

bio

October 12, 2009

References

- [1] R. Aebersold and M. Mann. Mass spectrometry-based proteomics. *Nature*, 422(6928):198–207, Mar 2003.
- [2] J. Aires-de Sousa and J. Gasteiger. Prediction of enantiomeric excess in a combinatorial library of catalytic enantioselective reactions. *J Comb Chem*, 7(2):298–301, 2005.
- [3] T. Akutsu, S. Miyano, and S. Kuhara. Inferring qualitative relations in genetic networks and metabolic pathways. *Bioinformatics*, 16(8):727–734, 2000.
- [4] M. Alexandersson, S. Cawley, and L. Pachter. SLAM: cross-species gene finding and alignment with a generalized pair hidden Markov model. *Genome Res.*, 13(3):496–502, Mar 2003.
- [5] C.F. Aliferis, D.P. Hardin, and P. Massion. Machine Learning Models For Lung Cancer Classification Using Array Comparative Genomic Hybridization. In *Proceedings of the 2002 American Medical Informatics Association (AMIA) Annual Symposium*, pages 7–11, 2002.
- [6] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25:3389–3402, 1997.
- [7] C. Ambroise and G.J. McLachlan. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc. Natl. Acad. Sci. USA*, 99(10):6562–6566, 2002.
- [8] D.C. Anderson, W. Li, D.G. Payan, and W.S. Noble. A new algorithm for the evaluation of shotgun peptide sequencing in proteomics: support vector machine classification of peptide MS/MS spectra and SEQUEST scores. *J Proteome Res*, 2(2):137–146, 2003.

- [9] Yindalon Aphinyanaphongs, Ioannis Tsamardinos, Alexander Statnikov, Douglas Hardin, and Constantin F Aliferis. Text categorization models for high-quality article retrieval in internal medicine. *J. Am. Med. Inform. Assoc.*, 12(2):207–16, 2005.
- [10] Rieko Arimoto, Madhu-Ashni Prasad, and Eric M Gifford. Development of CYP3A4 inhibition models: comparisons of machine-learning techniques and molecular descriptors. *J Biomol Screen*, 10(3):197–205, Apr 2005.
- [11] Tomasz Arodz, Marcin Kurdziel, Erik O D Sevre, and David A Yuen. Pattern recognition techniques for automatic detection of suspicious-looking anomalies in mammograms. *Comput. Methods Programs Biomed.*, 79(2):135–49, Aug 2005.
- [12] V. Atalay and R. Cetin-Atalay. Implicit motif distribution based hybrid computational kernel for sequence classification. *Bioinformatics*, 21(8):1429–1436, Apr 2005.
- [13] A. M. Bagirov, B. Ferguson, S. Ivkovic, G. Saunders, and J. Yearwood. New algorithms for multi-class cancer diagnosis using tumor gene expression signatures. *Bioinformatics*, 19(14):1800–7, Sep 2003.
- [14] P. Baldi, S. Brunak, P. Frasconi, G. Soda, and G. Pollastri. Exploiting the past and the future in protein secondary structure prediction. *Bioinformatics*, 15:937–946, 1999.
- [15] Mukesh Bansal, Vincenzo Belcastro, Alberto Ambesi-Impiombato, and Diego di Bernardo. How to infer gene networks from expression profiles. *Mol Syst Biol*, 3:78, 2007.
- [16] L. Bao and Z. Sun. Identifying genes related to drug anticancer mechanisms using support vector machine. *FEBS Lett.*, 521:109–114, 2002.
- [17] Lei Bao. Identifying genes related to chemosensitivity using support vector machine. *Methods Mol Med*, 111:233–40, 2005.
- [18] Lei Bao and Yan Cui. Prediction of the phenotypic effects of non-synonymous single nucleotide polymorphisms using structural and evolutionary information. *Bioinformatics*, 21(10):2185–90, May 2005.
- [19] A.-L. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286:509–512, 1999.
- [20] G. Baudat and F. Anouar. Generalized discriminant analysis using a kernel approach. *Neural Comput*, 12(10):2385–404, Oct 2000.
- [21] C. Baumgartner, C. Bohm, D. Baumgartner, G. Marini, K. Weinberger, B. Olgemoller, B. Liebl, and A. A. Roscher. Supervised machine learning techniques for the classification of metabolic disorders in newborns. *Bioinformatics*, 20(17):2985–2996, 2004.

- [22] A. Bazzani, A. Bevilacqua, D. Bollini, R. Brancaccio, R. Campanini, N. Lanconelli, A. Riccardi, and D. Romani. An SVM classifier to separate false signals from microcalcifications in digital mammograms. *Phys Med Biol*, 46(6):1651–63, Jun 2001.
- [23] M. J. Beal, F. Falciani, Z. Ghahramani, C. Rangel, and D. L. Wild. A Bayesian approach to reconstructing genetic regulatory networks with hidden factors. *Bioinformatics*, 21(3):349–356, Feb 2005.
- [24] N. Beerenwinkel, T. Lengauer, M. Daumer, R. Kaiser, H. Walter, K. Korn, D. Hoffmann, and J. Selbig. Methods for optimizing antiviral combination therapies. *Bioinformatics*, 19(Suppl. 1):i16–i25, 2003.
- [25] N. Beerenwinkel, B. Schmidt, H. Walter, R. Kaiser, T. Lengauer, D. Hoffman, K. Korn, and J. Selbig. Geno2pheno: Interpreting Genotypic HIV Drug Resistance Tests. *IEEE Intelligent Systems*, 6(6):35–41, 2001.
- [26] A. Ben-Dor, L. Bruhn, N. Friedman, I. Nachman, M. Schummer, and Z. Yakhini. Tissue classification with gene expression profiles. *J. Comput. Biol.*, 7(3-4):559–583, 2000.
- [27] A. Ben-Hur and D. Brutlag. Remote homology detection: a motif based approach. *Bioinformatics*, 19(Suppl. 1):i26–i33, 2003.
- [28] A. Ben-Hur and W. S. Noble. Kernel methods for predicting protein-protein interactions. *Bioinformatics*, 21(Suppl. 1):i38–i46, Jun 2005.
- [29] Monica Benito, Joel Parker, Quan Du, Junyuan Wu, Dong Xiang, Charles M Perou, and J. S. Marron. Adjustment of systematic microarray data biases. *Bioinformatics*, 20(1):105–14, Jan 2004.
- [30] M. Bern, D. Goldberg, W. H. McDonald, and III Yates, J. R. Automatic Quality Assessment of Peptide Tandem Mass Spectra. *Bioinformatics*, 20(Suppl. 1):i49–i54, 2004.
- [31] M. Bhasin and G. P. S. Raghava. Analysis and prediction of affinity of TAP binding peptides using cascade SVM. *Protein Sci.*, 13(3):596–607, Mar 2004.
- [32] M. Bhasin and G. P. S. Raghava. Classification of Nuclear Receptors Based on Amino Acid Composition and Dipeptide Composition. *J. Biol. Chem.*, 279(22):23262–23266, 2004.
- [33] M. Bhasin and G. P. S. Raghava. ESLpred: SVM-based method for sub-cellular localization of eukaryotic proteins using dipeptide composition and PSI-BLAST. *Nucl. Acids Res.*, 32(Suppl. 2):W414–419, 2004.
- [34] M. Bhasin and G. P. S. Raghava. GPCRpred: an SVM-based method for prediction of families and subfamilies of G-protein coupled receptors. *Nucl. Acids Res.*, 32(Suppl.2):W383–389, 2004.

