

Methods for Data Analysis and Mining

Christian Borgelt and Rudolf Kruse

Otto-von-Guericke- University of Magdeburg, Germany



JOHN WILEY & SONS, LTD

Contents

Preface	ix
1 Introduction	1
1.1 Data and Knowledge	2
1.2 Knowledge Discovery and Data Mining	5
1.2.1 The KDD Process	6
1.2.2 Data Mining Tasks	7
1.2.3 Data Mining Methods	8
1.3 Graphical Models	10
1.4 Outline of this Book	12
2 Imprecision and Uncertainty	15
2.1 Modeling Inferences	15
2.2 Imprecision and Relational Algebra	17
2.3 Uncertainty and Probability Theory	19
2.4 Possibility Theory and the Context Model	21
2.4.1 Experiments with Dice	22
2.4.2 The Context Model	27
2.4.3 The Insufficient Reason Principle	30
2.4.4 Overlapping Contexts	31
2.4.5 Mathematical Formalization	35
2.4.6 Normalization and Consistency	37
2.4.7 Possibility Measures	39
2.4.8 Mass Assignment Theory	43
2.4.9 Degrees of Possibility for Decision Making	45
2.4.10 Conditional Degrees of Possibility	47
2.4.11 Imprecision and Uncertainty	48
2.4.12 Open Problems	48
3 Decomposition	53
3.1 Decomposition and Reasoning	54
3.2 Relational Decomposition	55

3.2.1	A Simple Example.	55
3.2.2	Reasoning in the Simple Example.	57
3.2.3	Decomposability of Relations.	61
3.2.4	Tuple-Based Formalization.	63
3.2.5	Possibility-Based Formalization.	66
3.2.6	Conditional Possibility and Independence.	70
3.3	Probabilistic Decomposition	74
3.3.1	A Simple Example.	74
3.3.2	Reasoning in the Simple Example.	76
3.3.3	Factorization of Probability Distributions.	77
3.3.4	Conditional Probability and Independence.	78
3.4	Possibilistic Decomposition	82
3.4.1	Transfer from Relational Decomposition.	83
3.4.2	A Simple Example.	83
3.4.3	\sim Reasoning in the Simple Example.	84
3.4.4	Conditional Degrees of Possibility and Independence	85
3.5	Possibility versus Probability.	87
Graphical Representation		91
4.1	Conditional Independence Graphs.	92
4.1.1	Axioms of Conditional Independence.	92
4.1.2	Graph Terminology.	95
4.1.3	Separation in Graphs.	98
4.1.4	Dependence and Independence Maps.	100
4.1.5	Markov Properties of Graphs.	103
4.1.6	Graphs and Decompositions.	108
4.1.7	Markov Networks and Bayesian Networks.	114
4.2	Evidence Propagation in Graphs.	115
4.2.1	Propagation in Polytrees.	116
4.2.2	Join Tree Propagation.	125
4.2.3	Other Evidence Propagation Methods.	131
Computing Projections		133
5.1	Databases of Sample Cases.	134
5.2	Relational and Sum Projections.	135
5.3	Expectation Maximization.	137
5.4	Maximum Projections.	142
5.4.1	A Simple Example.	143
5.4.2	Computation via the Support.	145
5.4.3	Computation via the Closure.	146
5.4.4	Experimental Results.	149
5.4.5	Limitations.	150

6	Naive Classifiers	151
6.1	Naive Bayes Classifiers	151
6.1.1	The Basic Formula	151
6.1.2	Relation to Bayesian Networks	154
6.1.3	A Simple Example	155
6.2	A Naive Possibilistic Classifier	156
6.3	Classifier Simplification	158
6.4	Experimental Results	158
7	Learning Global Structure	161
7.1	Principles of Learning Global Structure	162
7.1.1	Learning Relational Networks	162
7.1.2	Learning Probabilistic Networks	171
7.1.3	Learning Possibilistic Networks	177
7.1.4	Components of a Learning Algorithm	186
7.2	Evaluation Measures	187
7.2.1	General Considerations	187
7.2.2	Notation and Presuppositions	191
7.2.3	Relational Evaluation Measures	193
7.2.4	Probabilistic Evaluation Measures	195
7.2.5	Possibilistic Evaluation Measures	221
7.3	Search Methods	223
7.3.1	Exhaustive Graph Search	223
7.3.2	Guided Random Graph Search	225
7.3.3	Conditional Independence Search	234
7.3.4	Greedy Search	240
7.4	Experimental Results	247
7.4.1	Learning Probabilistic Networks	247
7.4.2	Learning Possibilistic Networks	249
8	Learning Local Structure	253
8.1	Local Network Structure	253
8.2	Learning Local Structure	255
8.3	Experimental Results	259
9	Inductive Causation	261
9.1	Correlation and Causation	261
9.2	Causal and Probabilistic Structure	262
9.3	Stability and Latent Variables	264
9.4	The Inductive Causation Algorithm	266
9.5	Critique of the Underlying Assumptions	267
9.6	Evaluation	272

10 Applications	275
10.1 Application in Telecommunications	275
10.2 Application at Volkswagen	279
10.3 Application at DaimlerChrysler	282
A Proofs of Theorems	287
A.1 Proof of Theorem 4.1.2	287
A.2 Proof of Theorem 4.1.18	291
A.3 Proof of Theorem 4.1.20	292
A.4 Proof of Theorem 4.1.22	297
A.5 Proof of Theorem 4.1.24	302
A.6 Proof of Theorem 4.1.26	305
A.7 Proof of Theorem 4.1.27	307
A.8 Proof of Theorem 5.4.8	308
A.9 Proof of Theorem 7.3.2]..!	310
A.10 Proof of Lemma 7.2.2	311"
A.11 Proof of Lemma 7.2.4	313
A.12 Proof of Lemma 7.2.6	315
A.13 Proof of Theorem 7.3.4	316
A.14 Proof of Theorem 7.-3.5	317
A.15 Proof of Theorem 7.3.6	318
A.16 Proof of Theorem 7.3.8	321
B Software Tools	323
Bibliography	329
Index	349