

Practical Neural Network Recipes in C++

Timothy Masters



Academic Press
San Diego New York Boston
London Sydney Tokyo Toronto

Contents

Preface	xvii
1. Foundations	1
Motivation	2
New Life for Old Techniques	3
Perceptrons and Linear Separability	4
Neural Network Capabilities	6
Basic Structure of a Neural Network	8
Training	9
Validation	10
Leave- <i>k</i> -out Method	12
2. Classification	15
Binary Decisions	16
Making the Decision	17
Multiple Classes	18
Reject Category	18
Making the Decision	19
Other Encoding Schemes	19
Supervised versus Unsupervised Training	21
3. Autoassociation	23
Autoassociative Filtering	24
Code for Autoassociative Filtering	28
Noise Reduction	29
Learning a Prototype from Exemplars	31
Exposing Isolated Events	32
Pattern Completion	40
Error Correction	41
Encoding Words	42
Data Compression	44
4. Time-Series Prediction	47
The Basic Model	49
Input Data	50
Trend Elimination	51
Code for Detrending and Retrending	55
Seasonal Variation	58

Differencing	60
Scaling	60
Multiple Prediction	61
Multiple Predictors	62
Measuring Prediction Error	64
5. Function Approximation	67
Univariate Function Approximation	68
Inverse Modeling	72
Multiple Regression	74
6. Multilayer Feedforward Networks	77
Basic Architecture	78
Activation Functions	80
Example Network	82
Linear Output Neurons	84
Theoretical Discussion	85
Bibliography of Feedforward Network Theory	88
Algorithms for Executing the Network	90
Training the Network	94
Training by Backpropagation of Errors	100
Training by Conjugate Gradients	105
Minimizing along a Direction	106
Choosing the Direction for Minimization	110
Eluding Local Minima in Learning	111
Local Minima Happen Easily	112
Mistaken Minima	114
Other Means of Escape	115
When to Use a Multiple-Layer Feedforward Network	116
7. Eluding Local Minima I: Simulated Annealing	117
Overview	118
Choosing the Annealing Parameters	119
Implementation in Feedforward Network Learning	121
A Sample Program	122
A Sample Function	126
Random Number Generation	128
Going on from Here	132

8. Eluding Local Minima II: Genetic Optimization	135
Overview	136
Designing the Genetic Structure	138
Evaluation	140
Parent Selection	144
Reproduction	147
Mutation	148
A Genetic Minimization Subroutine	149
Some Functions for Genetic Optimization	155
Advanced Topics in Genetic Optimization	157
Gray Codes	157
Overinitialization	159
Two-Point Crossover	159
9. Regression and Neural Networks	165
Overview	166
Singular-Value Decomposition	167
Regression in Neural Networks	169
10. Designing Feedforward Network Architectures	173
How Many Hidden Layers?	174
How Many Hidden Neurons?	176
How Long Do I Train This Thing???	180
11. Interpreting Weights: How Does This Thing Work?	187
Features Used by Networks in General	190
Features Used by a Particular Network	191
Examination of Weight Vectors	191
Hinton Diagrams	192
Clustering	194
Sensitivity Analysis	195
Stereotypical Inputs	197
12. Probabilistic Neural Networks	201
Overview	202
Computational Aspects	208
Optimizing Sigma	209
Related Models	210
A Sample Program	211
Optimizing Sigma	213

Other Optimization Criteria	218
Bayesian Confidence Measures	219
Autoassociative Versions	220
When to Use a Probabilistic Neural Network	221
13. Functional Link Networks	223
Application to Nonlinear Approximation	226
Mathematics of the Functional Link Network	227
When to Use a Functional Link Network	229
14. Hybrid Networks	231
Functional Link Net as a Hidden Layer	232
Fast Bayesian Confidences	235
Training	238
Attention-based Processing	239
Factorable Problems	242
Training the Data Reduction Networks	243
Splitting Is Not Always Effective	244
15. Designing the Training Set	245
Number of Samples	246
Communal Random Errors	247
Overfitting	247
Network Size Affects Training Set Size	248
Stratified Training Data	249
Borderline Cases	249
Hidden Bias	250
Balancing the Classes	251
Fudging Cases	251
16. Preparing Input Data	253
General Considerations	254
Types of Measurements	255
Nominal Variables	255
Ordinal Variables	259
Interval Variables	262
Ratio Variables	266
Is Scaling Always Necessary?	266
Transformations	267
Circular Discontinuity	270

View Angles	271
Hue	272
Outliers	274
Discarding Data	275
Missing Data	276
17. Fuzzy Data and Processing	279
Treating Fuzzy Values as Nominal and Ordinal	281
Advantages of Fuzzy Set Processing	282
The Neural Network – Fuzzy Set Interface	283
Membership Functions	284
Continuous Variables	287
Multivariate Domains	288
Discrete Variables	289
Hedges	289
Negation, Conjunction, and Disjunction	290
Modus Ponens	292
Combining Operations	295
Defuzzification	299
Maximum Height Method	300
Centroid Method	301
Code for Fuzzy Set Operations	303
Constructors	303
Negation and Scaling	307
Conjunction and Disjunction	308
Centroid	314
Examples of Neural Network Fuzzy Preprocessing	316
Simplifying Interactions	316
Fuzzy One-of- n Coding	317
Examples of Neural Network Fuzzy Postprocessing	319
Simple Membership Output	319
Postprocessing with Defuzzification	320
18. Unsupervised Training	327
Input Normalization	330
Z-Axis Normalization	331
Training the Kohonen Network	332
Updating the Weights	334
Learning Rate	336
Measuring Network Error	337

Determining Convergence	338
Neurons That Refuse to Learn	339
Self-Organization	340
19. Evaluating Performance of Neural Networks	343
Overview	344
Mean Square Error	344
Problems with Mean Square Error	345
Relatives of Mean Square Error	346
Cost Functions	347
Confusion Matrix	348
ROC (Receiver Operating Characteristic) Curves	351
Computing the ROC Curve Area	354
Cost Functions and ROC Curves	357
Signal-to-Noise Ratio	359
20. Confidence Measures	361
Testing Individual Hypotheses	362
Computing Confidence	367
Confidence in the Null Hypothesis	368
Multiple Classes	369
Confidence in the Confidence	370
Example Programs	371
Sorting	372
Estimating the Distribution	373
Estimating Confidences	374
Bayesian Methods	376
Example Program	381
Multiple Classes	382
Hypothesis Testing versus Bayes' Method	384
21. Optimizing the Decision Threshold	389
22. Using the NEURAL Program	403
Output Models	405
CLASSIFY Model	405
AUTO Model	405
GENERAL Model	405
Building the Training Set	406
The LAYER Network Model	406

Initialization by Simulated Annealing	407
Initialization by Genetic Optimization	407
Learning	408
The KOHONEN Network Model	409
Initialization and Learning	410
Confusion Matrices	412
Saving Weights and Execution Results	412
Alphabetical Glossary of Commands	413
Verification of Program Operation	417
Appendix	423
Bibliography	479
Index	491