Index

$\chi^2$, see Chi square
2D-DIGE, see two-dimensional difference in-gel electrophoresis
2D-PAGE, see two-dimensional polyacrylamide gel electrophoresis

accuracy, 24
  balanced, 24, 194
adjacency matrix, 206–207
affycomp, 64
Affymetrix, 5, 52–53, 64
Agilent, 53–55
amplicon, 3
ANOVA, see one-way analysis of variance
arabadopsis, 47
ARD, see automatic relevance determination
area under the curve, 152, 169
ArrayAssist, 62
assortative mixing, 210, 220
assortativeness, 211
AUC, see area under the curve
automatic relevance determination, 156
BACC, see accuracy, balanced background correction, 55
backward elimination, 17, 155
bagging, see bootstrap aggregation
baseline subtraction, 82
beam search, 155
betweenness
  centrality, 208
distribution, 210
bias, 22
  experimenter, 14
  selection, 27, 177, 196
bias-variance trade-off, 23
biclustering, 124, 131, 135
Bioconductor, 62, 74, 98, 104, 112, 120
blocking, 43, 47
  complete block design, 43
blotting
  Northern, 4
  Southern, 4
  Western, 5
Bonferroni, 20, 134, 154, 232
bootstrap, 28, 179
  .632 bootstrap, 28, 179
  .632+ bootstrap, 180, 184
aggegration, 238
BRB-ArrayTools, 184

CAAT, 138
cabig, 98
calibrant, 84, 119
calibration, 81, 82
capacity control, 23
CART, see classification and regression trees
centroid, 193
Chi square, 152, 169
CIPHERGEN, 84, 93, 161
ClaNC, 197
class comparison, 41, 124
class discovery, 40, 124
class prediction, 124
classification, 10
classification and regression trees, 155, 166
classifier
   linear maximal margin, 189
   non-linear, 188
CLENCH, 134
cluster analysis, 40
clustering, 10, 124, 156, 158
   average, 209
   biclustering, 131
   coefficient, 207
   distribution, 210
   fuzzy, 137
   hierarchical, 46, 126, 128
      average linkage, 129
      complete linkage, 129
      single linkage, 129
   k-means, 126, 129, 158
   model-based, 130
   probabilistic, 158
   self-organizing maps, 130
   self-organizing tree algorithm, 130
   soft, 158
confidence interval, 24
confounding, 40, 47
connectedness, 40, 47
correlation
   filtering, 154
   jackknifed correlation coefficient, 127
   Pearson, 117, 126
   profile, 210
   Spearman rank, 127
covariate, 7
cross-hybridization, 53
cross-validation, 27, 191
   10-fold, 184
   5x2CV, 30
   external, 27, 181
   internal, 27, 181
   k-fold, 28, 178
   leave-k-out, 28
   leave-one-out, 28, 177, 184
curse of dimensionality, 2
Cy3, 52
Cy5, 52
D/S/A algorithm, 230
data mining, 2, 8
data re-scaling, 15
data transformation, 14
Daubechies wavelet, 87
DAVID, 134
dChip MBEI, 63
DDWT, see decimated discrete wavelet transform
decimated discrete wavelet transform, 86
deconvolution, 92
dendrogram, 10, 104
detector, 5
differential display, 4
discrete wavelet transform, 86
discriminant analysis, 41
distance metric, 135, 157
   City Block, see distance metric, Manhattan
correlation, 157
   Cosine, 157
   Euclidean, 105, 126, 157
      standardized, 157
   Hamming, 157
   Jaccard, 157
   Mahalanobis, 137, 157
   Manhattan, 112, 157
   Minkowski, 157
   Dunn-like indices, 131
dye-swap, 45, 47
EAM, see energy absorbing matrix
eGOn, 134
eigengene, 16
electrophoresis, 4
electrospray ionization, 5
embedded methods, 150, 155, 160
energy absorbing matrix, 79
entity recognition, 254
ER graph, see graph, Erdős-Rényi
test
   of prediction, 25
   rate, 24
   family-wise, 20
   comparison-wise, 20
error
   observed, 24-25, 181, 183
   true, 24-25, 181
   resubstitution estimate, 174
   selection, 177
split-sample estimate, 175
Type I, 18, 258
Type II, 19, 258
ESI, see electrospray ionization
EST, see expressed sequence tag
ETA, see experimental treatment assignment
experimental treatment assignment assumption, 235
expressed sequence tag, 5
F-measure, 258
false discovery rate, 20, 97, 134, 154, 233
false positive rate, 20
FatiGO, 134
FatiGOplus, 138
FDR, see false discovery rate feature, 7, 149
construction, 151
selection, 149–169, 190
recursive feature elimination, 190
filter, 150, 160, 169, 190
fingerprint, 7
Fisher score, 151, 152, 169
Fisher-like score, 18
FLD, see linear discriminant, Fisher forward selection, 16, 155
FPR, see false positive rate
FWER, see error rate, family-wise
Gaussian mixture, 158
GCRMA, 63
Gene Expression Omnibus, 53, 65, 69
Gene Ontology, 261
GeneChip, 5, 52
GenePix, 55, 60, 69
GeneSifter, 62
GeneSpring, 62
genetic algorithm, 150, 155
genomics, 1
functional, 2
GEO, see Gene Expression Omnibus
GO, see Gene Ontology
GoMiner, 134
goodness-of-fit, 30, 111
GOStat, 134
Gosurfer, 134
GOTM, 134
GOToolBox, 134
graph, 203, 206
directed, 205
Erdős-Rényi, 204, 207
k-scaffold, 211
random, 209
random modular graph, 204
scale-free, 204, 209
undirected, 206
weighted, 206
Graphviz, 219
heatmap, 104
Hidden Markov model, 125
high-throughput, 2
hill climbing, 150, 155
hyperplane, 200
maximal margin, 188
hyperplanes, 188
ICA, see independent component analysis
IE, see information extraction
iHOP, 252, 254, 260
Imagene, 55
Incogen, 99
independent component analysis, 16, 159
inference, 43
information extraction, 254, 268
retrieval, 253, 267
interquartile range, 59
inverse-probability-of-treatment-weighted transformation, 231
ion source, 5
IPTW, see inverse-probability-of-treatment-weighted transformation
IQR, see interquartile range
IR, see information retrieval
iterative signature algorithm, 131
J5-score, 152, 169
jackknife, 178
k-nearest neighbor, 26
k-NN, see k-nearest neighbor
Karhunen-Loève transform, see principal component analysis
kernel function, 189, 200

