

Inhaltsverzeichnis

Einführung

1 Das biologische Paradigma	3
1.1 Neuronale Netze als Berechnungsmodell	3
1.1.1 Natürliche und künstliche neuronale Netze	3
1.1.2 Entstehung der Berechenbarkeitsmodelle.....	5
1.1.3 Elemente eines Berechnungsmodells.....	9
1.2 Biologische neuronale Netze	10
1.2.1 Aufbau der Nervenzellen	10
1.2.2 Informationsübertragung in Neuronennetzen	12
1.2.3 Informationsverarbeitung an Membran und Synapse	18
1.2.4 Speicherung von Information – Lernen	20
1.3 Künstliche neuronale Netze als Funktionennetze	22
1.3.1 Atomare Bestandteile der Informationsverarbeitung	22
1.3.2 Neuronale versus Funktionennetze	24
1.4 Historische Anmerkungen	25

Teil I: Vorwärtsgerichtete Netze

2 Das Modell von McCulloch und Pitts	29
2.1 Netze von Funktionen	29
2.1.1 Vorwärtsgerichtete und rekursive Netze.....	29
2.1.2 Ein abstraktes Neuronenmodell	31
2.2 Aufbau von logischen Funktionen	34
2.2.1 Konjunktion und Disjunktion	34
2.2.2 Nichtmonotone Funktionen und die Negation.....	35
2.2.3 Geometrische Interpretation.....	36
2.3 Netze für beliebige logische Funktionen	38
2.3.1 Konstruktive Methode	38
2.3.2 Verknüpfungsbasen.....	39
2.4 Äquivalente Netze	40
2.4.1 Gewichtete und ungewichtete Netze.....	40
2.4.2 Absolute und relative Hemmung	42
2.4.3 Binäre Information.....	43
2.5 Rekursive Netze	44
2.5.1 Netze mit „Erinnerung“	44
2.5.2 Endliche Automaten.....	45
2.5.3 Äquivalenz endlicher Automaten und neuronaler Netze	47

2.5.4	Erste Klassifizierung von neuronalen Netzen.....	48
2.6	Historische Anmerkungen	50
3	Gewichtete Netze - Das Perzeptron.....	51
3.1	Perzeptronen und parallele Datenverarbeitung.....	51
3.1.1	Das Modell von Rosenblatt.....	51
3.1.2	Das klassische und das Minsky-Papert-Perzeptron	52
3.1.3	Parallele Algorithmen und unlösbare Probleme	53
3.2	Realisierung logischer Funktionen mit Perzeptronen	57
3.2.1	Einfaches Perzeptron und geometrische Interpretation	57
3.2.2	Gewichtete Netze mit einem kanonischen Baustein.....	58
3.2.3	Logische Funktionen und das XOR-Problem.....	59
3.3	Lineare Entscheidungsfunktionen.....	60
3.3.1	Lineare Trennbarkeit.....	60
3.3.2	Eingabe- und Gewichteraum - Dualität	61
3.3.3	Fehlerfunktion im Gewichteraum	62
3.3.4	Allgemeine Entscheidungskurven	64
3.4	Anwendungen und biologische Analogie	66
3.4.1	Kantenerkennung mit Perzeptronen.....	66
3.4.2	Die Verschaltung der Netz haut	68
3.4.3	Die „Silicon-Retina“ von Carver Mead	70
3.5	Historische Anmerkungen	71
4	Der Perzeptron-Lernalgorithmus	73
4.1	Lernalgorithmen für neuronale Netze	73
4.1.1	Lernen in parametrischen Systemen	73
4.1.2	Klassen von Lernalgorithmen	74
4.2	Lineare Trennbarkeit.....	76
4.2.1	Vektornotation	76
4.2.2	Absolute lineare Trennbarkeit.....	77
4.2.3	Fehlerflächen und Suchvorgang	79
4.3	Perzeptron-Lernen.....	83
4.3.1	Definition des Lernalgorithmus	83
4.3.2	Geometrische Visualisierung	84
4.3.3	Konvergenz des Algorithmus	86
4.3.4	Beschleunigung der Konvergenz – Delta-Regel.....	88
4.3.5	Komplexität des Perzeptron-Lernalgorithmus	90
4.4	Perzeptron-Lernen als lineares Programm.....	91
4.4.1	Lineare Optimierung und innere Punkte.....	91
4.4.2	Lineare Trennbarkeit als lineares Programm.....	93
4.4.3	Der Algorithmus von Karmarkar	94
4.5	Historische Anmerkungen	96

