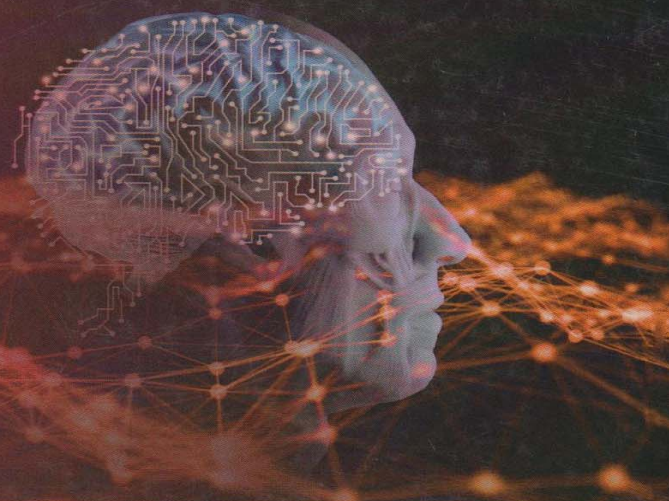


Texts in Computer Science


Rudolf Kruse ·  
Sanaz Mostaghim ·  
Christian Borgelt ·  
Christian Braune ·  
Matthias Steinbrecher



# Computational Intelligence

A Methodological Introduction

*Third Edition*

 Springer





2-006-143-1


Rudolf Kruse • Sanaz Mostaghim •  
Christian Borgelt • Christian Braune •  
Matthias Steinbrecher

# Computational Intelligence

A Methodological Introduction

Third Edition

With contributions from Frank Klawonn  
and Christian Moewes

 Springer

# Contents

<b>1</b>	<b>Introduction</b> . . . . .	<b>1</b>
1.1	Intelligent Systems . . . . .	1
1.2	Computational Intelligence . . . . .	2
1.3	About the Third Edition of this Book. . . . .	3
 <b>Part I Neural Networks</b>		
<b>2</b>	<b>Introduction to Artificial Neural Networks</b> . . . . .	<b>7</b>
2.1	Motivation . . . . .	7
2.2	Biological Background . . . . .	10
	References . . . . .	12
<b>3</b>	<b>Threshold Logic Units</b> . . . . .	<b>15</b>
3.1	Definition and Examples . . . . .	15
3.2	Geometric Interpretation . . . . .	17
3.3	Limitations . . . . .	19
3.4	Networks of Threshold Logic Units . . . . .	21
3.5	Training the Parameters . . . . .	24
3.6	Variants . . . . .	35
3.7	Training Networks . . . . .	35
	References . . . . .	36
<b>4</b>	<b>General Neural Networks</b> . . . . .	<b>39</b>
4.1	Structure of Neural Networks . . . . .	39
4.2	Operation of Neural Networks . . . . .	42
4.3	Scale Types and Encoding . . . . .	46
4.4	Training Neural Networks . . . . .	48
	References . . . . .	52
<b>5</b>	<b>Multi-layer Perceptrons</b> . . . . .	<b>53</b>
5.1	Definition and Examples . . . . .	53
5.2	Why Non-linear Activation Functions? . . . . .	60
5.3	Function Approximation . . . . .	61
5.4	Logistic Regression . . . . .	67



