## Index

### A

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A* state-space search algorithm</td>
<td>227</td>
</tr>
<tr>
<td>abstraction</td>
<td>2</td>
</tr>
<tr>
<td>activation function</td>
<td>172</td>
</tr>
<tr>
<td>Active Template Library</td>
<td>73</td>
</tr>
<tr>
<td>ActiveX Template Library</td>
<td>81</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>37, 57</td>
</tr>
<tr>
<td>AdaCost</td>
<td>37</td>
</tr>
<tr>
<td>adaptive learning rate</td>
<td>194</td>
</tr>
<tr>
<td>adaptive natural gradient learning</td>
<td>186</td>
</tr>
<tr>
<td>administrative application</td>
<td>232</td>
</tr>
<tr>
<td>adult database</td>
<td>102</td>
</tr>
<tr>
<td>algorithm-specific preprocessing</td>
<td>4</td>
</tr>
<tr>
<td>algorithms</td>
<td>14, 123</td>
</tr>
<tr>
<td>application domain</td>
<td>243</td>
</tr>
<tr>
<td>approximation sets</td>
<td>211</td>
</tr>
<tr>
<td>archaeological knowledge</td>
<td>263</td>
</tr>
<tr>
<td>artificial neural networks</td>
<td>8, 241, 244</td>
</tr>
<tr>
<td>association</td>
<td>6</td>
</tr>
<tr>
<td>attributes</td>
<td>268</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>backward propagation</td>
<td>172</td>
</tr>
<tr>
<td>bagging</td>
<td>54</td>
</tr>
<tr>
<td>Bayes() Theorem</td>
<td>109</td>
</tr>
<tr>
<td>Bayesian belief networks</td>
<td>115</td>
</tr>
<tr>
<td>Bayesian methods</td>
<td>8</td>
</tr>
<tr>
<td>Bayesian networks</td>
<td>244</td>
</tr>
<tr>
<td>Bayesian neural network</td>
<td>116</td>
</tr>
<tr>
<td>Bayesian perspective</td>
<td>109</td>
</tr>
<tr>
<td>binary reuse</td>
<td>75</td>
</tr>
<tr>
<td>blocking</td>
<td>113</td>
</tr>
<tr>
<td>boosting</td>
<td>28, 56</td>
</tr>
<tr>
<td>Bootstrap aggregating Algorithm</td>
<td>56</td>
</tr>
<tr>
<td>brand loyalty</td>
<td>210</td>
</tr>
<tr>
<td>brand switching</td>
<td>210</td>
</tr>
<tr>
<td>Bayesian perspective</td>
<td>109</td>
</tr>
<tr>
<td>Bayes() Theorem</td>
<td>109</td>
</tr>
<tr>
<td>backward propagation</td>
<td>172</td>
</tr>
<tr>
<td>bagging</td>
<td>54</td>
</tr>
<tr>
<td>Bayes() Theorem</td>
<td>109</td>
</tr>
<tr>
<td>Bayesian belief networks</td>
<td>115</td>
</tr>
<tr>
<td>Bayesian methods</td>
<td>8</td>
</tr>
<tr>
<td>Bayesian networks</td>
<td>244</td>
</tr>
<tr>
<td>Bayesian neural network</td>
<td>116</td>
</tr>
<tr>
<td>Bayesian perspective</td>
<td>109</td>
</tr>
<tr>
<td>binary reuse</td>
<td>75</td>
</tr>
<tr>
<td>blocking</td>
<td>113</td>
</tr>
<tr>
<td>boosting</td>
<td>28, 56</td>
</tr>
<tr>
<td>Bootstrap aggregating Algorithm</td>
<td>56</td>
</tr>
<tr>
<td>brand loyalty</td>
<td>210</td>
</tr>
<tr>
<td>brand switching</td>
<td>210</td>
</tr>
</tbody>
</table>

### C

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>case-based reasoning</td>
<td>228</td>
</tr>
<tr>
<td>chaotic time series</td>
<td>145</td>
</tr>
<tr>
<td>class factory</td>
<td>72</td>
</tr>
<tr>
<td>class of functions</td>
<td>144</td>
</tr>
<tr>
<td>classification</td>
<td>6, 14</td>
</tr>
<tr>
<td>classification algorithm</td>
<td>19</td>
</tr>
<tr>
<td>clinical Applications</td>
<td>231</td>
</tr>
<tr>
<td>Cluster analysis</td>
<td>208</td>
</tr>
<tr>
<td>clustering</td>
<td>6, 20</td>
</tr>
<tr>
<td>clustering algorithms</td>
<td>90, 208, 244</td>
</tr>
<tr>
<td>CN2</td>
<td>241</td>
</tr>
<tr>
<td>COM API</td>
<td>73</td>
</tr>
<tr>
<td>COM class</td>
<td>72</td>
</tr>
<tr>
<td>COM Servers</td>
<td>73</td>
</tr>
<tr>
<td>commercial neural network software</td>
<td>170</td>
</tr>
<tr>
<td>Common Object Request Broker</td>
<td>71</td>
</tr>
<tr>
<td>Component Object Model</td>
<td>68</td>
</tr>
<tr>
<td>component reuse</td>
<td>75</td>
</tr>
<tr>
<td>computational cost</td>
<td>197</td>
</tr>
<tr>
<td>computational experiments</td>
<td>198</td>
</tr>
<tr>
<td>computational time</td>
<td>19</td>
</tr>
<tr>
<td>condition attribute</td>
<td>268</td>
</tr>
<tr>
<td>conditional probabilities</td>
<td>127</td>
</tr>
<tr>
<td>confidence limits</td>
<td>128</td>
</tr>
<tr>
<td>connected component</td>
<td>95</td>
</tr>
<tr>
<td>consumer choice models</td>
<td>210</td>
</tr>
<tr>
<td>convenience-oriented shoppers</td>
<td>208</td>
</tr>
<tr>
<td>convex programming</td>
<td>14</td>
</tr>
<tr>
<td>cost-sensitive boosting</td>
<td>28</td>
</tr>
<tr>
<td>cost-sensitive classification</td>
<td>28</td>
</tr>
<tr>
<td>cost-sensitive tree induction</td>
<td>28</td>
</tr>
<tr>
<td>credit risk analysis</td>
<td>250</td>
</tr>
<tr>
<td>crossover</td>
<td>59</td>
</tr>
<tr>
<td>cultural algorithm</td>
<td>266</td>
</tr>
</tbody>
</table>
D

