Contents

PREFACE xxi

ACKNOWLEDGMENTS xxiii

Overview 1

O.1 Early Biomedical Systems 1
O.1.1 History 1
O.1.2 Medical Records 2
O.1.3 Drawbacks of Traditional Approaches 3
O.1.4 Numerical versus Symbolic Approaches 3
O.2 Medical and Biological Data 3
O.2.1 Binary data 4
O.2.2 Categorical data 4
O.2.3 Integer Data 4
O.2.4 Continuous Data 4
O.2.5 Fuzzy Data 4
O.2.6 Temporal Data 5
O.2.7 Time Series Data 6
O.2.8 Image Data 8
O.3 Organization of the Book 8
References 9

PART I NEURAL NETWORKS

CHAPTER 1 Foundations of Neural Networks 13

1.1 Objectives of Neural Networks 13
1.1.1 Modeling Biomedical Systems 13
1.1.2 Establishment of Decision-Making Systems 14

1.2 Biological Foundations of Neural Networks 14
1.2.1 Structure of the Neuron 14
1.2.2 Structure of the Central Nervous System 16
1.3 Early Neural Models 16
1.3.1 McCulloch and Pitts Neuron 16
1.3.2 Hebbian Learning 17
1.3.3 ADALINE 17
1.3.4 Rosenblatt Perceptron 17
1.3.5 Problems with Early Systems 18
1.4 Precursor to Current Models: Pattern Classification 19
1.4.1 Feature Extraction 19
1.4.2 Supervised Learning 21
1.4.3 Unsupervised Learning 21
1.4.4 Learning Algorithms 22
1.5 Resurgence of the Neural Network Approach 24
1.6 Basic Concepts 25
1.6.1 Artificial Neurons 25
1.6.2 Selection of Input Nodes 25
1.6.3 Network Structure 25
1.6.3.1 Feed-Forward Networks 25
1.6.3.2 Feed-Backward Networks 25
1.6.4 Learning Mechanism 26
1.6.5 Output 26
1.7 Summary 26
References 27

CHAPTER 2 Classes of Neural Networks 29

2.1 Basic Network Properties 29
2.1.1 Terminology 29
2.1.2 Structure of Networks 29
2.1.3 Computational Properties of Nodes 30
2.1.4 Algorithm Design 31
2.2 Classification Models 32
2.3 Association Models 32
2.3.1 Hopfield Nets 33
2.3.1.1 Theoretical Basis 33
2.3.2 Other Associative Memory Approaches 34
2.3.2.1 Theoretical Basis of Bidirectional Associative Memory (BAM) 34
2.3.3 Hamming Net 36
2.3.3.1 Theoretical Basis 36
2.3.4 Applications of Association Models 37
2.4 Optimization Models 38
2.4.1 Hopfield Net 38
2.4.2 Boltzmann Machines 39
2.4.2.1 Theoretical Basis 39
2.4.3 Applications of Optimization Models 40
2.5 Self-Organization Models 40

2.6 Radial Basis Functions (RBFs) 41
2.6.1 Theoretical Basis 41
2.6.2 Applications of Radial Basis Functions 42
2.7 Summary 43
References 43

CHAPTER 3 Classification Networks and Learning 45

3.1 Network Structure 45
3.1.1 Layer Definition 45
3.1.2 Input Layer 45
3.1.3 Hidden Layer 45
3.1.4 Output Layer 45
3.2 Feature Selection 46
3.2.1 Types of Variables 46
3.2.2 Feature Vectors 46
3.2.3 Image Data 47
3.2.4 Time Series Data 47
3.2.4.1 Chaotic Analysis of Time Series 48
3.2.4.2 Graphical Measures of Chaos 49
3.2.4.3 Numerical Measures of Chaos 49
3.2.5 Issues of Dimensionality 50
3.3 Types of Learning 51
3.3.1 Supervised Learning 51
3.3.1.1 Selection of Training and Test Sets 51
3.3.1.2 Selection of Learning Algorithm 52
3.3.2 Unsupervised Learning 52
3.3.3 Causal Models 55
3.4 Interpretation of Output 55
3.5 Summary 55
References 56