5.5	Gradient Descent . . . . .	71
5.6	Error Backpropagation . . . . .	85
5.7	Gradient Descent Examples . . . . .	86
5.8	Variants of Gradient Descent . . . . .	89
5.8.1	Manhattan Training . . . . .	90
5.8.2	Momentum Term . . . . .	90
5.8.3	Nesterov's Accelerated Gradient (NAG) . . . . .	91
5.8.4	Self-adaptive Error Backpropagation (SuperSAB) . . . . .	91
5.8.5	Resilient Backpropagation . . . . .	92
5.8.6	Quick Propagation . . . . .	93
5.8.7	Adaptive Subgradient Descent (AdaGrad) . . . . .	94
5.8.8	Root Mean Squared Gradient Descent (RMSProp) . . . . .	95
5.8.9	Adaptive Subgradient Descent Over Windows (AdaDelta) . . . . .	96
5.8.10	Adaptive Moment Estimation (Adam) . . . . .	96
5.8.11	Lifting the Derivative of the Activation Function . . . . .	98
5.8.12	Weight Decay . . . . .	98
5.8.13	Batch Normalization . . . . .	99
5.9	Examples for Some Variants . . . . .	100
5.10	Initializing the Parameters . . . . .	106
5.11	Number of Hidden Neurons . . . . .	107
5.12	Deep Learning . . . . .	109
5.13	Sensitivity Analysis . . . . .	120
	References . . . . .	122
<b>6</b>	<b>Radial Basis Function Networks . . . . .</b>	<b>125</b>
6.1	Definition and Examples . . . . .	125
6.2	Function Approximation . . . . .	130
6.3	Initializing the Parameters . . . . .	133
6.4	Training the Parameters . . . . .	139
6.5	Example of Training . . . . .	144
6.6	Generalized Form . . . . .	146
	References . . . . .	148
<b>7</b>	<b>Self-organizing Maps . . . . .</b>	<b>149</b>
7.1	Definition and Examples . . . . .	149
7.2	Learning Vector Quantization . . . . .	153
7.3	Neighborhood of the Output Neurons . . . . .	160
	References . . . . .	165
<b>8</b>	<b>Hopfield Networks . . . . .</b>	<b>167</b>
8.1	Definition and Examples . . . . .	167
8.2	Convergence of the Computations . . . . .	170
8.3	Associative Memory . . . . .	175



8.4	Solving Optimization Problems . . . . .	180
8.5	Simulated Annealing . . . . .	186
8.6	Boltzmann Machines . . . . .	188
	References . . . . .	192
<b>9</b>	<b>Recurrent Networks . . . . .</b>	<b>195</b>
9.1	Simple Examples . . . . .	195
9.2	Representing Differential Equations . . . . .	200
9.3	Vectorial Neural Networks . . . . .	202
9.4	Error Backpropagation Through Time . . . . .	205
9.5	Long Short-Term Memory . . . . .	207
	References . . . . .	211
<b>10</b>	<b>Neural Networks: Mathematical Remarks . . . . .</b>	<b>213</b>
10.1	Equations for Straight Lines . . . . .	213
10.2	Regression . . . . .	215
10.3	Activation Transformation . . . . .	220
 <b>Part II Evolutionary Algorithms</b>		
<b>11</b>	<b>Introduction to Evolutionary Algorithms . . . . .</b>	<b>225</b>
11.1	Metaheuristics . . . . .	225
11.2	Biological Evolution . . . . .	226
11.3	Simulated Evolution . . . . .	231
	11.3.1 Optimization Problems . . . . .	231
	11.3.2 Basic Notions and Concepts . . . . .	234
	11.3.3 Building Blocks of an Evolutionary Algorithm . . . . .	237
11.4	The n-Queens Problem . . . . .	239
11.5	Related Optimization Techniques . . . . .	244
	11.5.1 Gradient Ascent or Descent . . . . .	245
	11.5.2 Hill Climbing . . . . .	246
	11.5.3 Simulated Annealing . . . . .	247
	11.5.4 Threshold Accepting . . . . .	248
	11.5.5 Great Deluge Algorithm . . . . .	248
	11.5.6 Record-to-Record Travel . . . . .	249
11.6	The Traveling Salesman Problem . . . . .	250
	References . . . . .	253
<b>12</b>	<b>Elements of Evolutionary Algorithms . . . . .</b>	<b>255</b>
12.1	Encoding of Solution Candidates . . . . .	255
	12.1.1 Hamming Cliffs . . . . .	256
	12.1.2 Epistasis . . . . .	259
	12.1.3 Closedness of the Search Space . . . . .	260



