

TOPIC INDEX

- absorption bands, 188
- activation function, 11
- adulteration in olive oil, 59
- Akaike's Information Criteria (AIC), 42, 61
- alfalfa, 191
- all subsets, 175
- apparent error rate, 54
- architecture of an MLP, 70
- artificial neural networks, 1
- aspect angle dependent, 200
- asymptotic efficiency, 168
- atmospheric absorption, 187
- AVIRIS instrument, 187
- band interleaved by line, 115
- basis function expansion, 185
- Bayes error rate, 54
- Bayes' formula, 94
- Bayesian decision boundary, 37
- Bayesian Information Criteria BIC, 61
- Beale restarts, 256
- Bernoulli random variable, 39
- bias, 54
- bias and variance, 35
- binomial random variable, 39
- biological plausibility, 1
- black boxes, 3
- Boltzmann machine, 100
- bootstrap, 55
- boundary conditions, 182
- breakdown point, 170
- B-spline basis functions, 180
- B-splines, 185
- canonical link, 39
- canonical variate analysis, 20, 183
- centering the data, 203
- central limit theorems, 171
- Charles Darwin, 22
- Cholesky decomposition, 115
- classification trees, 71
- column space, 20
- compact support, 185
- condition number, 62
- confirmatory models, 54
- conjugate directions, 251
- conjugate gradient, 246, 250
- connectionism, 1
- consistency, 37
- consistency of weight decay, 203
- consistent bias, 205
- contingency tables, 42
- convergence, 33
- Cook's distance, 175
- cooling schedule, 97
- correlated features, 23
- covariance function, 180
- covariance matrix, 20
 - between classes, 20
 - within classes, 20
- cross-entropy penalty function, 12
- cross validation, 55

- cumulative probability distribution, 40
- curse of dimensionality, 58
- Cushing's syndrome, 175
- data are smooth curves, 184
- decision boundary, 15, 206
- degrees of freedom, 59
- determinant, 20
- diagnostic matrix T , 72
- diagnostic procedures, 172
- difference operator, 183
- Dirac delta function, 36, 181
- discrepancy, 126
- distribution-free, 122
- double precision, 149
- early stopping, 65
- eigen decomposition, 20
- eigenfunctions, 180
- eigenvalues, 180
- eigenvectors, 180
- ellipsoid classes, 169
- end-member materials, 189
- energy function, 99
- entropy, 46
- error surface, 144
- estimator, 122
- estimator (0.632), 55
- exponential family, 249
- exponential family of distributions, 39
- false color image, 76
- fanning into, 11
- fanning out, 11
- Fisher efficiency, 47
- Fisher K. A., 19
- Fisher's information matrix, 60, 249
- fixed-slope training, 220
- Fletcher-Reeves updating formula, 252
- forward difference operator, 183
- functional CVA, 183
- functional data analysis, 179
- functional MLP model, 184
- fraction C^* , 184
- fuzzy set membership function, 169
- Gâteaux differentiable, 123
- generalization, 53
- generalized eigen-decomposition, 20
- generalized linear models, 39
- Gebs random field, 98
- global maximum, 97
- gradient information, 205
- Gram-Schmidt algorithm, 251
- graphically selecting initial decision boundaries, 208
- gross error sensitivity, 126
- ground truth, 189
- Hampel estimators, 126
- Hampel's redescending tanh estimator, 160
- Hebbian learning rule, 160
- Hermitian matrix, 183
- Hessian matrix, 246
- hidden layer, 10-11
- hierarchical cluster analysis, 73
- Hilbert matrix, 90
- Hopfield network, 93
- Huber estimators, 126
- HyMap instrument, 183, 187
- hyper-spectral image, 160
- identity output function, 11
- ill-posed, 63
- illusion of progress, 7
- improper prior, 184
- independent test set, 53
- influence curve, 124
- influence matrix, 175
- influence of the position in feature space, 128
- influence of the residual, 125
- input layer, 11
- input units, 11
- integral equation, 181
- inverse transform, 205
- Ising spin glass model, 100
- a posteriori probability*, 36
- iterative conditional modes, 97
- iterative fitting procedure, 144
- iteratively re-weighted least squares, 40
- iteratively re-weighted least squares, 250
- iterative updating procedure, 93
- kernel density estimation, 185
- Kronecker delta, 36
- Kronecker product, 115
- Kullback-Liebler divergence, 46, 61
- Landsat MSS and TM data, 94
- LDA linear discriminant analysis, 19
- least Gaussian projection, 171
- leave-one-out estimator, 55
- Leptograpsus variegatus* - crabs data set, 22
- likelihood, 40
- linear approximation, 15
- linear discriminants, 20
- linear predictors, 159
- linear transformation, 203
- line search routine, 144, 254
- local minimum, 203
- local shift sensitivity, 126
- location and scale estimators, 169
- logistic activation function, 11
- logistic regression, 39, 159
- logistic sigmoid, 11
- log-linear models, 42
- low-dimensional projections, 171
- machine learning, 1
- magnitude of the norm of the weight vector, 219
- Mahalanobis distance, 34, 94, 146, 169
- Mallow's C_p , 61
- Markov random field, 93
- masked by other points, 148
- masking effect, 170
- MASS library, 18, 30, 50, 92, 177
- McNemar's test, 57
- mean field approximation, 107
- mean square classifier, 26
- median, 168

