Advances in Fuzzy Clustering and its Applications

EDITED BY

J. Valente de Oliveira
University of Algarve, Portugal

W. Pedrycz
University of Alberta, Canada
Systems Research Institute of the Polish Academy of Sciences, Poland

John Wiley & Sons, Ltd
Contents

List of Contributors xi
Foreword xv
Preface xvii

Part I  Fundamentals 1

1  Fundamentals of Fuzzy Clustering 3
   Rudolf Kruse, Christian Döring and Marie-Jeanne Lesot
   1.1 Introduction 3
   1.2 Basic Clustering Algorithms 4
   1.3 Distance Function Variants 14
   1.4 Objective Function Variants 18
   1.5 Update Equation Variants: Alternating Cluster Estimation 25
   1.6 Concluding Remarks 27
   Acknowledgements 28
   References 29

2  Relational Fuzzy Clustering 31
   Thomas A. Runkler
   2.1 Introduction 31
   2.2 Object and Relational Data 31
   2.3 Object Data Clustering Models 34
   2.4 Relational Clustering 38
   2.5 Relational Clustering with Non-spherical Prototypes 41
   2.6 Relational Data Interpreted as Object Data 45
   2.7 Summary 46
   2.8 Experiments 46
   2.9 Conclusions 49
   References 50

3  Fuzzy Clustering with Minkowski Distance Functions 53
   Patrick J.F. Groenen, Uzay Kaymak and Joost van Rosmalen
   3.1 Introduction 53
   3.2 Formalization 54
   3.3 The Majorizing Algorithm for Fuzzy C-means with Minkowski Distances 56
3.4 The Effects of the Robustness Parameter $\lambda$ 60
3.5 Internet Attitudes 62
3.6 Conclusions 65
References 66

4 Soft Cluster Ensembles 69
Kunal Punera and Joydeep Ghosh
4.1 Introduction 69
4.2 Cluster Ensembles 71
4.3 Soft Cluster Ensembles 75
4.4 Experimental Setup 78
4.5 Soft vs. Hard Cluster Ensembles 82
4.6 Conclusions and Future Work s
Acknowledgements 90
References 90

Part II Visualization 93

5 Aggregation and Visualization of Fuzzy Clusters Based on Fuzzy Similarity Measures 95
János Abonyi and Balázs Feil
5.1 Problem Definition 97
5.2 Classical Methods for Cluster Validity and Merging 99
5.3 Similarity of Fuzzy Clusters 100
5.4 Visualization of Clustering Results 103
5.5 Conclusions 116
Appendix 5A.1 Validity Indices 117
Appendix 5A.2 The Modified Sammon Mapping Algorithm 120
Acknowledgements 120
References 120

6 Interactive Exploration of Fuzzy Clusters 123
Bernd Wiswedel, David E. Patterson and Michael R. Berthold
6.1 Introduction 123
6.2 Neighborgram Clustering 125
6.3 Interactive Exploration 131
6.4 Parallel Universes 135
6.5 Discussion 136
References 136

Part III Algorithms and Computational Aspects 137

7 Fuzzy Clustering with Participatory Learning and Applications 139
Leila Roling Scariot da Silva, Fernando Gomide and Ronald Yager
7.1 Introduction 139
7.2 Participatory Learning 140
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3</td>
<td>Participatory Learning in Fuzzy Clustering</td>
<td>142</td>
</tr>
<tr>
<td>7.4</td>
<td>Experimental Results</td>
<td>145</td>
</tr>
<tr>
<td>7.5</td>
<td>Applications</td>
<td>148</td>
</tr>
<tr>
<td>7.6</td>
<td>Conclusions</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Acknowledgements</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>152</td>
</tr>
<tr>
<td>8</td>
<td>Fuzzy Clustering of Fuzzy Data</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td><em>Pierpaolo D'Urso</em></td>
<td></td>
</tr>
<tr>
<td>8.1</td>
<td>Introduction</td>
<td>155</td>
</tr>
<tr>
<td>8.2</td>
<td>Informational Paradigm, Fuzziness and Complexity in Clustering</td>
<td>156</td>
</tr>
<tr>
<td>8.3</td>
<td>Fuzzy Data</td>
<td>160</td>
</tr>
<tr>
<td>8.4</td>
<td>Fuzzy Clustering of Fuzzy Data</td>
<td>165</td>
</tr>
<tr>
<td>8.5</td>
<td>An Extension: Fuzzy Clustering Models for Fuzzy Data Time Arrays</td>
<td>176</td>
</tr>
<tr>
<td>8.6</td>
<td>Applicative Examples</td>
<td>180</td>
</tr>
<tr>
<td>8.7</td>
<td>Concluding Remarks and Future Perspectives</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>189</td>
</tr>
<tr>
<td>9</td>
<td>Inclusion-based Fuzzy Clustering</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td><em>Samia Nefti-Meziani and Mourad Oussalah</em></td>
<td></td>
</tr>
<tr>
<td>9.1</td>
<td>Introduction</td>
<td>193</td>
</tr>
<tr>
<td>9.2</td>
<td>Background: Fuzzy Clustering</td>
<td>195</td>
</tr>
<tr>
<td>9.3</td>
<td>Construction of an Inclusion Index</td>
<td>196</td>
</tr>
<tr>
<td>9.4</td>
<td>Inclusion-based Fuzzy Clustering</td>
<td>198</td>
</tr>
<tr>
<td>9.5</td>
<td>Numerical Examples and Illustrations</td>
<td>201</td>
</tr>
<tr>
<td>9.6</td>
<td>Conclusions</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>Acknowledgements</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>Appendix 9A.1</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>208</td>
</tr>
<tr>
<td>10</td>
<td>Mining Diagnostic Rules Using Fuzzy Clustering</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td><em>Giovanna Castellano, Anna M. Fanelli and Corrado Mencar</em></td>
<td></td>
</tr>
<tr>
<td>10.1</td>
<td>Introduction</td>
<td>211</td>
</tr>
<tr>
<td>10.2</td>
<td>Fuzzy Medical Diagnosis</td>
<td>212</td>
</tr>
<tr>
<td>10.3</td>
<td>Interpretability in Fuzzy Medical Diagnosis</td>
<td>213</td>
</tr>
<tr>
<td>10.4</td>
<td>A Framework for Mining Interpretable Diagnostic Rules</td>
<td>216</td>
</tr>
<tr>
<td>10.5</td>
<td>An Illustrative Example</td>
<td>221</td>
</tr>
<tr>
<td>10.6</td>
<td>Concluding Remarks</td>
<td>226</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>226</td>
</tr>
<tr>
<td>11</td>
<td>Fuzzy Regression Clustering</td>
<td>229</td>
</tr>
<tr>
<td></td>
<td><em>Mikal Sato-Ilic</em></td>
<td></td>
</tr>
<tr>
<td>11.1</td>
<td>Introduction</td>
<td>229</td>
</tr>
<tr>
<td>11.2</td>
<td>Statistical Weighted Regression Models</td>
<td>230</td>
</tr>
<tr>
<td>11.3</td>
<td>Fuzzy Regression Clustering Models</td>
<td>232</td>
</tr>
<tr>
<td>11.4</td>
<td>Analyses of Residuals on Fuzzy Regression Clustering Models</td>
<td>237</td>
</tr>
</tbody>
</table>
CONTENTS

