

Floris Ernst | Achim Schweikard

Fundamentals of Machine Learning

Support Vector Machines Made Easy

UVK Verlag · München

Contents

Preface.....	7
--------------	---

Part I | Support Vector Machines

1	Symbolic Classification and Nearest Neighbour Classification.....	11
1.1	Symbolic Classification.....	11
1.2	Nearest Neighbour Classification	11
2	Separating Planes and Linear Programming	15
2.1	Finding a Separating Hyperplane	16
2.2	Testing for feasibility of linear constraints	16
2.3	Linear Programming	19
	MATLAB example.....	22
2.4	Conclusion.....	24
3	Separating Margins and Quadratic Programming.....	25
3.1	Quadratic Programming.....	25
3.2	Maximum Margin Separator Planes.....	30
3.3	Slack Variables.....	32
4	Dualization and Support Vectors.....	37
4.1	Duals of Linear Programs	37
4.2	Duals of Quadratic Programs.....	38
4.3	Support Vectors.....	40
5	Lagrange Multipliers and Duality.....	43
5.1	Multidimensional functions	46
5.2	Support Vector Expansion.....	47
5.3	Support Vector Expansion with Slack Variables.....	48
6	Kernel Functions	53
6.1	Feature Spaces	53
6.2	Feature Spaces and Quadratic Programming.....	54
6.3	Kernel Matrix and Mercer's Theorem.....	58
6.4	Proof of Mercer's Theorem.....	59
	Step 1 – Definitions and Prerequisites	59
	Step 2 – Designing the right Hilbert Space.....	61
	Step 3 – The reproducing property.....	63

7	The SMO Algorithm	67
7.1	Overview and Principles.....	67
7.2	Optimisation Step.....	68
7.3	Simplified SMO	69
8	Regression	79
8.1	Slack Variables.....	80
8.2	Duality, Kernels and Regression.....	81
8.3	Deriving the Dual form of the QP for Regression.....	82
Part II Beyond Support Vectors		
9	Perceptrons, Neural Networks and Genetic Algorithms.....	89
9.1	Perceptrons.....	89
	Perceptron-Algorithm.....	91
	Perceptron-Lemma and Convergence.....	92
	Perceptrons and Linear Feasibility Testing.....	93
9.2	Neural Networks	94
	Forward Propagation	95
	Training and Error Backpropagation.....	96
9.3	Genetic Algorithms.....	98
9.4	Conclusion	99
10	Bayesian Regression	101
10.1	Bayesian Learning.....	101
10.2	Probabilistic Linear Regression.....	103
10.3	Gaussian Process Models.....	108
10.4	GP model with measurement noise	111
	Optimization of hyperparameters	112
	Covariance functions.....	113
10.5	Multi-Task Gaussian Process (MTGP) Models.....	114
11	Bayesian Networks.....	121
	Propagation of probabilities in causal networks	131
Appendix - Linear Programming.....		
A.1	Solving LP0 problems	140
A.2	Schematic representation of the iteration steps	146
A.3	Transition from LP0 to LP	148
A.4	Computing time and complexity issues	150
References.....		
		151
Index		
		153