- [35] M. Bhasin and G. P. S. Raghava. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine*, 22(23-24):3195–3204, 2004.
- [36] M. Bhasin and G. P. S. Raghava. SVM based method for predicting HLA-DRB1*0401 binding peptides in an antigen sequence. *Bioinformatics*, 20(3):421–423, 2004.
- [37] M. Bhasin and G. P. S. Raghava. GPCRclass: a web tool for the classification of amine type of G-protein-coupled receptors. *Nucleic Acids Res.*, 33(Web Server issue):W143–7, Jul 2005.
- [38] M. Bhasin and G. P. S. Raghava. Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. *Nucleic Acids Res*, 33(Web Server issue):W202–7, Jul 2005.
- [39] S. Bhavani, A. Nagargadde, A. Thawani, V. Sridhar, and N. Chandra. Substructure-based support vector machine classifiers for prediction of adverse effects in diverse classes of drugs. *J Chem Inf Model*, 46(6):2478–2486, 2006.
- [40] J. R. Bock and D. A. Gough. Predicting protein-protein interactions from primary structure. *Bioinformatics*, 17(5):455–460, 2001.
- [41] J. R. Bock and D. A. Gough. A New Method to Estimate Ligand-Receptor Energetics. *Mol Cell Proteomics*, 1(11):904–910, 2002.
- [42] J. R. Bock and D. A. Gough. Whole-proteome interaction mining. *Bioinformatics*, 19(1):125–134, 2003.
- [43] A. Boobis, U. Gundert-Remy, P. Kremers, P. Macheras, and O. Pelkonen. In silico prediction of ADME and pharmacokinetics. Report of an expert meeting organised by COST B15. *Eur J Pharm Sci*, 17(4-5):183–193, Dec 2002.
- [44] Andrew J Bordner and Ruben Abagyan. Statistical analysis and prediction of protein-protein interfaces. *Proteins*, 60(3):353–66, Aug 2005.
- [45] K.M. Borgwardt, C.S. Ong, S. Schönauer, S.V.N. Vishwanathan, A.J. Smola, and H.-P. Kriegel. Protein function prediction via graph kernels. *Bioinformatics*, 21(Suppl. 1):i47–i56, Jun 2005.
- [46] H. R. Bosshard. Molecular recognition by induced fit: how fit is the concept? *News Physiol Sci*, 16:171–173, Aug 2001.
- [47] Christopher Bowd, Kwokleung Chan, Linda M Zangwill, Michael H Goldbaum, Te-Won Lee, Terrence J Sejnowski, and Robert N Weinreb. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Invest Ophthalmol Vis Sci*, 43(11):3444–54, Nov 2002.

- [48] Christopher Bowd, Linda M Zangwill, Felipe A Medeiros, Jiucang Hao, Kwokleung Chan, Te-Won Lee, Terrence J Sejnowski, Michael H Goldbaum, Pamela A Sample, Jonathan G Crowston, and Robert N Weinreb. Confocal scanning laser ophthalmoscopy classifiers and stereophotograph evaluation for prediction of visual field abnormalities in glaucoma-suspect eyes. *Invest Ophthalmol Vis Sci*, 45(7):2255–62, Jul 2004.
- [49] James R Bradford and David R Westhead. Improved prediction of protein-protein binding sites using a support vector machines approach. *Bioinformatics*, 21(8):1487–94, Apr 2005.
- [50] Hans Briem and Judith Günther. Classifying ”kinase inhibitor-likeness” by using machine-learning methods. *Chembiochem*, 6(3):558–66, Mar 2005.
- [51] M. P. Brown, W. N. Grundy, D. Lin, N. Cristianini, C. W. Sugnet, T. S. Furey, M. Ares, and D. Haussler. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc. Natl. Acad. Sci. USA*, 97(1):262–7, Jan 2000.
- [52] V. Brusic, N. Petrovsky, G. Zhang, and V. B. Bajic. Prediction of promiscuous peptides that bind HLA class I molecules. *Immunol. Cell Biol.*, 80(3):280–285, Jun 2002.
- [53] H.-H. Bui, A. J. Schiewe, H. von Grafenstein, and I. S. Haworth. Structural prediction of peptides binding to MHC class I molecules. *Proteins*, 63(1):43–52, Apr 2006.
- [54] Huynh-Hoa Bui, John Sidney, Bjoern Peters, Muthuraman Sathiamurthy, Asabe Sinichi, Kelly-Anne Purton, Bianca R Mothé, Francis V Chisari, David I Watkins, and Alessandro Sette. Automated generation and evaluation of specific mhc binding predictive tools: Arb matrix applications. *Immunogenetics*, 57(5):304–314, Jun 2005.
- [55] R. Bunescu, R. Ge, R. J. Kate, E. M. Marcotte, R. J. Mooney, A. K. Ramani, and Y. W. Wong. Comparative experiments on learning information extractors for proteins and their interactions. *Artif. Intell. Med.*, 33(2):139–55, Feb 2005.
- [56] R. Burbidge, M. Trotter, B. Buxton, and S. Holden. Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Comput. Chem.*, 26(1):4–15, December 2001.
- [57] T. Burckin, R. Nagel, Y. Mandel-Gutfreund, L. Shiue, T. A. Clark, J.-L. Chong, T.-H. Chang, S. Squazzo, G. Hartzog, and M. Ares. Exploring functional relationships between components of the gene expression machinery. *Nat. Struct. Mol. Biol.*, 12(2):175–82, Feb 2005.

