
Bibliographie

[Apté *et al.*, 1994] C. Apté, F. Damerau, S. M. Weiss. Automated learning of decision rules for text categorization. *ACM Transactions on Information Systems*, 12(3), 233-251, 1994.

[Apté *et al.*, 1998] C. Apté, F. Damerau, S. M. Weiss. Text mining with decision rules and decision trees. *Proceedings of the Conference on Automated Learning and Discovery, Workshop 6: Learning from Text and the Web*, 1998.

[Akaike, 1974] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723, 1974.

[Arampatzis *et al.*, 2000] A. T. Arampatzis, Th. P. van der Weide, C. H. A. Koster, P. van Bommel. Term Selection for Filtering based on Distribution of Terms over Time. *Proceedings of the Sixth Conference on Content-Based Multimedia Information (RIAO 2000)*, 1221-1237, 2000.

[Bartlett, 1997] P.L. Bartlett. For valid generalization, the size of the weights is more important than the size of the network *Neural Information Processing Systems*, 9, Morgan Kaufmann, San Mateo CA, 1997.

[Beaulieu *et al.*, 1997] M. M. Beaulieu, M. Gatford, X. Huang, S. E. Robertson, S. Walker, P. Williams, P. Williams. Okapi at TREC-5. *Proceedings of the Fifth Text REtrieval Conference (TREC-5)*, NIST Special Publication 500-238, 143-166, 1997.

[Bishop, 1992] C. M. Bishop. Exact Calculation of the Hessian Matrix for the multilayer perceptron. *Neural Computation*, 4(4), 494-501, 1992.

[Bishop, 1995] C. M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.

[Björck, 1967] A. Björck. Solving linear least squares problems by Gram-Schmidt orthogonalization. *Nordisk Tidshrift for Informationsbehandling*, 7, 1-21, 1967.

[Boughanem *et al.*, 2000] M. Boughanem, C. Julien, J. Mothe, C. Soule-Dupuy. Mercure at TREC-8: Adhoc, Web, CLIR and Filtering tasks. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 431-444, 2000.

[Broyden, 1970] C. G. Broyden, The Convergence of a Class of Double-Rank Minimization Algorithms 2 : the New Algorithm. *Journal Institute of Mathematics and its Applications*, 6, 222-231, 1970.

[Buntine et Weigend, 1991] W. Buntine, A.S Weigend. Bayesian back-propagation. *Complex Systems*, 5, 603-643, 1991.

[Chen *et al.*, 1989] S. Chen, S. A. Billings, W. Luo. Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5), 1873-1896, 1989.

[Cohn *et al.*, 1996] D. Cohn, Z. Ghahramani, M. Jordan. Active learning with statistical models. *Journal of Artificial Intelligence Research*, 4, 129-145, 1996.

[Cohen, 1996] W. W. Cohen. Learning rules that classify e-mail. *Proceedings of the 1996 AAAI Spring Symposium on Machine Learning in Information Access*, 18-25, 1996.

[Cohen et Singer, 1996] W. W. Cohen, Y. Singer. Context-sensitive methods for text categorization. *Proceedings of the 19th Annual International Conference on Research and Development in Information Retrieval (SIGIR '96)*, 307-315, 1996.

[Cover, 1965] T. M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, 14, 326-334, 1965.

[Cybenko, 1989] G. Cybenko. Approximation by Superpositions of Sigmoidal Function. *Signals and Systems*, 2, 303-314, 1989.

[Deerwester *et al.*, 1990] S. Deerwester, S. Dumais, G. Furnas, T. Landauer, R. Harshman. Indexing by Latent Semantic Indexing. *Journal of the American Society for Information Science*, 41(6), 391-407, 1990.

[Dumais *et al.*, 1998] S. Dumais, J. Platt, D. Heckerman, M. Sahami. Inductive Learning Algorithms and Representations for Text Categorization. *Proceedings of the seventh International Conference on Information and Knowledge Management (CIKM' 98)*, 148-155, 1998.

