

References

- Mark S. Aldenderfer and Roger K. Blashfield, 1984. *Cluster Analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-044, Sage, Beverly Hills.
- Vladimir Batagelj and Anuška Ferligoj, 2000. Clustering relational data. In W. Gaul, O. Opitz and M. Schader (eds.), *Data Analysis: Scientific Modeling and Practical Application*, Springer, Berlin, Germany, pp. 3-15.
- Amir Ben-Dor, Ron Shamir and Zohar Yakhini, 1999. Clustering gene expression patterns. *Journal of Computational Biology*, 6(3/4):281–297.
- Marcelo Blatt, Shai Wiseman and Eytan Domany, 1997. Data clustering using a model granular magnet. *Neural Computation*, 9(8):1805-1842.
- Chris Brew and Sabine Schulte im Walde, 2002. Spectral clustering for German verbs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2002)*, pp. 117-124.
- Yuri Boykov Olga Veksler and Ramin Zabih, 1999. A new algorithm for energy minimization with discontinuities. In E. R. Hancock and M. Pelillo (eds.), *Energy Minimization Methods in Computer Vision and Pattern Recognition, the Second International Workshop, EMMCVPR'99*, York, UK. Lecture Notes in Computer Science 1654, Springer, pp. 205-220.
- Sugato Basu, Arindam Banerjee and Raymond Mooney, 2002. Semi-supervised clustering by seeding. In C. Sammut and A. G. Hoffmann (eds.), *Proceedings of the Nineteenth International Conference on Machine Learning (ICML-2002)*, Sydney, Australia. Morgan Kaufmann, pp. 19-26.
- Gal Chechik and Naftali Tishby, 2003. Extracting relevant structures with side information. In S. Becker, S. Thrun, and K. Obermayer (eds.), *Advances in Neural Processing Information Systems 15 (NIPS 2002)*, Vancouver, Canada. MIT Press, Cambridge, MA, pp. 857-864.
- Ido Dagan, Contextual Word Similarity, 2000. In R. Dale, H. Moisl and H. Somers (eds.), *Handbook of Natural Language Processing*, Marcel Dekker Inc., pp. 459-476.
- Ido Dagan, Shaul Marcus and Shaul Markovitch, 1995. Contextual word similarity and estimation from sparse data. *Computer Speech and Language*, 9(2):123-152.
- Ido Dagan, Zvika Marx and Eli Shamir, 2002. Cross-dataset clustering: Revealing corresponding themes across multiple corpora. In D. Roth and A. van den Bosch (eds.), *Proceedings of the 6th Conference on Natural Language Learning (CoNLL-2002)*, Taipei, Taiwan, pp. 15-21.

- Gautam Das, Heikki Mannila and Pirjo Ronkainen, 1998. Similarity of Attributes by External Probes. In R. Agrawal, P. E. Stolorz, G. Piatetsky-Shapiro (eds.), *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98)*, New York, NY. AAAI Press, pp. 23-29.
- Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer and Richard Harshman, 1990. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391-407.
- Arthur P. Dempster, Nan M. Laird and Donald B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1), 1-38.
- Inderjit S. Dhillon and Dharmendra S. Modha, 2001. Concept decompositions for large sparse text data using clustering. *Machine Learning*, 42(1):143-175.
- Richard O. Duda, Peter E. Hart and David G. Stork, 2001. *Pattern Classification* (2nd edition), Wiley-Interscience, John Wiley & Sons Inc., New-York.
- Vladimir Estivill-Castro, 2003. Why so many clustering algorithms – a position paper, *SIGKDD Explorations*, 4(1):65-75.
- Brian Falkenhainer, Kenneth D. Forbus and Dedre Gentner, 1989. The structure mapping engine: Algorithm and example. *Artificial Intelligence*, 41(1):1-63.
- Kenneth D. Forbus, Dedre Gentner and Keith Law, 1995. MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19(2):141-205.
- Robert M. French, 2002. The computational modeling of analogy-making. *Trends in Cognitive Sciences*, 6(5):200-205.
- Nir Friedman, Ori Mosenzon, Noam Slonim and Naftali Tishby, 2002. Multivariate information bottleneck. In J. S. Breese and D. Koller (eds.), *Proceedings of the 17th conference on Uncertainty in Artificial Intelligence (UAI-17)*, University of Washington, Seattle, WA. Morgan Kaufmann, pp. 152-161.
- Michael R. Garey and David S. Johnson, 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman & Co., New York, NY.
- Tomáš Gedeon, Albert E. Parker, and Alexander G. Dimitrov, 2003. Information distortion and neural coding. *Canadian Applied Mathematics Quarterly* 10(1):33-70.
- Stuart Geman and Donald Geman, 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721-741.
- Ran Gilad-Bachrach, Amir Navot and Naftali Tishby, 2003. An information theoretic tradeoff between complexity and accuracy. Unpublished Manuscript.