- [58] S. Busuttill, J. Abela, and G. J. Pace. Support vector machines with profile-based kernels for remote protein homology detection. *Genome Inform Ser Workshop Genome Inform*, 15(2):191–200, 2004.
- [59] E. Byvatov, U. Fechner, J. Sadowski, and G. Schneider. Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *J Chem Inf Comput Sci*, 43(6):1882–9, 2003.
- [60] E. Byvatov and G. Schneider. Support vector machine applications in bioinformatics. *Appl Bioinformatics*, 2(2):67–77, 2003.
- [61] Evgeny Byvatov and Gisbert Schneider. SVM-based feature selection for characterization of focused compound collections. *J Chem Inf Comput Sci*, 44(3):993–9, 2004.
- [62] C. Z. Cai, L. Y. Han, Z. L. Ji, X. Chen, and Y. Z. Chen. SVM-Prot: Web-based support vector machine software for functional classification of a protein from its primary sequence. *Nucleic Acids Res*, 31(13):3692–7, Jul 2003.
- [63] C.Z. Cai, L.Y. Han, Z.L. Ji, and Y.Z. Chen. Enzyme family classification by support vector machines. *Proteins*, 55(1):66–76, 2004.
- [64] C.Z. Cai, W.L. Wang, L.Z. Sun, and Y.Z. Chen. Protein function classification via support vector machine approach. *Math. Biosci.*, 185(2):111–122, 2003.
- [65] Y.-D. Cai, X.-J. Liu, X.-B. Xu, and K.-C. Chou. Support vector machines for prediction of protein subcellular location by incorporating quasi-sequence-order effect. *J. Cell. Biochem.*, 84(2):343–348, 2002.
- [66] Y.-D. Cai, X.-J. Liu, X.-B. Xu, and G.-P. Zhou. Support Vector Machines for predicting protein structural class. *BMC Bioinformatics*, 2(3):3, 2001.
- [67] Y.-D. Cai, X.-J. Liu, X.-B. Xu, and G.-P. Zhou. Prediction of protein structural classes by support vector machines. *Comput. Chem.*, 26(3):293–296, 2002.
- [68] Y.-D. Cai, G.-P. Zhou, and K.-C. Chou. Support Vector Machines for Predicting Membrane Protein Types by Using Functional Domain Composition. *Biophys. J.*, 84(5):3257–3263, 2003.
- [69] Y.D. Cai, K.Y. Feng, Y.X. Li, and K.C. Chou. Support vector machine for predicting alpha-turn types. *Peptides*, 24(4):629–630, 2003.
- [70] Y.D. Cai and S.L. Lin. Support vector machines for predicting rRNA-, RNA-, and DNA-binding proteins from amino acid sequence. *Biochim. Biophys. Acta*, 1648(1-2):127–133, 2003.

- [71] Y.D. Cai, S.L. Lin, and K.C. Chou. Support vector machines for prediction of protein signal sequences and their cleavage sites. *Peptides*, 24(1):159–161, 2003.
- [72] Y.D. Cai, X.J. Liu, Y.X. Li, X.B. Xu, and K.C. Chou. Prediction of beta-turns with learning machines. *Peptides*, 24(5):665–669, 2003.
- [73] Y.D. Cai, X.J. Liu, X.B. Xu, and K.C. Chou. Support vector machines for prediction of protein subcellular location. *Mol. Cell Biol. Res. Commun.*, 4(4):230–234, 2000.
- [74] Y.D. Cai, X.J. Liu, X.B. Xu, and K.C. Chou. Support Vector Machines for predicting HIV protease cleavage sites in protein. *J. Comput. Chem.*, 23(2):267–274, 2002.
- [75] Y.D. Cai, X.J. Liu, X.B. Xu, and K.C. Chou. Support vector machines for predicting the specificity of GalNAc-transferase. *Peptides*, 23:205–208, 2002.
- [76] Y.D. Cai, X.J. Liu, X.B. Xu, and K.C. Chou. Support vector machines for the classification and prediction of beta-turn types. *J. Pept. Sci.*, 8(7):297–301, 2002.
- [77] Y.D. Cai, X.J. Liu, X.B. Xu, and K.C. Chou. Support vector machines for prediction of protein domain structural class. *J. Theor. Biol.*, 221(1):115–120, 2003.
- [78] Y.D. Cai, P.W. Ricardo, C.H. Jen, and K.C. Chou. Application of SVM to predict membrane protein types. *J. Theor. Biol.*, 226(4):373–376, 2004.
- [79] Y.D. Cai, G.P. Zhou, C.H. Jen, S.L. Lin, and K.C. Chou. Identify catalytic triads of serine hydrolases by support vector machines. *J. Theor. Biol.*, 228(4):551–557, 2004.
- [80] Yu-Dong Cai and Andrew J Doig. Prediction of *Saccharomyces cerevisiae* protein functional class from functional domain composition. *Bioinformatics*, 20(8):1292–300, May 2004.
- [81] Renato Campanini, Danilo Dongiovanni, Emiro Iampieri, Nico Lanconelli, Matteo Masotti, Giuseppe Palermo, Alessandro Riccardi, and Matteo Roffilli. A novel featureless approach to mass detection in digital mammograms based on support vector machines. *Phys Med Biol*, 49(6):961–75, Mar 2004.
- [82] G. Camps-Valls, A.M. Chalk, A.J. Serrano-Lopez, J.D. Martin-Guerrero, and E.L. Sonnhammer. Profiled support vector machines for antisense oligonucleotide efficacy prediction. *BMC Bioinformatics*, 5(135):135, 2004.