[Evans et Lefferts, 1995] D. A. Evans, R. Lefferts. CLARIT-TREC experiments. *Information Processing and Management*, 31(3), 385-395.

[Flechter, 1970] R. Flechter. A New Approach to Variable Metric Algorithms. *The Computer Journal*, 13(3), 317-322, 1970.

[Funahashi, 1989] K. Funahashi. On the Approximate Realization of Continuous Mappings by Neural Networks. *Neural Networks*, 2, 183-192, 1989.

[Gallinari et Cibas, 1999] P. Gallinari, T. Cibas. Practical complexity control in multilayer perceptrons. *Signal Processing*, 74, 29-46, 1999.

[Geman *et al.*, 1992]. S. Geman, E. Bienenstock, R. Doursat. Neural Network and the Bias/Variance Dilemma. *Neural Computation*, 4, 1-58, 1992.

[Goldfarb, 1970] D. Goldfarb. A Family of Variable Metric Methods Derived by Variational Means. *Mathematics of Computation*, 24, 23-26, 1970.

[Grossman et Frieder, 1998] D. Grossman, O. Frieder. *Information Retrieval. Algorithms and Heuristics*. Kluwer Academic Publishers, 1998.

[Gull, 1988] S. F. Gull. Bayesian inductive inference and maximum entropy. G. J. Erickson, C.R. Smith eds. *Maximum-Entropy and Bayesian Methods in Science and Engineering, Vol. 1: Foundations*, 53-74, Dordrecht: Kluwer, 1988.

[Harman, 1991] D. Harman. How effective is suffixing ? *Journal of the American Society for Information Science*, 42(1), 321-331, 1991.

[Hinton, 1987] G.E. Hinton. Learning Translation Invariant Recognition in Massively Parallel Networks. *Proceedings PARLE Conference on Parallel Architectures and Languages Europe*, J. W. de Bakker, A. J. Nijman, P.C. Treleaven (Eds.), Springer-Verlag, Berlin, 1-13, 1987.

[Hornik *et al.*, 1994] K. Hornik, M. Stinchcombe, H. White & P. Auer. Degree of Approximation Results for Feedforward Networks Approximating Unknown Mappings and their Derivatives. *Neural Computation*, 6(6), 1262-1275, 1994.

[Hull, 1993] D. A . Hull. Using Statistical testing in the evaluation of the retrieval performance. *Proceedings of the 16th Annual International Conference on Research and Development in Information Retrieval (SIGIR'93)*, 329-338, 1993.

[Hull, 1994] D. A. Hull. *Information retrieval using statistical classification*. Ph.D. thesis, Stanford University, 1994.

[Hull, 1996] D. A. Hull. Stemming Algorithms – A Case Study for Detailed Evaluation. *Journal of the American Society for Information Science*, 47(1), 70-84, 1996.

[Hull, 1999]. D. A. Hull. The TREC-7 Filtering Track: Description and Analysis. *Proceedings of the Seventh Text REtrieval Conference (TREC-7)*. NIST Special Publication 500-242, 33-56, 1999.

[Hull et Robertson, 2000]. D. A. Hull, S. Robertson. The TREC-8 Filtering Track Final Report. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 35-56, 2000.

[Iyer *et al.*, 2000] R. D. Iyer, D. D. Lewis, R. E. Schapire, Y. Singer, A. Singhal. Boosting for Document Routing. *Proceedings of the ninth International Conference on Information and Knowledge Management (CIKM 2000)*, 2000.

[Jing et Tzoukermann, 1999] H. Jing, E. Tzoukermann. Information Retrieval Based on Context Distance and Morphology. *Proceedings of the 22nd Annual International Conference on Research and Development in Information Retrieval (SIGIR '99)*, 90-96, 1999.

[Joachims, 1998] T. Joachims. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. *Proceedings of the Tenth European Conference on Machine Learning (ECML'98)*, Springer Verlag, 137-142, 1998.

