

Index

A

- Adaptive boosting (AB), *see* Evolutionary programming (EP)
- Alternative kernel functions, 190
- Alzheimer's disease (AD)
 - ABeta, 108, 109
 - APOE4, 109
 - ASR, 113, 114
 - biomarkers, 112
 - causes of death, USA, 108
 - demographics, 110
 - description, 108
 - detection, from speech
 - experimental designs, 121–122
 - intuition, 114
 - parse trees (*see* Parse trees)
 - recorded goal-setting interviews, 114
 - signal-based features, 117–118
 - speech acquisition protocol, 115
 - speech samples, 115–116
 - speech to text, 118–120
 - syntactic complexity features, 123
 - text-based features, 120–121
 - diagnosis, 110–111
 - effects on speech, 112, 113
 - linguistic measures, 113
 - MMSE, 111
 - research, 109
 - speech-based test, 113
 - syndrome, 108
- Alzheimer's speech data, 206, 207
 - BN
 - approach, 214
 - confusion matrix, 217

- estimating conditional probabilities, 214–215
 - structure, 214
 - tenfold training, 216
 - machine design approach, 213
 - MI methods, 217
 - three-tier BN network structure, 217
- Apolipoprotein E4 allele (APOE4), 109, 110
- Area under curve (AUC), 171
- Area under the ROC curve (AUC), 145, 146, 149, 150
- Artificial neural network (ANN), 220
- Ask-an-expert approach, 175
- Automatic speech recognition (ASR), 113–116, 118, 120
- Automatic speech-to-text, 124

B

- Bank of SVMs
 - BN approach, 246
 - mammogram research project, 237–239
 - measures of performance, 242–243
 - processing design, 241
 - structure of, 239
 - SVM-Oracle hybrid, 247
 - three-tier BN, 243–246
- Bayes theorem, 169, 170
 - biomedical illustration cases, 171–174
 - exclusive and exhaustive events, 168
 - law of total probability, 168, 169
 - measures of performance, 171
- Bayes theory
 - CI, 162–167

- Bayes theory (*cont.*)
 conditional probability, 159–162
 principle of indifference, 159
 probability function/probability space, 158, 159
- Bayesian approach, 181
- Bayesian networks (BNs), 174
 ask-an-expert approach, 175
 chain rule of conditional probabilities, 179, 180
 Cooper-Herskovitz K2 metric, 177–179
 cyber security analysis, 175
 data-driven approach, 175
 dependency analysis approach, 176, 177
 design, 212
 functional flow–pruning, 175, 176
 machine intelligent approaches, 174
 MI application, 213
 MI method, 181–184
 search approach, 177
 template layers, 212–213
 three-layer construction, 213
- Bayesian probabilistic neural network (BPNN)
 benign and malignant, 194, 195
 cross-validation analysis, 198
 density function and misclassification costs, 198
 dynamic range, 192
 exponential algorithm, 192
 Gaussian density function, 197
 Gaussian kernel σ 's and multipliers, 200
 OWN sigma value, 192
 Parzen density functions, 197
 Parzen's mathematical conditions, 198
 PNN classification, 199
 PNN kernel, 196, 197
 ROC analysis, 193
 speech data example, 206, 207
 suitable and accurate, 197
 toy example, 200
 two-class classification problem, 193
- Beta Amyloid (ABeta), 108, 109
- Binary coded decimal (BCD) and Gray code, 13
- Blend crossover (BLX), 25, 39
- Breast cancer
 FNA (*see* Fine needle aspirate (FNA))
- C**
- CHC algorithm, 138, 146
 BLX- α crossover operator, 25
 cross-generational rank selection, 17, 18
 crossover operator, real number genes, 24, 25
- exploration/exploitation trade-off, 15
- halting mechanisms, 20, 21
- HUX, 15, 19
- limits on complexity
 interacting genes, 21
 NK landscapes, 21, 24
 search biases, 22
 SGA and RBC+, 22
 trial and error learning, 21
- Pseudo code, 16
- soft restart operation, 19, 20
- UX, 19
- CH-K2 theorem
 “B-greedy” algorithm, 179
 “classic” algorithm, 178
 decomposable scoring function, 179
 drawback, 179
 parent-ordering assumption, 179
 probability density function, 178
 variables, 178
- Classification accuracy, 144
- Clock drawing test, 112
- Colon cancer
 BN, 234
 conditional probabilities
 binary feature, 235
 CR, 236
 confusion matrix, 237
 CRC screening method, 228
 decision trees (DT), 231
 gastroenterology practice, 229
 medical pattern discovery, 228
 MI application, 236
 precondition features, 234
 prior probabilities, 234
- Colorectal cancer (CRC)
 risk factors, 231
 screening method, 228
- Computational linguistics, 114
- Computerized Propositional Idea Density Rater (CPIDR) program, 124
- Conditional independence (CI)
 BNs analysis, 165
 classes/probability spaces, 162, 163
 definition, 162
 indifference and counting arguments, 166, 167
 objects, 165
 sets, 166
 Venn diagram, 163, 164
- Conditional probability
 definition, 159
 fraction of items, 160

