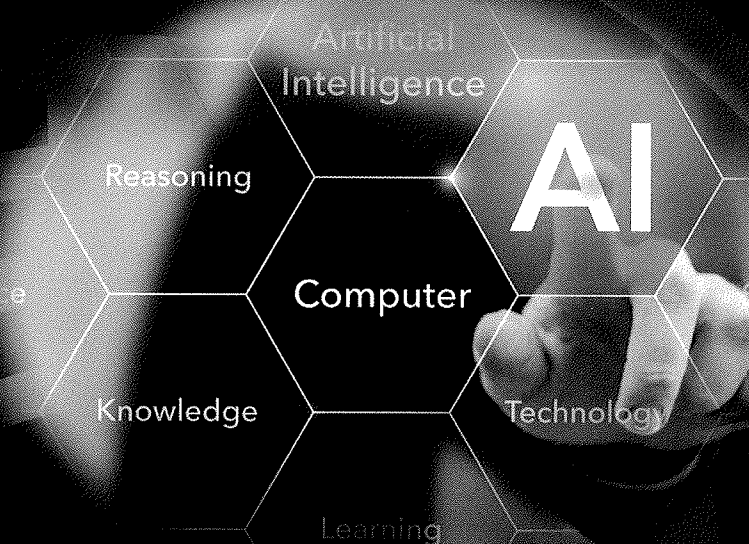


XIZHAO WANG ■ JUNHAI ZHAI

LEARNING WITH Uncertainty



LEARNING WITH
Uncertainty

LEARNING WITH
Uncertainty

XIZHAO WANG
JUNHAI ZHAI



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2017 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Printed at CPI on sustainably sourced paper
Version Date: 20161021

International Standard Book Number-13: 978-1-4987-2412-8 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Names: Wang, Xizhao, author. | Zhai, Junhai, author.
Title: Learning with uncertainty / Xizhao Wang, Junhai Zhai.
Description: Boca Raton : CRC Press, [2016] | Includes bibliographical references and index.
Identifiers: LCCN 2016030316 | ISBN 9781498724128 (acid-free paper) | ISBN 9781498724135 (e-book)
Subjects: LCSH: Machine learning. | Fuzzy decision making. | Decision trees.
Classification: LCC Q325.5 .W36 2016 | DDC 006.3/1--dc23
LC record available at <https://lccn.loc.gov/2016030316>

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Contents

Preface.....	ix
Symbols and Abbreviations	xi
1 Uncertainty.....	1
1.1 Randomness.....	1
1.1.1 Entropy.....	1
1.1.2 Joint Entropy and Conditional Entropy.....	3
1.1.3 Mutual Information.....	4
1.2 Fuzziness.....	5
1.2.1 Definition and Representation of Fuzzy Sets.....	6
1.2.2 Basic Operations and Properties of Fuzzy Sets.....	7
1.2.3 Fuzzy Measures.....	8
1.3 Roughness.....	9
1.4 Nonspecificity.....	10
1.5 Relationships among the Uncertainties.....	11
1.5.1 Entropy and Fuzziness.....	12
1.5.2 Fuzziness and Ambiguity.....	14
References.....	15
2 Decision Tree with Uncertainty	17
2.1 Crisp Decision Tree.....	17
2.1.1 ID3 Algorithm.....	18
2.1.2 Continuous-Valued Attributes Decision Trees.....	22
2.2 Fuzzy Decision Tree.....	28
2.3 Fuzzy Decision Tree Based on Fuzzy Rough Set Techniques.....	37
2.3.1 Fuzzy Rough Sets.....	37
2.3.2 Generating Fuzzy Decision Tree with Fuzzy Rough Set Technique.....	40
2.4 Improving Generalization of Fuzzy Decision Tree by Maximizing Fuzzy Entropy.....	44
2.4.1 Basic Idea of Refinement.....	44
2.4.2 Globally Weighted Fuzzy If-Then Rule Reasoning.....	44