- [83] E. Capriotti, P. Fariselli, and R. Casadio. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.*, 33(Web Server issue):W306–10, Jul 2005.
- [84] R. J. Carter, I. Dubchak, and S. R. Holbrook. A computational approach to identify genes for functional RNAs in genomic sequences. *Nucl. Acids Res.*, 29(19):3928–3938, 2001.
- [85] D. Cavaliere and C. De Filippo. Bioinformatic methods for integrating whole-genome expression results into cellular networks. *Drug Discov Today*, 10(10):727–34, 2005.
- [86] Ian Chan, William Wells, Robert V Mulkern, Steven Haker, Jianqing Zhang, Kelly H Zou, Stephan E Maier, and Clare M C Tempany. Detection of prostate cancer by integration of line-scan diffusion, T2-mapping and T2-weighted magnetic resonance imaging; a multichannel statistical classifier. *Med Phys*, 30(9):2390–8, Sep 2003.
- [87] Kwokleung Chan, Te-Won Lee, Pamela A Sample, Michael H Goldbaum, Robert N Weinreb, and Terrence J Sejnowski. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Trans Biomed Eng*, 49(9):963–74, Sep 2002.
- [88] Y. Chen and D. Xu. Understanding protein dispensability through machine-learning analysis of high-throughput data. *Bioinformatics*, 21:575–581, Mar 2005.
- [89] Y.C. Chen, Y.S. Lin, C.J. Lin, and J.K. Hwang. Prediction of the bonding states of cysteines using the support vector machines based on multiple feature vectors and cysteine state sequences. *Proteins*, 55(4):1036–1042, 2004.
- [90] Vladimir Cherkassky and Yunqian Ma. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw*, 17(1):113–26, Jan 2004.
- [91] K.-C. Chou. Prediction of protein signal sequences and their cleavage sites. *Protein. Struct. Funct. Genet.*, 42:136–139, 2001.
- [92] K.-C. Chou. Using subsite coupling to predict signal peptides. *Protein Eng.*, 14(2):75–79, 2001.
- [93] K.-C. Chou and Y.-D. Cai. Using Functional Domain Composition and Support Vector Machines for Prediction of Protein Subcellular Location. *J. Biol. Chem.*, 277(48):45765–45769, 2002.
- [94] M. L. Chow, E. J. Moler, and I. S. Mian. Identifying marker genes in transcription profiling data using a mixture of feature relevance experts. *Physiol. Genomics*, 5(2):99–111, Mar 2001.

- [95] Kai-Min Chung, Wei-Chun Kao, Chia-Liang Sun, Li-Lun Wang, and Chih-Jen Lin. Radius margin bounds for support vector machines with the RBF kernel. *Neural Comput*, 15(11):2643–81, Nov 2003.
- [96] Jason C Cole, Christopher W Murray, J. Willem M Nissink, Richard D Taylor, and Robin Taylor. Comparing protein-ligand docking programs is difficult. *Proteins*, 60(3):325–332, Aug 2005.
- [97] Nigel Collier and Koichi Takeuchi. Comparison of character-level and part of speech features for name recognition in biomedical texts. *J Biomed Inform*, 37(6):423–35, Dec 2004.
- [98] Ronan Collobert, Samy Bengio, and Yoshua Bengio. A parallel mixture of SVMs for very large scale problems. *Neural Comput*, 14(5):1105–14, May 2002.
- [99] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921, Feb 2001.
- [100] B. Coupez and R. A. Lewis. Docking and scoring—theoretically easy, practically impossible? *Curr Med Chem*, 13(25):2995–3003, 2006.
- [101] J. A. Cuff and G. J. Barton. Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Protein. Struct. Funct. Genet.*, 34:508–519, 1999.
- [102] Q. Cui, T. Jiang, B. Liu, and S. Ma. Esub8: A novel tool to predict protein subcellular localizations in eukaryotic organisms. *BMC Bioinformatics*, 5(66):66, 2004.
- [103] M. Cuturi and J.-P. Vert. A mutual information kernel for strings. In *Proceedings of IJCNN 2004*, pages 1904–1910, 2004.
- [104] M. Cuturi and J.-P. Vert. The context-tree kernel for strings. *Neural Network.*, 18(4):1111–1123, 2005.
- [105] Matthew N Davies and Darren R Flower. Harnessing bioinformatics to discover new vaccines. *Drug Discov Today*, 12(9-10):389–395, May 2007.
- [106] C. Debouck and P. N. Goodfellow. DNA microarrays in drug discovery and development. *Nat Genet*, 21(1 Suppl):48–50, Jan 1999.
- [107] S. Degroeve, B. De Baets, Y. Van de Peer, and P. Rouze. Feature subset selection for splice site prediction. *Bioinformatics*, 18(Suppl. 1):S75–S83, 2002.
- [108] S. Degroeve, Y. Saeys, B. De Baets, P. Rouze, and Y. Van de Peer. SpliceMachine: predicting splice sites from high-dimensional local context representations. *Bioinformatics*, 21:1332–1338, 2005.

