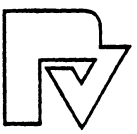
Fuzzy Classifier Design



Studies in Fuzziness and Soft Computing

Editor-in-chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw, Poland
E-mail: kacprzyk@ibspan.waw.pl
http://www.springer.de/cgi-bin/search_book.pl?series=2941

Vol. 3. A. Geyer-Schulz
Fuzzy Rule-Based Expert Systems and Genetic Machine Learning, 2nd ed. 1996
ISBN 3-7908-0964-0

Vol. 4. T. Onisawa and J. Kacprzyk (Eds.)
Reliability and Safety Analyses under Fuzziness, 1995
ISBN 3-7908-0837-7

Vol. 5. P. Bosc and J. Kacprzyk (Eds.)
Fuzziness in Database Management Systems, 1995
ISBN 3-7908-0858-X

Vol. 6. E.S. Lee and Q. Zhu
Fuzzy and Evidence Reasoning, 1995
ISBN 3-7908-0880-6

Vol. 7. B.A. Juliano and W. Bandler
Tracing Chains-of-Thought, 1996
ISBN 3-7908-0922-5

Vol. 8. F. Herrera and J.L. Verdegay (Eds.)
Genetic Algorithms and Soft Computing, 1996
ISBN 3-7908-0956-X

Vol. 9. M. Sato et al.
Fuzzy Clustering Models and Applications, 1997
ISBN 3-7908-1026-6

Vol. 10. L.C. Jain (Ed.)
Soft Computing Techniques in Knowledge-based Intelligent Engineering Systems, 1997
ISBN 3-7908-1035-5

Vol. 11. W. Mielczarski (Ed.)
Fuzzy Logic Techniques in Power Systems, 1998
ISBN 3-7908-1044-4

Vol. 12. B. Bouchon-Meunier (Ed.)
Aggregation and Fusion of Imperfect Information, 1998
ISBN 3-7908-1048-7

Vol. 13. E. Orłowska (Ed.)
Incomplete Information: Rough Set Analysis, 1998
ISBN 3-7908-1049-5

Vol. 14. E. Hisdal
Logical Structures for Representation of Knowledge and Uncertainty, 1998
ISBN 3-7908-1056-8

Vol. 15. G.J. Klir and M.J. Wierman
Uncertainty-Based Information, 2nd ed., 1999
ISBN 3-7908-1242-0

Vol. 16. D. Driankov and R. Palm (Eds.)
Advances in Fuzzy Control, 1998
ISBN 3-7908-1090-8

Vol. 17. L. Reznik, V. Dimitrov and J. Kacprzyk (Eds.)
Fuzzy Systems Design, 1998
ISBN 3-7908-1118-1

Vol. 18. L. Polkowski and A. Skowron (Eds.)
Rough Sets in Knowledge Discovery 1, 1998
ISBN 3-7908-1119-X

Vol. 19. L. Polkowski and A. Skowron (Eds.)
Rough Sets in Knowledge Discovery 2, 1998
ISBN 3-7908-1120-3

Vol. 20. J.N. Mordeson and P.S. Nair
Fuzzy Mathematics, 1998
ISBN 3-7908-1121-1

Vol. 21. L.C. Jain and T. Fukuda (Eds.)
Soft Computing for Intelligent Robotic Systems, 1998
ISBN 3-7908-1147-5

Vol. 22. J. Cardoso and H. Camargo (Eds.)
Fuzziness in Petri Nets, 1999
ISBN 3-7908-1158-0

Vol. 23. P.S. Szczepaniak (Ed.)
Computational Intelligence and Applications, 1999
ISBN 3-7908-1161-0

Vol. 24. E. Orłowska (Ed.)
Logic at Work, 1999
ISBN 3-7908-1164-5

continued on page 315

Ludmila I. Kuncheva

Fuzzy Classifier Design

With 113 Figures
and 81 Tables

Springer-Verlag Berlin Heidelberg GmbH

Dr. Ludmila I. Kuncheva
School of Informatics
University of Wales, Bangor
Dean Street, Bangor
Gwynedd LL57 1UT
United Kingdom
E-mail: L.I.Kuncheva@bangor.ac.uk

ISBN 978-3-7908-2472-8 ISBN 978-3-7908-1850-5 (eBook)
DOI 10.1007/978-3-7908-1850-5

Cataloging-in-Publication Data applied for
Die Deutsche Bibliothek – CIP-Einheitsaufnahme
Kuncheva, Ludmila I.: Fuzzy classifier design: with 81 tables / Ludmila I. Kuncheva. – Springer-Verlag Berlin Heidelberg GmbH
(Studies in fuzziness and soft computing; Vol. 49)

Originally published by Physica-Verlag Heidelberg in 2000
Softcover reprint of the hardcover 1st edition 2000

Hardcover Design: Erich Kirchner, Heidelberg

Preface

Fuzzy sets were first proposed by Lotfi Zadeh in his seminal paper [366] in 1965, and ever since have been a center of many discussions, fervently admired and condemned. Both proponents and opponents consider the arguments pointless because none of them would step back from their territory. And still, discussions burst out from a single sparkle like a conference paper or a message on some fuzzy-mail newsgroup. Here is an excerpt from an e-mail message posted in 1993 to fuzzy-mail@vexpert.dbai.twvien.ac.at, by somebody who signed "Dave".