5	Unüberwachtes Lernen	97
5.1	Lernen durch Konkurrenz	97
5.1.1	Klassen von unüberwachtem Lernen	97
5.1.2	Verallgemeinerung des Perzeptron-Problems.....	97
5.1.3	Unüberwachtes Lernen durch Konkurrenz	99
5.2	Konvergenzanalyse	101
5.2.1	Der eindimensionale Fall – Energiefunktion	101
5.2.2	Mehrdimensionaler Fall	104
5.2.3	Unüberwachtes Lernen als Minimierungsaufgabe.....	105
5.2.4	Beziehung zum Perzeptron-Lernen.....	109
5.2.5	Stabilität der Lösungen	110
5.3	Hauptkomponentenanalyse	113
5.3.1	Unüberwachtes Lernen mit Verstärkung	113
5.3.2	Konvergenz des Lernalgorithmus	115
5.3.3	Bestimmung zusätzlicher Hauptkomponenten	116
5.4	Beispiele.....	117
5.4.1	Mustererkennung	117
5.4.2	Selbstorganisation im menschlichen Gehirn.....	118
5.5	Historische Anmerkungen	119
6	Netze mit mehreren Schichten.....	121
6.1	Struktur und geometrische Visualisierung.....	121
6.1.1	Netzarchitekturen.....	121
6.1.2	Das XOR-Problem	123
6.1.3	Geometrische Visualisierung	126
6.2	Regionen im Eingabe- und Gewichteraum	128
6.2.1	Gewichteraumregionen für das XOR-Problem.....	128
6.2.2	Bipolarvektoren.....	130
6.2.3	Projektion der Lösungsregionen auf die Ebene	131
6.2.4	Geometrische Interpretation.....	135
6.3	Regionen für komplexere Netze	137
6.3.1	Regionen im Gewichteraum des XOR-Problems	137
6.3.2	Anzahl der Regionen im Allgemeinen.....	139
6.3.3	Konsequenzen	141
6.3.4	Die Vapnik-Chervonenkis-Dimension.....	142
6.3.5	Die Frage der lokalen Minima	145
6.4	Historische Anmerkungen	147
7	Der Backpropagation-Algorithmus	149
7.1	Lernen in Backpropagation-Netzen	149
7.1.1	Gradientenabstiegsverfahren.....	149
7.1.2	Differenzierbare Aktivierungsfunktionen.....	149

7.1.3	Regionen im Eingaberaum.....	151
7.1.4	Entstehung von lokalen Minima	152
7.2	Backpropagation in Funktionennetzen	154
7.2.1	Das Lernproblem	154
7.2.2	Berechnungen in einem Backpropagation-Netz	154
7.2.3	Die Kettenregel	156
7.2.4	Parallelschaltungen	157
7.2.5	Gewichtete Kanten.....	159
7.2.6	Die Fehlerfunktion	160
7.3	Lernen mit Backpropagation.....	161
7.3.1	Schritte des Algorithmus.....	161
7.3.2	Backpropagation in Matrixform	165
7.3.3	Backpropagation als lokale Informationsverarbeitung	166
7.3.4	Die Form der Fehlerfunktion für das XOR-Problem	168
7.3.5	Varianten des Backpropagation-Verfahrens	168
7.4	Historische Anmerkungen	172

Teil II: Theoretische Analyse

8	Backpropagation und statistische Regression.....	175
8.1	Statistische Funktionsanpassung.....	175
8.1.1	Annäherung versus Verallgemeinerung.....	175
8.1.2	Der lineare Assoziator.....	177
8.1.3	Lineare Regression.....	179
8.2	Nichtlineare Regression	181
8.2.1	Form der Ausgabefunktion	181
8.2.2	Die Logit-Transformation.....	182
8.2.3	Logistische Regression	184
8.3	Backpropagation in komplexen Netzen	186
8.3.1	Regression in einem Backpropagation-Netzwerk.....	186
8.3.2	Visualisierung der Lösungs-Regionen	188
8.3.3	Form der Fehlerfunktion	189
8.4	Regression in mehrschichtigen Netzen	192
8.4.1	Die Rolle der verborgenen Schicht	192
8.4.2	Tragweite der Nichtlinearität	193
8.4.3	Matrixinversion durch Backpropagation	194
8.5	Anwendungen	196
8.5.1	Die Lernmatrix	197
8.5.2	Datencodierung und Datenkompression	197
8.5.3	NETtalk.....	198
8.5.4	Prognose von Zeitreihen	200
8.6	Historische Anmerkungen	201

