

Subject Index

- ϵ -neighborhood, 243
- λ -gradient, 151
- τ_ϵ -approximation, 197
- k -means clustering, 12, 70, 104, 145, 150, 245, 262
 - Fuzzy, 70
 - Harmonic, 94
 - Kernel, 70, 79
- k -nearest neighbors, 70, 313
- .NET, 49

- AdaBoost, 150
- Alternating least squares (ALS), 334
- Amdahl's law, 453
- Application concurrency, 450
- Application programming interface (API), 69, 171
- Application-specific integrated circuit (ASIC), 16, 411
- Apriori algorithm, 83, 424
- Area under a curve (AUC), 363
- Arithmetic logic unit (ALU), 91
- Arnoldi factorization, 244
- ARPACK, 244
- Asynchronous distributed learning, 222
- Atomic operations, 199, 200
- Average filling (AF), 340

- Backpropagation, 299
- Backward elimination, 362
- Balanced minimum cut, 211
- Bandwidth limit, 285
- Basic linear algebra subprograms (BLAS), 7, 133
- Bayesian network, 14, 231

- Belief propagation (BP), 14, 190
 - Loopy, 190
 - Round-Robin, 199
 - Splash, 202
 - Synchronous, 195
 - Wildfire, 200
- Belief residuals, 200
- Bioinformatics, 421
- Blue Gene, 77
- Boosted decision tree, 12, 14, 25, 149
- Boosted regression tree, 152
 - Data-distributed, 155
 - Feature-distributed, 154

- Cache miss, 285
- Candidate generation, 423
- Chip multiprocessing (CMP), 421
- Cholesky decomposition, 63, 180
- Classification, 3, 128
 - Maximum margin, 128
- Classification and regression tree (CART) algorithm, 385
- Cloud computing, 1
- Clustering, 3
- Co-clustering, 263
 - Information-theoretic, 15, 265
 - Sequential, 15, 264
- Coalesced memory access, 96, 379, 460
- Coalitional classification, 343
- Collaborative filtering, 332
 - Cross system, 333
- Commutative monoid, 70
- Computational complexity, 10

- Compute unified device architecture (CUDA), 12, 90, 377, 400, 459
- Computer vision, 16
- Conditional probability table (CPT), 232
- Conditional random field (CRF), 362
- Conjugate gradient, 301
- Cooperative matrix factorization (CoMF), 333
 - Supervised, 345
- Cross-correlation, 379, 382

- Data parallelism, 6, 377, 434
 - Coarse-grained, 7
 - Fine-grained, 7, 13, 89
- Data sparsity problem, 331
- Data weaving, 264
- Data-flow computing, 407
- Datacenter, 1
- DataLoom, 264
- Dataset
 - 20 Newsgroups, 251, 275, 348
 - AdCorpus, 41
 - Adult census, 139, 185
 - California housing, 185
 - Caltech-101, 403
 - CoverType, 122
 - Cslogs, 437
 - Forest, 139
 - LabelMe, 375
 - Medline, 224
 - MNIST, 138, 405
 - Netflix, 276, 339
 - NORB, 138, 389, 416
 - PASCAL competition, 82
 - PicasaWeb, 251
 - RCV1, 122, 251, 276, 285, 365
 - Spambase, 185
 - Switchboard I (SWB), 317
 - Treebank, 427
 - UCI repository, 224, 364
 - UMASS, 416
 - Wikipedia, 224
- Decision tree, 12, 59, 70, 384
 - Bagging, 12, 28
 - Boosting, 28
- Deep belief network (DBN), 16, 373
- Delayed updates, 288
- Digital signal processor (DSP), 137
- Direct memory access (DMA), 413
- Distributed coalitional learning (DisCo), 333
- Double data computation, 144
- Dryad runtime, 49
- DryadLINQ, 12, 49
- Dynamic scheduling, 194

- Eclat algorithm, 84
- Eigendecomposition, 243, 311
 - Parallel, 248
- Embarrassingly parallel algorithm, 6
- Embedded subtree, 423
- Embedding list (EL), 424
- Ensemble learning, 28
- Entropy, 315
- Expectation-maximization (EM), 94, 308

- Factor graph, 191
- Factorized distribution, 191
- Feature class, 354, 356
- Feature hashing, 286
- Feature pooling layer, 401
- Feature selection, 16, 354
 - Filter, 354
 - Forward, 16, 355
 - Wrapper, 354
- Feature sharding, 287
- Field-programmable gate array (FPGA), 1, 128, 284, 407
- Filter bank layer, 400
- Forward-backward sampler, 234
- Forward-backward schedule, 193, 197
- Frequent pattern mining, 16, 70, 83
 - Subtree, 17, 421

