

Shigeo Abe

Pattern Classification

Neuro-fuzzy Methods and
Their Comparison



Springer

Contents

Part I. Pattern Classification

1. Introduction	3
1.1 Development of a Classification System	3
1.2 Optimum Features	4
1.3 Classifiers	7
1.3.1 Neural Network Classifiers	7
1.3.2 Conventional Fuzzy Classifiers	10
1.3.3 Fuzzy Classifiers with Learning Capability	12
1.4 Evaluation	17
1.5 Data Sets Used in the Book	19
2. Multilayer Neural Network Classifiers	21
2.1 Three-layer Neural Networks	21
2.2 Synthesis Principles	23
2.3 Training Methods	31
2.4 Training by the Back-propagation Algorithm	34
2.5 Training by Solving Inequalities	35
2.5.1 Setting of Target Values	36
2.5.2 Formulation of Training by Solving Inequalities	37
2.5.3 Determination of Weights by Solving Inequalities	39
2.6 Performance Evaluation	41
2.6.1 Iris Data	41
2.6.2 Numeral Data	42
2.6.3 Thyroid Data	43
2.6.4 Blood Cell Data	43
2.6.5 Hiragana Data	45
2.6.6 Discussions	45
3. Support Vector Machines	47
3.1 Support Vector Machines for Pattern Classification	47
3.1.1 Conversion to Two-class Problems	48
3.1.2 The Optimal Hyperplane	48
3.1.3 Mapping to a High-dimensional Space	54
3.2 Performance Evaluation	56

3.2.1	Iris Data	56
3.2.2	Numeral Data	56
3.2.3	Thyroid Data	58
3.2.4	Blood Cell Data	58
3.2.5	Hiragana Data	58
3.2.6	Discussions	60
4.	Membership Functions	63
4.1	One-dimensional Membership Functions	63
4.1.1	Triangular Membership Functions	64
4.1.2	Trapezoidal Membership Functions	66
4.1.3	Bell-shaped Membership Functions	68
4.2	Multi-dimensional Membership Functions	69
4.2.1	Extension to Multi-dimensional Membership Functions	70
4.2.2	Rectangular Pyramidal Membership Functions	71
4.2.3	Truncated Rectangular Pyramidal Membership Functions	73
4.2.4	Polyhedral Pyramidal Membership Functions	74
4.2.5	Truncated Polyhedral Pyramidal Membership Functions	76
4.2.6	Bell-shaped Membership Functions	77
4.2.7	Relations between Membership Functions	79
5.	Static Fuzzy Rule Generation	81
5.1	Classifier Architecture	81
5.2	Fuzzy Rules	82
5.2.1	Fuzzy Rules with Pyramidal Membership Functions ..	82
5.2.2	Polyhedral Fuzzy Rules	85
5.2.3	Ellipsoidal Fuzzy Rules	90
5.3	Class Boundaries	93
5.3.1	Fuzzy Rules with Pyramidal Membership Functions ..	93
5.3.2	Ellipsoidal Fuzzy Rules	95
5.3.3	Class Boundaries for the Iris Data	97
5.4	Training Architecture	102
5.4.1	Fuzzy Rule Generation by Preclustering	103
5.4.2	Fuzzy Rule Generation by Postclustering	105
6.	Clustering	109
6.1	Fuzzy c-means Clustering Algorithm	109
6.2	The Kohonen Network	111
6.3	Minimum Volume Clustering Algorithm	113
6.4	Fuzzy Min-max Clustering Algorithm	115
6.5	Overlap Resolving Clustering Algorithm	116
6.5.1	Approximation of Overlapping Regions	116
6.5.2	Extraction of Data from the Overlapping Regions	117
6.5.3	Clustering Algorithm	118

7.	Tuning of Membership Functions	119
7.1	Problem Formulation	119
7.2	Direct Methods	120
7.2.1	Tuning of Slopes	121
7.2.2	Tuning of Locations	129
7.2.3	Order of Tuning	139
7.3	Indirect Methods	139
7.3.1	Tuning of Slopes Using the Least-squares Method	140
7.3.2	Tuning by the Steepest Descent Method	140
7.4	Performance Evaluation	141
7.4.1	Performance Evaluation of the Fuzzy Classifier with Pyramidal Membership Functions	141
7.4.2	Performance Evaluation of the Fuzzy Classifier with Polyhedral Regions	148
7.4.3	Performance Evaluation of the Fuzzy Classifier with Ellipsoidal Regions	150
8.	Robust Pattern Classification	159
8.1	Why Robust Classification Is Necessary?	159
8.2	Robust Classification	162
8.2.1	The First Stage	162
8.2.2	The Second Stage	164
8.2.3	Tuning Slopes near Class Boundaries	165
8.2.4	Upper and Lower Bounds Determined by Correctly Classified Data	167
8.2.5	Range of the Interclass Tuning Parameter that Re- solves Misclassification	168
8.3	Performance Evaluation	169
8.3.1	Classification Performance without Outliers	169
8.3.2	Classification Performance with Outliers	170
9.	Dynamic Fuzzy Rule Generation	177
9.1	Fuzzy Min-max Classifiers	177
9.1.1	Concept	177
9.1.2	Approximation of Input Regions	178
9.1.3	Fuzzy Rule Extraction	179
9.1.4	Performance Evaluation	181
9.2	Fuzzy Min-max Classifiers with Inhibition	184
9.2.1	Concept	184
9.2.2	Fuzzy Rule Extraction	185
9.2.3	Fuzzy Rule Inference	188
9.2.4	Performance Evaluation	191

