# CONTENTS

Preface xi  
Acknowledgments xv  

## 1 Introduction  
1.1 The Machine Intelligence Research 1  
1.2 The Two-Fold Objectives: Data-Driven and Biologically Inspired Approaches 4  
1.3 How to Read This Book 8  
1.3.1 Part I: Data-Driven Approaches for Machine Intelligence (Chapters 2, 3, and 4) 8  
1.3.2 Part II: Biologically-Inspired Approaches for Machine Intelligence (Chapters 4, 5, and 6) 9  
1.4 Summary and Further Reading 10  
References 10  

## 2 Incremental Learning  
2.1 Introduction 13  
2.2 Problem Foundation 13  
2.3 An Adaptive Incremental Learning Framework 14  
2.4 Design of the Mapping Function 19  
2.4.1 Mapping Function Based on Euclidean Distance 19  
2.4.2 Mapping Function Based on Regression Learning Model 20  
2.4.3 Mapping Function Based on Online Value System 23  
2.4.3.1 A Three-Curve Fitting (TCF) Technique 23  
2.4.3.2 System-Level Architecture for Online Value Estimation 26
3.3.3 Kernel Modification Methods for Imbalanced Learning 70
3.3.4 Active Learning Methods for Imbalanced Learning 71
3.3.5 Additional Methods for Imbalanced Learning 73
3.4 Assessment Metrics for Imbalanced Learning 75
  3.4.1 Singular Assessment Metrics 75
  3.4.2 Receiver Operating Characteristics (ROC) Curves 77
  3.4.3 Precision-Recall (PR) Curves 79
  3.4.4 Cost Curves 80
  3.4.5 Assessment Metrics for Multiclass Imbalanced Learning 80
3.5 Opportunities and Challenges 82
3.6 Case Study 84
  3.6.1 Nonlinear Normalization 84
  3.6.2 Data Sets Distribution 88
  3.6.3 Simulation Results and Discussions 92
3.7 Summary 98
References 100

4 Ensemble Learning 108
4.1 Introduction 108
4.2 Hypothesis Diversity 108
  4.2.1 Q-Statistics 109
  4.2.2 Correlation Coefficient 110
  4.2.3 Disagreement Measure 110
  4.2.4 Double-Fault Measure 110
  4.2.5 Entropy Measure 111
  4.2.6 Kohavi–Wolpert Variance 111
  4.2.7 Interrater Agreement 111
  4.2.8 Measure of Difficulty 112
  4.2.9 Generalized Diversity 112
4.3 Developing Multiple Hypotheses 114
  4.3.1 Bootstrap Aggregating 114
  4.3.2 Adaptive Boosting 114
CONTENTS

4.3.3 Subspace Learning 119
4.3.4 Stacked Generalization 120
4.3.5 Mixture of Experts 122

4.4 Integrating Multiple Hypotheses 123

5 Adaptive Dynamic Programming for Machine Intelligence 140
5.1 Introduction 140
5.2 Fundamental Objectives: Optimization and Prediction 141
5.3 ADP for Machine Intelligence 143
5.3.1 Hierarchical Architecture in ADP Design 143
5.3.2 Learning and Adaptation in ADP 146
5.3.2.1 The Action Network 148
5.3.2.2 The Reference Network 150
5.3.2.3 The Critic Network 152
5.3.3 Learning Strategies: Sequential Learning and Cooperative Learning 154

5.4 Case Study 155
5.5 Summary 160

6 Associative Learning 165
6.1 Introduction 165
6.2 Associative Learning Mechanism 165
6.2.1 Structure Individual Processing Elements 166
6.2.2 Self-Determination of the Function Value 167
6.2.3 Signal Strength for Associative Learning 168

References 137

56

155

160

161

165
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.4</td>
<td>The Associative Learning Principle</td>
<td>169</td>
</tr>
<tr>
<td>6.3</td>
<td>Associative Learning in Hierarchical Neural Networks</td>
<td>173</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Network Structure</td>
<td>173</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Network Operation</td>
<td>174</td>
</tr>
<tr>
<td>6.3.2.1</td>
<td>Feedforward Operation</td>
<td>174</td>
</tr>
<tr>
<td>6.3.2.2</td>
<td>Feedback Operation</td>
<td>177</td>
</tr>
<tr>
<td>6.4</td>
<td>Case Study</td>
<td>180</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Hetero-Associative Application</td>
<td>180</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Auto-Associative Application</td>
<td>182</td>
</tr>
<tr>
<td>6.4.2.1</td>
<td>Panda Image Recovery</td>
<td>183</td>
</tr>
<tr>
<td>6.4.2.2</td>
<td>Chinese Character Recognition and Recovery</td>
<td>184</td>
</tr>
<tr>
<td>6.4.2.3</td>
<td>Associative Memory for Online Incremental Learning</td>
<td>186</td>
</tr>
<tr>
<td>6.5</td>
<td>Summary</td>
<td>187</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>188</td>
</tr>
<tr>
<td>7</td>
<td>Sequence Learning</td>
<td>190</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>190</td>
</tr>
<tr>
<td>7.2</td>
<td>Foundations for Sequence Learning</td>
<td>190</td>
</tr>
<tr>
<td>7.3</td>
<td>Sequence Learning in Hierarchical Neural Structure</td>
<td>194</td>
</tr>
<tr>
<td>7.4</td>
<td>Level 0: A Modified Hebbian Learning Architecture</td>
<td>195</td>
</tr>
<tr>
<td>7.5</td>
<td>Level 1 to Level N: Sequence Storage, Prediction, and Retrieval</td>
<td>198</td>
</tr>
<tr>
<td>7.5.1</td>
<td>Sequence Storage</td>
<td>198</td>
</tr>
<tr>
<td>7.5.2</td>
<td>Sequence Prediction</td>
<td>201</td>
</tr>
<tr>
<td>7.5.2.1</td>
<td>Prediction Mechanism</td>
<td>201</td>
</tr>
<tr>
<td>7.5.2.2</td>
<td>Activation of Prediction Neuron</td>
<td>205</td>
</tr>
<tr>
<td>7.5.2.3</td>
<td>Time-Controlled Multiplexer</td>
<td>205</td>
</tr>
<tr>
<td>7.5.3</td>
<td>Sequence Retrieval</td>
<td>207</td>
</tr>
<tr>
<td>7.6</td>
<td>Memory Requirement</td>
<td>207</td>
</tr>
<tr>
<td>7.7</td>
<td>Learning and Anticipation of Multiple Sequences</td>
<td>208</td>
</tr>
<tr>
<td>7.8</td>
<td>Case Study</td>
<td>211</td>
</tr>
<tr>
<td>7.9</td>
<td>Summary</td>
<td>212</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>213</td>
</tr>
</tbody>
</table>
CONTENTS

8 Hardware Design for Machine Intelligence 217

8.1 A Final Comment 217

References 220

List of Abbreviations 222

Index 227