[Kindermann et Leopold, 2000] J. Kindermann, E. Leopold. Classification of Texts with Support Vector Machines. An estimation of the efficiency of Kernels and Data-Transformations. *Proceedings of the 24th Annual Conference of the Gessellschaft für Klassifikation*, 2000.

[Koller et Sahami, 1996] D. Koller, M. Sahami. Toward optimal feature selection. *Proceedings of the Thirteenth International Conference on Machine Learning*, 1996.

[Krogh et Hertz, 1992] A. Krogh, J.A. Hertz. A Simple Weight Decay Can Improve Generalization. *Advances in Neural Information Processing Systems*, 4, J.E. Moody, S.J. Hanson and R.P. Lippmann, eds., Morgan Kauffmann Publishers, San Mateo CA, 950-957, 1992.

[Krovetz, 1993] R. Krovetz. Viewing morphoogy as an inference process. *Proceedings of the 16th Annual International Conference on Research and Development in Information Retrieval (SIGIR'93)*, 191-203, 1993.

[Kwok *et al.*, 2000] K. L. Kwok, L. Grunfeld, M. Chan. TREC-8 Ad-Hoc, Query and Filtering Track Experiments using PIRCS. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 217-228, 2000.

[Landau *et al.*, 1993] M.-C. Landau, F. Sillion, F. Vichot. Exoseme: a Thematic Document Filtering System based on conceptual graphs. *Proceedings of the International Conference on Conceptual Structures (ICCS'93)*, 171-187, 1993.

[Lewis, 1992] D. D. Lewis. *Representation and Learning in Information Retrieval*. Ph.D. thesis, Graduate School of the University of Massachusetts, 1992.

[Lewis et Ringuette, 1994] D. D. Lewis, M. Ringuette. Comparison of two learning algorithms for text categorization. *Proceedings of Third Annual Symposium on Document Analysis and Information Retrieval (SDAIR'94)*, 181-93, 1994.

[Lewis et al., 1996] D. D. Lewis, R. E. Schapire, J. P. Callan, R. Papka. Training algorithms for linear text classifiers. *Proceedings of the 19th Annual International Conference on Research and Development in Information Retrieval (SIGIR '96)*, 298-306, 1996.

[Lovins, 1968] J. Lovins. Development of a stemming algorithm. *Mechanical Translation and Computational Linguistics*, 11, 22-31, 1968.

[MacFarlane et Robertson, 2000] A. MacFarlane, S. E. Robertson. PLIERS at TREC-8. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 241-252, 2000.

[MacKay, 1992a] D. J. C. MacKay. Bayesian interpolation. *Neural Computation*, 4(3), 415-447, 1992.

[MacKay, 1992b] D. J. C. MacKay. A Practical Bayesian Framework for Backpropagation Networks. *Neural Computation*, 4(3), 448-472, 1992.

[MacKay, 1992c] D. J. C. MacKay. The Evidence Framework Applied to Classification Networks. *Neural Computation*, 4(5), 698-714, 1992.

[MacKay, 1999] D. J. C. MacKay. Comparison of Approximate Methods for Handling Hyperparameters. *Neural Computation*, 11(5), 1035-1068, 1999.

[Manning et Schütze, 1999] C.D. Manning, H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.

[McCallum et Nigam, 1998a] A. McCallum, K. Nigam. A Comparison of Event Models for Naive Bayes Text Classification. *AAAI-98 Workshop on Learning for Text Categorization*, 41-48, 1998.

[McCallum et Nigam, 1998b] A. McCallum, K. Nigam. Employing EM and Pool-Based Active Learning for Text Classification. *Proceedings of the Fifteenth International Conference on Machine Learning (ICML'98)*, 350-358, 1998.

[Miller *et al.*, 1990] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, K. J. Miller. Introduction to WordNet: An on-line lexical database. *Journal of Lexicography*, 3(4), 235-244, 1990.

