

Contents

Introduction — 1

1 Introduction of machine learning — 3

Introduction — 3

1.1 Machine learning — 3

1.1.1 Basic concepts — 3

1.1.2 Definition and significance — 4

1.1.3 History of machine learning — 5

1.1.4 Machine learning system — 6

1.1.5 Basic elements of the machine learning system — 6

1.1.6 Category of machine learning — 7

1.1.6.1 Classification based on learning strategies — 7

1.1.6.2 Classification based on the representation of acquired knowledge — 9

1.1.6.3 Classification based on application area — 11

1.1.6.4 Comprehensive classification — 11

1.1.7 Current research field — 13

1.2 Statistical pattern recognition — 14

1.2.1 Problem representation — 15

1.2.2 Experience risk minimization — 16

1.2.3 Complexity and generalization — 17

1.3 Core theory of statistical learning — 19

1.3.1 Consistency condition of the learning process — 19

1.3.2 Generalization bounds — 19

1.3.3 Structural risk minimization — 22

Summary — 24

2 PAC Model — 25

Introduction — 25

2.1 Basic model — 25

2.1.1 Introduction of PAC — 25

2.1.2 Basic concepts — 26

2.1.3 Problematic — 26

2.2 Sample complexity in the PAC model — 27

2.2.1 Sample complexity in finite space — 27

2.2.2 Sample complexity in infinite space — 29

3 Decision tree learning — 33

Introduction — 33

3.1 Overview of a decision tree — 33

3.1.1	Decision tree —	34
3.1.2	Property —	36
3.1.3	Application —	36
3.1.4	Learning —	37
3.2	Design of decision tree —	37
3.2.1	Characteristics of decision trees —	37
3.2.2	Generation of decision trees —	38
	Summary —	45
4	Bayesian learning —	47
	Introduction —	47
4.1	Bayesian learning —	47
4.1.1	Bayesian formula —	47
4.1.2	Minimum error decision —	48
4.1.3	Normal probability density —	49
4.1.4	Maximum likelihood estimation —	50
4.2	Naive Bayesian principle and application —	51
4.2.1	Bayesian best hypothesis —	51
4.2.2	Naive Bayesian classification —	52
4.2.3	Text classification based on Naive Bayes —	53
4.3	Hidden Markov model and application —	56
4.3.1	Markov property —	56
4.3.2	Markov chain —	56
4.3.3	Transition probability matrix —	57
4.3.4	Hidden Markov model and application —	57
	Summary —	60
5	Support vector machines —	63
	Introduction —	63
5.1	Support vector machines —	63
5.2	Implementation algorithm —	69
5.3	SVM example —	71
5.4	Multi-class SVM —	73
	Summary —	74
6	AdaBoost —	75
	Introduction —	75
6.1	AdaBoost and object detection —	75
6.1.1	AdaBoost algorithm —	75
6.1.2	AdaBoost initialization —	77
6.2	Robust real-time object detection —	80
6.2.1	Rectangular feature selection —	80

6.2.2	Integral image — 81
6.2.3	Training result — 82
6.2.4	Cascade — 82
6.3	Object detection using statistical learning theory — 85
6.4	Random forest — 86
6.4.1	Principle description — 86
6.4.2	Algorithm details — 86
6.4.3	Algorithms analysis — 86
	Summary — 87
7	Compressed sensing — 89
	Introduction — 89
7.1	Theory framework — 89
7.2	Basic theory and core issues — 91
7.2.1	Mathematical model — 91
7.2.2	Signal sparse representation — 91
7.2.3	Signal observation matrix — 92
7.2.4	Signal reconstruction algorithm — 93
7.3	Application and simulation — 94
7.3.1	Application — 94
7.3.2	Face recognition — 95
	Summary — 97
8	Subspace learning — 99
	Introduction — 99
8.1	Feature extraction based on PCA — 99
8.2	Mathematical model — 102
8.3	Mathematical calculation of PCA — 103
8.3.1	Conclusions of linear algebra — 103
8.3.2	Eigenvalue decomposition based on the covariance matrix — 104
8.3.3	PCA — 104
8.4	Property of PCA — 105
8.5	Face recognition based on PCA — 107
	Summary — 107
9	Deep learning and neural networks — 109
	Introduction — 109
9.1	Neural network — 109
9.1.1	Forward neural network — 109
9.1.2	Perceptron network — 109
9.1.3	Three-layer forward neural network — 112
9.1.4	BP algorithm — 112

9.2	Deep learning —	116
9.2.1	Overview of deep learning —	116
9.2.2	Auto-Encoder algorithm —	117
9.2.3	Auto-Encoder deep network —	118
9.2.4	Convolution neural network —	119
9.3	Applications of deep learning —	124
9.3.1	Binarized convolutional networks for classification —	124
9.3.2	Time-series recognition —	124
	Summary —	125

10 Reinforcement learning — 127

	Introduction —	127
10.1	Overview of reinforcement learning —	127
10.2	Process of reinforcement learning —	128
10.2.1	Markov property —	128
10.2.2	Reward —	129
10.2.3	Value function —	129
10.2.4	Dynamic programming —	130
10.2.5	MC method —	130
10.2.6	Temporal difference learning —	131
10.2.7	Q-learning —	132
10.2.8	Improved Q-learning —	135
10.3	Code implementation —	137

Bibliography — 141

Index — 143