

REFERENCES

- [1] T. Tsai, W. Chou, W. Shih-Hung, S. Ting-Yi, S. Sarawagi, H. Jieh, H. Wen-Lian, "Integrating Linguistic Knowledge into a Conditional Random Field Framework to Identify Biomedical Named Entities," *Expert Systems with Applications*, 2006, pp. 117-128.
- [2] S. Irena, A. Sophia, M. John and K. Anand, "Text mining and ontology in biomedicine: making sense of raw text," *Briefings in Bioinformatics*, Vol. 6, No. 3. , 01 September 2005, pp. 239-251.
- [3] B. Markus, D. Mathaeus, S. Martin, T. Volker and K. Hans-Peter, "Extraction of semantic biomedical relations from text using conditional random fields," *BMC Bioinformatics*, 2008, **9**:207.
- [4] California Institute of Technology, TexPresso, [Online]. Available <http://www.textpresso.org/>, (22 June 2008).
- [5] Chilibot, Mining PubMed for relationships, [Online]. Available <http://www.chilibot.net/>, (22 June 2008).
- [6] UBC Bioinformatics Center, protein database, [Online]. Available. <http://bioinformatics.ubc.ca/>, (23 June 2008).
- [7] NCBI, literature mining, [Online]. Available. <http://www.ncbi.nlm.nih.gov/pubmed/16842628>, 23 June 2008).
- [8] B. C. Andrew, "Computational-Linguistic Approaches to Biological Text Mining", School of Crystallography Birkbeck, University of London, Malet Street London, United Kingdom, May, 2008.

- [9] M. Krallinger and A. Valencia, "Text-mining and information-retrieval services for molecular biology", *Genome Biology*, 2005, 6:224.
- [10] S. Staab, C. Blaschke, C. Nedellec, J. Park, B. Schatz, A. Valencia, L. Bernardi, E. Ratsch, R. Kania, J. Saric, I. Rojas, S. Staab, "Mining information for functional genomics", *IEEE Intelligent Systems* 2002, 17:66-80.
- [11] A. M. Cohen and W. R. Hersh, "A survey of current work in biomedical text mining", *BRIEFINGS IN BIOINFORMATICS*, 2005 VOL 6. NO 1, pp.57–71.
- [12] L. Tanabe and W. J. Wilbur, "Tagging gene and protein names in biomedical text", *Bioinformatics*, 2002, Vol. 18(8), pp. 1124–1132.
- [13] B. Settles, "Biomedical Named Entity Recognition Using Conditional Random Fields and Rich Feature Sets," *In Proceedings of the COLING 2004 International Joint Workshop on Natural Language Processing in Biomedicine and its Applications (NLPBA)*. Geneva, Switzerland, 2004, pp.104-107.
- [14] B. Settles, "ABNER: an open source tool for automatically tagging genes, Proteins and other entity names in text", *BIOINFORMATICS APPLICATIONS NOTE*, April 28, 2005, ;21(14):3191-2.
- [15] Y. Miyao, T. Ohta, K. Masuda, Y. Tsuruoka, K. Yoshida, T. Ninomiya, J. Tsujii, "Semantic retrieval for the accurate identification of relational concepts in massive textbases", *In Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics*, 2006, Vol.2, pp. 1017–1024.
- [16] J. Cowie, and W. Lehnert, "Information extraction," *Commun. ACM*; 1996, Vol.39: pp.80-91.

- [17] U.Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledgediscovery: An Overview", *In Advances in Knowledge Discovery and Data Mining, U MIT Press,Cambridge,* Mass., 1996, pp.1-36.
- [18] I. Spasic, S. Ananiadou, J. McNaught, et al, "Text mining and ontologies in biomedicine: making sense of raw text," *Briefings in Bioinformatics*, Vol. 6:239, 2005.
- [19] K. Papineni, "Why inverse document frequency?," *In Proceeding of North American Association for Computational Linguistics*, 2001, pp. 25–32.
- [20] G. Coombe, "Eigenvalues/Eigenvectors, PCA, and SVD", the University of North Carolina at Chapel Hill, 31 Oct 2006, pp.26-30.
- [21] D. D. Lewis, and J. Catlett, "Heterogeneous uncertainty sampling for supervised learning," *In Proceeding ICML-94, 11th Int. Conf. on Machine Learning* (ed. W. W. Cohen and H. Hirsh), San Francisco, CA: Morgan Kaufmann, 1994, pp. 148–156.
- [22] M. W. Berry, and M. Browne, "Understanding Search Engines: Mathematical Modeling and Text Retrieval," *Society for Industrial and Applied Mathematics*. Philadelphia, PA: 1999.
- [23] P. Baldi, P. Frasconi, P. Smyth. "Modeling the Internet and the Web Probabilistic Methods and Algorithms, 2003, pp. 93.
- [24] G. Salton, and M. J. McGill, "Introduction to modern information retrieval," *McGraw-Hill*, 1983.
- [25] T. M. Cover, and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inform.Theory* Vol.13, 1967, pp.21–27.