- [109] Jayne L Dennis and Karin A Oien. Hunting the primary: novel strategies for defining the origin of tumours. *J Pathol*, 205(2):236–47, Jan 2005.
- [110] M. Deshpande and G. Karypis. Evaluation of Techniques for Classifying Biological Sequences. In *PAKDD '02: Proceedings of the 6th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining*, pages 417–431. Springer Verlag, 2002.
- [111] Vikas Dhingra, Mukta Gupta, Tracy Andacht, and Zhen F Fu. New frontiers in proteomics research: a perspective. *Int J Pharm*, 299(1-2):1–18, Aug 2005.
- [112] Casey Diekman, Wei He, Nagabhushana Prabhu, and Harvey Cramer. Hybrid methods for automated diagnosis of breast tumors. *Anal Quant Cytol Histol*, 25(4):183–90, Aug 2003.
- [113] Frank Dieterle, Silvia Mller-Hagedorn, Hartmut M Liebich, and Gnter Gauglitz. Urinary nucleosides as potential tumor markers evaluated by learning vector quantization. *Artif. Intell. Med.*, 28(3):265–79, Jul 2003.
- [114] C.H.Q. Ding and I. Dubchak. Multi-class protein fold recognition using support vector machines and neural networks. *Bioinformatics*, 17:349–358, 2001.
- [115] Chris Ding and Hanchuan Peng. Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol*, 3(2):185–205, Apr 2005.
- [116] Marko Djordjevic, Anirvan M Sengupta, and Boris I Shraiman. A biophysical approach to transcription factor binding site discovery. *Genome Res.*, 13(11):2381–90, Nov 2003.
- [117] P.D. Dobson and A.J. Doig. Distinguishing enzyme structures from non-enzymes without alignments. *J. Mol. Biol.*, 330(4):771–783, 2003.
- [118] P.D. Dobson and A.J. Doig. Predicting enzyme class from protein structure without alignments. *J. Mol. Biol.*, 345(1):187–199, Jan 2005.
- [119] I. Donaldson, J. Martin, B. de Bruijn, C. Wolting, V. Lay, B. Tuekam, S. Zhang, B. Baskin, G.D. Bader, K. Michalickova, T. Pawson, and C.W.V. Hogue. PreBIND and Textomy - mining the biomedical literature for protein-protein interactions using a support vector machine. *BMC Bioinformatics*, 4(1):11, Mar 2003.
- [120] S. Doniger, T. Hofmann, and J. Yeh. Predicting CNS permeability of drug molecules: comparison of neural network and support vector machine algorithms. *J. Comput. Biol.*, 9(6):849–864, 2002.
- [121] P. Dönnes and A. Elofsson. Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics*, 3(1):25, Sep 2002.

- [122] P. Dönnes and O. Kohlbacher. Integrated modeling of the major events in the MHC class I antigen processing pathway. *Protein Sci.*, 14:2132–2140, Jun 2005.
- [123] Pierre Dönnes and Arne Elofsson. Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics*, 3:25, Sep 2002.
- [124] Iriini A Doytchinova, Pingping Guan, and Darren R Flower. Identifying human MHC supertypes using bioinformatic methods. *J Immunol*, 172(7):4314–4323, Apr 2004.
- [125] S. Dreiseitl, L. Ohno-Machado, H. Kittler, S. Vinterbo, H. Billhardt, and M. Binder. A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *J Biomed Inform*, 34(1):28–36, Feb 2001.
- [126] G. Dror, R. Sorek, and R. Shamir. Accurate identification of alternatively spliced exons using support vector machine. *Bioinformatics*, 21(7):897–901, Apr 2005.
- [127] Anshul Dubey, Matthew J Realf, Jay H Lee, and Andreas S Bommarius. Support vector machines for learning to identify the critical positions of a protein. *J Theor Biol*, 234(3):351–61, Jun 2005.
- [128] Justis P Ehlers and J. William Harbour. NBS1 expression as a prognostic marker in uveal melanoma. *Clin. Cancer Res.*, 11(5):1849–53, Mar 2005.
- [129] S. Ekins, B. Boulanger, P. W. Swaan, and M. A. Z. Hupcey. Towards a new age of virtual ADME/TOX and multidimensional drug discovery. *J Comput Aided Mol Des*, 16(5-6):381–401, 2002.
- [130] I. El-Naqa, Y. Yang, M. N. Wernick, N. P. Galatsanos, and R. M. Nishikawa. A support vector machine approach for detection of microcalcifications. *IEEE Trans Med Imaging*, 21(12):1552–63, Dec 2002.
- [131] B. E. Engelhardt, M. I. Jordan, K. E. Muratore, and S. E. Brenner. Protein Molecular Function Prediction by Bayesian Phylogenomics. *PLoS Comput. Biol.*, 1(5):e45, Oct 2005.
- [132] J. Felsenstein. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, 17:368–376, 1981.
- [133] S. Fields, Y. Kohara, and D. J. Lockhart. Functional genomics. *Proc. Natl. Acad. Sci. USA*, 96:8825–8826, August 1999.
- [134] J. H. Fong, A. E. Keating, and M. Singh. Predicting specificity in bZIP coiled-coil protein interactions. *Genome Biol.*, 5(R11), 2004.
- [135] C. C. Friedel, K. H. V. Jahn, S. Sommer, S. Rudd, H. W. Mewes, and I. V. Tetko. Support vector machines for separation of mixed plant-pathogen EST collections based on codon usage. *Bioinformatics*, 21:1383–1388, 2005.

- [136] N. Friedman, M. Linial, I. Nachman, and D. Pe'er. Using Bayesian networks to analyze expression data. *J. Comput. Biol.*, 7(3-4):601–620, 2000.
- [137] B. Fritz, F. Schubert, G. Wrobel, C. Schwaenen, S. Wessendorf, M. Nessling, C. Korz, R. J. Rieker, K. Montgomery, R. Kucherlapati, G. Mechttersheimer, R. Eils, S. Joos, and P. Lichter. Microarray-based Copy Number and Expression Profiling in Dedifferentiated and Pleomorphic Liposarcoma. *Cancer Res.*, 62(11):2993–2998, 2002.
- [138] T. S. Furey, N. Cristianini, N. Duffy, D. W. Bednarski, M. Schummer, and D. Haussler. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, 16(10):906–914, Oct 2000.
- [139] C. Furlanello, M. Serafini, S. Merler, and G. Jurman. Entropy-based gene ranking without selection bias for the predictive classification of microarray data. *BMC Bioinformatics*, 4(54), 2003.
- [140] Roberto Fdez Galn, Silke Sachse, C. Giovanni Galizia, and Andreas V M Herz. Odor-driven attractor dynamics in the antennal lobe allow for simple and rapid olfactory pattern classification. *Neural Comput*, 16(5):999–1012, May 2004.
- [141] Rajeev Gangal and Pankaj Sharma. Human pol II promoter prediction: time series descriptors and machine learning. *Nucleic Acids Res*, 33(4):1332–6, 2005.
- [142] J. L. Gardy, M. R. Laird, F. Chen, S. Rey, C. J. Walsh, M. Ester, and F. S. L. Brinkman. PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics*, 21(5):617–623, Mar 2005.
- [143] A. Garg, M. Bhasin, and G.P. Raghava. SVM-based method for subcellular localization of human proteins using amino acid compositions, their order and similarity search. *J. Biol. Chem.*, 280(15):14427–32, Apr 2005.
- [144] Deon Garrett, David A Peterson, Charles W Anderson, and Michael H Thaut. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Trans Neural Syst Rehabil Eng*, 11(2):141–4, Jun 2003.
- [145] T. Gärtner, J.W. Lloyd, and P.A. Flach. Kernels and distances for structured data. *Mach. Learn.*, 57(3):205–232, 2004.
- [146] S. Gaudan, H. Kirsch, and D. Rebholz-Schuhmann. Resolving abbreviations to their senses in Medline. *Bioinformatics*, Jul 2005.
- [147] Xijin Ge, Shuichi Tsutsumi, Hiroyuki Aburatani, and Shuichi Iwata. Reducing false positives in molecular pattern recognition. *Genome Inform Ser Workshop Genome Inform*, 14:34–43, 2003.