10. Comparison of Classifier Performance	197
10.1 Evaluation Conditions	197
10.2 Iris Data	198
10.3 Numeral Data	198
10.4 Thyroid Data	198
10.5 Blood Cell Data	200
10.6 Hiragana Data	200
10.7 Discussions	202
11. Optimizing Features	205
11.1 Scaling Features	205
11.1.1 Scale and Translation Invariance of the Fuzzy Classifier with Pyramidal Membership Functions	206
11.1.2 Invariance of the Fuzzy Classifiers with Hyperbox Regions	207
11.1.3 Invariance of the Fuzzy Classifier with Ellipsoidal Regions	208
11.2 Feature Extraction	209
11.2.1 Principal Component Analysis	210
11.2.2 Discriminant Analysis	212
11.2.3 Neural-network-based Feature Extraction	212
11.3 Feature Selection	215
11.3.1 Criteria for Feature Selection	216
11.3.2 Monotonicity of the Selection Criteria	222
11.3.3 Selection Search	224
11.4 Performance Evaluation of Feature Selection	227
11.4.1 Evaluation Conditions	227
11.4.2 Selected Features	228
11.4.3 Iris Data	228
11.4.4 Numeral Data	230
11.4.5 Thyroid Data	232
11.4.6 Blood Cell Data	234
11.4.7 Discussions	236
12. Generation of Training and Test Data Sets	239
12.1 Resampling Techniques	239
12.1.1 Estimation of Classification Performance	239
12.1.2 Effect of Resampling on Classifier Performance	241
12.2 Division of Data by Pairing	241
12.3 Similarity Measures	243
12.3.1 Relative Difference of Center Norms	243
12.3.2 Relative Difference of Covariance Matrix Norms	244
12.4 Performance Evaluation	245
12.4.1 Evaluation Conditions	245
12.4.2 Blood Cell Data	245

12.4.3 Hiragana-50 Data	246
12.4.4 Discussions	246

Part II. Function Approximation

13. Introduction	251
13.1 Function Approximators	251
13.1.1 Neural Networks	251
13.1.2 Conventional Fuzzy Function Approximators	252
13.1.3 Fuzzy Function Approximators with Learning Capability	254
14. Fuzzy Rule Representation and Inference	257
14.1 Fuzzy Rule Representation	257
14.2 Defuzzification Methods	260
15. Fuzzy Rule Generation	263
15.1 Fuzzy Rule Generation by Preclustering	263
15.1.1 Clustering of Input Space	263
15.1.2 Clustering of Input and Output Spaces	264
15.2 Fuzzy Rule Generation by Postclustering	264
15.2.1 Concept	265
15.2.2 Rule Generation for FACG	267
15.2.3 Rule Generation for FALC	269
15.3 Fuzzy Rule Tuning	270
15.3.1 Concept	270
15.3.2 Evaluation Functions	271
15.3.3 Tuning of Parameters	272
15.4 Performance Evaluation	274
15.4.1 Mackey-Glass Differential Equation	274
15.4.2 Water Purification Plant	279
15.4.3 Discussions	285
16. Robust Function Approximation	287
16.1 Introduction	287
16.2 Fuzzy Function Approximator with Ellipsoidal Regions	289
16.2.1 Fuzzy Rule Representation	289
16.2.2 Fuzzy Rule Generation	290
16.3 Robust Parameter Estimation	290
16.3.1 Robust Estimation by the Least-median-of-squares Method	290
16.3.2 Robust Estimation by the Least-squares Method with Partial Data	291
16.4 Performance Evaluation	292

16.4.1 Determination of the Multiplier	292
16.4.2 Performance Comparison	294
16.5 Discussions	297

Part III. Appendices

A. Conventional Classifiers	301
A.1 Bayesian Classifiers	301
A.2 Nearest Neighbor Classifiers	302
A.2.1 Classifier Architecture	302
A.2.2 Performance Evaluation	303
B. Matrices	307
B.1 Matrix Properties	307
B.2 Least-squares Method and Singular Value Decomposition	309
B.3 Covariance Matrix	311
References	315
Index	323