L1-metric, see distance metric, Manhattan

L2-metric, see distance metric, Euclidean

Lagrangian, 199

latent vectors, 16

learning

supervised, 40–41

unsupervised, 9, 40

learning by rote, 24

LIBSVM, 197

lift, 24

linear discriminant

Fisher, 159

loess, 60, 245

print-tip, 61

LOOCV, see cross-validation, leave-one-out

m/z, see mass-to-charge ratio

MA-plot, 60–61, 76

MAC, see maximum allowed absolute correlation

MALDI, see matrix-assisted laser desorption/ionization

margin, 198

Markov blanket filtering, 17

MAS 5.0, 64

mass analyzer, 5

mass spectrometry, 5, 79–99, 180

mass-to-charge ratio, 5, 80

MATLAB, 98

matrix-assisted laser desorption and ionization, 79

matrix-assisted laser desorption/ionization, 5

maximum allowed absolute correlation, 164

MDS, see multidimensional scaling

Medical Subject Headings, 252

Medline, 255

MeSH, see Medical Subject Headings

microarray, 5, 39, 51–76

cDNA, 53, 55

single-channel, 52

spotted cDNA array, 5

two-channel, 52, 59

microarray sample pool, 64

mismatch, 54

missing value handling, 13

MM, see mismatch model, 6

assessment, 173, 191

construction, 182

selection, 173, 181, 182

modifications

posttranslational, 1, 5

modularity, 203

Monte Carlo cross-validation, see sampling, repeated random subsampling

Monte Carlo permutation, 21

MS, see mass spectrometry

MSP, see microarray sample pool

MUDWT, 96, 97

multi-array probe-level model, 57

multidimensional scaling, 12, 105, 159

multiple hypotheses testing, 19

mutual information, 152, 169

mzXML, 99

natural language processing, 255

negative predictive value, 24

neighborhood divergence, 259

network, 203

   cellular, 204

      hierarchical, 210

   ii, 206

   motif, 211

NLP, see natural language processing

No Free Lunch theorem, 10
	node, 207

   average degree, 208

   degree, 207

   indegree, 207

   outdegree, 207

   normalization, 51, 56

   between-slide, 61

   loess, 61

   of mass spectra, 82

   print-tip loess, 61, 76

   quantile, 57, 61, 63, 75

   variance stabilization, 64

   within-slide, 61

NPV, see negative predictive value

nuisance parameter, 229
NUSE, see standard error, normalized unscaled

Occam's razor, 10
one-versus-all, 18
one-way analysis of variance, 18
Onto-Express, 134
overfitting, 23-25, 150

Pajek, 219
partial least squares, 16, 159
path, 209
  average length, 209
PCA, see principal component analysis
PCR, see polymerase chain reaction
peak
  detection, 82
  matching, 82
  quantification, 82
peptide/protein chips, 6
perceptual mapping, see multidimensional scaling
perfect match, 53
phage display, 6
phase
  application, 22
  learning, 22
  test, 22
  training, 22
  validation, 22
PLIER, 64
PLM, see multi-array probe-level model
PLS, see partial least squares
PM, see perfect match
polysemny, 256
population, 42
positive predictive value, 24
PPV, see positive predictive value
pre-processing, 13
precision, 258
predictor, 7
prevalence, 24, 97
principal component analysis, 12, 15, 108, 159
probe, 6
probe set, 53
  spike-in probe set, 58
probes, 53