9 Die Komplexität des Lernens.....	203
9.1 Funktionen als Netze.....	203
9.1.1 Lernalgorithmen für mehrschichtige Netze	203
9.1.2 Hilberts Dreizehntes Problem.....	204
9.1.3 Der Satz von Kolmogorov	205
9.2 Funktionsannäherung.....	207
9.2.1 Der eindimensionale Fall.....	207
9.2.2 Der mehrdimensionale Fall.....	210
9.3 Komplexität von Lernproblemen.....	212
9.3.1 Das Lernproblem für neuronale Netze.....	212
9.3.2 Komplexitätsklassen	213
9.3.3 NP-vollständige Lernprobleme.....	215
9.3.4 Komplexität des Lernens bei AND-OR-Netzen	217
9.3.5 Vereinfachungen der Netzarchitektur – der Kortex.....	220
9.3.6 Lernen mit Hinweisen.....	222
9.4 Historische Anmerkungen	224
10 Fuzzy-Logik und neuronale Netze	225
10.1 Fuzzy-Mengen und Fuzzy-Logik.....	225
10.1.1 Unscharfe Daten und Regeln	225
10.1.2 Der Begriff der unscharfen Mengen	226
10.1.3 Geometrische Darstellung von unscharfen Mengen	229
10.1.4 Mengenlehre, logische Operatoren und Geometrie	233
10.1.5 Familien von Fuzzy-Operatoren	234
10.2 Berechnung von Fuzzy-Inferenzen	238
10.2.1 Unscharfes Schließen.....	238
10.2.2 Unscharfe Zahlen und inverse Operation	239
10.3 Kontrolle mit Fuzzy-Systemen	242
10.3.1 Fuzzy-Regler.....	242
10.3.2 Fuzzy-Netze	244
10.3.3 Funktionsapproximation mit Fuzzy-Methoden.....	245
10.3.4 Das Auge als <i>fuzzifier</i> — Farbensehen.....	246
10.4 Historische Anmerkungen	248

Teil III: Rekursive Netze

11 Assoziativspeicher.....	251
11.1 Grundlagen der assoziativen Speicherung	251
11.1.1 Rekursive Netze	251
11.1.2 Klassen von Assoziativspeichern.....	252
11.1.3 Struktur des Assoziativspeichers	253
11.1.4 Die Eigenvektorenmaschine	255
11.1.5 Bipolarvektoren und die Vorzeichen-Funktion	257

11.2	Lernen in Assoziativspeichern	258
11.2.1	<i>Hebbian-Learning</i> - Die Korrelationsmatrix	258
11.2.2	Geometrische Deutung der Hebb-Regel	262
11.2.3	Netze als dynamische Systeme - Resultate	263
11.3	Die Pseudoinverse	268
11.3.1	Orthogonale Projektionen	268
11.3.2	Eigenschaften der Pseudoinversen - Berechnung	272
11.3.3	Holographische Speicher	275
11.4	Historische Anmerkungen	276
12	Das Hopfield-Modell	277
12.1	Synchrone und asynchrone Netze	277
12.1.1	Rekursive Netze mit stochastischer Aktivierung	277
12.1.2	Der bidirektionale Assoziativspeicher	278
12.1.3	Definition der Energiefunktion	280
12.2	Definition der Hopfield-Netze	282
12.2.1	Asynchrone Netze	282
12.2.2	Der Ansatz von Hopfield	284
12.2.3	Isomorphie zwischen Hopfield- und Ising-Modell	290
12.3	Konvergenz des Modells	291
12.3.1	Dynamik des Hopfield-Netzes	291
12.3.2	Konvergenzanalyse	294
12.3.3	Die Hebb-Regel	296
12.4	Äquivalenz des Lernverfahrens mit Perzeptron-Lernen	298
12.4.1	Perzeptron-Lernen bei Hopfield-Netzen	298
12.4.2	Komplexität des Lernens bei Hopfield-Modellen	300
12.5	Historische Anmerkungen	301
13	Kombinatorische Optimierung und Parallelität	303
13.1	Parallele Algorithmen	303
13.1.1	<i>NP</i> -vollständige Probleme	303
13.1.2	Das Multi-Flop-Problem	304
13.1.3	Das Acht-Türme-Problem	305
13.1.4	Das Acht-Damen-Problem	307
13.1.5	Das Problem des Handlungsreisenden	308
13.2	Theoretische Betrachtungen	310
13.2.1	Die Klasse co- <i>NP</i>	310
13.2.2	Die Grenzen von Hopfield-Netzen	312
13.3	Implementierung des Hopfield-Netzes	313
13.3.1	Elektrische Realisierung	313
13.3.2	Optische Realisierung	314
13.4	Historische Anmerkungen	315