- Gaussian mixture model (GMM), 455
- Generalized boosted models (GBM), 38
- Gentle boost algorithm, 384
- Gibbs sampling, 14, 218, 391
 - Collapsed, 218
- Gini coefficient, 385
- Global memory, 91
- Global training, 298
- Gomory-Hu tree, 324
- Google file system (GFS), 246
- Gradient boosting (GB), 152, 172
- Gradient descent, 131, 357
 - Delayed, 288
 - Minibatch, 300
 - Online, 283
 - Stochastic, 283
- Grafting, 357
 - Parallel, 361
- Graph partitioning, 211
- Graphical model, 14, 190
- Graphics processing unit (GPU), 1, 89, 127, 229, 284, 373, 377, 459
 - General purpose, 90

- Hadoop, 18, 23, 69, 195
- Hardware description language (HDL), 411

SUBJECT INDEX

473

- Hessian matrix, 111
 Heteroscedastic linear discriminant analysis (HLDA), 462
 Hidden Markov model (HMM), 14, 233, 310, 447
 Hierarchical Dirichlet process (HDP), 14, 220
 High performance computing (HPC), 1, 77, 186
 Histogram of oriented gradient (HoG), 388, 405

 Image denoising problem, 192
 Incomplete Cholesky factorization (ICF), 13, 112
 Independent matrix factorization (IMF), 340
 Induced subtree, 423
 Induction, 2
 Inference, 3
 Approximate, 192
 Distributed, 220
 Information bottleneck (IB), 265
 Sequential, 269
 Information retrieval, 14, 151
 Information-theoretic clustering (ITC), 263
 Instance sharding, 286
 Interior point method (IPM), 13, 111
 Iterative conditional modes (ICM), 266

 Jacket software, 105
 Job descriptor, 432
 Job pool (JP), 432

 Karush-Kuhn-Tucker conditions (KKT), 127, 315
 kd-tree, 319
 Kernel acceleration, 141
 Kernel function, 79, 111
 Kernel matrix, 111
 Kinect, 66
 Kolmogorov's superposition theorem, 176
 Kullback-Leibler divergence, 312

 LambdaMART, 148, 151
 Data-distributed, 155
 Feature-distributed, 154
 LambdaRank, 151
 Language integrated queries (LINQ), 49
 Laplacian matrix, 240
 Normalized, 241
 LASSO, 362
 Latent Dirichlet allocation (LDA), 14, 19, 218, 276
 Approximate distributed, 220
 Latent semantic analysis (LSA), 218
 Probabilistic, 332
 Latent variable model, 217
 Learning rate, 283
 Linear regression, 14, 70
 Linear regression tree (LRT), 178
 Load balancing, 17, 190, 422, 431
 Over-partitioning, 212
 Partitioning, 211
 Local linear embedding (LLE), 312
 Local training, 294
 Logistic regression, 16, 353
 Longest common subsequence (LCS), 428
 Loss function, 171, 357
 Hinge, 172, 311
 Smoothed hinge, 345
 Squared, 25, 171, 311
 Low-level features, 375

 Machine translation, 5
 Manycore processor, 446
 MapReduce, 7, 12, 23, 56, 69, 94, 150, 195, 240, 332, 358, 453
 Map phase, 25
 Reduce phase, 25
 Shuffle phase, 25
 Markov blanket, 232
 Markov chain Monte Carlo (MCMC), 13, 218
 Markov logic network (MLN), 192
 Markov random field (MRF), 192
 MART, 152, 184
 Master-slave communication model, 6, 271
 Matlab, 49, 105, 236
 Max pooling, 382
 Maximum likelihood estimation, 353
 Measure propagation, 15, 312
 Memory bandwidth, 422
 Memory-conscious Trips (MCT), 427
 Message locking, 199
 Message passing interface (MPI), 7, 77, 158, 186, 213, 229, 240, 275, 335
 Micro parallelization, 136
 Micro-averaged accuracy, 275
 MiLDe, 130, 136
 MineBench, 102
 Minimum phone error (MPE), 462
 Model-based clustering, 99
 MPICH2, 77, 213, 246
 Multi-modal clustering, 263
 Multicore parallelization, 290, 454
 Multicore processor, 290, 408, 422, 446