- Maya Gorodetsky, 2001. *Methods for Revealing Semantic Relations between Words Based on Cooccurrence Patterns in Corpora*. MSc Thesis (in Hebrew). The Hebrew University, Jerusalem, Israel.
- Thomas Hofmann, 2001. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 41(1):177-196.
- Thomas Hofmann, Jan Puzicha and Michael I. Jordan, 1999. Learning from dyadic data. In M.S. Kearns, S.A. Solla and D.A. Cohn (eds.), *Advances in Neural Information Processing Systems 11 (NIPS*1998)*, Denver, CO, pp. 466-472.
- Douglas R. Hofstadter and the Fluid Analogies Research Group, 1995. *Fluid Concepts and Creative Analogies*. Basic Books, New-York, NY.
- Keith J. Holyoak and Paul Thagard, 1989. Analogical mapping by constraint satisfaction. *Cognitive Science*, 13:295-355.
- John E. Hummel and Keith J. Holyoak, 1997. Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, 104:427-466.
- Anil K. Jain , M. Narasimha Murty and Patrick J. Flynn, 1999. Data clustering: A review, *ACM Computing Surveys* 31(3):264-323.
- Edwin T. Jaynes, 1982. On the rationale of maximum-entropy methods. *Proceedings of the IEEE* 70(9):939-952.
- Scott Kirkpatrick, C. Daniel Gelatt and Mario P. Vecchi, 1983. Optimization by simulated annealing, *Science* 220(4598):671-680.
- Jon Kleinberg, 2003. An impossibility theorem for clustering. In S. Becker, S. Thrun, and K. Obermayer (eds.), *Advances in Neural Information Processing Systems 15 (NIPS 2002)*, Vancouver, Canada. MIT Press, Cambridge, MA, pp. 446-453.
- Boicho N. Kokinov and Alexander A. Petrov, 2001. Integration of memory and reasoning in analogy-making: The AMBR model. In Gentner, D., Holyoak, K., Kokinov, B. (eds.), *The Analogical Mind*, Cambridge, MA: MIT Press.
- Teuvo Kohonen, 1988. *Self-Organization and Associative Memory*, Third Edition, Springer-Verlag, Berlin.
- Anna Korhonen, Yuval Krymolowski. and Zvika Marx, 2003. Clustering polysemic subcategorization frame distributions semantically. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL 2003)*, Sapporo, Japan, pp. 64-71.
- Dekang Lin. Automatic retrieval and clustering of similar words, 1998. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics(COLING-ACL '98)*, Montreal, Canada, pp. 768-774