[Minoux, 1983] M. Minoux. *Programmation mathématique, théorie et algorithmes*. Ed. Dunod, Tome 1, 1983.

[Monari, 1999] G. Monari. *Sélection de modèles non linéaires par leave-one-out. Etude théorique et application des réseaux de neurones au procédé de soudage par points*. Thèse de l'université Paris VI, 1999.

[Mouliner, 1997] I. Moulinier. *Apprentissage et Acquisition de Connaissances*. Thèse de l'université Paris VI, 1997.

[Muc7, 1999] *Proceedings of the Seventh Message Understanding Conference*. Morgan Kaufmann, 1999.

[Neal, 1992] R. M. Neal. Bayesian Training of Backpropagation Networks by the Hybrid Monte Carlo Method. Technical Report CRG-TR-92-1, Department of Computer Science, University of Toronto, 1992.

[Neal, 1994] R. M. Neal. *Bayesian Learning for Neural Networks*. Ph.D. thesis, University of Toronto, 1994.

[Neal, 1996] R. M. Neal. *Bayesian Learning for Neural Networks*. New York : Springer-Verlag, 1996.

[Ng *et al.*, 2000] H. T. Ng, H. T. Ang, W. M. Soon. DSO at TREC-8: A Hybrid Algorithm for the Routing Task. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 267-274, 2000.

[Oard et Wang, 2000] D. W. Oard, J. Wang. TREC-8 Experiments at Maryland: CLIR, QA and Routing. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 623-636, 2000.

[Porter, 1980] M. F. Porter. An algorithm for suffix stripping. *Program* 14, 3, 130-1367, 1980.

[Press *et al.*, 1992] W. H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery. *Numerical Recipes in C : The Art of Scientific Computing*. second Edition, Cambridge university Press, 1992.

[Quinlan, 1996] J. R. Quinlan. Bagging, boosting, and C4.5. *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, 725-730, 1996.

[Richard et Lippman, 1991] M. D. Richard, R. P. Lippmann. Neural Network classifiers estimate Bayesian a-posteriori probabilities. *Neural Computation*, 3(4), 461-483, 1991.

[Robertson et Sparck Jones, 1976] S. Robertson, K. Sparck Jones. Relevance weighting of search terms. *Journal of American Society for Information Science*, 27(3), 129-146, 1976.

[Robertson et Walker, 2000] S. Robertson, S. Walker. OKAPI/Keenbow at TREC-8. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 151-162, 2000.

[Rocchio, 1971] J. J. Rocchio. Relevance Feedback in Information Retrieval. *The SMART Information Retrieval System*. G. Salton, Ed. Prentice Hall, Englewood Cliffs, NJ, 313-323, 1971.

[Rumelhart *et al.*, 1986] D.E. Rumelhart, G.E. Hinton, R.J. Williams. *Learning Internal Representations by Error Propagation*, Parallel Distributed Processing, MIT Press, Cambridge MA, 318-362, 1986.

[Sahami, 1998] M. Sahami. *Using Machine Learning to Improve Information Access*. Ph.D. Dissertation, Stanford, 1998.

[Sahami *et al.*, 1998] M. Sahami, S. Dumais, D. Heckerman, E. Horvitz. A Bayesian approach to filtering junk e-mail. *AAAI 1998 Workshop on Text Categorization*, 1998.

[Salton *et al.*, 1975] G. Salton, A. Wong, C. S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18, 613-620, 1975.

[Salton, 1989] G. Salton. *Automatic Text Processing, The Transformation, Analysis, and Retrieval of Information by Computer*. Addison Wesley, Reading, USA, 1989.

[Salton et Buckley, 1990] G. Salton, C. Buckley. Improving retrieval performance by relevance feedback. *Journal of the American Society for Information Science*, 41(4), 288-287, 1990.