independence, 161
 Venn diagrams, 161
 Cookie theft picture, 115, 125, 128
 Crossover operator for real number genes, 24, 25
 Cross-validation, 148

D

Data-driven approach, 175
 Data normalization, 139–141
 Decision trees (DT) approach, 234
 Demographics, AD, 110
 Dependency analysis approach
 difficulties, 176
 PC algorithm, 176
 Developmental Level, 124
 Diagnosis, AD, 109–112
 Directed acyclic graph (DAG), 181
 Dual Lagrangian, 68, 69

E

Effects on speech, AD, 112–113
 Electroencephalogram (ECG), 141
 Empirical risk minimization (ERM), 46, 47,
 50, 52
 Ensemble processing (EP), 88
 EP-AB hybrid, breast cancer
 Adaboost algorithm, 93
 ANN, 93
 Duke University mammogram data base, 92
 non-optimized weak learner, 95
 optimized ANN, 94
 pseudo code, 93
 weak classifiers, 94
 Estimating conditional probabilities, 212
 Estimating sample size, 99
 Evolutionary computation (EC)
 aerodynamic design problems, 2
 allocations, 3
 awareness, late 1970s, 13
 binary encodings, 12, 13
 biological, 15
 cost functions, 14
 Eshelman's CHC GA, 14
 essential ingredients, 2
 IHC, 15
 k slot machines, 3
 Koza's GP, 13
 machine learning tasks, 2
 NFL theorem, 14
 operations, 2
 programming, 2

Pseudo code, simple genetic algorithm, 3
 random search (RS), 15
 schema sampling theorem (*see* Schema
 sampling theorem)
 simulated annealing, 15
 trade-off, 3
 Evolutionary programming (EP) approach,
 220–221
 and adaptive boosting technologies, 92
 EP-AB hybrid
 breast cancer malignancy prediction (*see*
 EP-AB hybrid, breast cancer)
 GRNN oracle architectures, 97
 Extension of area under ROC, 145
 Extensive testing, 176
 Eye–brain system, 219

F

False positive (FP), 171, 174
 Feature space, 70, 72
 Feature subset selection, 137, 138, 144, 150
 Fine needle aspirate (FNA)
 benign data, 220 (*see also* Malignant
 and benign data)
 BN
 build-and-test approach, 226
 probability density functions, 225
 processing, 218
 Wolberg data, 224
 data description, 218
 EP approach
 GNARL, 220
 process and design, 221
 RNN, 221
 evolved architecture, 222
 GNARL, 217
 malignant data, 219 (*see also* Malignant
 and benign data)
 malignant vs. benign, 227
 performance evaluation, 222
 screening/classification performance, 227
 Wolberg data, 218, 228
 Function $f(X)$, expected risk, 48–50
 Future risk (FR), 228, 232

G

GA-SSS algorithm, 42
 GA-SVM hybrid approach
 AD speech data
 cross-validation, feature set, 148
 data set, 147, 148

GA-SVM hybrid approach (*cont.*)
 GRNN oracle, 153, 154
 picking the winners, 149, 150
 AUC concept, 145, 146, 149, 150
 data normalization, 139–141
 multi-objective selection, 141–144
 GA-SVM-Oracle approach, 214, 215
 and BN, 261, 262
 Generalized acquisition of recurrent links
 (GNARL), 217, 220
 Generalized regression neural network
 (GRNN) oracle, 153, 249
 CAS approach, 78, 87
 data sets, lung cancer (*see* Lung cancer
 data set)
 ensemble combination, 82–85
 ensemble processor, 89
 ensemble-expert architectures, 78
 inferior models, 88
 kernel density estimation, 79, 80
 modified combination design, 85–87
 multivariate probability density
 approximation, 81
 and PNN, 80
 predicted squared error, 88
 prediction model, 88
 ROC curve, 89–91
 statistical learning systems, 78
 sum–integration order, 81, 82
 trouble-makers (*see* Trouble-makers)
 weighted Euclidean distance, 88
 Genetic algorithms (GAs), 73, 74, 182
 CHC (*see* CHC algorithm)
 crossover for subset selection tasks (*see*
 Subset selection)
 evolutionary computation (*see* Evolutionary
 computation (EC))
 MMX operator, 42
 MMX_SSS approach, 42
 problem set, 43
 Genetic algorithm–support vector machine
 hybrid (GA-SVM), 252
 Genetic Algorithm–Support Vector Machine-
 Oracle hybrid (GA-SVM-Oracle),
 212–215, 217
 GNARL neural network classification
 (diagnostic) approach, 219