data analysis 173
data cleaning 4
data compression 236
data conversion 4
data description 4
data extraction 1
data matrix 211
data mining 1,109,122,209,225, 241
data preparation 245
data preprocessing 189
data set 89
data sources 243
data transformation 4
databases 1,241
datasets 13
decision attribute 268
decision rule 271
decision system 212
decision system construction 271
decision table 269
Decision tree induction (B 130
Decision Trees 8, 244
dense point 96
dependency modeling 6
digital medical imaging 230
dispersion 136
disproportionate stratification 134
dissimilarity coefficient 90
distributed COM 73
domain knowledge 263
domain-specific data preprocessing 4
dynamic processes 264
dynamic sampling 136

E

economic shoppers 208
electronic record keeping systems 230
emerging technology 78
ensemble of classifiers 54
error function 192
ethical shoppers 208
evaluation 236,245
evaluation 4
evolutionary learning tools 263
expected cost criterion 42

F

false nearest neighbour 160
feature selection 6
Feedforward Artificial Neural Networks 8
filters 152
fragmentation 132
functional mapping 186
Fuzzy logic 228

G

GA stacking 61
Gamma histogram 156
Gamma statistic 144
Gamma test 142
genetic search algorithm 227
Global Logical-Combinatorial Clustering 96
Globally Unique Identifier 72
gradient descent learning method 188

H

Hessian matrix 195
heuristic 227
heuristic function 59
heuristic methods 194
heuristic search engines 79
human genome research project 230

I

increasing embedding 160
incremental mining 235
information search 210
information system 211
informative features 14
initial population 60
instance weighting 29
interconnectivity 94
interface definition language 75
interface identifier 72
internal classifier 49
irregular embedding 162
K
knowledge discovery 241
knowledge discovery from data 89
knowledge discovery in databases 241, 242
knowledge discovery in databases 1
knowledge discovery process 263
knowledge extraction 242

L
language Independence 70
large data sets 89, 186
Large Scale Medical Data Mining 234
LDSCA 89
learning procedures 123
learning process 189
learning speed 199
legal liability 229
linguistic representation 241

M
Machine Learning 241
Maintenance Paths 77
market segmentation 208
marketing research 209
Markov Chain 114
mask 152
medical discovery 234
metaclassifier 55
Microsoft Foundation Classes 73
misclassification cost 30
mixed incomplete data 90
model evaluation 5
model representation 5
model search 5
Multi Layer Perceptron Neural Network 241
multiple languages 77
multivariate decision trees 8
mutation 59

N
natural gradient 197
natural gradient learning 197
network architecture 191
Network Initialization 193
neural networks 169, 186, 209, 228
noise 143, 230, 245
non-hierarchical cluster analysis 210, 215
non-integral dimension 148
non-linear data analysis 149

O
object descriptions 89
optimization 1
over-parameterisation 170

P
pattern recognition 13
patterns 2, 243
Personalizing shoppers 208
plateau problem 188
population converges 60
posterior distribution 118
prediction 6
predictions 13
preprocessing stage 188
price sensitivity 210
prior group membership 209
prior probability in 112
project management 77
proportion estimates 124, 125
proportionate stratification 134
protein-folding problem 164
pseudo-distance function 90

Q
quantitative methods 252
Quasi-Newton method 195

R
recreational shoppers 208
reduced version problems 70
refining 13
representative sample 243
representative set of samples 241
reproduction 59
rough clustering 209, 213
rough set 228
rough set concepts 263
rough sets theory 209
rough sets theory 209

S
sample size 125, 138
sampling 123
sampling error 148
sampling fraction 125
scientific theory induction 246
search 2, 5
search operators 59
search space 264
search-space reduction 2
selection 4, 13
semi-distance function 90
shopping orientations 208
similarity functions 90
similarity matrix 89
simple random sampling 124
skeleton objects 97
small data sets 170
small disjuncts 136
smooth systems 142
stacked generalization 58
stacking 54
standardisation 230
stratification 134
subjective methods 251
summarization 6
support vector machines 8

symbolic learning algorithms 241

T
Tabu search 227
target function complexity 136
test environment 79
time-series modelling 159, 170
training data set 172
TREPAN algorithm 241

U
uncertainty 266
unconditional prior probabilities 127
univariate decision trees 8

V
Valid 243
validation 246
validation data set 172
vendor independence 70
very large data set 89
visualisation techniques 229
VLDSCA 89

W
weight initialization 42
winGamma© 162
wrappers 152