- David J.C. MacKay, 2003. *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press.
- Alexandros S. Manos, 1996. *A Study on Out-of-Vocabulary Word Modeling for a Segment-based Keyword Spotting System*. M.S. thesis, MIT Department of Electrical Engineering and Computer Science, June 1996. MIT-LCS-TR-694.
- Willey J. R. Martin, Bernard P. F. Al and Piet J. G. van Sterkenburg, 1983. On the processing of text corpus: From textual data to lexicographical information. In R. R. K. Hartmann (ed.), *Lexicography: Principles and Practice*, Academic Press Inc., London, U.K, pp. 77-87.
- Zvika Marx and Ido Dagan, 2001. Conceptual mapping through keyword coupled clustering. *Mind & Society* 4(2):59-85.
- Zvika Marx, Ido Dagan and Joachim M. Buhmann, 2001. Coupled clustering: A method for detecting structural correspondence. In C. E. Brodley and A. P. Danyluk, (eds.), *Proceedings of the Eighteenth International Conference on Machine Learning (ICML-2001)*, Williams College, Williamstown, MA. Morgan Kaufmann Publishers, San Francisco, CA, pp. 353-360.
- Zvika Marx, Ido Dagan, Joachim M. Buhmann and Eli Shamir, 2002. Coupled clustering: A method for detecting structural correspondence. *Journal of Machine Learning Research* 3, 747-780.
- Zvika Marx, Ido Dagan and Eli Shamir, 2002. Cross-component clustering for template induction. In M. Grobelnik, N. Milic-Frayling and D. Mladenić (eds.), *Proceedings of the ICML-2002 Workshop on Text Learning*, The University of New South Wales, Sydney, Australia, pp. 66-75.
- Zvika Marx, Ido Dagan and Eli Shamir, 2004. Identifying structure across pre-partitioned data. In S. Thrun, L. Saul, and B. Schölkopf (eds.), *Advances in Neural Information Processing Systems 16 (NIPS 2003)*, Vancouver, Canada. MIT Press, Cambridge, MA, pp. 489-496.
- Marina Meila, 2003. Comparing clusterings by the variation of information. In B. Schölkopf and M. K. Warmuth (eds.), *Proceeding of Computational Learning Theory and Kernel Machines, 16th Annual Conference on Computational Learning Theory and 7th Kernel Workshop (COLT/Kernel 2003)*, Washington, DC. Springer, pp. 173-187.
- Michael Ortega, Yong Rui, Kaushik Chakrabarti, Sharad Mehrotra and Thomas S. Huang, 1998. Supporting ranked boolean similarity queries in mars. *IEEE Transactions on Knowledge and Data Engineering*, 10(6):905-925.
- Ferando C. Pereira, Naftali Tishby and Lillian J. Lee, 1993. Distributional clustering of English words. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics ACL' 93*, Columbus, OH, pp. 183-190.
- George Polya, 1957. *How to Solve It*, (2nd edition), Princeton University Press, Princeton, NJ.

- Jan Puzicha, Thomas Hofmann and Joachim M. Buhmann, 2000. A theory of proximity based clustering: Structure detection by optimization. *Pattern Recognition*, 33(4): 617-634.
- Noam Slonim and Naftali Tishby, 2000. Unsupervised document clustering using word clusters via the information bottleneck method. In N. J. Belkin, P. Ingwersen, and M. Leong (eds.), *SIGIR-2000: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Athens, Greece. ACM Press, New York, pp. 208-215.
- Noam Slonim, Nir Friedman and Naftali Tishby, 2002. Unsupervised document classification using sequential information maximization. In *SIGIR-2002: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tampere, Finland. ACM Press, New York, pp. 129-136.
- Frank Smadja, 1993. Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1):143-177.
- Ninian Smart, 1996. *Dimensions of the Sacred: An Anatomy of the World's Beliefs*. University of California Press, Berkeley and Los Angeles, CA.
- Naftali Tishby, Fernando C. Pereira and William Bialek, 1999. The information bottleneck method. In *37th Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, Illinois, pp. 368-379.
- Kedar Nath Tiwari, 1992. *Comparative Religion*. Jainendra Prakash Jain at Shri Jainendra Press, New Delhi, India.
- Cornelis Joost van Rijsbergen, 1979. *Information Retrieval*. Butterworths, London.
- Tony Veale, Diarmuid O'Donoghue and Mark Keane, 1999. Computability as a limiting cognitive constraint: Complexity concerns in metaphor comprehension about which cognitive linguists should be aware. In M. Hiraga, C. Sinha and S. Wilcox (eds.), *Cultural, Psychological and Typological Issues in Cognitive Linguistics*, John Benjamins Publishers, Amsterdam/Philadelphia., pp. 129-155
- Kiri Wagstaff, Claire Cardie, Seth Rogers and Stefan Schroedl, 2001. Constrained k-means clustering with background knowledge. In C. E. Brodley and A. P. Danyluk, (eds.), *Proceedings of the Eighteenth International Conference on Machine Learning (ICML-2001)*, Williams College, Williamstown, MA. Morgan Kaufmann Publishers, San Francisco, CA, pp. 577-584.