14 Stochastische Netze	317
14.1 Varianten des Hopfield-Modells.....	317
14.1.1 Beschränkungen des Standardmodells.....	317
14.1.2 Das Hopfield-Modell mit stetiger Aktivierungsfunktion.....	318
14.2 Stochastische Algorithmen und Systeme.....	320
14.2.1 Simulated Annealing.....	320
14.2.2 Stochastische neuronale Netze.....	321
14.2.3 Markov-Ketten.....	323
14.2.4 Die Boltzmann-Verteilung.....	326
14.2.5 Physikalische Bedeutung der Boltzmann-Verteilung	329
14.3 Lernverfahren und Anwendungen	331
14.3.1 Lernen in Boltzmann-Maschinen.....	331
14.3.2 Anwendungen in der kombinatorischen Optimierung.....	333
14.4 Historische Anmerkungen	334

Teil IV: Selbstorganisation und Neurohardware

15 Kohonens topologieerhaltende Abbildungen	339
15.1 Selbstorganisation.....	339
15.1.1 Kartierung des Eingaberaums	339
15.1.2 Sensorische Karten im Gehirn	340
15.2 Kohonens Modell.....	343
15.2.1 Der Lernalgorithmus	343
15.2.2 Projektion auf niedrigere Dimensionen	346
15.3 Konvergenzanalyse	348
15.3.1 Potentialfunktion - Eindimensionaler Fall	348
15.3.2 Zweidimensionaler Fall	351
15.3.3 Auswirkung der Nachbarschaft eines Neurons	353
15.4 Anwendungen	355
15.4.1 Kartierung von Funktionen	355
15.4.2 Kartierung von Räumen	356
15.4.3 Automatische Anpassung an Hindernisse	358
15.5 Historische Anmerkungen	359
16 Hybride Modelle	361
16.1 Netzketten	361
16.1.1 Die Grossberg-Schicht	361
16.1.2 ART-Architekturen	363
16.2 Netze mit einer Kohonen-Schicht	367
16.2.1 Counterpropagation-Netz	367
16.2.2 Kohonen-Schicht und lineare Assoziatoren	369
16.2.3 Radiale Funktionen	370
16.2.4 Wechselwirkung zwischen Schichten	373

16.3 Historische Anmerkungen	375
17 Genetische Algorithmen.....	377
17.1 Codierung und Operatoren.....	377
17.1.1 Optimierung durch Evolutionsstrategien	377
17.1.2 Methoden der stochastischen Optimierung.....	379
17.1.3 Genetische Codierung.....	381
17.1.4 Informationsaustausch durch genetische Operatoren	383
17.2 Eigenschaften von genetischen Algorithmen.....	384
17.2.1 Konvergenzanalyse.....	384
17.2.2 Genetisches Driften.....	389
17.2.3 Gradientenmethoden versus genetische Algorithmen	391
17.3 Optimierung von Funktionen.....	392
17.3.1 Genetische Algorithmen für neuronale Netze.....	392
17.3.2 Andere Anwendungen von genetischen Algorithmen	396
17.4 Historische Anmerkungen	398
18 Hardware für neuronale Netze.....	399
18.1 Klassen von Neurohardware	399
18.1.1 Implementierung künstlicher neuronaler Netze.....	399
18.1.2 Taxonomie der Hardwarearchitekturen	401
18.2 Neuronale Netze in Analog-Technik	403
18.2.1 Codierung der Signale.....	403
18.2.2 VLSI-Transistor-Schaltungen	404
18.2.3 Transistoren mit eingebautem Ladungsspeicher.....	406
18.2.4 Pulsierende CCD-Schaltungen.....	407
18.3 Der digitale Ansatz	409
18.3.1 Analog versus Digital	409
18.3.2 Numerische Darstellung der Netzparameter und Signale.....	410
18.3.3 Vektor- und Signalprozessoren.....	411
18.3.4 Systolische Felder	412
18.3.5 Eindimensionale Strukturen	414
18.4 Innovative Rechnerarchitekturen	417
18.4.1 VLSI-Mikroprozessoren für neuronale Netze	417
18.4.2 Optische Rechner	420
18.4.3 Pulscodierte Netze	423
18.5 Historische Anmerkungen	425
Literaturverzeichnis	427
Index.....	443