CHAPTER 4 Supervised Learning 59

4.1 Decision Surfaces 59
4.2 Two-Category Separation, Linearly Separable Sets 60
4.2.1 Fisher's Linear Discriminant 60
4.2.2 Gradient Descent Procedures 61
4.2.3 Perceptron Algorithm 62
4.2.4 Relaxation Procedures 62
4.2.5 Potential Functions 63
4.3 Nonlinearly Separable Sets 64
4.3.1 Nonlinear Discriminant Functions 64
4.3.2 Hypernet, A Nonlinear Potential Function Algorithm 64
4.3.3 Categorization of Nonlinearly Separable Sets 65
4.3.3.1 Minimum Squared Error Procedures (MSEs) 65
4.3.3.2 Ho–Kashyap Procedure 66
4.4 Multiple Category Classification Problems 66
4.4.1 Extension of Fisher Discriminant 66

4.4.2 Kesler Construction 67
4.4.3 Backpropagation 68
4.5 Relationship to Neural Network Models 69
4.6 Comparison of Methods 70
4.6.1 Convergence and Stability 70
4.6.2 Training Time 70
4.6.3 Predictive Power 70
4.7 Applications 71
4.7.1 Single-Category Classification 71
4.7.2 Multicategory Classification 72
4.7.3 Reduction of Nodes 74
4.8 Summary 74
References 76

CHAPTER 5 Unsupervised Learning 79

5.1 Background 79
5.2 Clustering 79
5.2.1 Basic Isodata 79
5.2.2 Similarity Measures 80
5.2.3 Criterion Functions 80
5.2.3.1 Sum of Squared Error Criteria 80
5.2.3.2 Minimum Error Criteria 81
5.2.3.3 Scattering Criteria 81
5.2.3.4 Iterative Optimization 81
5.2.4 Hierarchical Clustering 82
5.2.5 Metrics 82
5.3 Kohonen Networks and Competitive Learning 83
5.4 Hebbian Learning 85
5.5 Adaptive Resonance Theory (ART) 86
5.6 Applications 88
5.6.1 Dimensionality Reduction 88
5.6.1.1 Minimization of Criteria Functions 88
5.6.1.2 Clustering and Dimensionality Reduction 89
5.6.2 Biomedical Applications 89
5.6.3 Diagnosis of CAD as a Clustering Problem 89
5.6.4 Other Biomedical Applications 90
5.7 Summary 91
References 92

CHAPTER 6 Design Issues 95

6.1 Introduction 95
6.1.1 Objective of the Model 95
6.1.2 Information Sources 95
6.2 Input Data Types 96
6.2.1 Extracting Information from the Medical Record 96
6.2.2 Using Information from Data Collection Sheets 97
6.2.2.1 Coding Multiple Responses 98
6.2.2.2 Ordering Categorical Data 99

6.2.2.3 Changing Categorical Data into Binary Data 99
6.2.2.4 Considering the Use of Fuzzy Input Data 99
6.2.2.5 Missing Data 100
6.2.3 Time Series Data 100
6.2.4 Image Data 101
6.3 Structure of Networks 101
6.3.1 Number of Layers 101
6.3.2 Connectivity 101
6.4 Implications of Network Structures 101
6.4.1 Classification Potential 101
6.4.2 Training 103
6.4.2.1 Number of Cases Required 103
6.4.3 Reduction in Number of Nodes 104
6.4.3.1 Expert Intervention 104
6.4.3.2 Statistical Significance 104
6.4.3.3 Threshold Pruning 104
6.4.3.4 Other Techniques 105
6.4.4 Output 105
6.5 Choice of Learning Algorithm 105
6.6 Summary 106
References 106