- [26] E. H. Han, G. Karypis, and V. Kumar, "Text categorization using weight-adjusted k -nearest neighbor classification," *In Proc. PAKDD-01, 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining* (ed. D. Cheung, Q. Li and G. Williams). *Lecture Notes in Computer Science Series*, Springer, Vol. 2035, 2001, pp. 53–65.
- [27] T. Hastie, R. Tibshirani, and J. Friedman, "Elements of Statistical Learning: Data Mining, Inference, and Prediction," The second edition, Springer, 2001.
- [28] A.Y. Ng, and M. I. Jordan, "On discriminative vs generative classifiers: a comparison of logistic regression and Naive Bayes," *Advances in Neural Information Processing Systems Proc. Neural Information Processing Systems (NIPS) Conference*. 200, pp.169-187.
- [29] P. Domingos and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss", *Machine Learning*, Vol.29, 1997, pp.103–130.
- [30] A. McCallum, K. Nigam, and L. H. Ungar, "Efficient clustering of high-dimensional data sets with application to reference matching," *In Proc. 6th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, NewYork: ACM Press, 2000b, pp. 169–178.
- [31] A. McCallum, and K. Nigam, "A comparison of event models for naive Bayes text classification," *AAAI/ICML-98 Workshop on Learning for Text Categorization*, MenloPark, CA: AAAI Press, 1998, pp. 41–48.
- [32] D. D. Lewis, "Naive Bayes at forty: the independence assumption in information retrieval," *In Proc. 10th Eur. Conf. on Machine Learning*, Springer, 1998 pp. 4–15.

- [33] N. Friedman, and M. Goldszmidt, "Learning Bayesian networks with local structure," *In Proc. 12th Conf. on Uncertainty in Artificial Intelligence, Portland, Oregon* (ed. E. Horwitz and F. Jensen), San Francisco, CA: Morgan Kaufmann, 1996, pp. 274–282.
- [34] M. Pazzani, 1996 "Searching for dependencies in Bayesian classifiers", *In Proc. 5th Int. Workshop on Artificial Intelligence and Statistics*, pp. 239–248.
- [35] C. Cortes, and V. N. Vapnik, "Support vector networks," *Machine Learning* , Vol.20, 1995, pp.1–25.
- [36] T.Evgeniou, M. Pontil, and T. Poggio, "Regularization networks and support vector machines," *Adv. Comput. Math.*, 2000, Vol.13, pp.1–50.
- [37] V. N. Vapnik, "Statistical Learning Theory," *JohnWiley & Sons, Ltd/Inc*,1998.
- [38] B. Schölkopf and A. Smola, "Learning with Kernels," Cambridge, MA: MIT Press, 2002.
- [39] P., M. Pontil, and P. Frasconi, "From margins to probabilities in multiclass learning problems". *In Proc. 15th Eur. Conf. on Artificial Intelligence* (ed. F. van Harmelen). Amsterdam: IOS Press. 2002 p.146.
- [40] T. Mitchell, "Machine Learning," McGraw-Hill, New York, 1997.
- [41] A.K. Jain, R. Duin, J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4–37 (2000).
- [42] D.Mladenic, and M. Grobelnik, "Feature selection for unbalanced class distribution and Naïve Bayes," *In Proceeding of the 16th Int. Conf. on Machine Learning*. San Francisco: Morgan Kaufmann, 1999, pp. 258–267.

- [43] M.J Zaki, N. Ramakrishnan and S. Parthasarathy, “Biological data mining. Scientific Programming,” *IOS Press*, 2008, Vol.16: pp.3–3.
- [44] M Frederick, L.W. David, “Applied Bayesian and Classical Inference: the Case of the Federalist Papers. *Springer*, New York, 1984.
- [45] B.J Stapley, G Benoit, “Biobibliometrics: information retrieval and visualization form co-occurrences of gene names in Medline abstracts,” *In Proceedings of the Pacific Symposium on Biocomputing*, 2000, pp.529-540.
- [46] R. Hoffmann and A. Valencia, “A Gene Network for Navigating the Literature,” *A. Nature Genetics*, Valencia, 2004, Vol.36: 664.
- [47] E.M. Marcotte, I. Xenarios, D. Eisenberg , “Mining literature for protein-protein interactions,” *Bioinformatics*, Vol.17-4, 2001, pp.359-363.
- [48] Z. GuoDong, S. Jian, “Exploring deep knowledge resources in biomedical name recognition,” *In Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Application*, 2004, pp.96–99.
- [49] J.D. Kim, “Introduction to the Biological Entity Recognition Task of the JNLPBA workshop” *In Proc. of the JNPBA*, 2004, pp.70-76.
- [50] Y. Sasaki, Y. Tsuruoka¹, J. McNaught, “How to make the most of NE dictionaries in statistical NER,” *BMC bioinformatics*, Vol.9: S5, 2008.
- [51] H. M. Wallach, “Conditional random fields: An introduction”, Technical Report MS-CIS-04-01, University of Pennsylvania, 2004.
- [52] U. Leser, J Hakenberg, “What makes a gene name? Named entity recognition in the biomedical literature,” *Briefings in Bioinformatics*, 2005, Vol.6: 357.

