

Contents

Preface	v
1. Introduction	1
1.1 Background	1
1.2 Neuron Model	2
1.3 Historical Remarks	4
1.4 Network architecture	6
1.4.1 Supervised Neural Networks	6
1.4.1.1 McCulloch and Pitts Model	7
1.4.1.2 The Perceptron Model	11
1.4.1.3 Multi-layer Feedforward Network	14
1.4.1.4 Recurrent Networks	15
1.4.2 Unsupervised Neural Networks	17
1.5 Modeling and Learning Mechanism	19
1.5.1 Determination of Parameters	20
1.5.2 Gradient Descent Searching Method	26
Exercises	28
2. Learning Performance and Enhancement	31
2.1 Fundamental of Gradient Descent Optimization	32
2.2 Conventional Backpropagation Algorithm	35
2.3 Convergence Enhancement	42
2.3.1 Extended Backpropagation Algorithm	44
2.3.2 Least Squares Based Training Algorithm	47
2.3.3 Extended Least Squares Based Algorithm	55
2.4 Initialization Consideration	59
2.4.1 Weight Initialization Algorithm I	61
2.4.2 Weight Initialization Algorithm II	64
2.4.3 Weight Initialization Algorithm III	67
2.5 Global Learning Algorithms	69
2.5.1 Simulated Annealing Algorithm	70

2.5.2	Aloplex Algorithm	71
2.5.3	Reactive Tabu Search	72
2.5.4	The NOVEL Algorithm	73
2.5.5	The Heuristic Hybrid Global Learning Algorithm	74
2.6	Concluding Remarks	82
2.6.1	Fast Learning Algorithms	82
2.6.2	Weight Initialization Methods	83
2.6.3	Global Learning Algorithms	84
Appendix 2.1	85
Exercises	87
3.	Generalization and Performance Enhancement	91
3.1	Cost Function and Performance Surface	93
3.1.1	Maximum Likelihood Estimation	94
3.1.2	The Least-Square Cost Function	95
3.2	Higher-Order Statistic Generalization	98
3.2.1	Definitions and Properties of Higher-Order Statistics	99
3.2.2	The Higher-Order Cumulants based Cost Function	101
3.2.3	Property of the Higher-Order Cumulant Cost Function	105
3.2.4	Learning and Generalization Performance	108
3.2.4.1	Experiment one: Henon Attractor	109
3.2.4.2	Experiment Two: Sunspot time-series	116
3.3	Regularization for Generalization Enhancement	117
3.3.1	Adaptive Regularization Parameter Selection (ARPS) Method	120
3.3.1.1	Stalling Identification Method	121
3.3.1.2	λ Selection Schemes	122
3.3.2	Synthetic Function Mapping	124
3.4	Concluding Remarks	126
3.4.1	Objective function selection	128
3.4.2	Regularization selection	129
Appendix 3.1	Confidence Upper Bound of Approximation Error	131
Appendix 3.2	Proof of the Property of the HOC Cost Function	133
Appendix 3.3	The Derivation of the Sufficient Conditions of the Regularization Parameter	136
Exercises	137
4.	Basis Function Networks for Classification	139
4.1	Linear Separation and Perceptions	140
4.2	Basis Function Model for Parametric Smoothing	142
4.3	Radial Basis Function Network	144
4.3.1	RBF Networks Architecture	144
4.3.2	Universal Approximation	146

4.3.3 Initialization and Clustering	149
4.3.4 Learning Algorithms.	152
4.3.4.1 Linear Weights Optimization.	152
4.3.4.2 Gradient Descent Optimization	154
4.3.4.3 Hybrid of Least Squares and Penalized Optimization	155
4.3.5 Regularization Networks.	157
4.4 Advanced Radial Basis Function Networks	159
4.4.1 Support Vector Machine.	159
4.4.2 Wavelet Network	161
4.4.3 Fuzzy RBF Controllers.	164
4.4.4 Probabilistic Neural Networks	167
4.5 Concluding Remarks.	169
Exercises	170
5. Self-organizing Maps	173
5.1 Introduction	173
5.2 Self-Organizing Maps	177
5.2.1 Learning Algorithm	178
5.3 Growing SOMs	182
5.3.1 Cell Splitting Grid	182
5.3.2 Growing Hierarchical Self-Organizing Quadtree Map	185
5.4 Probabilistic SOMs	188
5.4.1 Cellular Probabilistic SOM	188
5.4.2 Probabilistic Regularized SOM	193
5.5 Clustering of SOM	202
5.6 Multi-Layer SOM for Tree-Structured data	205
5.6.1 SOM Input Representation	207
5.6.2 MLSOM Training	210
5.6.3 MLSOM visualization and classification	212
Exercises	216
6 Classification and Feature Selection	219
6.1 Introduction	219
6.2 Support Vector Machines (SVM)	223
6.2.1 Support Vector Machine Visualization (SVMV)	224
6.3 Cost Function	229
6.3.1 MSE and MCE Cost Functions	230
6.3.2 Hybrid MCE-MSE Cost Function	232
6.3.3 Implementing MCE-MSE	236
6.4 Feature Selection	239
6.4.1 Information Theory	241
6.4.1.1 Mutual Information	241

- 6.4.1.2 Probability density function (pdf) estimation 243
- 6.4.2 MI Based Forward Feature Selection 245
 - 6.4.2.1 MIFS and MIFS-U 247
 - 6.4.2.2 Using quadratic MI 248
- Exercises 253

- 7. Engineering Applications 255
 - 7.1 Electric Load Forecasting 255
 - 7.1.1 Nonlinear Autoregressive Integrated Neural Network Model . . 257
 - 7.1.2 Case Studies 261
 - 7.2 Content-based Image Retrieval Using SOM 266
 - 7.2.1 GHSOQM Based CBIR Systems 267
 - 7.2.1.1 Overall Architecture of GHSOQM-Based CBIR System 267
 - 7.2.1.2 Image Segmentation, Feature Extraction and Region-Based Feature Matrices 268
 - 7.2.1.3 Image Distance 269
 - 7.2.1.4 GHSOQM and Relevance Feedback in the CBIR System 270
 - 7.2.2 Performance Evaluation 274
 - 7.3 Feature Selection for cDNA Microarray 278
 - 7.3.1 MI Based Forward Feature Selection Scheme 279
 - 7.3.2 The Supervised Grid Based Redundancy Elimination 280
 - 7.3.3 The Forward Gene Selection Process Using MIIO and MISF . . 281
 - 7.3.4 Results 282
 - 7.3.4.1 Prostate Cancer Classification Dataset 284
 - 7.3.4.2 Subtype of ALL Classification Dataset 288
 - 7.3.5 Remarks 294

- Bibliography 291

- Index 305