

# Contents

## Part I Machine Learning Fundamentals

<b>1 Introduction</b> . . . . .	3
References. . . . .	7
<b>2 Machine Learning</b> . . . . .	9
2.1 Introduction . . . . .	10
2.2 Machine Learning Categorization According to the Type of Inference . . . . .	11
2.2.1 Model Identification . . . . .	12
2.2.2 Shortcoming of the Model Identification Approach. . . . .	16
2.2.3 Model Prediction. . . . .	17
2.3 Machine Learning Categorization According to the Amount of Inference . . . . .	21
2.3.1 Rote Learning . . . . .	22
2.3.2 Learning from Instruction . . . . .	22
2.3.3 Learning by Analogy . . . . .	22
2.4 Learning from Examples . . . . .	23
2.4.1 The Problem of Minimizing the Risk Functional from Empirical Data . . . . .	24
2.4.2 Induction Principles for Minimizing the Risk Functional on Empirical Data . . . . .	26
2.4.3 Supervised Learning . . . . .	26
2.4.4 Unsupervised Learning . . . . .	29
2.4.5 Reinforcement Learning . . . . .	31
2.5 Theoretical Justifications of Statistical Learning Theory . . . . .	32
2.5.1 Generalization and Consistency. . . . .	34
2.5.2 Bias-Variance and Estimation-Approximation Trade-Off . . . . .	36
2.5.3 Consistency of Empirical Minimization Process . . . . .	38
2.5.4 Uniform Convergence . . . . .	40

2.5.5	Capacity Concepts and Generalization Bounds . . . . .	42
2.5.6	Generalization Bounds . . . . .	46
	References. . . . .	50
<b>3</b>	<b>The Class Imbalance Problem . . . . .</b>	<b>51</b>
3.1	Nature of the Class Imbalance Problem . . . . .	51
3.2	The Effect of Class Imbalance on Standard Classifiers . . . . .	55
3.2.1	Cost Insensitive Bayes Classifier. . . . .	55
3.2.2	Bayes Classifier Versus Majority Classifier. . . . .	59
3.2.3	Cost Sensitive Bayes Classifier . . . . .	66
3.2.4	Nearest Neighbor Classifier. . . . .	72
3.2.5	Decision Trees . . . . .	73
3.2.6	Neural Networks . . . . .	74
3.2.7	Support Vector Machines . . . . .	76
	References. . . . .	76
<b>4</b>	<b>Addressing the Class Imbalance Problem . . . . .</b>	<b>79</b>
4.1	Resampling Techniques . . . . .	79
4.1.1	Natural Resampling. . . . .	80
4.1.2	Random Over-Sampling and Random Under-Sampling. . . . .	80
4.1.3	Under-Sampling Methods . . . . .	80
4.1.4	Over-Sampling Methods . . . . .	83
4.1.5	Combination Methods . . . . .	87
4.2	Cost Sensitive Learning . . . . .	88
4.2.1	The MetaCost Algorithm. . . . .	89
4.3	One Class Learning . . . . .	91
4.3.1	One Class Classifiers. . . . .	91
4.3.2	Density Models . . . . .	95
4.3.3	Boundary Methods . . . . .	97
4.3.4	Reconstruction Methods . . . . .	100
4.3.5	Principal Components Analysis. . . . .	102
4.3.6	Auto-Encoders and Diabolo Networks . . . . .	103
	References. . . . .	105
<b>5</b>	<b>Machine Learning Paradigms . . . . .</b>	<b>107</b>
5.1	Support Vector Machines . . . . .	107
5.1.1	Hard Margin Support Vector Machines. . . . .	107
5.1.2	Soft Margin Support Vector Machines . . . . .	113
5.2	One-Class Support Vector Machines . . . . .	118
5.2.1	Spherical Data Description . . . . .	118
5.2.2	Flexible Descriptors . . . . .	123
5.2.3	$\nu$ - SVC . . . . .	124
	References. . . . .	128

**Part II Artificial Immune Systems**

<b>6 Immune System Fundamentals.</b>	133
6.1 Introduction	133
6.2 Brief History and Perspectives on Immunology	134
6.3 Fundamentals and Main Components	137
6.4 Adaptive Immune System.	139
6.5 Computational Aspects of Adaptive Immune System	140
6.5.1 Pattern Recognition	140
6.5.2 Immune Network Theory	141
6.5.3 The Clonal Selection Principle	144
6.5.4 Immune Learning and Memory	146
6.5.5 Immunological Memory as a Sparse Distributed Memory	149
6.5.6 Affinity Maturation	151
6.5.7 Self/Non-self Discrimination	154
References.	156
<b>7 Artificial Immune Systems</b>	159
7.1 Definitions	159
7.2 Scope of AIS	162
7.3 A Framework for Engineering AIS.	164
7.3.1 Shape-Spaces	166
7.3.2 Affinity Measures	168
7.3.3 Immune Algorithms	181
7.4 Theoretical Justification of the Machine Learning Ability of the Adaptive Immune System	183
7.5 AIS-Based Clustering	191
7.5.1 Background Immunological Concepts	191
7.5.2 The Artificial Immune Network (AIN) Learning Algorithm	192
7.5.3 AiNet Characterization and Complexity Analysis	196
7.6 AIS-Based Classification	199
7.6.1 Background Immunological Concepts	200
7.6.2 The Artificial Immune Recognition System (AIRS) Learning Algorithm.	202
7.6.3 Source Power of AIRS Learning Algorithm and Complexity Analysis	212
7.7 AIS-Based Negative Selection	214
7.7.1 Background Immunological Concepts	217
7.7.2 Theoretical Justification of the Negative Selection Algorithm	217
7.7.3 Real-Valued Negative Selection with Variable-Sized Detectors.	220

7.7.4	AIS-Based One-Class Classification . . . . .	223
7.7.5	V-Detector Algorithm . . . . .	225
	References. . . . .	229
<b>8</b>	<b>Experimental Evaluation of Artificial Immune System-Based Learning Algorithms. . . . .</b>	<b>237</b>
8.1	Experimentation . . . . .	237
8.1.1	The Test Data Set. . . . .	238
8.1.2	Artificial Immune System-Based Music Piece Clustering and Database Organization. . . . .	242
8.1.3	Artificial Immune System-Based Customer Data Clustering in an e-Shopping Application. . . . .	248
8.1.4	AIS-Based Music Genre Classification . . . . .	254
8.1.5	Music Recommendation Based on Artificial Immune Systems. . . . .	299
	References. . . . .	321
<b>9</b>	<b>Conclusions and Future Work . . . . .</b>	<b>325</b>

Machine Learning Paradigms

Artificial Immune Systems and their Applications in  
Software Personalization

Sotiropoulos, D.; Tsihrintzis, G.A.

2017, XVI, 327 p. 71 illus., 18 illus. in color., Hardcover

ISBN: 978-3-319-47192-1