CHAPTER 7 Comparative Analysis 109

7.1 Introduction 109
7.2 Input Data Considerations 109
7.3 Supervised Learning Algorithms 110
7.3.1 Gradient Descent Procedures 111
7.3.2 Extensions to Nonlinear Decision Functions 111
7.3.3 Extensions to Multiple Categories 113
7.4 Unsupervised Learning 114
7.4.1 Clustering Methods 114
7.4.1.1 Choice of Method 114
7.4.1.2 Choice of Distance Measure 114
7.4.1.3 Limitations 115
7.4.2 Self-Organization Networks 115
7.4.2.1 Kohonen Networks 115
7.4.2.2 ART Networks 115
7.5 Network Structures 115
7.5.1 Number of Categories 115
7.5.2 Connectivity 115
7.6 Interpretation of Results 116
7.6.1 Supervised Learning 116
7.6.1.1 Training 116
7.6.1.2 Evaluation 116
7.6.1.3 Classification 117
7.6.2 Unsupervised Learning 117
7.6.2.1 Training 117
7.6.2.2 Evaluation 118
7.6.2.3 Classification 118

7.6.3 Data Scaling and Normalization 118
7.6.4 Dependence on Training Data 118
7.7 Summary 118
References 119

CHAPTER 8 Validation and Evaluation 121

8.1 Introduction 121
8.2 Data Checking 121
8.2.1 Verification of Accuracy of Data for Training 121
8.2.1.1 Chart Review 121
8.2.1.2 Databases 122
8.2.1.3 Prospective Studies 122
8.2.2 Appropriateness of Data for Training 122
8.2.3 Use of Gold Standards in Supervised Learning 123
8.3 Validation of Learning Algorithm 123
8.3.1 Technical Integrity of Algorithm 123
8.3.2 Appropriateness of Algorithm for Given Application 123
8.3.2.1 Theoretical Considerations 123
8.3.2.2 Practical Considerations 123
8.3.2.3 Comparison to Other Approaches 123
8.4 Evaluation of Performance 124
8.4.1 Supervised Learning Algorithms 124
8.4.1.1 Performance of Algorithm on Test Set 124
8.4.1.2 Relevance of Contributing Parameters 124
8.4.1.3 Comparison to Other Methods 124
8.4.2 Unsupervised Learning 124
8.4.2.1 Applicability to Other Data Sets 125
8.5 Summary 126
References 127

PART II ARTIFICIAL INTELLIGENCE

CHAPTER 9 Foundations of Computer-Assisted Decision Making 131

9.1 Motivation for Computer-Assisted Decision Making 131
9.2 Databases and Medical Records 131
9.2.1 The First Decade (1970–1980) 131
9.2.2 The Second Decade (1980–1990) 134
9.2.3 Current Approaches to Medical Databases 134
9.3 Mathematical Modeling and Simulation 135
9.4 Pattern Recognition 136
9.5 Bayesian Analysis 137
9.5.1 Early Bayesian Systems 137
9.5.2 Bayesian Belief Networks 138
9.6 Decision Theory 138
9.7 Symbolic Reasoning Techniques 139
9.7.1 Early Expert Systems 139
9.7.2 Second-Generation Expert Systems 142

9.7.2.1 Causal Systems 143
9.7.2.2 Reasoning with Uncertainty 143
9.7.2.3 Hybrid Systems 144
9.8 Summary 144
References 145