12.2	Fitness and Selection . . . . .	263
12.2.1	Fitness Proportionate Selection . . . . .	264
12.2.2	The Dominance Problem . . . . .	265
12.2.3	Vanishing Selective Pressure . . . . .	266
12.2.4	Adapting the Fitness Function . . . . .	267
12.2.5	The Variance Problem . . . . .	269
12.2.6	Rank-Based Selection . . . . .	270
12.2.7	Tournament Selection . . . . .	271
12.2.8	Elitism . . . . .	272
12.2.9	Environmental Selection . . . . .	273
12.2.10	Niche Techniques . . . . .	273
12.2.11	Characterization of Selection Methods . . . . .	274
12.3	Genetic Operators . . . . .	275
12.3.1	Mutation Operators . . . . .	275
12.3.2	Crossover Operators . . . . .	278
12.3.3	Multi-parent Operators . . . . .	282
12.3.4	Characteristics of Recombination Operators . . . . .	283
12.3.5	Interpolating and Extrapolating Recombination . . . . .	284
	References . . . . .	285
<b>13</b>	<b>Fundamental Evolutionary Algorithms . . . . .</b>	<b>287</b>
13.1	Genetic Algorithms . . . . .	287
13.1.1	The Schema Theorem . . . . .	289
13.1.2	The Two-Armed Bandit Argument . . . . .	296
13.1.3	The Principle of Minimal Alphabets . . . . .	298
13.2	Evolution Strategies . . . . .	299
13.2.1	Selection . . . . .	300
13.2.2	Global Variance Adaptation . . . . .	301
13.2.3	Local Variance Adaptation . . . . .	303
13.2.4	Covariances . . . . .	304
13.2.5	Recombination Operators . . . . .	309
13.3	Genetic Programming . . . . .	310
13.3.1	Initialization . . . . .	313
13.3.2	Genetic Operators . . . . .	315
13.3.3	Application Examples . . . . .	316
13.3.4	The Problem of Introns . . . . .	321
13.3.5	Extensions . . . . .	321
13.4	Multi-objective Optimization . . . . .	322
13.4.1	Weighted Combination of Objectives . . . . .	323
13.4.2	Pareto-Optimal Solutions . . . . .	323
13.4.3	Finding Pareto Frontiers with Evolutionary Algorithms . . . . .	325



13.5	Special Applications and Techniques . . . . .	330
13.5.1	Behavioral Simulation . . . . .	330
13.5.2	Parallelization . . . . .	336
	References . . . . .	339
<b>14</b>	<b>Computational Swarm Intelligence . . . . .</b>	<b>343</b>
14.1	Introduction . . . . .	343
14.2	Basic Principles of Computational Swarm Intelligence . . . . .	344
14.2.1	Swarms in Known Environments . . . . .	347
14.2.2	Swarms in Unknown Environments . . . . .	348
14.3	Particle Swarm Optimization . . . . .	349
14.3.1	Influence of the Parameters . . . . .	351
14.3.2	Turbulence Factor . . . . .	352
14.3.3	Boundary Handling . . . . .	352
14.4	Multi-objective Particle Swarm Optimization . . . . .	353
14.4.1	Leader Selection Mechanism . . . . .	354
14.4.2	Archiving . . . . .	357
14.5	Many-objective Particle Swarm Optimization . . . . .	360
14.5.1	Ranking Non-dominated Solutions . . . . .	361
14.5.2	Distance Based Ranking . . . . .	361
14.6	Ant Colony Optimization . . . . .	362
	References . . . . .	368

### Part III Fuzzy Systems

<b>15</b>	<b>Introduction to Fuzzy Sets and Fuzzy Logics . . . . .</b>	<b>373</b>
15.1	Natural Languages and Formal Models . . . . .	373
15.2	Fuzzy Sets . . . . .	374
15.3	Representation of Fuzzy Sets . . . . .	376
15.3.1	Definition Based on Functions . . . . .	377
15.3.2	$\alpha$ -Cuts . . . . .	379
15.4	Fuzzy Logic . . . . .	381
15.4.1	Propositions and Truth Values . . . . .	381
15.4.2	t-Norms and t-Conorms . . . . .	385
15.4.3	Aggregation Functions . . . . .	390
15.5	Semantics of Membership Degrees . . . . .	392
15.5.1	Membership Degrees as Truth Degrees . . . . .	392
15.5.2	Membership Degrees as Similarity to a Reference Value . . . . .	393
15.5.3	Membership Degrees as Preferences . . . . .	394
15.5.4	Membership Degrees as Possibility . . . . .	394
15.5.5	Consistent Interpretations of Fuzzy Sets in Applications . . . . .	394



