

Contents

1	Introduction	1
1.1	Terminology	1
1.1.1	What Is a Pattern?	1
1.1.2	Why Pattern Representation?	2
1.1.3	What Is Pattern Representation?	2
1.1.4	How to Represent Patterns?	2
1.1.5	Why Represent Patterns as Vectors?	2
1.1.6	Notation	3
1.2	Proximity Function	3
1.2.1	Distance Function	3
1.2.2	Similarity Function	4
1.2.3	Relation Between Dot Product and Cosine Similarity	5
1.3	Classification	6
1.3.1	Class	6
1.3.2	Representation of a Class	6
1.3.3	Choice of $G(X)$	7
1.4	Classifiers	7
1.4.1	Nearest Neighbor Classifier (NNC)	7
1.4.2	K-Nearest Neighbor Classifier (KNNC)	7
1.4.3	Minimum-Distance Classifier (MDC)	8
1.4.4	Minimum Mahalanobis Distance Classifier	9
1.4.5	Decision Tree Classifier: (DTC)	10
1.4.6	Classification Based on a Linear Discriminant Function	12
1.4.7	Nonlinear Discriminant Function	12
1.4.8	Naïve Bayes Classifier: (NBC)	13
1.5	Summary	14
	References	14

2	Linear Discriminant Function	15
2.1	Introduction	15
2.1.1	Associated Terms	15
2.2	Linear Classifier	17
2.3	Linear Discriminant Function	19
2.3.1	Decision Boundary	19
2.3.2	Negative Half Space	19
2.3.3	Positive Half Space	19
2.3.4	Linear Separability	20
2.3.5	Linear Classification Based on a Linear Discriminant Function	20
2.4	Example Linear Classifiers	23
2.4.1	Minimum-Distance Classifier (MDC)	23
2.4.2	Naïve Bayes Classifier (NBC)	23
2.4.3	Nonlinear Discriminant Function	24
	References	25
3	Perceptron	27
3.1	Introduction	27
3.2	Perceptron Learning Algorithm	28
3.2.1	Learning Boolean Functions	28
3.2.2	W Is Not Unique	30
3.2.3	Why Should the Learning Algorithm Work?	30
3.2.4	Convergence of the Algorithm	31
3.3	Perceptron Optimization	32
3.3.1	Incremental Rule	33
3.3.2	Nonlinearly Separable Case	33
3.4	Classification Based on Perceptrons	34
3.4.1	Order of the Perceptron	35
3.4.2	Permutation Invariance	37
3.4.3	Incremental Computation	37
3.5	Experimental Results	38
3.6	Summary	39
	References	40
4	Linear Support Vector Machines	41
4.1	Introduction	41
4.1.1	Similarity with Perceptron	41
4.1.2	Differences Between Perceptron and SVM	42
4.1.3	Important Properties of SVM	42
4.2	Linear SVM	43
4.2.1	Linear Separability	43
4.2.2	Margin	44
4.2.3	Maximum Margin	46
4.2.4	An Example	47

4.3	Dual Problem	49
4.3.1	An Example	50
4.4	Multiclass Problems	51
4.5	Experimental Results	52
4.5.1	Results on Multiclass Classification	52
4.6	Summary	54
	References	56
5	Kernel-Based SVM	57
5.1	Introduction	57
5.1.1	What Happens if the Data Is Not Linearly Separable?	57
5.1.2	Error in Classification	58
5.2	Soft Margin Formulation	59
5.2.1	The Solution.	59
5.2.2	Computing b	60
5.2.3	Difference Between the Soft and Hard Margin Formulations	60
5.3	Similarity Between SVM and Perceptron	60
5.4	Nonlinear Decision Boundary	62
5.4.1	Why Transformed Space?	63
5.4.2	Kernel Trick.	63
5.4.3	An Example	64
5.4.4	Example Kernel Functions	64
5.5	Success of SVM	64
5.6	Experimental Results	65
5.6.1	Iris Versicolour and Iris Virginica	65
5.6.2	Handwritten Digit Classification	66
5.6.3	Multiclass Classification with Varying Values of the Parameter C	66
5.7	Summary	67
	References	67
6	Application to Social Networks.	69
6.1	Introduction	69
6.1.1	What Is a Network?	69
6.1.2	How Do We Represent It?	69
6.2	What Is a Social Network?	72
6.2.1	Citation Networks	73
6.2.2	Coauthor Networks	73
6.2.3	Customer Networks	73
6.2.4	Homogeneous and Heterogeneous Networks.	73
6.3	Important Properties of Social Networks.	74
6.4	Characterization of Communities	75
6.4.1	What Is a Community?	75
6.4.2	Clustering Coefficient of a Subgraph	76

6.5	Link Prediction	77
6.5.1	Similarity Between a Pair of Nodes.	78
6.6	Similarity Functions.	79
6.6.1	Example	80
6.6.2	Global Similarity.	81
6.6.3	Link Prediction based on Supervised Learning	82
6.7	Summary	83
	References	83
7	Conclusion	85
	Glossary	89
	Index	91

Support Vector Machines and Perceptrons
Learning, Optimization, Classification, and Application to
Social Networks

Murty, M.N.; Raghava, R.

2016, XIII, 95 p. 25 illus., Softcover

ISBN: 978-3-319-41062-3