

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

Rafał Scherer

Multiple Fuzzy Classification Systems

 Springer

Author

Dr. Rafał Scherer
Department of Computer Engineering
Czestochowa University of Technology
Poland

ISSN 1434-9922

e-ISSN 1860-0808

ISBN 978-3-642-30603-7

e-ISBN 978-3-642-30604-4

DOI 10.1007/978-3-642-30604-4

Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012938369

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Exploratory data analysis is a vital set of methods used in various engineering, scientific and business applications. Fuzzy classifiers are important tools in this growing field. They use fuzzy rules and do not require assumptions common to statistical classification.

Rough set theory is useful when data sets are incomplete. It defines a formal approximation of crisp sets by providing the lower and the upper approximation of the original set. Systems based on rough sets have natural ability to work on such data and we do not have to preprocess incomplete vectors before classification.

To achieve better performance than existing machine learning systems, we combine them in ensembles. Such ensembles consists of a finite set of learning models, usually weak learners. In the book two popular methods are applied – boosting and negative correlation learning. Both of them help to achieve better accuracy and are not prone to over-fitting.

This book combines three aforementioned fields – fuzzy systems, rough sets and ensembling techniques. As the trained ensemble should represent a single hypothesis, a lot of attention is placed on possibility to combine fuzzy rules from fuzzy systems being members of classification ensemble. We build ensembles of various neuro-fuzzy systems with certain modifications to let the fuzzy rule bases to be activated at the same level. Furthermore, a lot of emphasis is placed on ensembles that can work on incomplete data, thanks to rough set theory. Presented ensembles give information how the number of missing features influences classification accuracy.

First of all I would express my sincere thanks to my friends Dr Marcin Korytkowski and Prof. Robert Nowicki who provided me help in research and to Prof. Leszek Rutkowski, who introduced me to scientific work and supported me in a friendly manner. I am also grateful to the Department of Computer Engineering at Częstochowa University of Technology for providing a scholarly environment for both teaching and research. Finally, I am truly grateful to my wife Magda for her love and patience and to my mother for raising me in the way that she did.

Acknowledgements

This book was partly supported by the Foundation for Polish Science – TEAM project 2010-2014 (TEAM/2009-4/9) and by the Polish Ministry of Science and Higher Education (Habilitation Project 2007-2010 Nr N N516 1155 33).

Contents

1	Introduction	1
1.1	A Brief Summary of the Content	2
1.2	Datasets Used in Experiments	3
1.2.1	Glass Identification	3
1.2.2	Ionosphere	3
1.2.3	Pima Indians Diabetes Problem	4
1.2.4	Wisconsin Breast Cancer Database	4
	References	4
2	Introduction to Fuzzy Systems	7
2.1	Elements of the Theory of Fuzzy Sets	7
2.2	Operations on Fuzzy Sets	12
2.3	Triangular Norms and Negations	14
2.4	Fuzzy Relations and Their Properties	16
2.5	Approximate Reasoning	18
2.5.1	Basic Rules of Inference in Binary Logic	18
2.5.2	Basic Rules of Inference in Fuzzy Logic	18
2.5.3	Inference Rules for the Mamdani Model	19
2.5.4	Inference Rules for the Logical Model	20
2.6	Fuzzy Inference Systems	22
2.6.1	Rule Base	22
2.6.2	Fuzzification Block	24
2.6.3	Inference Block	24
2.6.4	Defuzzification Block	25
2.7	Fuzzy Classification	26
2.8	Fuzzy c-Means	27
2.9	Summary and Discussion	28
	References	28

3	Ensemble Techniques	29
3.1	Bagging	29
3.2	Boosting	30
3.3	Negative Correlation Learning	32
3.4	Backpropagation Learning in Boosting Neuro-fuzzy Ensembles ..	33
3.5	Fuzzy c-Means Modification for Boosting Initialization	35
3.6	Summary and Discussion	37
	References	37
4	Relational Modular Fuzzy Systems	39
4.1	Basic Relational Fuzzy Model	39
4.2	Generalized Relational Fuzzy Model	42
4.3	Relational Boosting Ensemble	44
4.4	Experimental Results	46
4.5	Summary and Discussion	48
	References	49
5	Ensembles of the Mamdani Fuzzy Systems	51
5.1	Mamdani Neuro-fuzzy Systems	51
5.2	Interpretable Boosting Ensembles of the Mamdani Systems	53
5.3	Negative Correlation Learning Ensembles of the Mamdani Systems	56
5.4	Experimental Results	57
5.5	Summary and Discussion	58
	References	58
6	Logical Type Fuzzy Systems	61
6.1	Neuro-fuzzy Systems of Logical Type	61
6.2	Interpretable Boosting Ensembles of Logical Models	66
6.3	Negative Correlation Learning Ensembles of Logical Models	68
6.4	Experimental Results	68
6.5	Summary and Discussion	71
	References	71
7	Takagi-Sugeno Fuzzy Systems	73
7.1	Takagi-Sugeno Systems	73
7.2	Normalization of Rule Bases within Ensemble	74
7.3	Experimental Results	77
7.4	Summary and Discussion	79
	References	79
8	Rough-neuro-fuzzy Ensembles for Classification with Missing Data	81
8.1	Basic Definitions of the Rough Set Theory	81
8.2	Fuzzy Classifiers	83
8.2.1	T-CA Fuzzy Classifier	88
8.2.2	T-DCOG Fuzzy Classifier	88

8.2.3	S/D–DCOG Fuzzy Classifier	89
8.2.4	R–DCOG Fuzzy Classifier	89
8.2.5	QL–DCOG Fuzzy Classifier	90
8.3	Rough Fuzzy Classifiers	90
8.3.1	T–CA Rough Fuzzy Classifier	91
8.3.2	T–DCOG Rough Fuzzy Classifier	94
8.3.3	S/D–DCOG Rough Fuzzy Classifier	102
8.3.4	R–DCOG Rough Fuzzy Classifier	111
8.3.5	QL–DCOG Rough Fuzzy Classifier	120
8.3.6	MICOG Rough Fuzzy Classifier	120
8.4	Ensembles	121
8.5	Experimental Results	122
8.6	Summary and Discussion	125
	References	125
9	Concluding Remarks and Challenges for Future Research	129
	References	130
	Index	131