- [148] U. Gether. Uncovering molecular mechanisms involved in activation of g protein-coupled receptors. *Endocr Rev*, 21(1):90–113, Feb 2000.
- [149] Girosi. An Equivalence Between Sparse Approximation and Support Vector Machines. *Neural Comput*, 10(6):1455–80, Jul 1998.
- [150] Dimitris Glotsos, Panagiota Spyridonos, Dionisis Cavouras, Panagiota Ravazoula, Petroula-Arampantoni Dadioti, and George Nikiforidis. Automated segmentation of routinely hematoxylin-eosin-stained microscopic images by combining support vector machine clustering and active contour models. *Anal Quant Cytol Histol*, 26(6):331–40, Dec 2004.
- [151] Dimitris Glotsos, Panagiota Spyridonos, Panagiotis Petalas, Dionisis Cavouras, Panagiota Ravazoula, Petroula-Arampatoni Dadioti, Ioanna Lekka, and George Nikiforidis. Computer-based malignancy grading of astrocytomas employing a support vector machine classifier, the WHO grading system and the regular hematoxylin-eosin diagnostic staining procedure. *Anal Quant Cytol Histol*, 26(2):77–83, Apr 2004.
- [152] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver. Life with 6000 genes. *Science*, 274:546–567, October 1996.
- [153] Michael H Goldbaum, Pamela A Sample, Kwokleung Chan, Julia Williams, Te-Won Lee, Eytan Blumenthal, Christopher A Girkin, Linda M Zangwill, Christopher Bowd, Terrence Sejnowski, and Robert N Weinreb. Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. *Invest Ophthalmol Vis Sci*, 43(1):162–9, Jan 2002.
- [154] S. M. Gomez, W. S. Noble, and A. Rzhetsky. Learning to predict protein-protein interactions from protein sequences. *Bioinformatics*, 19(15):1875–1881, 2003.
- [155] L. Gordon, A. Y. Chervonenkis, A. J. Gammerman, I. A. Shahmuradov, and V. V. Solovyev. Sequence alignment kernel for recognition of promoter regions. *Bioinformatics*, 19(15):1964–1971, 2003.
- [156] S. Goto, T. Nishioka, and M. Kanehisa. LIGAND: chemical database for enzyme reactions. *Bioinformatics*, 14:591–599, 1998.
- [157] S. Goto, Y. Okuno, M. Hattori, T. Nishioka, and M. Kanehisa. LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res.*, 30:402–404, 2002.
- [158] W. N. Grundy. Family-based Homology Detection via Pairwise Sequence Comparison. In *Proceedings of the Second Annual International Conference on Computational Molecular Biology, March 22-25*, pages 94–100, 1998.

- [159] N. Guelzim, S. Bottani, P. Bourguine, and F. Képès. Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.*, 31:60–63, 2002.
- [160] Y. Guermeur. Combining Discriminant Models with New Multi-Class SVMs. *Pattern Anal. Appl.*, 5(2):168–179, 2002.
- [161] Y. Guermeur, A. Lifschitz, and R. Vert. A kernel for protein secondary structure prediction. In B. Schölkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 193–206. MIT Press, 2004.
- [162] Y. Guermeur, G. Pollastri, A. Elisseeff, D. Zelus, H. Paugam-Moisy, and P. Baldi. Combining protein secondary structure prediction models with ensemble methods of optimal complexity. *Neurocomputing*, 56:305–327, 2004.
- [163] Guodong Guo and Charles R Dyer. Learning from examples in the small sample case: face expression recognition. *IEEE Trans Syst Man Cybern B Cybern*, 35(3):477–88, Jun 2005.
- [164] Hong Guo, Lindsay B Jack, and Asoke K Nandi. Feature generation using genetic programming with application to fault classification. *IEEE Trans Syst Man Cybern B Cybern*, 35(1):89–99, Feb 2005.
- [165] J. Guo, H. Chen, Z. Sun, and Y. Lin. A novel method for protein secondary structure prediction using dual-layer SVM and profiles. *Proteins*, 54(4):738–743, 2004.
- [166] Ting Guo, Yanxin Shi, and Zhirong Sun. A novel statistical ligand-binding site predictor: application to ATP-binding sites. *Protein Eng Des Sel*, 18(2):65–70, Feb 2005.
- [167] T. Gururaja, W. Li, W.S. Noble, D.G. Payan, and D.C. Anderson. Multiple functional categories of proteins identified in an in vitro cellular ubiquitin affinity extract using shotgun peptide sequencing. *J Proteome Res*, 2(394-404):394–404, 2003.
- [168] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Mach. Learn.*, 46(1/3):389–422, Jan 2002.
- [169] S. B. Gktrk, C. Tomasi, B. Acar, C. F. Beaulieu, D. S. Paik, R. B. Jeffrey, J. Yee, and S. Napel. A statistical 3-D pattern processing method for computer-aided detection of polyps in CT colonography. *IEEE Trans Med Imaging*, 20(12):1251–60, Dec 2001.
- [170] Bernard Haasdonk. Feature space interpretation of SVMs with indefinite kernels. *IEEE Trans Pattern Anal Mach Intell*, 27(4):482–92, Apr 2005.