11.5 Numerical Examples 242
11.6 Conclusion 245
References 245

12 Implementing Hierarchical Fuzzy Clustering in Fuzzy Modeling
Using the Weighted Fuzzy C-means 247
George E. Tsekouras

12.1 Introduction 247
12.2 Takagi and Sugeno’s Fuzzy Model 248
12.3 Hierarchical Clustering-based Fuzzy Modeling 249
12.4 Simulation Studies 256
12.5 Conclusions 261
References 261

13 Fuzzy Clustering Based on Dissimilarity Relations Extracted from Data 265
Mario G.C.A. Cimino, Beatrice Lazzerini and Francesco Marcelloni

13.1 Introduction 265
13.2 Dissimilarity Modeling 267
13.3 Relational Clustering 275
13.4 Experimental Results 280
13.5 Conclusions 281
References 281

14 Simultaneous Clustering and Feature Discrimination
with Applications 285
Hichem Frigui

14.1 Introduction 285
14.2 Background 287
14.3 Simultaneous Clustering and Attribute Discrimination (SCAD) 289
14.4 Clustering and Subset Feature Weighting 296
14.5 Case of Unknown Number of Clusters 298
14.6 Application 1: Color Image Segmentation 298
14.7 Application 2: Text Document Categorization and Annotation 302
14.8 Application 3: Building a Multi-modal Thesaurus from Annotated Images 305
14.9 Conclusions 309
Appendix 14A.1 310
Acknowledgements 311
References 311

Part IV Real-time and Dynamic Clustering 313

15 Fuzzy Clustering in Dynamic Data Mining – Techniques and Applications 315
Richard Weber

15.1 Introduction 315
15.2 Review of Literature Related to Dynamic Clustering 315
15.3 Recent Approaches for Dynamic Fuzzy Clustering 317
16 Fuzzy Clustering of Parallel Data Streams
Jürgen Beringer and Eyke Hüllermeier

16.1 Introduction
16.2 Background
16.3 Preprocessing and Maintaining Data Streams
16.4 Fuzzy Clustering of Data Streams
16.5 Quality Measures
16.6 Experimental Validation
16.7 Conclusions
References

17 Algorithms for Real-time Clustering and Generation of Rules from Data
Dimitar Filev and Plamer Angelov

17.1 Introduction
17.2 Density-based Real-time Clustering
17.3 FSPC: Real-time Learning of Simplified Mamdani Models
17.4 Applications
17.5 Conclusion
References

Part V Applications and Case Studies

18 Robust Exploratory Analysis of Magnetic Resonance Images using FCM with Feature Partitions
Mark D. Alexiuk and Nick J. Pizzi

18.1 Introduction
18.2 FCM with Feature Partitions
18.3 Magnetic Resonance Imaging
18.4 FMRI Analysis with FCMP
18.5 Data-sets
18.6 Results and Discussion
18.7 Conclusion
Acknowledgements
References

19 Concept Induction via Fuzzy C-means Clustering in a High-dimensional Semantic Space
Dawei Song, Guihong Cao, Peter Bruza and Raymond Lau

19.1 Introduction
19.2 Constructing a High-dimensional Semantic Space via Hyperspace Analogue to Language
19.3 Fuzzy C-means Clustering