

---

# Index

- $\alpha$  helix, 7
- $\beta$  sheet, 8
- 2D gel electrophoresis, 26
- 3'-end, 4
- 5'-end, 4
  
- affinity, 32
- alignment, 22
- alternative splicing, 16
- amino acid, 7
- apoptosis, 17
- archaea, 4
  
- Bayes point machines, 58
- binding pocket, 8
- biological network, 26
  
- C-terminus, 7
- cancer diagnosis, 311
- canonical correlation analysis, 212
  - integrated, 216
  - kernel, 89, 186, 212
  - multiple, 215
  - regularization, 214
- CCA, *see* canonical correlation analysis
- cDNA, *see* complementary DNA
- cell, 3
- chaperone, 7, 17
- chemical compound, 165
- cheminformatics, 32
- chemoinformatics, 32
- chromatin, 5
- chromosome, 5
- codon, 15
- compartment, 12
- cross-validation, 62
  
- curse of dimensionality, 300, 322
  
- data integration, 210
- diagnosis, 30
- diffusion, 174
- diploid, 5
- DNA, 4
  - complementary, 5
- docking, 32
- downstream, 5
- drug design, 323, 354
  
- EM algorithm, 90, 266, 309
- empirical kernel map, 69, 141, 250
- empirical risk minimization, 52
- enzyme, 9, 11, 183
- EST, *see* expressed sequence tag
- eubacteria, 4
- eukarya, 4
- exon, 15, 281
- expressed sequence tag, 24
  
- FDR, *see* Fisher discriminant ratio
- feature construction, 324
- feature selection, 300, 324
  - backward elimination, 304
  - embedded, 305
  - filter, 301, 326
  - forward selection, 304
  - lasso, 305
  - R2W2, 328
  - ranking, 326
  - recursive feature elimination, 304
  - wrapper, 303, 326
  - zero-norm minimization, 305
- feature space, 40
- Fisher discriminant ratio, 301

- fold, 7
- function prediction, 32, 255
- gene, 6, 280
  - expression, 16, 184, 221
  - finding, 29
  - network, 183
  - selection, 299, 322
- genetic marker, 30
- genetic network, 32
- genome, 4
  - comparison, 29
- genotype, 18
- graph
  - $p$ -regular tree, 178
  - closed chain, 178
  - complete, 177
  - diffusion, 175
  - labeled, 157
  - Laplacian, 175, 219
  - product, 179
- grid search, 62
- haploid, 5
- hidden Markov model, 23
  - pair HMM, 136
- hinge loss, 52
- histone, 5
- HMM, *see* hidden Markov model
- homology, 19
- horizontal gene transfer, 19
- hybridization, 5, 23
- hydropathy profile, 249
- hydrophobicity, 10
- induced fit, 32
- information geometry, 264
- interactions, 26
- intron, 15, 281
- kernel
  - alignment, 198
  - Bhattacharya, 169
  - combination, 226
  - completion, 262
  - composition, 74
  - conditionally positive definite, 67
  - convolution, 135
  - diagonal dominance, 139
  - diffusion, 175, 219, 251, 256
  - expected likelihood, 169
  - exponential, 168
  - FFT, 250
  - Fisher, 65, 72
  - gappy, 99
  - Gaussian RBF, 41, 63, 174
  - geometric, 168
  - heat, 175
  - label sequence, 158
  - linear, 39, 63
  - local alignment, 137
  - locality improved, 286
  - marginalized, 74, 157
  - metric space, 189
  - mismatch, 78, 98
  - motif, 75
  - normalization, 61
  - on groups, 67
  - operations, 67
  - pairwise comparison, 75, 250, 256, 288
  - Pfam, 250
  - polynomial, 63, 285
  - positive definite, 38
  - probability product, 169
  - rational, 169
  - Riemannian manifold, 189
  - Schur product, 67
  - sigmoid, 64
  - spectral translation, 141
  - spectrum, 77, 98, 288
  - string, 64, 115
  - substitution, 100
  - translation-invariant, 67
  - tree, 116
  - weighted degree, 287
  - wildcard, 101
- kernel methods, 35
- kernel trick, 44

