List of Recommended Articles

Articles are sorted chronologically rather than by topics. They should be available on the Web or via the library – ask for a copy only if you really cannot access a paper.

- C.A.R. Hoare: An Axiomatic Basis for Computer Programming. Communications of the ACM 12 (10) 576–580 (1969) – About proving properties (correctness etc.) of programs by the axiomatic method. As the paper is old, a review should relate the paper to more recent developments.
- B.H. Bloom: Space/Time Trade-offs in Hash Coding with Allowable Errors. Communications of the ACM 13 (7) 422–426 (1970) – Error-tolerant hashing with a sample application. The method became popular as Bloom filter and has been further developed and applied.
- C.L. Liu, J.W. Layland: Scheduling Algorithms for Multiprogramming in a Hard-Real-Time Environment. Journal of the ACM 20 (1) 46–61 (1973) – Despite the age of this work, topics are not outdated, and much research has been done in the field since then.
- J. Ziv, Ab. Lempel: A Universal Algorithm for Sequential Data Compression. IEEE Transactions on Information Theory 23 (3), 337–343 (1977) – The classical Ziv-Lempel compression.
- L.G. Valiant: A Theory of the Learnable. Communications of the ACM 27 (11) 1134–1142 (1984) – Pioneering paper in machine learning that introduced the concept known as "probably almost correct" (PAC) learning.
- J.R. Quinlan: Induction of Decision Trees. Machine Learning 1 (1) 81–106 (1986) – A classical approach in machine learning, for representing and inferring concepts.
- G. Salton, C. Buckley: Term-Weighting Approaches in Automatic Text Retrieval. Information Processing and Management 24 (5) 513–523 (1988). – An early summary of the performance of term weighting schemes for information retrieval from text.
- J.L. Elman: Finding Structure in Time. Cognitive Science 14 (2) 179–211 (1990) – About representing the dimension of time in neural networks, with applications to language processing.

- B.E. Boser, I. Guyon, Vl. Vapnik: A Training Algorithm for Optimal Margin Classifiers. 5th ACM Conf. on Computational Learning Theory COLT 1992, 144–152

- R. Agrawal, T. Imielinski, A.N. Swami: Mining Association Rules between Sets of Items in Large Databases. ACM SIGMOD Int. Conf. on Management of Data 1993, 207–216, and
  R. Agrawal, R. Srikant: Fast Algorithms for Mining Association Rules in Large Databases. Conf. on Very Large Data Bases VLDB 1994, 487–499 – How to discover relevant patterns and knowledge in relational databases.

- Y. Freund, R.E. Schapire: Experiments with a New Boosting Algorithm. Int. Conf. Machine Learning ICM 1996, 148–156 – Boosting is a general method to combine the strengths of different machine learning algorithms.

- H. Hoppe: Progressive Meshes. 23rd Conf. on Computer Graphics SIGGRAPH 1996, 99–108. Important concept in computer graphics and simulations.

- M. Ester, H.P. Kriegel, J. Sander, X. Xu: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. 2nd Int. Conf. on Knowledge Discovery and Data Mining KDD 1996, 226–231. – Introcuded a popular clustering algorithm called DBSCAN.

- P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman: Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Transactinos on Pattern Analysis and Machine Intelligence 19(7) 711–720 (1997) – Robust face recognition via linear algebra.

- T. Joachims: Text Categorization with Support Vector Machines: Learning with Many Relevant Features. 10th European Conf. on Machine Learning ECML 1998, 137–142 – Proposes to use SVMs to divide texts into categories.

- L. Page, S. Brin, R. Motwani, T. Winograd: The PageRank Citation Ranking: Bringing Order to the Web. Technical Report, Stanford InfoLab (1998), and
  J.M. Kleinberg: Authoritative Sources in a Hyperlinked Environment. Journal of the ACM 46 (5) 604–632 (1999) – How to find important web pages based on the network structure of links.

- A.L. Barabási, R. Albert: Emergence of Scaling in Random Networks. Science 286, 509–512 (1999) – Explains the structure of networks, both biological networks and the WWW, by a random self-organization mechanism.

- S.T. Roweis, L.K. Saul: Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science 290, 2323–2326 (2000) – Brief description of an idea to compress high-dimensional data sets to facilitate data analysis and visualization; should be compared to later developments in machine learning.

- J. Shi, J. Malik: Normalized Cuts and Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) 888–905 (2000) – Graph algorithm approaches to the partitioning of images into areas of objects.

- Y. Boykov, O. Veksler, R. Zabih: Fast Approximate Energy Minimization via Graph Cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (11) 1222–1239 (2001), conference version in: International Conference on Computer Vision ICCV 1999, pp. 377–384 (1999) – The title does not say it, but again, a graph-theoretic approach is used for image processing tasks.

- L. Breiman: Random Forests. Machine Learning 45 (1) 5–32 (2001) – This machine learning techniques reduces classification errors by combining several decision trees.

- R.L. Rivest, A. Shamir, Y. Tauman: How to Leak a Secret. 7th Int. Conf. on the Theory and Applications of Cryptology and Information Security ASIACRYPT 2001, 552–565 – A cryptosystem that solves some problems in multiparty communication of secrets.

- K. Papineni, S. Roukos, T. Ward, W.J. Zhu: Bleu: a Method for Automatic Evaluation of Machine Translation. 40th Annual Meeting of the Association for Computational Linguistics ACL 2002, 311–318

- D.G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. Int. Journal of Computer Vision 60 (2) 91–110 (2004) – Presents a method to extract image features for image comparison and object recognition.

- N. Dalal, B. Triggs: Histograms of Oriented Gradients for Human Detection. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition CVPR 2005, 886–893 – Short report about feature selection for support vector machines detecting humans in pictures.
- D.L. Donoho: Compressed sensing. IEEE Transactions on Information Theory 52 (4) 1289–1306 (2006) – Mathematical foundation of storing and reconstructing sparse signals (for instance, compressed images). Meanwhile the field is extensively studied and has found numerous applications.
- E.J. Candes, J.K. Romberg, T. Tao: Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. IEEE Transactions on Information Theory 52 (2) 489–509 (2006)