> '... Why then the "logic" in "fuzzy logic"? I don't think anyone has successfully used fuzzy sets for logical inference, nor do I think anyone will. In my admittedly neophyte opinion, "fuzzy logic" is a misnomer, an oxymoron. (I would be delighted to be proven wrong on that.)... I came to the fuzzy literature with an open mind (and open wallet), high hopes and keen interest. I am very much disillusioned with "fuzzy" per se, but I did happen across some extremely interesting things along the way."

Dave, thanks for the nice quote! Enthusiastic on the surface, are not many of us suspicious deep down? In some books and journals the word *fuzzy* is religiously avoided: fuzzy set theory is viewed as a second-hand cheap trick whose aim is nothing else but to devalue good classical theories and open up the way to lazy ignorants and newcomers. This view is sometimes attributed to the conservatism of the "western culture" compared to the enterprising and practical "eastern culture" that took fuzzy sets on board right from the very start [342]. The other face of this coin is the "fuzzy" euphoria going under the motto "fuzzify the unfuzzifiable", sometimes with little or no rationale. Despite the unfinished philosophic debate there are a remarkable number of research papers and engineering applications involving fuzzy sets – some good and some bad, like in any other branch of science.

What are fuzzy classifiers and where do they stand in this stream? Do we need fuzzy classifiers or are we satisfied with pattern recognition "classics"? Are fuzzy classifiers better in some sense? In which cases and by how much are they better? Are fuzzy classifiers just "fuzzified" versions of non-fuzzy classifiers or are they an original contribution?

The difficulty in writing a coherent text on fuzzy classifiers roots in the diversity of the field itself. A paragraph taken from the Introduction by John Hartigan to the book "Clustering and Classification" [20], 1996, reads:

> 'Let us agree that classification does not fit easily into any of the academic disciplines in which its practitioners live. We remain too diffuse and dispersed in our methods and applications to form a coherent academic discipline ourselves.'

The author speaks about clustering and classification, but insert *fuzzy* in front of these two words, and the statement is enhanced by orders of magnitude.

Fuzzy Classifier Design is an attempt to address and answer some of the above questions without giving a comprehensive account of the area. My intention was not to take sides in the pro-con argument but to highlight some issues which would let you, my reader, make your own educated choice about fuzzy classifier design.

Bangor, Wales
February, 2000

Ludmila Ilieva Kuncheva

Contents

1. Introduction 1
- 1.1 What are fuzzy classifiers? 1
  - 1.1.1 Three "fuzzy" definitions of a fuzzy classifier 1
  - 1.1.2 Why should we use fuzzy classifiers? 3
  - 1.1.3 What obstructs using fuzzy classifiers? 4
- 1.2 The data sets used in this book 5
  - 1.2.1 Small synthetic data sets 5
  - 1.2.2 Two benchmark synthetic data sets 6
  - 1.2.3 Two real data sets 8
- 1.3 Notations and acronyms 10
- 1.4 Organization of the book 10
- 1.5 Acknowledgements 12

2. Statistical pattern recognition 15
- 2.1 Class, feature, feature space 15
- 2.2 Classifier, discriminant functions, classification regions 17
- 2.3 Clustering 19
  - 2.3.1 Hard c-means 19
  - 2.3.2 Single linkage clustering 20
- 2.4 Prior probabilities, class-conditional probability density functions, posterior probabilities 22
- 2.5 Minimum error and minimum risk classification. Loss matrix. 23
- 2.6 Performance estimation 24
  - 2.6.1 Classification accuracy, error rate, training, testing, and generalization 25
- 2.7 Experimental comparison of classifiers 27
  - 2.7.1 Error rate comparison 28
  - 2.7.2 Experiment design 31
  - 2.7.3 Classifier complexity 32
  - 2.7.4 Comments and recommendations 33
- 2.8 A taxonomy of classifier design methods 34

3. Statistical classifiers 37
3.1 Parametric classifiers 37
3.1.1 Normal distribution. Dependent features 37
3.1.2 Linear discriminant classifier (LDC) 39
3.1.3 Quadratic discriminant classifier (QDC) 41
3.2 Nonparametric classifiers 42
3.2.1 Preliminaries 42
3.2.2 Method of histograms 44
3.2.3 The k-nearest neighbor rule 47
3.2.4 Parzen classifier 50
3.3 Finding k-nn prototypes 52
3.3.1 Edited k-nn (prototype selection) 54
3.3.2 Calculating prototypes from data (prototype extraction) 56
3.4 Neural networks 60
3.4.1 What are they? 60
3.4.2 Neuron: the processing unit 62
3.4.3 The Rosenblatt perceptron 64
3.4.4 Multi-layer perceptron (MLP) 65
3.4.5 Backpropagation training of MLP 66
3.4.6 Radial basis function (RBF) networks 71
3.4.7 Learning vector quantization (LVQ) 74
3.4.8 An experimental example 75