CHAPTER 10 Knowledge Representation 151

10.1 Production Rules 151
10.1.1 General Structure 151
10.1.2 Methods of Confirming Conditions 152
10.1.2.1 Matching 152
10.1.2.1.1 Natural Language Processing 152
10.1.2.1.2 Partial Matching 152
10.1.2.1.3 Data-Driven Approaches 153
10.1.2.2 Conflict Resolution 153
10.1.3 Rule Searching Strategies 153
10.1.4 Expanded Production Rule Systems 155
10.1.5 Certainty Factors 156
10.1.6 Advantages of Production Systems 158
10.1.7 Disadvantages of Production Systems 158
10.1.8 Areas of Applications of Production Rule Systems 160
10.2 Frames 160
10.2.1 General Structure 160
10.2.2 Inheritance 161
10.2.3 Example of a Frame-Based Decision-Support System 161
10.3 Databases 162
10.3.1 Relational Databases 163
10.3.2 Query Languages 165
10.3.3 Object-Oriented Databases 165
10.4 Predicate Calculus and Semantic Nets 166
10.4.1 Syntax 166
10.4.2 Method of Proof 167
10.4.3 Semantic Trees 167
10.4.4 Semantic Interpretation 167
10.4.5 Applications 167
10.5 Temporal Data Representations 168
10.5.1 Knowledge-Based Systems 168
10.5.2 Data-Based Systems 169
10.6 Summary 169
References 171

CHAPTER 11 Knowledge Acquisition 173

11.1 Introduction 173
11.2 Expert Input 173
11.2.1 Acquisition for a New System 173
11.2.1.1 Use of Existing Material 173
11.2.1.2 Techniques for Knowledge Elicitation 175
11.2.1.3 Dealing with Multiple Experts 176

11.2.1.4 Consistency 177
11.2.1.5 Completeness 177
11.2.2 Updating of Information 177
11.3 Learned Knowledge 178
11.3.1 Rote Learning 178
11.3.2 Learning by Being Told 178
11.3.3 Learning from Examples 178
11.3.3.1 Pattern Classification and Neural Networks 178
11.3.3.2 Learning Single Concepts 178
11.3.3.3 Learning Multiple Concepts 179
11.4 Meta-Knowledge 181
11.5 Knowledge Base Maintenance 182
11.5.1 Assuring Accuracy 182
11.5.1.1 Permission to Change Knowledge Base 182
11.5.1.2 Verifying Sources of Information 182
11.5.2 Maintaining Consistency 183
11.6 Summary 183
References 183

CHAPTER 12 Reasoning Methodologies 185

12.1 Introduction 185
12.2 Problem Representations 185
12.2.1 Graphs and Trees 185
12.2.2 Binary Trees 187
12.2.3 State-Space Representations 188
12.2.4 Problem Reduction Representations 189
12.2.5 Game Trees 190
12.3 Blind Searching 190
12.3.1 Depth-First Search 190
12.3.2 Breadth-First Search 191
12.4 Ordered Search 191
12.4.1 Uniform Cost Method 192
12.4.2 Using Heuristic Information 192
12.5 AND/OR Trees 193
12.5.1 Breadth-First Search and Depth-First of AND/OR Tree 194
12.5.2 Costs of Solution Trees 194
12.5.3 Ordered Searching of AND/OR Trees 195
12.6 Searching Game Trees 196
12.6.1 Minimax 196
12.6.2 Alpha-Beta 196
12.7 Searching Graphs 196
12.7.1 Breadth-First Graph Searching 197
12.7.2 Uniform Cost Algorithm 197
12.7.3 AND/OR Graphs 197
12.8 Rule Base Searching 197
12.8.1 Backward-Chaining 197
12.8.2 Forward-Chaining 198
12.9 Higher-Level Reasoning Methodologies 198
12.9.1 Inference Engines 198

12.9.1.1 Binary Logic Engines 198
12.9.1.2 Approximate Reasoning Engines 199
12.9.1.3 Fuzzy Logic Engines 199
12.9.1.4 Probability-Based Engines 199
12.9.2 Cognitive Models 200
12.9.3 Automatic Deduction 200
12.9.4 Natural Language Processing 200
12.10 Examples in Biomedical Expert Systems 201
12.11 Summary 201
References 203