2.4.3	Refinement Approach to Updating the Parameters.....	49
2.4.3.1	Maximum Fuzzy Entropy Principle.....	50
References.....		55
3	Clustering under Uncertainty Environment	59
3.1	Introduction.....	59
3.2	Clustering Algorithms Based on Hierarchy or Partition.....	60
3.2.1	Clustering Algorithms Based on Hierarchy.....	60
3.2.2	Clustering Algorithms Based on Partition.....	66
3.3	Validation Functions of Clustering.....	70
3.4	Feature Weighted Fuzzy Clustering.....	71
3.5	Weighted Fuzzy Clustering Based on Differential Evolution.....	73
3.5.1	Differential Evolution and Dynamic Differential Evolution....	73
3.5.1.1	Basic Differential Evolution Algorithm.....	73
3.5.1.2	Dynamic Differential Evolution Algorithm.....	77
3.5.2	Hybrid Differential Evolution Algorithm Based on Coevolution with Multi-Differential Evolution Strategy.....	78
3.6	Feature Weight Fuzzy Clustering Learning Model Based on MEHDE.....	80
3.6.1	MEHDE-Based Feature Weight Learning: MEHDE-FWL.....	81
3.6.2	Experimental Analysis.....	82
3.6.2.1	Comparison between MEHDE-FWL and GD-FWL Based on FCM.....	84
3.6.2.2	Comparisons Based on SMTTC Clustering.....	88
3.6.2.3	Efficiency Analysis of GD-, DE-, DDE-, and MEHDE-Based Searching Techniques.....	90
3.7	Summary.....	95
References.....		96
4	Active Learning with Uncertainty	99
4.1	Introduction to Active Learning.....	99
4.2	Uncertainty Sampling and Query-by-Committee Sampling.....	102
4.2.1	Uncertainty Sampling.....	102
4.2.1.1	Least Confident Rule.....	102
4.2.1.2	Minimal Margin Rule.....	103
4.2.1.3	Maximal Entropy Rule.....	103
4.2.2	Query-by-Committee Sampling.....	103
4.3	Maximum Ambiguity-Based Active Learning.....	105
4.3.1	Some Concepts of Fuzzy Decision Tree.....	106
4.3.2	Analysis on Samples with Maximal Ambiguity.....	107
4.3.3	Maximum Ambiguity-Based Sample Selection.....	109
4.3.4	Experimental Results.....	111

4.4	Active Learning Approach to Support Vector Machine.....	120
4.4.1	Support Vector Machine.....	122
4.4.2	SVM Active Learning.....	123
4.4.3	Semisupervised SVM Batch Mode Active Learning.....	124
4.4.4	IALPSVM: An Informative Active Learning Approach to SVM.....	125
4.4.5	Experimental Results and Discussions.....	128
4.4.5.1	Experiments on an Artificial Data Set by Selecting a Single Query Each Time.....	128
4.4.5.2	Experiments on Three UCI Data Sets by Selecting a Single Query Each Time.....	131
4.4.5.3	Experiments on Two Image Data Sets by Selecting a Batch of Queries Each Time.....	136
	References.....	146
5	Ensemble Learning with Uncertainty	149
5.1	Introduction to Ensemble Learning.....	149
5.1.1	Majority Voting and Weighted Majority Voting.....	150
5.1.2	Approach Based on Dempster–Shafer Theory of Evidence.....	151
5.1.3	Fuzzy Integral Ensemble Approach.....	152
5.2	Bagging and Boosting.....	153
5.2.1	Bagging Algorithm.....	153
5.2.2	Boosting Algorithm.....	154
5.3	Multiple Fuzzy Decision Tree Algorithm.....	154
5.3.1	Induction of Multiple Fuzzy Decision Tree.....	155
5.3.2	Experiment on Real Data Set.....	168
5.4	Fusion of Classifiers Based on Upper Integral.....	170
5.4.1	Extreme Learning Machine.....	170
5.4.2	Multiple Classifier Fusion Based on Upper Integrals.....	172
5.4.2.1	Upper Integral and Its Properties.....	173
5.4.2.2	A Model of Classifier Fusion Based on Upper Integral.....	175
5.4.2.3	Experimental Results.....	179
5.5	Relationship between Fuzziness and Generalization in Ensemble Learning.....	186
5.5.1	Classification Boundary.....	186
5.5.1.1	Boundary and Its Estimation Given by a Learned Classifier.....	186
5.5.1.2	Two Types of Methods for Training a Classifier.....	188
5.5.1.3	Side Effect of Boundary and Experimental Verification.....	189

5.5.2	Fuzziness of Classifiers.....	192
5.5.2.1	Fuzziness of Classifier.....	193
5.5.2.2	Relationship between Fuzziness and Misclassification.....	193
5.5.2.3	Relationship between Fuzziness and Classification Boundary.....	195
5.5.2.4	Divide and Conquer Strategy.....	198
5.5.2.5	Impact of the Weighting Exponent m on the Fuzziness of Fuzzy K -NN Classifier.....	198
5.5.3	Relationship between Generalization and Fuzziness.....	199
5.5.3.1	Definition of Generalization and Its Elaboration.....	199
5.5.3.2	Classifier Selection.....	201
5.5.3.3	Explanation Based on Extreme (max/min) Fuzziness.....	202
5.5.3.4	Experimental Results.....	205
	References.....	216
	Index	221