

Table of Contents

Part I. Causal Models

1. Statistics, Causality, and Graphs

J. Pearl	3
1.1 A Century of Denial	3
1.2 Researchers in Search of a Language	5
1.3 Graphs as a Mathematical Language	8
1.4 The Challenge	13
References	14

2. Causal Conjecture

Glenn Shafer	17
2.1 Introduction	17
2.2 Variables in a Probability Tree	18
2.3 Causal Uncorrelatedness	19
2.4 Three Positive Causal Relations	20
2.5 Linear Sign	22
2.6 Causal Uncorrelatedness Again	26
2.7 Scored Sign	27
2.8 Tracking	28
References	32

3. Who Needs Counterfactuals?

A. P. Dawid	33
3.1 Introduction	33
3.1.1 Decision-Theoretic Framework	33
3.1.2 Unresponsiveness and Insensitivity	34
3.2 Counterfactuals	35
3.3 Problems of Causal Inference	36
3.3.1 Causes of Effects	36
3.3.2 Effects of Causes	36

3.4	The Counterfactual Approach	37
3.4.1	The Counterfactual Setting	37
3.4.2	Counterfactual Assumptions	38
3.5	Homogeneous Population	39
3.5.1	Experiment and Inference	40
3.6	Decision-Analytic Approach	43
3.7	Sheep and Goats	45
3.7.1	ACE	45
3.7.2	Neyman and Fisher	45
3.7.3	Bioequivalence	46
3.8	Causes of Effects	47
3.8.1	A Different Approach?	48
3.9	Conclusion	48
	References	49

4. Causality: Independence and Determinism

	Nancy Cartwright	51
4.1	Introduction	51
4.2	Conclusion	61
	References	63

Part II. Intelligent Data Management

5. Intelligent Data Analysis and Deep Understanding

	David J. Hand	67
5.1	Introduction	67
5.2	The Question: The Strategy	68
5.3	Diminishing Returns	74
5.4	Conclusion	78
	References	79

6. Learning Algorithms in High Dimensional Spaces

	A. Gammerman and V. Vovk	81
6.1	Introduction	81
6.2	SVM for Pattern Recognition	82
6.2.1	Dual Representation of Pattern Recognition	83
6.3	SVM for Regression Estimation	84
6.3.1	Dual Representation of Regression Estimation	84
6.3.2	SVM Applet and Software	85
6.4	Ridge Regression and Least Squares Methods in Dual Variables	86
6.5	Transduction	87
6.6	Conclusion	88
	References	88

7. Learning Linear Causal Models by MML Sampling	
Chris S. Wallace and Kevin B. Korb.....	89
7.1 Introduction	89
7.2 Minimum Message Length Principle.....	90
7.3 The Model Space	92
7.4 The Message Format	93
7.5 Equivalence Sets.....	95
7.5.1 Small Effects.....	96
7.5.2 Partial Order Equivalence	97
7.5.3 Structural Equivalence	97
7.5.4 Explanation Length.....	98
7.6 Finding Good Models	98
7.7 Sampling Control.....	102
7.8 By-products	102
7.9 Prior Constraints	102
7.10 Test Results	103
7.11 Remarks on Equivalence	106
7.11.1 Small Effect Equivalence	106
7.11.2 Equivalence and Causality	107
7.12 Conclusion.....	110
References	110
8. Game Theory Approach to Multicommodity Flow Network Vulnerability Analysis	
Y. E. Malashenko, N. M. Novikova and O. A. Vorobeichikova	112
References	118
9. On the Accuracy of Stochastic Complexity Approximations	
Petri Kontkanen, Petri Myllymäki, Tomi Silander, and Henry Tirri .	120
9.1 Introduction	120
9.2 Stochastic Complexity and Its Applications	122
9.3 Approximating the Stochastic Complexity in the Incomplete Data Case	124
9.4 Empirical Results.....	125
9.4.1 The Problem.....	125
9.4.2 The Experimental Setting	127
9.4.3 The Algorithms	129
9.4.4 Results.....	130
9.5 Conclusion.....	132
References	134

10. AI Modelling for Data Quality Control

Xiaohui Liu	137
10.1 Introduction	137
10.2 Statistical Approaches to Outliers	137
10.3 Outlier Detection and Analysis	139
10.4 Visual Field Test	139
10.5 Outlier Detection	141
10.5.1 Self-Organising Maps (SOM).....	141
10.5.2 Applications of SOM.....	142
10.6 Outlier Analysis by Modelling ‘Real Measurements’	143
10.7 Outlier Analysis by Modelling Noisy Data	145
10.7.1 Noise Model I: Noise Definition.....	145
10.7.2 Noise Model II: Construction	146
10.7.3 Noise Elimination	147
10.8 Concluding Remarks	147
References	148

11. New Directions in Text Categorization

Richard S. Forsyth	151
11.1 Introduction	151
11.2 Machine Learning for Text Classification	153
11.3 Radial Basis Functions and the Bard	156
11.4 An Evolutionary Algorithm for Text Classification	158
11.5 Text Classification by Vocabulary Richness	161
11.6 Text Classification with Frequent Function Words	163
11.7 Do Authors Have Semantic Signatures?	164
11.8 Syntax with Style	166
11.9 Intermezzo.....	167
11.10 Some Methods of Textual Feature-Finding	168
11.10.1 Progressive Pairwise Chunking	169
11.10.2 Monte Carlo Feature Finding	170
11.10.3 How Long Is a Piece of Substring?.....	173
11.10.4 Comparative Testing	175
11.11 Which Methods Work Best? – A Benchmarking Study	177
11.12 Discussion	180
11.12.1 In Praise of Semi-Crude Bayesianism	180
11.12.2 What’s So Special About Linguistic Data?.....	180
References	181