- [171] Torsten Haferlach, Alexander Kohlmann, Susanne Schnittger, Martin Dugas, Wolfgang Hiddemann, Wolfgang Kern, and Claudia Schoch. AML M3 and AML M3 variant each have a distinct gene expression signature but also share patterns different from other genetically defined AML subtypes. *Genes Chromosomes Cancer*, 43(2):113–27, Jun 2005.
- [172] Torsten Haferlach, Alexander Kohlmann, Susanne Schnittger, Martin Dugas, Wolfgang Hiddemann, Wolfgang Kern, and Claudia Schoch. A global approach to the diagnosis of leukemia using gene expression profiling. *Blood*, 106(4):1189–1198, Aug 2005.
- [173] J. Hakenberg, S. Schmeier, A. Kowald, E. Klipp, and U. Leser. Finding kinetic parameters using text mining. *OMICS*, 8(2):131–152, 2004.
- [174] Jrg Hakenberg, Steffen Bickel, Conrad Plake, Ulf Brefeld, Hagen Zahn, Lukas Faulstich, Ulf Leser, and Tobias Scheffer. Systematic feature evaluation for gene name recognition. *BMC Bioinformatics*, 6 Suppl 1:S9, 2005.
- [175] L.Y. Han, C.Z. Cai, Z.L. Ji, Z.W. Cao, J. Cui, and Y.Z. Chen. Predicting functional family of novel enzymes irrespective of sequence similarity: a statistical learning approach. *Nucl. Acids Res.*, 32(21):6437–6444, 2004.
- [176] L.Y. Han, C.Z. Cai, Z.L. Ji, and Y.Z. Chen. Prediction of functional class of novel viral proteins by a statistical learning method irrespective of sequence similarity. *Virology*, 331(1):136–143, 2005.
- [177] L.Y. Han, C.Z. Cai, S.L. Lo, M.C. Chung, and Y.Z. Chen. Prediction of RNA-binding proteins from primary sequence by a support vector machine approach. *RNA*, 10(3):355–368, 2004.
- [178] Sangjo Han, Byung-Chul Lee, Seung Taek Yu, Chan-Seok Jeong, Soyoung Lee, and Dongsup Kim. Fold recognition by combining profile-profile alignment and support vector machine. *Bioinformatics*, 21(11):2667–73, Jun 2005.
- [179] D. Hanisch, A. Zien, R. Zimmer, and T. Lengauer. Co-clustering of biological networks and gene expression data. *Bioinformatics*, 2002.
- [180] J. Harborth, S. M. Elbashir, K. Vandenburgh, H. Manninga, S. A. Scaringe, K. Weber, and T. Tuschl. Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid. Drug. Dev.*, 13(2):83–105, Apr 2003.
- [181] D. Haussler. Convolution Kernels on Discrete Structures. Technical Report UCSC-CRL-99-10, UC Santa Cruz, 1999.

- [182] D. Heckerman. A tutorial on learning with Bayesian networks. In M. Jordan, editor, *Learning in graphical models*, pages 301–354. MIT Press, Cambridge, MA, USA, 1999.
- [183] C. Helma, T. Cramer, S. Kramer, and L. De Raedt. Data mining and machine learning techniques for the identification of mutagenicity inducing substructures and structure activity relationships of noncongeneric compounds. *J. Chem. Inf. Comput. Sci.*, 44(4):1402–11, 2004.
- [184] S. Hochreiter and K. Obermayer. Gene selection for microarray data. In B. Schlkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 319–355. MIT Press, 2004.
- [185] Oliver Hofmann and Dietmar Schomburg. Concept-based annotation of enzyme classes. *Bioinformatics*, 21(9):2059–66, May 2005.
- [186] Y. Hou, W. Hsu, M. L. Lee, and C. Bystroff. Efficient remote homology detection using local structure. *Bioinformatics*, 19(17):2294–2301, 2003.
- [187] Y. Hou, W. Hsu, M. L. Lee, and C. Bystroff. Remote homolog detection using local sequence-structure correlations. *Proteins*, 57(3):518–530, 2004.
- [188] C. Hu, X. Li, and J. Liang. Developing optimal non-linear scoring function for protein design. *Bioinformatics*, 20(17):3080–3098, 2004.
- [189] H.J. Hu, Y. Pan, R. Harrison, and P.C. Tai. Improved protein secondary structure prediction using support vector machine with a new encoding scheme and an advanced tertiary classifier. *IEEE Trans. Nanobioscience*, 3(4):265–271, 2004.
- [190] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty. Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8):1509–1515, Apr 2005. To appear.
- [191] S. Hua and Z. Sun. A Novel Method of Protein Secondary Structure Prediction with High Segment Overlap Measure: Support Vector Machine Approach. *J. Mol. Biol.*, 308(2):397–407, April 2001.
- [192] S. Hua and Z. Sun. Support vector machine approach for protein subcellular localization prediction. *Bioinformatics*, 17(8):721–728, 2001.
- [193] J. Huan, W. Wang, A. Washington, J. Prins, R. Shah, and A. Tropsha. Accurate classification of protein structural families using coherent subgraph analysis. In *Proceedings of the Pacific Symposium on Biocomputing 2002*, pages 411–422, 2004.
- [194] Jing Huang and Feng Shi. Support vector machines for predicting apoptosis proteins types. *Acta Biotheor.*, 53(1):39–47, 2005.

