

Contents

1	Introduction	1
1.1	Unsupervised Exploratory Data Analysis	1
1.2	Projection Methods	3
1.3	Clustering	3
1.4	Structure of the Book	4
2	Review of Clustering Algorithms	7
2.1	Distance Measures	8
2.2	Hierarchical Clustering	9
2.2.1	Single-Linkage versus Complete-Linkage	11
2.3	Partitioning Clustering	14
2.3.1	K-Means Clustering Algorithm	14
2.3.2	K-Means++ Clustering Algorithm	15
2.3.3	K-Medoids (or Partition Around Medoids - PAM) Clustering Algorithm	16
2.3.4	Fuzzy C-Means Clustering Algorithm (FCM)	17
2.3.5	Soft K-Means Clustering Algorithm	18
2.3.6	K-Harmonic Means Clustering Algorithm (KHM)	18
2.3.7	Kernel K-Means Clustering Algorithm	19
2.3.8	Spectral Clustering Algorithm	21
2.4	Topology Preserving Mapping	22
2.4.1	Self-Organizing Map (SOM)	22
2.4.2	Generative Topographic Map (GTM)	24
2.4.3	Topographic Product of Experts (ToPoE)	26
2.4.4	Harmonic Topographic Mapping (HaToM)	27
2.5	Conclusion	28

3	Review of Linear Projection Methods.....	29
3.1	Linear Projection Methods	29
3.1.1	Principal Component Analysis	29
3.1.2	Exploratory Projection Pursuit	31
3.1.3	Independent Component Analysis	34
3.1.4	Canonical Correlation Analysis.....	38
3.1.5	Deflationary Orthogonalization Methods	39
3.2	Kernel Methods	40
3.2.1	Kernel Principal Component Analysis.....	40
3.2.2	Kernel Canonical Correlation Analysis	43
3.3	Latent Variable Models	44
3.3.1	Density Modeling and Latent Variables	45
3.3.2	Probabilistic Principal Component Analysis	47
3.4	Conclusions.....	48
4	Non-standard Clustering Criteria.....	49
4.1	A Family of New Algorithms.....	49
4.1.1	Weighted K-Means Algorithm (WK).....	50
4.1.2	Inverse Weighted K-Means Algorithm (IWK)	52
4.1.3	The Inverse Weighted Clustering Algorithm	56
4.1.4	Inverse Exponential K-Means Algorithm 1 (IEK1).....	64
4.1.5	Inverse Exponential K-Means Algorithm 2 (IEK2).....	67
4.1.6	Simulations.....	67
4.1.7	Summary	69
4.2	Spectral Clustering Algorithm.....	70
4.3	Conclusion	72
5	Topographic Mappings and Kernel Clustering	73
5.1	A Topology Preserving Mapping	73
5.1.1	Simulations.....	74
5.2	Kernel Clustering Algorithms	79
5.2.1	Kernel Inverse Weighted Clustering Algorithm (KIWC).....	80
5.2.2	Kernel K-Harmonic Means Algorithm (KKHM)	80
5.2.3	Kernel Inverse Weighted K-Means Algorithm (KIWK)	81
5.2.4	Simulations.....	81
5.3	Conclusion	84
6	Online Clustering Algorithms and Reinforcement Learning.....	85
6.1	Online Clustering Algorithms	85
6.1.1	Online K-Means Algorithm	85

6.1.2	IWK Online Algorithm v1 (IWK01)	86
6.1.3	IWK Online Algorithm v2 (IWK02)	87
6.1.4	K-Harmonic Means – Online Mode Algorithm (KHMO)	89
6.1.5	Inverse-Weighted K-Means (Online) Topology-Preserving Mapping (IKoToM)	91
6.2	Reinforcement Learning	92
6.2.1	Immediate Reward Reinforcement Learning	93
6.2.2	Global Reinforcement Learning in Neural Networks with Stochastic Synapses	95
6.2.3	Temporal Difference Learning	97
6.2.4	Evolutionary Algorithms for Reinforcement Learning	99
6.3	Clustering with Reinforcement Learning	101
6.3.1	New Algorithm RL1	102
6.3.2	New Algorithm RL2	103
6.3.3	New Algorithm RL3	103
6.3.4	Simulations	104
6.3.5	Topology Preserving Mapping	106
6.4	Conclusion	108
7	Connectivity Graphs and Clustering with Similarity Functions	109
7.1	Different Similarity Graphs (or Connectivity Graphs)	109
7.1.1	The ϵ -Neighborhood Graph	109
7.1.2	k-Nearest Neighbor Graphs	109
7.1.3	New Similarity Graph	110
7.2	Simulations	112
7.3	Clustering with Similarity Functions	113
7.3.1	Exponential Function as Similarity Function	115
7.3.2	Simulations	117
7.3.3	Inverse Weighted Clustering with Similarity Function Topology Preserving Mapping (IWCSFToM)	119
7.4	Conclusion	122
8	Reinforcement Learning of Projections	123
8.1	Projection with Immediate Reward Learning	123
8.1.1	An Example: Independent Component Analysis	124
8.1.2	Multiple Components with Immediate Reward Reinforcement Learning - PCA	130
8.1.3	Simulation: Canonical Correlation Analysis	133
8.1.4	Deflationary Orthogonalization for Kernel Methods – Kernel PCA	135
8.2	Projections with Stochastic Synapses	137

8.2.1	Linear Projection Methods with Stochastic Weights	138
8.2.2	Kernel Methods with Stochastic Weights	140
8.3	Projection with Temporal Difference Learning	142
8.3.1	Linear Projection with Q-Learning	143
8.3.2	Non-linear Projection with Sarsa Learning	144
8.4	Conclusion	148
9	Cross Entropy Methods	151
9.1	The Cross Entropy Method	151
9.1.1	Rare-Event Simulation via Cross Entropy	151
9.1.2	Combinatorial Optimization via Cross Entropy	154
9.2	ICA as Associated Stochastic Problem	156
9.3	Linear Projection with Cross Entropy Method	159
9.3.1	Principal Component Analysis	160
9.3.2	Exploratory Projection Pursuit	161
9.3.3	Canonical Correlation Analysis	162
9.4	Cross Entropy Latent Variable Models	164
9.4.1	Probabilistic Principal Component Analysis	164
9.4.2	Independent Component Analysis	166
9.4.3	Topology Preserving Manifolds	169
9.5	Deep Architectures in Unsupervised Data Exploration	171
9.5.1	Multilayer Topology Preserving Manifolds	171
9.6	Conclusion	173
10	Artificial Immune Systems	175
10.1	Artificial Immune Systems	175
10.1.1	Clonal Selection Algorithm	176
10.1.2	Artificial Immune Network	178
10.2	Projection with Immune-Inspired Algorithms	180
10.2.1	Linear Projections with the Modified CLONALG Algorithm	180
10.2.2	Multiple Components	185
10.3	Combining Adaptation Methods	188
10.3.1	Artificial Immune System with Cross Entropy	188
10.3.2	TD Learning with Artificial Immune Systems	190
10.4	Ensembles of the Non-standard Adaptation Methods	193
10.4.1	Bootstrapping and Bagging	194
10.4.2	Non-standard Adaptation Methods with Bagging	194
10.5	Conclusion	196

11 Conclusions	199
11.1 Rationale.....	199
11.2 Summary and Remarks	201
11.3 Further Research	204
References	207
Index	221