Multiple Kernel CCA .....	175
7.3.3	Weighted Multiple Kernel CCA .....	177
7.4	Computational Issue .....	178
7.4.1	Standard Eigenvalue Problem for WMKCCA .....	178
7.4.2	Incomplete Cholesky Decomposition .....	179
7.4.3	Incremental Eigenvalue Solution for WMKCCA .....	180
7.5	Learning from Heterogeneous Data Sources by WMKCCA .....	181
7.6	Experiment .....	183
7.6.1	Classification in the Canonical Spaces .....	183
7.6.2	Efficiency of the Incremental EVD Solution .....	185

7.6.3	Visualization of Data in the Canonical Spaces . . . . .	185
7.7	Summary . . . . .	189
	References . . . . .	190
<b>8</b>	<b>Cross-Species Candidate Gene Prioritization with MerKator . . . . .</b>	<b>191</b>
8.1	Introduction . . . . .	191
8.2	Data Sources . . . . .	192
8.3	Kernel Workflow . . . . .	194
8.3.1	Approximation of Kernel Matrices Using Incomplete Cholesky Decomposition . . . . .	194
8.3.2	Kernel Centering . . . . .	195
8.3.3	Missing Values . . . . .	197
8.4	Cross-Species Integration of Prioritization Scores . . . . .	197
8.5	Software Structure and Interface . . . . .	200
8.6	Results and Discussion . . . . .	201
8.7	Summary . . . . .	203
	References . . . . .	204
<b>9</b>	<b>Conclusion . . . . .</b>	<b>207</b>
	<b>Index . . . . .</b>	<b>209</b>