PROcess, 98, 114
profile, 7
  array, 7
  gene expression, 7
  protein expression, 7
projection pursuit, 113
ProteinChip, 89, 93, 114
proteomics, 1
PubMed, 252
qRT-PCR, 3, see quantitative real-time reverse transcriptase PCR
QT-Clust, 136
quantitative real-time reverse transcriptase PCR, 3
randomization, 46
recall, 258
receiver operating characteristic, 169
reference design, 45
reference RNA, 45
regression, 10
  least angle, 121
  regularized logistic, 166
regularization, 155
relative log expression, 59, 76
replicate
  biological, 43
  technical, 43
replication, 42
reverse transcriptase, 3
ribonuclease, 3
ribonuclease protection assay, 3
RLE, see relative log expression
RLR, see regression, regularized logistic
RMA, see robust multi-chip analysis
robust multi-chip analysis, 56-57, 62, 74
ROC, see receiver operating characteristic
RPA, see ribonuclease protection assay
RProteomics, 98
S+Analyzer, 62
S2N, see signal-to-noise
SAGE, see serial analysis of gene expression
SAM, see significance analysis of microarrays
SAM scoring criterion, 152, 169
SAMBA, 131
Sammon mapping, 111
sample, 1, 42
sampling, 173–185, 196
  bootstrapping, see bootstrap
  k-fold random subsampling, 28
  random subsampling, 27
  repeated random subsampling, 178
  single hold-out method, 27
  split-sample, 175
  two-fold nested resampling, 181
Savitzky-Golay, 90
scale-freeness, 203
scaling
  metric, 107
  nonmetric, 109
ScanAlyze, 55
segmentation, 55
SELDI-TOF, see surface-enhanced laser desorption/ionization
time-of-flight
self-organizing maps, 113, 126, 130
self-organizing tree algorithm, 126, 128, 130
sensitivity, 24, 97
serial analysis of gene expression, 4
set
  learning, 25, 176
  test, 25, 176
  training, 191
  validation, 25, 181, 191
shrunken centroid classifier, 181
signal-to-noise, 18, 59, 82, 90, 190
significance analysis of microarrays, 154
silhouette coefficient, 131
simulated annealing, 150, 155
singular value decomposition, 15
SiZer plot, 88
small-n-large-p problem, 2
small-world pattern, 203, 209
SOMs, see self-organizing map
SOTA, see self-organizing tree algorithm
specificity, 24
spectrum, 7
Spot, 55
SSH, see suppression subtractive hybridization
  standard error
normalized unscaled, 58, 75
stress
  function, 106
  squared, 109
  weighted, 108
study
  experimental, 40
  observational, 40
subtractive hybridization, 4
SUDWT, 96–97
summarization, 53
support vector machine, 11, 156, 187–200
suppression subtractive hybridization, 4
surface-enhanced laser desorption/ionization time-of-flight, 6, 79, 104, 161, 194
SVD, see singular value decomposition
SVM, see support vector machine
SVMLight, 197
SW pattern, see small-world pattern
synonymy, 256
t-statistic, 17, 169
tag, 4
target, 6
test
  Anderson-Darling, 11
  ANOVA, 18
  Bartlett, 17
  Benjamini and Hochberg, 21, 233
  Brown and Forsythe, 19
  Cochran, 19
  Duncan, 19
  Dunnett, 19
  F-test, 18
  Hochberg, 21
  Holm, 20, 134
  Kruskal-Wallis, 19
  Levene, 17
  McNemar, 30
  post-hoc, 19
  random permutation, 21, 152, 153, 183
  Storey and Tibshirani, 21
  Student, 19
t-test, 17, 152, 169
  Tukey, 19
  variance-corrected resampled, 30–31
Welch, 19
Wilcoxon rank-sum, 152
testing, 25, 182
text mining, 32, 251–270
  full text mining, 257
TIC, see total ion current
time resolution, 79
time series analysis, 227–247
time-of-flight, 5, 79
TOF, see time-of-flight
topological overlap analysis, 211
total ion current, 89
training, 25, 182
transcriptomics, 2
truly alternative, 20
truly null, 20
two-dimensional difference in-gel electrophoresis, 5
two-dimensional polyacrylamide gel electrophoresis, 5
UCSF Spot, 55
UDWT, see undecimated discrete wavelet transform
undecimated discrete wavelet transform, 83
validating, 25
validation, 182
variance, 22
VSN, see normalization, variance stabilization
Welch-Satterthwaite, 17
Wolfe dual, 199
wrapper, 150, 155, 160, 190
yeast two-hybrid, 6
z-score transformation, 15