- Multimedia extension instructions (MMX), 134
- Multinode parallelization, 291
- Multiple-data multiple-thread (MDMT), 90
- Mutual information (MI), 265
- Naïve Bayes, 294
- Natural language processing (NLP), 17, 192
- Neocognitron, 403
- Neural network, 176
 - Convolutional, 16, 145, 400
- Nonlinearity layer, 401
- Nonparametric model, 220
- Normalized discounted cumulative gain (NDCG), 151, 158
- Normalized mutual information (NMI), 252
- NVIDIA, 90, 102
- Nyström method, 243
- Object recognition, 404
- Occurrence-based support, 423
- Online learning, 15, 283
- OpenCL, 92
- OpenCV, 387
- OpenGL, 92
- OpenMP, 229, 451
- Optical character recognition (OCR), 402
- Overfitting, 2, 44
- Parallel machine learning toolbox (PML), 12, 69, 171
- Parallel spectral clustering (PSC), 241
- Parallel support vector machine (PSVM), 109
- PARPACK, 248
- Pattern growth mechanism, 424
- Peer-to-peer (P2P) networks, 18
- Perceptron, 283
 - Multi-layered, 316
- Perceptual linear prediction (PLP), 462
- Perplexity, 225
- PLANET, 29
- Predictive sparse decomposition (PSD), 403
- Principal component analysis (PCA), 70, 218
 - Kernel, 70
- ProbE data mining engine, 170
- Processing tile (PT), 411
- Protein side-chain prediction, 205
- Pyramid match kernel (PMK), 405
- Quadratic programming (QP), 111, 127
- Radial basis function (RBF), 79, 139, 313
- Real-time performance, 16
- Recommender system, 16, 331
- Recursive doubling algorithm, 249
- Regression, 3
- Regression clustering, 99
- Regression k -means, 12
- Regression tree, 27, 152, 170
- Remote procedure call (RPC), 40
- Representation power, 293
- Reproducing kernel Hilbert space (RKHS), 110
- Residual scheduling, 201
- Restricted Boltzmann machine (RBM), 390, 403
- Root mean square error (RMSE), 341
- Rooted tree, 423
- Round-robin tournament, 273
- Scale invariant feature transform (SIFT), 375, 399
- Self-training, 308
- Semi-supervised learning, 3, 15, 307
 - Graph-based, 15, 308
- Sequential clustering, 263
- Sequential minimal optimization (SMO), 13, 111, 130
- Share no write variables (SNWV), 93
- Shared cache memory, 380
- Shared memory, 91
- Shared memory platform (SMP), 77
- Shared nothing architecture, 84
- Sherman-Morrison-Woodbury formula (SMW), 114
- Single feature optimization (SFO), 16, 355
- Single-instruction multiple-data (SIMD), 17, 134, 460
- Singular value decomposition (SVD), 62
- Sliding window object detection, 383
- SLIQ, 44, 81
- Sparsification, 240
- Spectral clustering, 14, 240
- Spectral graph transduction (SGT), 311
- Speech recognition, 15, 17, 446
- Splash algorithm, 14, 202
 - Multicore, 206
 - Sequential, 205
- Spread kernel, 134
- Stream processor (SP), 378
- Streaming SIMD extensions (SSE), 134
- Strong scalability, 186
- Subtree matching, 428
- Sufficient statistics, 97
- Supervised learning, 2
- Support counting, 423

Cambridge University Press

978-0-521-19224-8 - Scaling Up Machine Learning: Parallel and Distributed Approaches

Edited by Ron Bekkerman, Mikhail Bilenko and John Langford

Index

[More information](#)

SUBJECT INDEX

475

- Support vector machine (SVM), 13, 18, 69,
 - 109, 127
 - Dual, 129
 - Laplacian, 311
 - Primal, 129
 - Transductive, 308
- Switchboard transcription project (STP), 317
- Symmetric multiprocessing (SMP), 45, 309

- Task parallelism, 7
 - Adaptive, 433
- Task scheduling service, 436
- Text classification, 16
- Thread blocks, 378
- Thread pool (TP), 432
- Topic model, 14, 217
- Transaction-based support, 423
- Transductive learning, 307
- Transfer learning, 15, 331, 343

- Transform regression, 14, 170
- Tree architecture, 292
- TreeMiner, 424
- Trips algorithm, 425

- Unmanned aerial vehicles (UAV), 399
- Unsupervised learning, 3

- Variational inference, 230
- Vector processing, 377
- Viterbi algorithm, 448
- Vowpal Wabbit software, 285

- Weak scalability, 186
- Web search ranking, 24, 148
- Weighted finite state transducer (WFST), 448
- Work sharing, 431
- Working set selection, 130
- Working set size, 422