- kinase, 11
- Laplacian prior, 308
- leave-one-out machine, 58
- ligand, 8, 32
- linear discriminant analysis, 303
- linear programming machine, 59
- LP machine, *see* linear programming machine
- MAR, *see* matrix attachment region
- margin, 52
- Markov model, 284
- Markov random field, 234
- mass spectrometry, 26
- matrix attachment region, 5
- membrane, 10
- metabolic pathway, 26, 183, 221
- metabolite, 9, 12
- microarray, 23, 83, 184, 251, 320
- mismatch tree, 101
- mitochondria, 4
- model organism, 27
- molecular interactions, 26
- motif, 7
- mRNA, *see* messenger RNA
- MS, *see* mass spectrometry
- mutation, 18
- N-terminus, 7
- northern blotting, 25
- novelty detection, 58
- nucleosome, 5
- nucleotide, 4
- nucleus, 4
- oligonucleotides, 5
- oligopeptide, 7
- one-against-all, 61
- operon, 220
- organelle, 12
- ortholog, 19
- pairwise data, 354
- paralog, 19
- pattern discovery algorithm, 165
- pattern recognition, 51
- PCA, *see* principal component analysis
- PCR, *see* polymerase chain reaction
- penalized kernel logistic regression, 305
- peptide, 7
- peptide bonds, 7
- pharmacophore, 33
- phenotype, 18
- phylogenetic profile, 80
- point mutation, 5
- poly(A), 15
- polymerase chain reaction, 21
  - quantitative, 25
- polymorphism, 18
- polypeptide, 7
- polyploid, 5
- population genetics, 30
- position-specific scoring matrix, 23
- posttranslational modification, 15
- predictive toxicology, 33, 166
- principal component analysis, 48, 303
  - kernel PCA, 50, 218
- prokarya, 4
- promoter, 79, 281
- protein
  - backbone, 7
  - classification, 268
  - domain, 8
  - expression, 16
  - function, 8
  - membrane, 248
  - primary structure, 7
  - secondary structure, 7, 194
  - sidechain, 7
  - structural, 11
  - tertiary structure, 7, 8
- protein-protein interaction, 91
- PSSM, *see* position-specific scoring matrix

- qPCR, *see* quantitative polymerase chain reaction
- QSAR, *see* quantitative structure-activity relationships
- quantitative structure-activity relationships, 32
  
- random walk, 158, 173
- reading frame, 15
- receptor, 11, 18
- recursive feature elimination, 87, 328
- regularization, 43
  - with diffusion, 180
- regulation, 16, 29
- regulatory interactions, 26
- relevance vector machine, 60, 306
- remote homology detection, 72, 142
- representer theorem, 47
- reproducing kernel Hilbert space, 43
- residue, 7
- ribosome, 10
- ribozyme, 6
- RNA, 5
  - functional, 82
  - messenger, 6, 11, 15
  - ribosomal, 10, 11
  - transfer, 11
- rRNA, *see* ribosomal RNA
  
- saccharide, 10
- SAGE, *see* serial analysis of gene expression
- SDP, *see* semidefinite programming
- secondary structure, 82
- selection, 18
- semidefinite programming, 89, 236
- sequence
  - alignment, 133
  - local alignment, 134
  - spectrum, 96
- sequencing, 21
- sequential minimal optimization, 336
- serial analysis of gene expression, 24
- shortest-path distance, 172
  
- signal molecule, 12, 17
- signal peptide cleavage sites, 81
- single nucleotide polymorphism, 18
- Smith-Waterman score, 134, 250, 256, 288
- SNP, *see* single nucleotide polymorphism
- sparse logistic regression, 306
- sparse probit regression, 306
- spectral clustering, 218
- spectral variants, 263
- splice sites, 81, 282
- spliceosome, 282
- splicing, 15, 281
- structure prediction, 31
- subcellular localization, 14, 80
- substitution matrix, 200
- subunit, 10
- suffix tree, 119
- support vector, 57
- support vector machine, 50, 328
  - $\nu$ -SVM, 59
  - 0-Norm, 328
  - implementations, 60
  - multiclass, 60, 195
  - one-norm, 306
  - potential, 335
  - sphered, 332
- SVM, *see* support vector machine
- systems biology, 30
  
- target finding, 30
- therapy optimization, 30
- transcription, 15
  - factor, 11, 15
- translation, 15
  - start sites, 81
- translocation, 15
- transposon, 19
- tRNA, *see* transfer RNA
  
- upstream, 4
  
- virtual high-throughput screening, 32