

Contents

Preface	xi
1. INTRODUCTION	1
1.1 Inductive Learning	3
1.1.1 Learning and Regression	4
1.1.2 Polynomial Models	5
1.1.3 Inductive Computation Machinery	5
1.2 Why Polynomial Networks?	7
1.2.1 Advantages of Polynomial Networks	8
1.2.2 Multilayer Polynomial Networks	9
1.3 Evolutionary Search	16
1.3.1 STROGANOFF and its Variants	17
1.4 Neural Network Training	21
1.5 Bayesian Inference	22
1.6 Statistical Model Validation	23
1.7 Organization of the Book	23
2. INDUCTIVE GENETIC PROGRAMMING	25
2.1 Polynomial Neural Networks (PNN)	26
2.1.1 PNN Approaches	27
2.1.2 Tree-structured PNN	29
2.2 IGP Search Mechanisms	35
2.2.1 Sampling and Control Issues	36
2.2.2 Biological Interpretation	36
2.3 Genetic Learning Operators	38
2.3.1 Context-preserving Mutation	38
2.3.2 Crossover Operator	40

2.3.3	Size-biasing of the Genetic Operators	41
2.3.4	Tree-to-Tree Distance	42
2.4	Random Tree Generation	46
2.5	Basic IGP Framework	48
2.6	IGP Convergence Characteristics	50
2.6.1	Schema Theorem of IGP	50
2.6.2	Markov Model of IGP	51
2.7	Chapter Summary	54
3.	TREE-LIKE PNN REPRESENTATIONS	55
3.1	Discrete Volterra Series	56
3.2	Mapping Capabilities of PNN	57
3.3	Errors of Approximation	59
3.3.1	Approximation Error Bounds	59
3.3.2	Empirical Risk	60
3.4	Linear Polynomial Networks	62
3.4.1	Horizontal PNN Models	62
3.4.2	Kernel PNN Models	66
3.5	Nonlinear Polynomial Networks	68
3.5.1	Block PNN Models	68
3.5.2	Orthogonal PNN Models	69
3.5.3	Trigonometric PNN Models	71
3.5.4	Rational PNN Models	75
3.5.5	Dynamic PNN Models	77
3.6	Chapter Summary	80
4.	FITNESS FUNCTIONS AND LANDSCAPES	81
4.1	Fitness Functions	83
4.1.1	Static Fitness Functions	84
4.1.2	Dynamic Fitness Functions	91
4.1.3	Fitness Magnitude	94
4.2	Fitness Landscape Structure	95
4.3	Fitness Landscape Measures	96
4.3.1	Statistical Correlation Measures	96
4.3.2	Probabilistic Measures	102
4.3.3	Information Measures	104
4.3.4	Quantitative Measures	107
4.4	Chapter Summary	109

5. SEARCH NAVIGATION	111
5.1 The Reproduction Operator	112
5.1.1 Selection Strategies	113
5.1.2 Replacement Strategies	117
5.1.3 Implementing Reproduction	118
5.2 Advanced Search Control	119
5.2.1 Macroevoolutionary Search	119
5.2.2 Memetic Search	120
5.2.3 Search by Genetic Annealing	122
5.2.4 Stochastic Genetic Hillclimbing	124
5.2.5 Coevolutionary Search	125
5.2.6 Distributed Search	128
5.3 Performance Examination	128
5.3.1 Fitness Evolvability	129
5.3.2 Convergence Measures	130
5.3.3 Diversity Measures	133
5.3.4 Measures of Self-Organization	139
5.4 Chapter Summary	146
6. BACKPROPAGATION TECHNIQUES	147
6.1 Multilayer Feed-forward PNN	148
6.2 First-Order Backpropagation	149
6.2.1 Gradient Descent Search	150
6.2.2 First-Order Error Derivatives	151
6.2.3 Batch Backpropagation	157
6.2.4 Incremental Backpropagation	158
6.2.5 Control of the Learning Step	159
6.2.6 Regularized Delta Rule	162
6.3 Second-Order Backpropagation	163
6.3.1 Second-Order Error Derivatives	164
6.3.2 Newton's Method	169
6.3.3 Pseudo-Newton Training	170
6.3.4 Conjugate Gradients	170
6.3.5 Levenberg-Marquardt Method	171
6.4 Rational Backpropagation	172
6.5 Network Pruning	176
6.5.1 First-Order Network Pruning	176
6.5.2 Second-Order Network Pruning	177

6.6	Chapter Summary	179
7.	TEMPORAL BACKPROPAGATION	181
7.1	Recurrent PNN as State-Space Models	182
7.2	Backpropagation Through Time	184
7.2.1	Derivation of BPTT Algorithms	185
7.2.2	Real-Time BPTT Algorithm	189
7.2.3	Epochwise BPTT Algorithm	190
7.3	Real-Time Recurrent Learning	191
7.4	Improved Dynamic Training	198
7.4.1	Teacher Enforced Training	198
7.4.2	Truncating in Time	199
7.4.3	Subgrouping	199
7.4.4	Common Temporal Training Problem	200
7.5	Second-Order Temporal BP	200
7.6	Recursive Backpropagation	204
7.7	Recurrent Network Optimization	206
7.7.1	Regularization	207
7.7.2	Recurrent Network Pruning	207
7.8	Chapter Summary	208
8.	BAYESIAN INFERENCE TECHNIQUES	209
8.1	Bayesian Error Function	211
8.2	Bayesian Neural Network Inference	212
8.2.1	Deriving Hyperparameters	215
8.2.2	Local vs. Global Regularization	217
8.2.3	Evidence Procedure for PNN Models	218
8.2.4	Predictive Data Distribution	221
8.2.5	Choosing a Weight Prior	221
8.3	Bayesian Network Pruning	222
8.4	Sparse Bayesian Learning	224
8.5	Recursive Bayesian Learning	229
8.5.1	Sequential Weight Estimation	229
8.5.2	Sequential Dynamic Hessian Estimation	230
8.5.3	Sequential Hyperparameter Estimation	232
8.6	Monte Carlo Training	234
8.6.1	Markov Chain Monte Carlo	235
8.6.2	Importance Resampling	237

8.6.3 Hybrid Sampling Resampling	237
8.7 Chapter Summary	239
9. STATISTICAL MODEL DIAGNOSTICS	241
9.1 Deviations of PNN Models	242
9.2 Residual Bootstrap Sampling	243
9.3 The Bias/Variance Dilemma	244
9.3.1 Statistical Bias and Variance	244
9.3.2 Measuring Bias and Variance	245
9.4 Confidence Intervals	248
9.4.1 Interval Estimation by the Delta Method	248
9.4.2 Bootstrapping Confidence Intervals	252
9.5 Prediction Intervals	254
9.5.1 Analytical Prediction Intervals	255
9.5.2 Empirical Learning of Prediction Bars	256
9.6 Bayesian Intervals	262
9.6.1 Analytical Bayesian Intervals	263
9.6.2 Empirical Bayesian Intervals	265
9.7 Model Validation Tests	267
9.8 Chapter Summary	271
10. TIME SERIES MODELLING	273
10.1 Time Series Modelling	274
10.2 Data Preprocessing	276
10.3 PNN vs. Linear ARMA Models	277
10.4 PNN vs. Genetically Programmed Functions	279
10.5 PNN vs. Statistical Learning Networks	281
10.6 PNN vs. Neural Network Models	283
10.7 PNN vs. Kernel Models	285
10.8 Recurrent PNN vs. Recurrent Neural Networks	288
10.9 Chapter Summary	290
11. CONCLUSIONS	291
References	295
Index	313