[Saporta, 1990] G. Saporta. *Probabilités, analyse des données et statistiques*. Editions Technip, 1990.

[Schapire *et al.*, 1998] R. E. Schapire, Y. Singer, A. Singhal. Boosting and Rocchio Applied to Text Filtering. *Proceedings of the 21st Annual International Conference on Research and Development in Information Retrieval (SIGIR '98)*, 215-223, 1998.

[Schmidt, 1994] H. Schmidt. Probabilistic part-of-speech tagging using decision trees. *International Conference on New Methods in Language Processing*, 44-49, 1994.

[Schütze et Pedersen, 1995] H. Schütze, J. O. Pedersen. Information Retrieval Based on Word Senses. *Fourth Annual Symposium on Document Analysis and Information Retrieval*, 161-175, 1995.

[Schütze *et al.*, 1995] H. Schütze, D. A. Hull, J. O. Pedersen. A Comparison of Classifiers and Document Representations for the Routing Problem. *Proceedings of the 18th Annual International Conference on Research and Development in Information Retrieval (SIGIR'95)*, 229-238, 1995.

[Schwartz, 1978] G. Schwartz. Estimating the dimension of a model. *Annals of statistics*, 6(2), 461-464, 1978.

[Shanno, 1970] D.F. Shanno. Conjugate gradient methods with inexact searches. *Mathematics of operations Research* 3 (3), 244-256, 1978.

[Singhal, 1996] A. Singhal. Pivoted Length Normalization. *Proceedings of the 19th Annual International Conference on Research and Development in Information Retrieval (SIGIR'96)*, 21-29, 1996.

[Singhal, 1998] A. Singhal. AT&T at TREC-6. *Proceedings of the Sixth Text REtrieval Conference (TREC-6)*. NIST Special Publication 500-240, 215-226, 1998.

[Sjöberg, 1994] J. Sjöberg, L. Ljung. Overtraining, Regularization and Searching for Minimum in Neural Networks. Technical Report LiTH-ISY-R-1567, Department of electrical Engineering, Linköping University, 1994.

[Sparck Jones, 1999] K. Sparck Jones. Information retrieval and artificial intelligence. *Artificial Intelligence*, 114, 257-281, 1999.

[Stoica et Evans, 2000] E. Stoica, D. A. Evans. Dynamic Term Selection in Learning a Query from Examples. *Proceedings of the Sixth Conference on Content-Based Multimedia Information (RIAO 2000)*, 1703-1719, 2000.

[Stoppiglia, 1997] H. Stoppiglia. *Méthodes statistiques de sélection de modèles neuronaux; applications financières et bancaires*. Thèse de l'université Paris VI, 1997.

[Stricker et Haré, 1998] M. Stricker, S. Haré. Réseaux de neurones et analyse marketing. Rapport Technique DTA, Informatique-CDC, 1998.

[Stricker *et al.*, 2000a] M. Stricker, F. Vichot, F. Wolinski, G. Dreyfus. Towards the Automatic Design of Efficient Custom Filters. *Reconnaissance des formes et Intelligence Artificielle (RFIA 2000)*, 129-137, 2000.

[Stricker *et al.*, 2000b] M. Stricker, F. Vichot, G. Dreyfus, F. Wolinski. Two Step Feature Selection for the TREC-8 Routing. *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 425-430, 2000.

[Stricker *et al.*, 2001] M. Stricker, F. Vichot, G. Dreyfus, F. Wolinski. Training Context-Sensitive Neural Networks With Few Relevant Examples for the TREC-9 Routing. *Proceedings of the Ninth Text REtrieval Conference (TREC-9)*. NIST Special Publication, A paraître, 2001.

[Thodberg, 1996] H. H. Thodberg. A Review of Bayesian Networks with an Application to Near Infrared Spectroscopy. *IEE Transactions of Neural Networks*, 7(1), 56-72, 1996.

[van Rijsbergen, 1979] C. J. van Rijsbergen. *Information Retrieval*. Butterworths, London, second edition, 1979.