- [53] C. M. Friedrich, T. Revillion, M. Hofmann, “Biomedical and chemical named entity recognition with conditional random fields: The advantage of dictionary features,” *In Proceedings of the Second International Symposium on Semantic Mining in Biomedicine (SMBM 2006)*, 2006, pp.85–89.
- [54] C. Haruechiyasak, “Word Segmentation and Named Entities Recognition”, *A Driving Force for National Science and Technology Capability*, 2008.
- [55] H. Wang, T. Zhao, H. Tan, et al., “Biomedical Named Entity Recognition Based on Classifiers Ensemble,” *International Journal of Computer Science and Applications*, 2008, Vol.5:1–11.
- [56] S. Ananiadou, C. Friedman and J. Tsujii, “J Biomed Inform, Special issue: Named entity recognition in biomedicine”, 2004, Vol. 37 (6), 2004.
- [57] A. Ekbal and S. Bandyopadhyay, “Named Entity Recognition using Support Vector Machine: A Language Independent Approach,” *International Journal of Electrical and Electronics Engineering*, 2010.
- [58] L.R. Rabiner and B.H. Juang, “A tutorial on hidden Markov models,” *Proceedings of the IEEE*, Vol. 77, 1989, pp. 257–286.
- [59] P. Blunsom, “Maximum Entropy Markov Models for Semantic Role Labelling,” *In Proceedings of the Australasian Language Technology Workshop*, 2004, pp.109-116.
- [60] J.D. Lafferty, A. McCallum, and F.C.N. Pereira, “Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data,” *Proceedings of the Eighteenth International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., 2001, pp. 282-289.
- [61] C. Sutton and A. McCallum, “An Introduction to Conditional Random Fields for

- Relational Learning,” *Introduction to Statistical Relational Learning*, MIT Press, 2006.
- [62] C. Zhu, R.H. Byrd, P. Lu, and J. Nocedal, “L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization,” *ACM Transactions on Mathematical Software*, Vol. 23, 1997, pp. 550–560.
- [63] H. Wallach, “Efficient training of conditional random fields,” *Proc. 6th Annual CLUK Research Colloquium*, 2002.
- [64] V. Della Pietra and J. Lafferty, “Inducing Features of Random Fields,” *Arxiv preprint cmlg9506014*, pp. 34.
- [65] R.C. Berwick, S.P. Abney, and C. Tenny, *Principle-based parsing: computation and psycholinguistics*, Springer, 1991.
- [66] A. McCallum, D. Freitag, and F. Pereira, “Maximum entropy Markov models for information extraction and segmentation,” *Proceedings of the Seventeenth International Conference on Machine Learning*, 2000, pp. 591–598.
- [67] A. Culotta and A. McCallum, “Confidence estimation for information extraction,” *Proceedings of Human Language Technology Conference and North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, 2004.
- [68] D. Pinto, A. McCallum and Wei X, “Table extraction using conditional random fields”, *In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, 2003, pp.235–242.
- [69] J.D. Kim, “GENIA corpus semantically annotated corpus for Biotextmining,” *Bioinformatics* 19, 2003, i180–i182.
- [70] R. McDonald, F. Pereira, “Identifying gene and protein mentions in text using