4. Fuzzy sets 79
4.1 Fuzzy logic, an oxymoron? 79
4.2 Basic definitions 80
4.2.1 Fuzzy set, membership function 80
4.2.2 Support, core, height, level-set 83
4.2.3 Cardinality, complement, measures of fuzziness 84
4.3 Operations on fuzzy sets 85
4.3.1 Intersections and unions, t-norms and t-conorms 85
4.3.2 Aggregation operations 92
4.3.3 Aggregation of more than two fuzzy sets 95
4.3.4 Ordered weighted averaging (OWA) 97
4.3.5 Fuzzy integral 98
4.3.6 Using consensus in fuzzy aggregation 102
4.3.7 Equivalence, inclusion, similarity, and consistency 105
4.3.8 Fuzzy relations 107
4.4 Determining membership functions 108
4.4.1 Modeling issues, approaches and difficulties 108
4.4.2 Modeling methods 110

5. Fuzzy if-then classifiers 117
5.1 Fuzzy if-then systems 117
5.1.1 Fuzzification 118
5.1.2 Fuzzy inference 119
5.1.3 Defuzzification 123
5.2 Function approximation with fuzzy if-then systems 125
5.2.1 MA approximation 125
5.2.2 TSK approximation 129
5.3 Fuzzy if-then classifiers 131
5.3.1 Function approximation and classification 131
5.3.2 The general fuzzy if-then classifier model 132
5.3.3 Rules as prototypes 133
5.3.4 Types of class labels 133
5.3.5 MA fuzzy if-then classifiers 134
5.3.6 TSK fuzzy if-then classifiers 135
5.4 Universal approximation and equivalences of fuzzy if-then classifiers 138
5.4.1 Exact match of the classification boundary 138
5.4.2 Universal approximation 143
5.4.3 Equivalence between fuzzy TSK classifiers and 1-nn and Parzen classifier 147
5.4.4 A caveat: fuzzy classifiers are look-up tables with hyperbox cells 148
5.4.5 Why fuzzy if-then classifiers then? 154

6. Training of fuzzy if-then classifiers 157
6.1 Expert opinion or data analysis? 157
6.2 Tuning the consequents 158
6.2.1 Crisp consequents 158
6.2.2 Fuzzy consequents 159
6.2.3 Distributed fuzzy rules 165
6.3 Tuning the antecedents 166
6.3.1 Means method 166
6.3.2 Histogram method 167
6.3.3 Min-max (hyperbox) fuzzy designs 169
6.4 Tuning antecedents *and* consequents using clustering 181
6.5 Genetic algorithms for tuning fuzzy if-then classifiers 185
6.5.1 Genetic algorithms (GAs) 185
6.5.2 GAs for tuning fuzzy classifiers 189
6.6 Fuzzy classifiers and neural networks: hybridization or identity? 194
6.7 Forget interpretability and choose a model 195

7. Non if-then fuzzy models 197
7.1 Early ideas 197
7.1.1 Fuzzy and probabilistic models 197
7.1.2 Fuzzy linear discriminant analysis 198
7.1.3 Fuzzy relational classifiers 199
7.2 Fuzzy k-nearest neighbors (k-nn) designs 204
7.2.1 Combinations of soft labels 205
7.2.2 Pre-labeling the data set 211
7.2.3 k-nn variants 214
7.2.4 Fuzzy k-nn or non-fuzzy k-nn? 221
7.3 Generalized nearest prototype classifier (GNPC) 222
7.3.1 Fuzzy prototype classification, 1-nn, and nearest mean 222
7.3.2 GNPC definition 222
7.3.3 The five families of GNPCs 227

8. Combinations of multiple classifiers using fuzzy sets 233
8.1 Combining classifiers: the variety of paradigms 233
8.2 Classifier selection 237
8.3 Classifier fusion 240
8.3.1 Majority vote 241
8.3.2 "Naive"-Bayes combination 246
8.3.3 Behavior-knowledge space (BKS) 249
8.3.4 Simple fuzzy connectives for classifier fusion 250
8.3.5 Probabilistic product 252
8.3.6 Fuzzy integral 253
8.3.7 A brute-force approach to classifier fusion 255
8.3.8 Decision templates 257
8.3.9 Dempster-Shafer combination 260
8.4 Experimental results 260
8.4.1 Methods and experimental setup 260
8.4.2 Discussion 263

9. Conclusions: What to choose? 269

A. Appendix: Numerical results 271
A.1 Cone-torus data 271
A.2 Normal mixtures data 278
A.3 Phoneme data 285
A.4 Satimage data 288

References 291

Index 311