[Vichot *et al.*, 1997] F. Vichot, F. Wolinski, J. Tomeh, S. Guennou, B. Dillet, S. Aidjan. High Precision Hypertext Navigation Based on NLP Automatic Extractions. *Proceedings of Hypertext Information Retrieval Multimedia (HIM'97)*, 161-174, 1997.

[Vichot *et al.*, 1999] F. Vichot, F. Wolinski, H.-C. Ferri, D. Urbani. Feeding a Financial Decision Support System with Textual Information. *Journal of Intelligent and Robotic Systems*, 26, 157-166, 1999.

[Voorhees, 1993] E. M. Voorhees. Using WordNet to disambiguate words senses for text retrieval. *Proceedings of the 16th Annual International Conference on Research and Development in Information Retrieval (SIGIR'93)*, 171-180, 1993.

[Voorhees et Harman, 2000] E. M. Voorhees, D. K. Harman. Overview of the eighth Text REtrieval Conference (TREC-8). *Proceedings of the Eighth Text REtrieval Conference (TREC-8)*. NIST Special Publication 500-246, 1-24, 2000.

[Wiener, 1993] E. D. Wiener. *A Neural Network Approach to Topic Spotting in Text*. Ph.D. Dissertation, Stanford, 1993.

[Wiener *et al.*, 1995] E. D. Wiener, J. O. Pedersen, A. S. Weigend. A Neural Network Approach for Topic Spotting. *Proceedings of the Fourth Annual Symposium on Document Analysis and Information Retrieval (SDAIR'95)*, 317-332, 1995.

[Williams, 1995] P.M. Williams. Bayesian regularization and pruning using a Laplace prior. *Neural Computation*, 7(1), 117-143, 1995.

[Wolinski *et al.*, 1995] F. Wolinski, F. Vichot, B. Dillet. Automatic Processing of Proper Names in Texts. *Proceedings of the seventh Conference of the European Chapter of the Association for Computational Linguistics (EACL'95)*, 23-30, 1995.

[Wolinski *et al.*, 1998] F. Wolinski, F. Vichot, O. Grémont. Producing NLP-based On-line Contentware. *Natural Language Processing & Industrial Applications*, 253-259, 1998.

[Wolinski *et al.*, 2000] F. Wolinski, F. Vichot, M. Stricker. Using Learning-Based Filters to Detect Rule-Based Filtering Obsolescence. *Proceedings of the Sixth Conference on Content-Based Multimedia Information (RIAO 2000)*, 1208-1220, 2000.

[Wolinski et Vichot, 2001] F. Wolinski, F. Vichot. Des multi agents pour produire des applications de contenu en ligne. *Technique et Science Informatique*, A paraître, 2001.

[Yang et Chute, 1994] Y. Yang, C. G. Chute. An example-based mapping method for text categorization and retrieval. *ACM Transactions on Information Systems*, 12(3), 252-277, 1994.

[Yang et Pedersen, 1997] Y. Yang, J. O. Pedersen. A Comparative Study on Feature Selection in Text Categorization. *Proceedings of the Fourteenth International Conference on Machine Learning (ICML'97)*, 412-420, 1997.

[Yang et Liu, 1999] Y. Yang, X. Liu. A re-examination of text categorization methods. *Proceedings of the 22nd Annual International Conference on Research and Development in Information Retrieval (SIGIR '99)*, 42-49, 1999.

[Yarowsky, 1995] D. Yarowsky. Unsupervised word sense disambiguation rivaling supervised methods. *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics*, 189-196, 1995.

[Zaragoza, 1999] H. Zaragoza. *Modèles dynamiques d'apprentissage numérique pour l'accès à l'information textuelle*. Thèse de l'université Paris VI, 1999.

[Zipf, 1949] G. K. Zipf. *Human Behavior and the Principle of Least Effort*. Addison-Wesley, 1949.