15.6	Operations on Fuzzy Sets . . . . .	395
15.6.1	Intersection . . . . .	395
15.6.2	Union . . . . .	397
15.6.3	Complement . . . . .	398
15.6.4	Covering and Partition . . . . .	399
15.6.5	Linguistic Modifiers . . . . .	399
15.7	Fuzzy Sets of Type 2 . . . . .	400
	References . . . . .	404
<b>16</b>	<b>The Extension Principle . . . . .</b>	<b>407</b>
16.1	Mappings of Fuzzy Sets . . . . .	407
16.2	Mappings of $\alpha$ -Cuts . . . . .	409
16.3	Cartesian Product and Cylindrical Extension . . . . .	410
16.4	Extension Principle for Multivariate Mappings . . . . .	412
	References . . . . .	413
<b>17</b>	<b>Fuzzy Relations . . . . .</b>	<b>415</b>
17.1	Crisp Relations . . . . .	415
17.2	Application of Relations . . . . .	417
17.3	Logical Deduction with Relations . . . . .	419
17.4	Simple Fuzzy Relations . . . . .	421
17.5	Composition of Fuzzy Relations . . . . .	425
17.6	Fuzzy Relational Equations . . . . .	427
	References . . . . .	428
<b>18</b>	<b>Similarity Relations . . . . .</b>	<b>429</b>
18.1	Similarity . . . . .	429
18.2	Fuzzy Sets and Extensional Hulls . . . . .	430
18.3	Scaling Concepts . . . . .	432
18.4	Fuzzy Sets and Similarity Relations . . . . .	435
	References . . . . .	439
<b>19</b>	<b>Approximate Reasoning . . . . .</b>	<b>441</b>
19.1	Linguistic Variables . . . . .	441
19.2	Computing with Words . . . . .	442
19.3	Generalized Logical Inference . . . . .	443
19.4	Approximation of Functions Using Linguistic If-Then Rules . . . . .	445
19.4.1	Approximation of Functions by Using Rules as Constraints . . . . .	445
19.4.2	Approximation of Functions by Solving Fuzzy Relational Equations . . . . .	447
19.4.3	Approximation of Functions by Interpolation Between Fuzzy Points . . . . .	448
	References . . . . .	449



<b>20</b>	<b>Fuzzy Control</b>	451
20.1	Mamdani Fuzzy Controller	452
20.2	Design of a Mamdani Fuzzy Controller	457
20.3	Mamdani Controller and Similarity Relations	461
20.3.1	Interpretation of a Mamdani Controller Using Similarity Relations	461
20.3.2	Construction of a Mamdani Controller Using Similarity Relations	463
20.4	Takagi–Sugeno Controller	465
	References	468
<b>21</b>	<b>Hybrid Systems for Tuning and Learning Fuzzy Systems</b>	471
21.1	Neuro-Fuzzy Control	471
21.1.1	Models with Supervised Learning Methods	474
21.1.2	Models with Reinforcement Learning	477
21.2	Evolutionary Fuzzy Control	481
21.2.1	Structure of an Evolutionary Fuzzy Controller	481
21.2.2	Optimizing Parameters of an Evolutionary Fuzzy Controller	483
21.2.3	Example	485
	References	486
<b>22</b>	<b>Fuzzy Data Analysis</b>	489
22.1	Fuzzy Methods in Data Analysis	489
22.2	Fuzzy Clustering	490
22.2.1	Clustering	490
22.2.2	Presuppositions and Notation	491
22.2.3	Classical <i>c</i> -Means Clustering	492
22.2.4	Fuzzification by Membership Transformation	494
22.2.5	Fuzzification by Membership Regularization	497
22.2.6	Comparison	501
22.3	Analysis of Precise Data with Possibility Theory	502
22.4	Analysis of Imprecise Data Using Random Sets	505
22.5	Analysis of Fuzzy Data by Fuzzy Random Variables	508
	References	510
<b>Part IV Bayesian Networks</b>		
<b>23</b>	<b>Bayesian Networks</b>	517
<b>24</b>	<b>Elements of Probability and Graph Theory</b>	523
24.1	Probability Theory	523
24.1.1	Random Variables and Random Vectors	526
24.1.2	Independences	530