H

Half-uniform crossover (HUX), 19
 Halting mechanisms, 21
 Hard and soft limiter classifiers, 66

Hessian matrix, 70
 Hierarchical multi-objective selection, 40

I

Incest prevention, 18–20, 22, 38
 Integral cross-validation, 139, 148
 Interval schemata, 25

K

K2 algorithm, *see* CH-K2 algorithm
 Kernel, 71–72
 Kernel-based probability density estimations,
 79, 80
 Kernel trick, 70, 71
 Koza's Genetic Programming, 13
 Kullback-Leibler divergence, 251

L

Lake picture, 115, 116, 129
 Learning classifier, 138, 139
 Learning/smart regression analysis, 47
 Linear separable problem, 57–59
 Lung cancer data set
 CAS algorithm processors, 101
 CASs, 97
 confidence interval *vs.* required minimum
 sample size, 100
 diagnosis problems, 101
 GRNN oracle theory, 97
 intra- and inter-institutional variability
 errors, 99
 margin of error, 99, 101
 multi-factor analysis, 102
 option, 99
 processing parameters, 99
 pseudo code, 100
 Raponi data set, 97
 resulting curves, 100
 ROC curves, 97, 98
 sample size, 99
 Shedden data set, 97
 targeted pre-processing, 102
 tools, 102
 validation, 101

M

Machine intelligence (MI) method
 Bayes probability of black, 183, 184
 conditional probability, 184

- counting arguments, 182, 183
 - GA-SVM-ORACLE hybrid approach, 182
 - principle of indifference, 182, 183
 - sets, 182
 - Magnetic resonance imaging (MRI), 109
 - Malignant and benign data
 - diagnostic/classification information, 219
 - EP results, 219
 - eye-brain system, 219
 - snowflakes, 220
 - Measure of performance (MOP) method, 218, 255
 - MI approach, 212
 - Mild cognitive impairment (MCI), 111
 - Mini-mental state exam (MMSE), 110, 111, 113–115, 148, 150, 153, 154, 252
- N**
- Negative emotion words, 131
 - Negative predicted value (NPV), 171, 183, 242
 - Nested learning models, 51
 - Neutral emotion words, 130
 - No Free Lunch (NFL) theorem, 14
 - N-sided polygon and spline methods
 - circular neighborhoods, 261
 - heuristic approach, 257
 - min_theta*, 256, 257, 260
 - MOPs, 259
 - polygon and spline, 259
 - polygonal uncertainty area, 257, 260
 - spline-fitting software, 261
 - uncertainty neighborhood, 256
- P**
- Parse trees, 123, 124
 - Parzen's windowing theorem, 189
 - Pause measurement, 117–119, 122, 123
 - Polygonal uncertainty area, 257, 260
 - Positive emotion words, 129
 - Positive predictive value (PPV), 171, 183, 242
 - Prediction-causal (PC) algorithm, 176
 - Primal Lagrangian, 68
 - Principle component regression (PCR), 228
 - Probabilistic neural network (PNN)
 - architecture, 191–192
 - Bayesian classification, 189
 - “Bayesian probability”, 188
 - description, 187
 - kernel and smoothing parameter
 - consistent estimator, 190
 - Gaussian function, 190
 - heuristic methods, 191
 - Parzen density estimator, 188
 - predictor variables, 188
 - properties, 189
 - Probabilistic neural networks (PNNs), 80, 97
 - Probability density functions, 48, 171, 172, 174, 178, 183, 184
- Q**
- Quadratic programming (QP) problem, 70
- R**
- Radcliffe's RAR (RRARw), 31
 - Random respectful recombination (RRR)
 - operator
 - assorting, 29
 - expected behaviors, 31, 32
 - MMXm crossover operator, 32
 - MRARw, 32, 33
 - operation, 30
 - SRARw, 32, 33
 - variant set, 31
 - Receiver operator characteristic (ROC) curve, 159, 171, 242
 - data, 91
 - issues, 90
 - learning classifier system, 89, 90
 - lung cancer
 - Raponi clinical data set, 98
 - Shedden clinical data set, 98
 - prediction performance, thresholds, 89
 - prediction scatter plot, 90
 - Recurrent neural networks (RNNs), 218, 221
 - Region of uncertainty
 - n-sided polygon and spline methods (*see* N-sided polygon and spline methods)
 - TM
 - boundary circles, 254
 - case prediction, 254
 - MOP, 255
 - nearest neighbors, 254
 - simplest circular, 255
 - Regularization parameter, 67, 70
 - Risk minimization principle (RMP), 47, 48
- S**
- Schema sampling theorem
 - behaviors, 5
 - bias, one-point crossover operator, 6
 - chromosomes strings, 4
 - crossover and mutation, 4

