

Contents

Preface	xi
1 Artificial Neural Networks	1
1.1 Introduction	1
1.2 Feed-forward networks generally	1
1.3 Neurons	2
1.3.1 Linear threshold units	3
1.3.2 Sigmoid units	4
1.3.3 Polynomial threshold units	4
1.4 Networks	5
1.5 Additional remarks and bibliographical notes	7
2 Boolean Functions	9
2.1 Introduction	9
2.2 Boolean functions and the hypercube	9
2.2.1 The Boolean hypercube	9
2.2.2 Geometry of the Boolean hypercube	9
2.2.3 The Boolean hypercube as a graph	10
2.2.4 The Boolean hypercube as a partially ordered set	10
2.3 Boolean formulas	10
2.3.1 Representing Boolean functions by formulas	10
2.3.2 DNF and CNF representations	11
2.3.3 Dualization	13
2.3.4 Classes of Boolean functions	14
2.4 Monotonicity and regularity	15
2.4.1 Increasing functions	15
2.4.2 Unate functions	15
2.4.3 Regular and 2-monotonic functions	15
2.4.4 MFPs and MTPs of increasing functions	16
2.5 Decision lists	17
2.5.1 Definition of decision lists	17
2.5.2 Decision lists and other Boolean function classes	18
2.6 Additional remarks and bibliographical notes	18

3	Threshold Functions	21
3.1	Introduction	21
3.2	Threshold functions	21
3.2.1	Boolean threshold functions	21
3.2.2	Real threshold functions	22
3.3	Properties of threshold functions	22
3.3.1	Closure under projection	22
3.3.2	2-monotonicity	23
3.4	Geometrical interpretation of threshold functions	25
3.4.1	Linear separability	25
3.4.2	Convexity and threshold functions	25
3.5	Threshold functions and asummability	26
3.6	1-Decision lists are threshold functions	28
3.7	Polynomial threshold functions	29
3.7.1	Real polynomial threshold functions	29
3.7.2	Boolean polynomial threshold functions	30
3.7.3	Closure under projections	30
3.8	Polynomial representation of Boolean functions	31
3.9	Geometrical interpretation of polynomial threshold functions	32
3.10	Asummability and polynomial threshold functions	32
3.11	Additional remarks and bibliographical notes	33
4	Number of Threshold Functions	35
4.1	Introduction	35
4.2	Number of threshold functions: Upper bounds	35
4.3	Number of threshold functions: Lower bound	38
4.4	Asymptotic number of threshold functions	39
4.5	Number of polynomial threshold functions: Upper bounds	40
4.6	Number of polynomial threshold functions: Lower bounds	40
4.7	Additional remarks and bibliographical notes	42
5	Sizes of Weights for Threshold Functions	43
5.1	Introduction	43
5.2	A general upper bound on size of weights	43
5.3	Exponential lower bound	44
5.3.1	Existence of large-weight threshold functions	44
5.3.2	A specific example	45
5.3.3	Tightness of general upper bound	46
5.4	Additional remarks and bibliographical notes	47
6	Threshold Order	49
6.1	Introduction	49
6.2	Algorithms and complexity	49
6.2.1	Introduction	49
6.2.2	Efficient algorithms and "hardness" of problems	49
6.3	Complexity of determining threshold order	50
6.3.1	The problem MEMBERSHIP(C)	50
6.3.2	Complexity of determining threshold order	52
6.4	Threshold synthesis	52

6.4.1	Introduction	52
6.4.2	Overview of the method	53
6.4.3	Recognition algorithm for 2-monotonic functions	53
6.4.4	Dualization algorithm for regular functions	54
6.4.5	Linear programming to determine separability	55
6.4.6	Time complexity of synthesis procedure	56
6.5	Expected threshold order	57
6.6	Additional remarks and bibliographical notes	58
7	Threshold Networks and Boolean Functions	61
7.1	Introduction	61
7.2	DNFs and threshold networks	61
7.3	A geometric approach: "Chopping"	62
7.3.1	The chopping procedure	62
7.3.2	"Chopping" and threshold decision lists	62
7.3.3	Threshold decision lists and threshold networks	63
7.4	Universal networks	64
7.4.1	Universal networks for Boolean functions	64
7.4.2	Sizes of universal networks	64
7.5	Computing the parity function	66
7.5.1	A lower bound on the size of a net computing parity	66
7.5.2	Partially ordered sets and Sperner's theorem	66
7.5.3	Proof of Theorem 7.6	67
7.6	Additional remarks and bibliographical notes	69
8	Specifying Sets	71
8.1	Introduction	71
8.2	Upper bounds on specification numbers of threshold functions	71
8.2.1	Introduction	71
8.2.2	The boundary approach	72
8.2.3	Using the partial order	73
8.3	A lower bound and its attainment	75
8.4	Signatures	77
8.5	Projections	77
8.6	The expected specification number	79
8.6.1	A general bound	79
8.6.2	Expected specification number of T_n	81
8.7	Additional remarks and bibliographical notes	81
9	Neural Network Learning	83
9.1	Introduction	83
9.2	Supervised learning	83
9.2.1	A learning framework	83
9.2.2	Consistent algorithms	84
9.3	Consistent algorithms for the perceptron	84
9.3.1	Linear programming method	84
9.3.2	The incremental perceptron learning algorithm	85
9.4	Efficiency of Boolean perceptron learning	87
9.4.1	The linear programming approach	87

9.4.2	The incremental approach	87
9.5	Additional remarks and bibliographical notes	88
10	Probabilistic Learning	89
10.1	Introduction	89
10.2	The PAC learning model	89
10.2.1	The learning framework	89
10.2.2	Probability and approximation	89
10.2.3	The finite case	90
10.3	The growth function and VC-dimension	91
10.3.1	The growth function of a hypothesis space	91
10.3.2	VC-dimension	92
10.4	Relating growth function and VC-dimension	93
10.5	VC-dimension and PAC learning	94
10.5.1	Upper bound on sample length	94
10.5.2	Proof of Theorem 10.7	95
10.5.3	Lower bound on sample length	98
10.5.4	VC-dimension characterizes PAC learnability	99
10.6	Additional remarks and bibliographical notes	99
11	VC-dimensions of Neural Networks	101
11.1	Introduction	101
11.2	VC-dimension and linear dimension	101
11.3	VC-dimension of the perceptron	103
11.4	VC-dimension of class of polynomial threshold functions	104
11.5	VC-dimension of threshold networks	106
11.6	VC-dimension of sigmoid networks	107
11.7	Additional remarks and bibliographical notes	107
12	The Complexity of Learning	109
12.1	Introduction	109
12.2	Efficient PAC learning algorithms	109
12.2.1	Definition of efficient PAC algorithm	109
12.2.2	Sufficient conditions for efficient algorithms	110
12.3	PAC algorithms and randomized consistent algorithms	110
12.4	Learning can be hard	112
12.5	Additional remarks and bibliographical notes	113
13	Boltzmann Machines and Combinatorial Optimization	115
13.1	Introduction	115
13.2	Boltzmann machines	115
13.3	Combinatorial optimization	117
13.4	Maximum cuts	117
13.5	The traveling salesman problem	118
13.6	Additional remarks and bibliographical notes	118
	Bibliography	119
	Index	127