- conditional random fields,” *BMC bioinformatics*, Vol. 6:S6, 2005.
- [71] I. Spasic, S. Ananiadou, J. McNaught, et al, “Text mining and ontologies in biomedicine: making sense of raw text,” *Briefings in Bioinformatics*, Vol. 6:239, 2005.
- [72] J. Finkel, S. Dingare, H. Nguyen, et al., “Exploiting Context for Biomedical Entity Recognition: From Syntax to the Web[A],” *In Proceedings of the Joint Workshop on Natural Language Processing in Biomedicine and its Applications (JNLPBA-2004) [C]*, Geneva, Switzerland, 2004.
- [73] A. Baronchella, E. Caglioti, V. Loreto, E. Pizzic, Dictionary-based methods for information extraction, *Physica A* 342 , 2004, pp.294 – 300.
- [74] S. Ananiadou, S. Pyysalo, J. I. Tsujii and D. B. Kell, “Event extraction for systems biology by text mining the literature”, *Trends in Biotechnology*, Volume 28, Issue 7, 2010, pp.381-390.
- [75] A. N      , W. Kim, W. J. Wilbur and Z. Lu, “Exploring two biomedical text genres for disease recognition”, *BioNLP '09 Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, 2009, pp. 144-152.
- [76] J. Bj     , F. Ginter , J. Heimonen , S. Pyysalo and T. Salakoski, “Learning to Extract Biological Event and Relation Graphs”, *In the Proceedings of NODALIDA*, 2009, pp. 18–25.
- R. Hudson, “Language Networks: the New Word Grammar”. Oxford: Oxford University Press. Pp xii + 275. ISBN-13: 978-0-19-926730-9, 2007.
- [77] Q. Uwe, W. Christian, “The Poisson collocation measure and its applications,” *In Proceedings International Workshop on Computational Approaches to Collocations*, 2002.

- [78] F. Smadja, "Retrieving collocations from text: Xtract," *Computational linguistics*, vol. 19, 1993, pp.143–177.
- [79] J. P. Goldman E. Wehrli, "FipsCo: Asyntax-based system for terminology extraction," *In Proceedings of Grammar and Natural Language Processing Conference*, 2001.
- [80] W. Wong, W. Liu, M. Bennamoun, "A probabilistic framework for automatic term recognition," *Intelligent Data Analysis IOS Press (13)*, 2009, pp.499–539.
- [81] Sareneva T, Matikainen S, Vanhatalo J, Mel'en K, Pelkonen J, Julkunen I, Kinetics of cytokine and NFAT gene expression in human interleukin-2-dependent T lymphoblasts stimulated via T-cell receptor. *Immunology* **93**, 1998, 350.
- [82] Liu S, Yu Y, Zhang M, Wang W, Cao X (2001) The involvement of NFalpha-related apoptosis-inducing ligand in the enhanced cytotoxicity of IFN-beta-stimulated human dendritic cells to tumor cells. *The Journal of Immunology* **166**, 5407–5415.
- [83] Hamilton T. A, Fishman M (1982) Activated macrophages selectively bind both normal and neoplastic lymphoblasts but not quiescent lymphocytes. *Cellular Immunology*, 72.
- [84] Lippman M. E, Yarbrow G. K, Leventhal B. G, "Clinical implications of glucocorticoid receptors in human leukemia", *Cancer research*, 1978, pp.38.
- [85] Golikov P. P, Bobkova A. S(1981) Glucocorticoid receptor level in the blood leukocytes in different acute diseases. *Vopr Med Khim* **27**, 654. [96]

- [86] Armanini D, Strasser T, Weber P. C (1985) Characterization of aldosterone binding sites in circulating human mononuclear leukocytes. The American journal of physiology.
- [87] Jehle P. M, Lutz M. P, Fussgaenger R. D (1996) High affinity binding sites for proinsulin in human IM-9 lymphoblasts. Diabetologia 39, 421-432.
- [88] Akane Yakushiji, Yuka Tateisi, Yusuke Miyao, Use of a Full Parser for Information Extraction in Molecular Biology Domain, Genome Informatics 11: 446–447 (2000).
- [89] K. Fundel, R. Küffner, R. Zimmer and Satoru Miyano, RelEx–Relation extraction using dependency parse trees, Bioinformatics 2007, Vol. 23 no. 3 2007, pages 365–371.
- [90] H. Y. Kao, Y. T. Tang, J. F. Wang, Evolutional Dependency Parse Trees for Biological Relation Extraction Bioinformatics and Bioengineering (BIBE), 2011 IEEE 11th International Conference on Date of Conference: 24-26 Oct. 2011, pp:167 – 174.
- [91] J. D. Kim, T. Ohta, S. Pyysalo, Y. Kano, J. Tsujii, “Overview of BioNLP'09 Shared Task on Event Extraction,” *In Proceedings of the BioNLP 2009 Workshop Companion Volume for Shared Task (June 2009)*, 2009. pp. 1-9.
- [92] M. EACL and A. Moschitti, “Making tree kernels practical for natural language learning” *In Proceedings of the Eleventh International Conference on European Association for Computational Linguistics*, Trento, Italy, 2006.
- [93] A. Moschitti, D. Pighin and R. Basili, “Semantic Role Labeling via Tree Kernel joint inference”, *In Proceedings of the 10th Conference on Computational Natural Language Learning*, New York, USA, 2006, pp.61-68.