CHAPTER 13 Validation and Evaluation 205

13.1 Introduction 205
13.2 Algorithmic Evaluation 205
13.2.1 Searching Algorithm 205
13.2.1.1 AND/OR Trees 205
13.2.1.2 Control by Meta Rules 206
13.2.1.3 Data-Driven Search 206
13.2.1.4 Hierarchical Search 206
13.2.2 Inference Engine 206
13.2.2.1 Binary Engines 206
13.2.2.2 Complex Engines 207
13.2.3 Aggregation of Evidence 207
13.2.4 Use of Meta Knowledge 207
13.3 Knowledge Base Evaluation 207
13.3.1 Expert-Derived Information 207
13.3.1.1 Individual Units of Knowledge 208
13.3.1.2 Combination of Knowledge 208
13.3.2 Learned Information 208
13.3.3 Meta Knowledge 208
13.4 System Evaluation 208
13.4.1 Evaluation with Original Knowledge Base 209
13.4.2 Evaluation with Updated Knowledge Base 209
13.4.3 Evaluation of User Interaction Parameters 209
13.4.3.1 Interpretation of Input 209
13.4.3.2 Explanation Capabilities 210
13.4.4 Validation with Clinical Data 210
13.4.4.1 Sources of Clinical Data 210
13.4.4.1.1 Retrospective Studies 210
13.4.4.2 Selection of Cases 210
13.4.4.3 Actual Outcome 211
13.4.4.4 Results from Clinical Decision-Support System 211
13.5 Summary 212
References 213

PART III ALTERNATIVE APPROACHES

CHAPTER 14 Genetic Algorithms 217

14.1 Foundations 217
14.2 Representation Schemes 218

14.3 Evaluation Functions 218
14.4 Genetic Operators 218
14.4.1 Mutation 218
14.4.2 Crossover 218
14.5 Evolution Strategies 219
14.5.1 Genetic Algorithms 219
14.5.2 Optimization Strategies 219
14.5.2.1 Hill Climbing 219
14.5.2.2 Simulated Annealing 219
14.5.3 Genetic Search 220
14.6 Biomedical Examples 221
14.6.1 Literature References 221
14.7 Summary 222
References 223

CHAPTER 15 Probabilistic Systems 225

15.1 Introduction 225
15.2 Bayesian Approaches 225
15.2.1 Bayes' Rule 225
15.2.2 Bayes' Decision Theory 226
15.2.3 Risk Analysis 226
15.2.4 Supervised Bayesian Learning 228
15.2.5 Decision Trees 229
15.3 Parameter Estimation 231
15.3.1 Maximum Likelihood Estimation 231
15.3.2 Bayesian Estimation 232
15.4 Discriminant Analysis 233
15.5 Statistical Pattern Classification 234
15.6 Unsupervised Learning 235
15.6.1 Parzen Windows 235
15.6.2 Nearest Neighbor Algorithms 236
15.6.3 Mixture Densities and Maximum Likelihood Estimates 236
15.6.4 Unsupervised Bayesian Learning 237
15.7 Regression Analysis 238
15.8 Biomedical Applications 240
15.9 Summary 240
References 241

CHAPTER 16 Fuzzy Systems 243

16.1 Introduction 243
16.2 Fuzzy Information 243
16.2.1 Input Data 243
16.2.2 Fuzzy Logic and Fuzzy Set Theory 244
16.2.3 Representation of Fuzzy Variables 244
16.2.4 Membership Functions 245

16.3 Fuzzy Neural Networks 245
16.4 Fuzzy Approaches for Supervised Learning Networks 246
16.4.1 Pre-Processing of Fuzzy Input 246
16.4.2 Propagation of Results 247
16.4.2.1 Max–Min Networks 247
16.4.2.2 Learning Algorithms for Interval Data 247
16.4.2.3 Analogue Models 248
16.5 Fuzzy Generalizations of Unsupervised Learning Methods 248
16.5.1 Fuzzy Associative Memories 248
16.5.2 Fuzzy Clustering 248
16.6 Reasoning with Uncertain Information 249
16.6.1 Uncertainty in Input Data 250
16.6.1.1 Degree of Presence Input 250
16.6.1.2 Linguistic Input 250
16.6.2 Uncertainty in Knowledge Base 250
16.6.3 Inference Engines for Uncertain Information 251
16.6.3.1 Binary Logic Engines 251
16.6.3.2 Fuzzy Logic Engines 252
16.6.4 Evidential Reasoning 252
16.6.4.1 Possibility Theory 252
16.6.4.2 Probabilistic Approaches 253
16.6.4.3 Dempster–Shafer Belief Theory 253
16.6.5 Compatibility Indices 254
16.6.6 Approximate Reasoning 255
16.7 Pre-Processing and Post-Processing Using Fuzzy Techniques 256
16.8 Applications in Biomedical Engineering 257
16.9 Summary 258
References 259

