

---

## Contents

<b>1</b>	<b>Introduction</b> .....	1
1.1	Basic Statistical Learning Problems .....	2
1.2	Categorizations of Machine Learning Techniques .....	4
1.2.1	Unsupervised vs. Supervised .....	4
1.2.2	Generative Models vs. Discriminative Models .....	4
1.2.3	Models for Simple Data vs. Models for Complex Data ..	6
1.2.4	Model Identification vs. Model Prediction .....	7
1.3	Multimedia Content Analysis .....	8

---

### Part I Unsupervised Learning

---

<b>2</b>	<b>Dimension Reduction</b> .....	15
2.1	Objectives .....	15
2.2	Singular Value Decomposition .....	16
2.3	Independent Component Analysis .....	20
2.3.1	Preprocessing .....	23
2.3.2	Why Gaussian is Forbidden .....	24
2.4	Dimension Reduction by Locally Linear Embedding .....	26
2.5	Case Study .....	30
	Problems .....	34
<b>3</b>	<b>Data Clustering Techniques</b> .....	37
3.1	Introduction .....	37
3.2	Spectral Clustering .....	39
3.2.1	Problem Formulation and Criterion Functions .....	39
3.2.2	Solution Computation .....	42

3.2.3	Example	46
3.2.4	Discussions	50
3.3	Data Clustering by Non-Negative Matrix Factorization	51
3.3.1	Single Linear NMF Model	52
3.3.2	Bilinear NMF Model	55
3.4	Spectral vs. NMF	59
3.5	Case Study: Document Clustering Using Spectral and NMF	
	Clustering Techniques	61
3.5.1	Document Clustering Basics	62
3.5.2	Document Corpora	64
3.5.3	Evaluation Metrics	64
3.5.4	Performance Evaluations and Comparisons	65
	Problems	68

---

**Part II Generative Graphical Models**

---

<b>4</b>	<b>Introduction of Graphical Models</b>	73
4.1	Directed Graphical Model	74
4.2	Undirected Graphical Model	77
4.3	Generative vs. Discriminative	79
4.4	Content of Part II	80
<b>5</b>	<b>Markov Chains and Monte Carlo Simulation</b>	81
5.1	Discrete-Time Markov Chain	81
5.2	Canonical Representation	84
5.3	Definitions and Terminologies	88
5.4	Stationary Distribution	91
5.5	Long Run Behavior and Convergence Rate	94
5.6	Markov Chain Monte Carlo Simulation	100
5.6.1	Objectives and Applications	100
5.6.2	Rejection Sampling	101
5.6.3	Markov Chain Monte Carlo	104
5.6.4	Rejection Sampling vs. MCMC	110
	Problems	112
<b>6</b>	<b>Markov Random Fields and Gibbs Sampling</b>	115
6.1	Markov Random Fields	115
6.2	Gibbs Distributions	117

6.3	Gibbs – Markov Equivalence . . . . .	120
6.4	Gibbs Sampling . . . . .	123
6.5	Simulated Annealing . . . . .	126
6.6	Case Study: Video Foreground Object Segmentation by MRF . . . . .	133
6.6.1	Objective . . . . .	134
6.6.2	Related Works . . . . .	134
6.6.3	Method Outline . . . . .	135
6.6.4	Overview of Sparse Motion Layer Computation . . . . .	136
6.6.5	Dense Motion Layer Computation Using MRF . . . . .	138
6.6.6	Bayesian Inference . . . . .	140
6.6.7	Solution Computation by Gibbs Sampling . . . . .	141
6.6.8	Experimental Results . . . . .	143
	Problems . . . . .	146
<b>7</b>	<b>Hidden Markov Models . . . . .</b>	<b>149</b>
7.1	Markov Chains vs. Hidden Markov Models . . . . .	149
7.2	Three Basic Problems for HMMs . . . . .	153
7.3	Solution to Likelihood Computation . . . . .	154
7.4	Solution to Finding Likeliest State Sequence . . . . .	158
7.5	Solution to HMM Training . . . . .	160
7.6	Expectation-Maximization Algorithm and its Variances . . . . .	162
7.6.1	Expectation-Maximization Algorithm . . . . .	162
7.6.2	Baum-Welch Algorithm . . . . .	164
7.7	Case Study: Baseball Highlight Detection Using HMMs . . . . .	167
7.7.1	Objective . . . . .	167
7.7.2	Overview . . . . .	167
7.7.3	Camera Shot Classification . . . . .	169
7.7.4	Feature Extraction . . . . .	172
7.7.5	Highlight Detection . . . . .	173
7.7.6	Experimental Evaluation . . . . .	174
	Problems . . . . .	175
<b>8</b>	<b>Inference and Learning for General Graphical Models . . . . .</b>	<b>179</b>
8.1	Introduction . . . . .	179
8.2	Sum-product algorithm . . . . .	182
8.3	Max-product algorithm . . . . .	188

8.4	Approximate inference .....	189
8.5	Learning .....	191
	Problems .....	196

---

### Part III Discriminative Graphical Models

---

<b>9</b>	<b>Maximum Entropy Model and Conditional Random Field .....</b>	<b>201</b>
9.1	Overview of Maximum Entropy Model .....	202
9.2	Maximum Entropy Framework .....	204
9.2.1	Feature Function .....	204
9.2.2	Maximum Entropy Model Construction .....	205
9.2.3	Parameter Computation .....	208
9.3	Comparison to Generative Models .....	210
9.4	Relation to Conditional Random Field .....	213
9.5	Feature Selection .....	215
9.6	Case Study: Baseball Highlight Detection Using Maximum Entropy Model .....	217
9.6.1	System Overview .....	218
9.6.2	Highlight Detection Based on Maximum Entropy Model .....	220
9.6.3	Multimedia Feature Extraction .....	222
9.6.4	Multimedia Feature Vector Construction .....	226
9.6.5	Experiments .....	227
	Problems .....	232
<b>10</b>	<b>Max-Margin Classifications .....</b>	<b>235</b>
10.1	Support Vector Machines (SVMs) .....	236
10.1.1	Loss Function and Risk .....	237
10.1.2	Structural Risk Minimization .....	237
10.1.3	Support Vector Machines .....	239
10.1.4	Theoretical Justification .....	243
10.1.5	SVM Dual .....	244
10.1.6	Kernel Trick .....	245
10.1.7	SVM Training .....	248
10.1.8	Further Discussions .....	255
10.2	Maximum Margin Markov Networks .....	257
10.2.1	Primal and Dual Problems .....	257

10.2.2 Factorizing Dual Problem .....	259
10.2.3 General Graphs and Learning Algorithm .....	262
10.2.4 Max-Margin Networks vs. Other Graphical Models ....	262
Problems .....	264
<b>A Appendix</b> .....	267
<b>References</b> .....	269
<b>Index</b> .....	275