- median absolute deviation (mad), 168
- M-estimator, 124
- method of scoring, 249
- mineral mapping, 189
- minimization techniques, 245
- Minkowski penalty function, 160
- MLP notation, 11
- model diagnostics, 159
- morphological variables, 22
- most interesting projection, 171
- multinomial model, 41
- multiple minima, 143
- multi-spectral image, 69
- natural pairing, 45
- nearest neighbor, 191
- neighboring pixels, 93
- neighbor MLP model, 103
- neighbor sub-network, 103
- network information criteria - NIC, 61
- Newton-Raphson method, 247
- Newton's method, 245
- non-smooth optimization, 100
- normalizing the data, 203
- normal vector, 148
- numerical accuracy, 250
- numerical precision, 149
- objective function, 36
- one of Q encoding, 12
- one of Q encoding, 36
- order statistics, 160
- orthogonal search directions, 250
- output units, 11
- overfitting, 53
- over-parameterized, 59
- overtraining, 164
- parameter, 122
- parametric, 122
- partially receding IC function, 146
- partition function, 99
- penalized discriminant analysis, 23, 179, 183
- penalized ridge, 63
- penalty matrix, 201
- perturbations, 143
- pixels, 76
- plant spectra, 188
- Poisson distribution, 42
- Folak-Ribiere algorithm, 256
- polar coordinates, 220
- posterior probabilities, 15, 35
- potato, 191
- Predicted Squared Error, 61
- pre-processing steps, 203
- principal components, 180-181
- probabilistic basis function network, 185
- probit model, 40
- projection matrix, 20
- projection-pursuit regression, 41
- projection pursuit regression, 171
- pruning, 71
- pruning schedule, 69
- QDA quadratic discrimination analysis, 25
- QR decomposition, 73
- quadratic form, 250
- quasi-Newton methods, 250
- quickprop, 233
- radial basis function network, 185
- radially symmetric basis functions, 185
- rda library, 201
- region of active influence, 146
- region of attraction, 212
- regularized discriminant analysis, 116
- rejection point, 127
- relative entropy, 46
- remotely sensed data, 76
- residuals, 160
- resistant, 122, 159
- resubstitution error rate, 54
- ridge regression, 62, 220
- robust, 122
- robust logistic regression, 162
- robustness of efficiency, 122
- rocks versus mines, 200
- salinity scanner, 187
- sampling a smooth function, 179
- San Luis Valley, 190
- scaled conjugate gradient, 246
- scaled residual, 168
- scaling the data, 203
- second derivative operator, 182
- self-adjoint matrix, 183
- semi-classified image, 101
- semi-classified image, 107
- sensitivity, 143
- sensitivity curve - SC, 143
- sensitivity peak, 151
- shape of the error surface, 211
- sigmoid, 13
- signal regression, 179
- simplex algorithm, 212, 254
- simulated scene, 114, 223
- single layer-learning network, 70
- single-precision arithmetic, 149
- singular covariance matrices, 187
- smoothed canonical variate analysis, 179
- smoothed discriminant analyses, 134
- smoothed error rate, 56
- smoothness penalty, 179
- smoothness penalty term, 186
- softmax activation function, 11
- sonar signals, 209
- spatial correlation, 94
- spatial dependency, 94
- spectral measurements, 94
- spectroscopy, 187
- squashing activation functions, 37
- starting weights, 70, 143
- STATLOG project, 19
- steepest descent, 245
- studentized residuals, 172
- sum of squares penalty function, 12
- super-smoother, 191
- surrogate Poisson model, 42

- SVD singular value decomposition, 21
- target matrix, 11
- tasks by units list, 74
- Taylor series, 246
- template fitting technique, 190
- tessellated cap, 188
- test for redundant units, 60
- texture information, 101
- Tikhonov regularization, 63, 186
- TM wavelengths, 77
- tolerance, 164
- tolerance distribution, 40
- truncated Gaussians, 185
- Tukey's biweight function, 160
- typicality probability, 34, 92
- unbiased estimators, 38
- uniform L^2 convergence, 37
- units by tasks list, 74
- universal approximators, 37
- unnixing, 189
- validation set, 65
- variable selection, 184
- vegetation mapping, 189
- Wald test, 60
- wavelengths, 76
- weak continuity, 123
- weight decay, 64, 184
- weight elimination, 64, 184
- weighting procedure, 166
- weight simplification, 110
- white noise, 184
- XOR problems, 5, 253
- zero influence, 205