CHAPTER 17 Hybrid Systems 261

17.1 Hybrid Systems Approaches 261
17.2 Components of Hybrid Systems 261
17.2.1 Knowledge-Based Approaches 262
17.2.1.1 Knowledge Base 262
17.2.1.2 Input Data 263
17.2.1.3 Decision-Making Algorithm 264
17.2.1.4 Output Data 264
17.2.2 Data-Based Approaches 264
17.2.2.1 Neural Networks 265
17.2.2.2 Genetic Algorithms 266
17.2.3 General Methodologies 266
17.2.3.1 Fuzzy Logic 266
17.2.3.2 Statistical Approaches 267
17.3 Use of Complex Data Structures 267
17.3.1 Time Series Data 267
17.3.1.1 Automatic ECG Analyzers 267
17.3.1.2 Summary Methods 268
17.3.2 Image Data 268
17.4 Design Methodologies 268
17.4.1 System Structure 268

17.4.2 User Interfaces 269
17.4.3 Pre-Processing 269
17.4.4 Post-Processing 269
17.4.5 Presentation of Results 270
17.5 Summary 270
References 270

CHAPTER 18 HyperMerge, a Hybrid Expert System 273

18.1 Introduction 273
18.2 Knowledge-Based Component 273
18.2.1 Crisp Implementation 273
18.2.2 Partial Substantiation of Antecedents 274
18.2.3 Weighted Antecedents and Partial Substantiation of Rules 275
18.2.4 Handling of Temporal Data 276
18.3 Neural Network Component 276
18.3.1 Learning Algorithm 276
18.3.2 Special Data Types 277
18.3.2.1 Temporal Data 277
18.3.2.2 Crisp versus Fuzzy Temporal Data 277
18.3.2.3 Time Series Data 278
18.4 Analysis of Time Series Data 278
18.4.1 Chaos Theory 278
18.4.2 Continuous versus Discrete Chaotic Modeling 279
18.4.3 Difference Equations and Graphs 280
18.4.4 Central Tendency Measure 281
18.5 Combined System 282
18.5.1 Weighting of Antecedents 282
18.5.2 Determination of Thresholds 282
18.5.3 Neural Network with Symbolic Layer 283
18.6 Application: Diagnosis of Heart Disease 284
18.6.1 Categories of Heart Disease 284
18.6.2 Knowledge-Based Information 285
18.6.3 Data-Based Information 285
18.6.4 Chaotic Data 285
18.6.5 Sample System 288
18.7 Summary 289
References 289

CHAPTER 19 Future Perspectives 291

19.1 Introduction 291
19.2 Effects of Hardware Advances 291
19.2.1 Faster Computing Speeds 291
19.2.2 Increased Memory 292
19.2.3 Parallel Machines 292
19.2.4 Miniaturization 292
19.2.5 Organic Semiconductors 292
19.3 Effects of Increase in Knowledge 293
19.3.1 Information Explosion 293
19.3.2 Human Genome Project 293

19.3.3 Proliferation of Databases 293
19.3.4 Communication of Information 294
19.4 The Future of Software 294
19.4.1 Hybrid Systems 294
19.4.2 Parallel Systems 294
19.4.3 Nontextual Data 295
19.4.4 Neural Network Models 295
19.4.5 Artificial Intelligence Approaches 295
References 295

INDEX 297

ABOUT THE AUTHORS 305