- Schema sampling theorem (*cont.*)
 - epistasis, 4
 - fitness of chromosome, 8
 - fitness-proportional reproduction, 9
 - FRs, 6
 - k -armed bandits problems, 7
 - maximum bias, 6
 - next generation, 9
 - offspring of parents, 8
 - optimization, over many generations, 9–12
 - predictions, 9
 - product of the probabilities, 5
 - properties, 4
 - Simple RAR (SRARw), 31
 - Size, subset selection
 - chromosome format, MMX_SSS, 38
 - copy common genes, 39
 - hierarchical multi-objective selection, 40
 - high-heeled boot curve, 41
 - incest prevention, 38
 - initial steps, MMX_SSS crossover, 39
 - MMX_SSS, 42
 - popsiz*e and *ss*s_max parameters, 40
 - redistribute unique genes, 39
 - ss*s gene values, 41
 - ss*s genes crossover, 39
 - Slack variables, 67, 68
 - Soft restart, 19, 20
 - Speech acquisition protocol, 115
 - Statistical learning theory (SLT), 47, 50, 51, 56
 - nested learning machines, 51, 52
 - optimal choice of model, 52
 - shattering, 54, 55
 - SRM, 51
 - SVM minimizing true risk, 55, 56
 - VC dimension, 53
 - Stochastic neighborhood embedding (SNE), 250
 - Alzheimer’s speech study, 253
 - dimension reduction, 252
 - Kullback-Leibler divergence, 251
 - MI method, 253
 - probability density, 250
 - reduced-dimensional space, 251
 - Shepard diagram, 253
 - speech samples, 252
 - straightforward modifications, 252
 - t -SNE algorithm, 251, 252
 - 2D t -SNE, 254
 - Structural risk minimization (SRM), 51
 - Student’s t distribution SNE (t -SNE), 250
 - Subset selection
 - first empirical tests, 29
 - outcomes, 29
 - representation
 - bit string crossover operators, 27
 - CHC, 26
 - index representations, 28
 - NK landscapes, 26
 - reciprocal values, 27
 - RRR operator, 29–32, 34
 - second empirical tests, 34–38
 - size, 38–42
 - Support vector machine (SVM), 182
 - ANN training, 56
 - components, 56
 - GA code, 138, 139
 - kernel selection, 73
 - Lagrangian formulation
 - linear non-separable case, 66–70
 - linear separable case, 60–66
 - non-separable and non-linear case, 70–72
 - learning classifier, 138
 - linear separable problem, 57–59
 - optimal choice of model, 52
 - SLT, 56
 - two separating hyperplane equations, 60
 - Support vector machines (SVMs), 97, 237
 - SVM-Oracle hybrid, 244, 247
- T**
- Three-level BN model, 236, 244
 - Trouble-makers (TMs), 249, 250, 252–254
 - complicated data sets, 92
 - definition, 91
 - EP-AB hybrid
 - breast cancer (*see* EP-AB hybrid, breast cancer)
 - and GRNN oracle architectures, 97
 - GRNN oracle, 95, 96
 - predictive models, 91
 - True negative (TN), 171, 172
 - True negative rate (TNR), 171
 - True positive (TP), 172
 - True positive rate (TPR), 171, 242
- U**
- Uncertainty area
 - GRNN oracle, 249
 - reductions of high-dimensional data, 250
 - region of uncertainty (*see* Region of uncertainty)
 - static pattern recognition problem, 249
 - t -SNE algorithm, 250

Uniform crossover (UX), 19
Unix Operating system, 125, 126

V

Vapnik and Chervonenkis (VC), 50, *see* VC
dimension
Vapnik's key learning theorem, 50
VC dimension, 50–57
Vocabulary richness, 125
Voice Activity Detector (VAD), 118, 119, 122

W

Weighting parameter, 68
Wolberg data, 225, 226
Wrapper approach, 137–139

Y

YAAPT (Yet Another Algorithm for Pitch
Tracking), 122