

## Referências bibliográficas

- [1] GOLDSCHMIDT R. & PASSOS, E. Data Mining: um guia prático. Editora Campus, Rio de Janeiro: Elsevier, 2005.
- [2] INMON, W.H. and HACKATHORN, R.D. Using the Data Warehouse, John Wiley & Son's, ISBN 0-471-05966-8. 1994.
- [3] KIMBALL, R. and ROSS, M. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (Second Edition), John Wiley & Sons, ISBN 0-471-20024-7. 2002.
- [4] CHEN, H. Knowledge management systems: a text mining perspective. University of Arizona (Knowledge Computing Corporation), Tucson, Arizona, 2001.
- [5] TAN, A.-H. Text mining: The state of the art and the challenges. In Proceedings, PAKDD'99 workshop on Knowledge Discovery from Advanced Databases, Beijing, pp.65-70, 1999.
- [6] KAO, A. and POTEET, S.R. (eds.) Natural Language Processing and Text Mining, Springer, New York, 2006.
- [7] MANNING, C.D. and SCHFITZE, H. Foundations of Statistical Natural Language Processing, MIT Press, Cambridge, MA, 1999.
- [8] GRISHAM, R. Computational Linguistics – an Introduction. Cambridge: Cambridge University Press, 1992.
- [9] OTHERO, G. de A. (2003) Introdução à lingüística computacional. Comunicação apresentada na *III Semana Acadêmica da Faculdade de Letras da Pontifícia Universidade Católica do Rio Grande do Sul* – PUCRS.
- [10] LUGER, G.F. Cognitive Science. Academic Press, 1994.
- [11] WILSON, R.A. and KEIL, F.C. (Eds.). The MIT Encyclopedia of the Cognitive Sciences. MIT Press, Cambridge, Massachusetts, 1999.
- [12] BAEZA-YATES, R. and BERTIER, R.N. Modern Information Retrieval. Harlow: Addison-Wesley, 1999.
- [13] <http://pt.wikipedia.org/wiki/Estatística>.

- [14] BRILL, E. Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part of Speech Tagging. Computational Linguistics, December, 1995.
- [15] CHAKRABARTI, S. Mining the Web: discovering knowledge from hypertext data. Morgan Kaufmann Publishers, London, 2003.
- [16] OLIVEIRA JUNIOR, H.A. e (Coordenador). Inteligência computacional: aplicada à administração, economia e engenharia em Matlab. São Paulo: Thomson Learning, 2007. 370p.
- [17] HERRERA, F. and VERDEGAY, J.L. Genetic Algorithms and Soft Computing, Editora: Physica-Verlag, vol. 1, 709p.
- [18] HERTZ, J.; KROGH, A. and PALMER, R.G. Introduction to The Theory of Neural Computation, Editora: Addison Wesley, Vol. 1, 327p.
- [19] WASSERMAN, P.D. Neural Computing - Theory and Practice, Editora: VNR Vol. 1, 230p.
- [20] EBECKEN, N.F.F.; LOPES, M.C.S.; COSTA, M.C.A. Mineração de Textos. In: Sistemas Inteligentes: Fundamentos e Aplicações. Barueri, SP: Manole, Cap. 13, 2003.
- [21] PANG, B.; LEE, L. and VAITHYANATHAN, S. Thumbs Up? Sentiment classification using machine learning techniques, In Proc. of the EMNLP, 2002.
- [22] OGURI, P. Aprendizado de Máquina para o Problema de Sentiment Classification, PUC-RIO, 2007. Dissertação de Mestrado. Rio de Janeiro : PUCRio, Departamento de Informática, 2007.
- [23] YI, J.; NASUKAWA, T.; BUNESCU, R. and NIBLAC, W. Sentiment analyser: Extracting sentiments about a given topic using natural language processing techniques, In IEEE Intl. Conf. on Data Mining (ICDM), 2003.
- [24] Text Mining: Predictive Methods for Analyzing Unstructured Information – Estou repetindo essa referencia. Pq, agora quero falar de um capítulo específico, p. 167.
- [25] FERNANDES, A.M.R.; BEPPLER, M.D. Aplicação de text mining para a extração de conhecimento jurisprudencial. Hífen, Uruguaiana, RS, v. 29, n. 55/56, p. 139-146, 2005.
- [26] Handbook of Medical Informatics – J.H. van Bommel and Musen-Springer 1997.
- [27] Knowledge management, Data mining, and Text mining in medical informatics – Chen, Fuller and Hersh, Springer 2005.
- [28] Information Retrieval: A Health and Biomedical Perspective – Hersh-Springer 2002.

- [29] ARANHA, C.N. Uma abordagem de pré-processamento automático para mineração de textos em português: sob o enfoque da inteligência computacional, Tese de Doutorado, Departamento de Engenharia Elétrica, PUC/Rio, 2007.
- [30] PETERSON, R.E. Eight internet search engines compared. First Monday.[http://www.firstmonday.dk/issues/issue2\\_2/peterson/](http://www.firstmonday.dk/issues/issue2_2/peterson/), 1997.
- [31] CHAKRABARTI, S.; BERG, M. van den.; DOM, B. Focused crawling: a new approach to topic-specific Web resource discovery , Computer Networks, Amsterdam, Netherlands, 1999.
- [32] IMAMURA, C. e REZENDE, S.O. Pré-processamento para extração de conhecimento de texto, Departamento de Ciências de Computação e Estatística, Instituto de Ciências Matemáticas de São Paulo.
- [33] KONCHADY. Text Mining Application Programming, Thomson, 2006.
- [34] REIS, D.C. e FONSECA, B.M. O fantástico mundo da distância de edição, 03 de julho de 2002.
- [35] YANG, Y.; PEDERSEN, J.O. A Comparative Study on Feature Selection in Text Categorization, Proceedings of the Fourteenth International Conference on Machine Learning, p. 412-420, July 08-12, 1997.
- [36] CHURCH, K.W. and HANKS, P. Word association norms, mutual information and lexicography. In Proceedings of the 27th Annual Conference of the Association of Computational Linguistics, pp. 76-82, 1989.
- [37] ARANHA, C.N. e PASSOS, E.P.L. A Tecnologia de Mineração de Textos. RESI. Revista Eletrônica de Sistemas de Informação, v. 2, p. 2, 2006.
- [38] SAEED, J.L. (1997). Semantics. Oxford: Blackwell, 1997.
- [39] WEISS, S.; INDUSKHYA, N.; ZHANG, T. and DAMERU, F. Text Mining: Predictive Methods for Analyzing Unstructured Information, Springer, New York, NY, 2005.
- [40] RABINER, L.R. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, 77 (2), p. 257-286, February 1989.
- [41] CAPPÉ, O.; MOULINES, E. and RYDÉN, T. Inference in Hidden Markov Models, Springer, ISBN 0-387-40264-0, 2005.
- [42] SEYMORE, K.; MCCALLUM, A. and ROSENFELD, R. Learning Hidden Markov Model Structure for Information Extraction. AAAI 99 Workshop on Machine Learning for Information Extraction, 1999.

- [43] SANTOS, C.N. dos e MILIDIÚ, R.L. Probabilistic Classifications with TBL. CICLing, pp. 196-207, 2007.
- [44] SCOTT, S. and MATWIN, S. Feature Engineering for Text Classification, Proceedings of the Sixteenth International Conference on Machine Learning, p. 379-388, June 27-30, 1999.
- [45] [http://pt.wikipedia.org/wiki/Árvore\\_B](http://pt.wikipedia.org/wiki/Árvore_B).
- [46] XHITS: estendendo o algoritmo HITS para extração de tópicos WWW / Francisco Benjamim Filho ; orientadores: Ruy Luiz Milidui, Raúl Rentería. – Rio de Janeiro : PUCRio, Departamento de Informática, 2005.
- [47] KLEINBERG, J. Authoritative sources in a hyperlinked environment. In Proc. Ninth Ann. ACM-SIAM Symp. Discrete Algorithms, pp. 668-677, ACM Press, New York, 1998.
- [48] PAGE, L.; BRIN, S.; MOTWANI, R. and WINOGRAD, T. The PageRank citation ranking: Bringing order to the Web, 1999.
- [49] GEISSER, S. The predictive sample reuse method with applications, Journal of the American Statistical Association, 70(350), 1975.
- [50] BAYES, T. An Essay Towards Solving a Problem in the Doctrine of Chances, Philos. Trans. R. Soc. London, 1763.
- [51] RUSSEL, S. and NORVIG, P. Artificial Intelligence: A Modern Approach, Prentice Hall, Englewood Cliffs, NJ, 4, 4.1.2 , 6, 1995.
- [52] GOMES, R. e PASSOS, E.P.L. Estudo Orientado Sobre Teoria em Mineração de Textos; Relatório Técnico do ICA de 2007.
- [53] SLONIM, N. and TISHBY, N. The power of word clusters for text classification. In Proceedings of ECIR-01, 23rd European Colloquium on Information Retrieval Research, Darmstadt, Germany, 2001.
- [54] HARTIGAN, J.A. and WONG, M.A. A K-Means Clustering Algorithm. Applied Statistics 28 (1): 100–108, 1979.
- [55] BARALDI, A. and BLONDA, P.A Survey of Fuzzy Clustering Algorithms for Pattern Recognition – Part II. In IEEE Transactions on systems, Man, and Cybernetics – Part B: Cybernetics, vol. 29, n. 6, 1999.
- [56] MENEZES, P.F.B. Linguagens Formais e Autômatos (Série Matemática da Computação e Processamento Paralelo) Sagra Luzzatto, UFRGS, Referência acessível, em português, com algumas diferenças significativas nas definições e notações, 1997.
- [57] SIPSER, M. Introdução à Teoria da Computação, Thomson, 2007.

- [58] HOPCROFT, J.E.; ULLMAN, J.D. and MOTWANI, R. Introdução à Teoria de Autômatos, Linguagens e Computação, Ed. Campus, 2002.
- [59] Tradução da segunda edição americana da referência, extensamente reformulada em relação à primeira edição. As respostas de diversos exercícios (e outras informações sobre o livro) se encontram em <http://www-db.stanford.edu/~ullman/ialc.html>.
- [60] JARGAS, A.M. Expressões Regulares - Uma Abordagem Divertida Editora: Novatec, ISBN: 85-7522-100-0.
- [61] SILVA, E.F.A.; BARROS, F.A. e PRUDÊNCIO, R.B.C. Uma Abordagem de Aprendizagem Híbrida para Extração de Informação em Textos Semi-Estruturados, ENIA, 2005.
- [62] GAMMA, E.; HELM, R.; JOHNSON, R. and VLISSIDES, J. Design Patterns: Elements of Reusable Object-Oriented Design Reading, Addison-Wesley, 1995.
- [63] RILLOF, E.; WIEB, J. and PHILIPS, W. Exploiting subjectivity classification to improve information extraction, In AAI, 2005.
- [64] PANG, B. and LEE, L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts Proc. of the ACL, 2004.