- [195] N. Huang, H. Chen, and Z. Sun. CTKPred: an SVM-based method for the prediction and classification of the cytokine superfamily. *Protein Eng. Des. Sel.*, Jun 2005.
- [196] Shao-Wei Huang and Jenn-Kang Hwang. Computation of conformational entropy from protein sequences using the machine-learning method—application to the study of the relationship between structural conservation and local structural stability. *Proteins*, 59(4):802–9, Jun 2005.
- [197] T. M. Huang and V. Kecman. Gene extraction for cancer diagnosis by support vector machines—An improvement. *Artif. Intell. Med.*, Jul 2005.
- [198] B. Hutter, C. Schaab, S. Albrecht, M. Borgmann, N. A. Brunner, C. Freiberg, K. Ziegelbauer, C. O. Rock, I. Ivanov, and H. Loferer. Prediction of Mechanisms of Action of Antibacterial Compounds by Gene Expression Profiling. *Antimicrob. Agents Chemother.*, 48(8):2838–2844, Aug 2004.
- [199] A. Ifantis and S. Papadimitriou. The nonlinear predictability of the electrotelluric field variations data analyzed with support vector machines as an earthquake precursor. *Int J Neural Syst*, 13(5):315–32, Oct 2003.
- [200] S. Imoto, T. Goto, and S. Miyano. Estimation of genetic networks and functional structures between genes by using Bayesian networks and nonparametric regression. *Pac. Symp. Biocomput.*, pages 175–186, 2002.
- [201] S. Imoto, S. Kim, T. Goto, S. Miyano, S. Aburatani, K. Tashiro, and S. Kuhara. Bayesian network and nonparametric heteroscedastic regression for nonlinear modeling of genetic network. *J. Bioinform. Comput. Biol.*, 1(2):231–252, Jul 2003.
- [202] S. Imoto, K. Sunyong, T. Goto, S. Aburatani, K. Tashiro, S. Kuhara, and S. Miyano. Bayesian network and nonparametric heteroscedastic regression for nonlinear modeling of genetic network. *Proc. IEEE Comput. Soc. Bioinform. Conf.*, 1:219–227, 2002.
- [203] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA*, 98(8):4569–4574, 2001.
- [204] T. Ito, K. Tashiro, S. Muta, R. Ozawa, T. Chiba, M. Nishizawa, K. Yamamoto, S. Kuhara, and Y. Sakaki. Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl. Acad. Sci. USA*, 93(3):1143–1147, 2000.
- [205] T. Jaakkola, M. Diekhans, and D. Haussler. A Discriminative Framework for Detecting Remote Protein Homologies. *J. Comput. Biol.*, 7(1,2):95–114, 2000.

- [206] T. S. Jaakkola, M. Diekhans, and D. Haussler. Using the Fisher Kernel Method to Detect Remote Protein Homologies. In *Proceedings of the Seventh International Conference on Intelligent Systems for Molecular Biology*, pages 149–158. AAAI Press, 1999.
- [207] T. S. Jaakkola and D. Haussler. Exploiting generative models in discriminative classifiers. In *Proc. of Tenth Conference on Advances in Neural Information Processing Systems*, 1999.
- [208] Martin Jambon, Anne Imberty, Gilbert Delage, and Christophe Geourjon. A new bioinformatic approach to detect common 3d sites in protein structures. *Proteins*, 52(2):137–145, Aug 2003.
- [209] R. Jansen, H. Yu, D. Greenbaum, Y. Kluger, N.J. Krogan, S. Chung, A. Emili, M. Snyder, J.F. Greenblatt, and M. Gerstein. A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science*, 302(5644):449–453, 2003.
- [210] Barbara Jarzab, Malgorzata Wiench, Krzysztof Fujarewicz, Krzysztof Simek, Michal Jarzab, Malgorzata Oczko-Wojciechowska, Jan Wloch, Agnieszka Czarniecka, Ewa Chmielik, Dariusz Lange, Agnieszka Pawlaczek, Sylwia Szpak, Elzbieta Gubala, and Andrzej Swierniak. Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications. *Cancer Res.*, 65(4):1587–97, Feb 2005.
- [211] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411:41–42, 2001.
- [212] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási. The large-scale organization of metabolic networks. *Nature*, 407:651–654, 2000.
- [213] S. Jiang-Ning, L. Wei-Jiang, and X. Wen-Bo. Cooperativity of the oxidation of cysteines in globular proteins. *J. Theor. Biol.*, 231(1):85–95, 2004.
- [214] R. N. Jorissen and M. K. Gilson. Virtual screening of molecular databases using a support vector machine. *J Chem Inf Model*, 45(3):549–61, 2005.
- [215] Svava Osk Jnsdttir, Flemming Steen Jrgensen, and Sren Brunak. Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates. *Bioinformatics*, 21(10):2145–2160, May 2005.
- [216] M. Kanehisa. A database for post-genome analysis. *Trends Genet.*, 13:375–376, 1997.
- [217] M. Kanehisa, S. Goto, S. Kawashima, and A. Nakaya. The KEGG databases at GenomeNet. *Nucleic Acids Res.*, 30:42–46, 2002.

- [218] Matthias Kaper, Peter Meinicke, Ulf Grosse-kathoefer, Thomas Lingner, and Helge Ritter. BCI Competition 2003–Data set IIb: support vector machines for the P300 speller paradigm. *IEEE Trans Biomed Eng*, 51(6):1073–6, Jun 2004.
- [219] Izet M Kapetanovic, Simon Rosenfeld, and Grant Izmirlian. Overview of commonly used bioinformatics methods and their applications. *Ann N Y Acad Sci*, 1020:10–21, May 2004.
- [220] R. Karchin, K. Karplus, and D. Haussler. Classifying G-protein coupled receptors with support vector machines. *Bioinformatics*, 18:147–159, 2002.
- [221] R. Karchin, L. Kelly, and A. Sali. Improving functional annotation of non-synonomous SNPs with information theory. *Pac Symp Biocomput*, pages 397–408, 2005.
- [222] Y. Karklin, R. F. Meraz, and S.R. Holbrook. Classification of non-coding RNA using graph representations of secondary structure. *Pac. Symp. Biocomput.*, pages 4–15, 2005.
- [223] P. D. Karp, C. A. Ouzounis, C. Moore-Kochlacs, L. Goldovsky, P. Kaipa, D. Ahren, S. Tsoka, N. Darzentas, V. Kunin, and N. Lopez-Bigas. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Res*, 33(19):6083–9, 2005.
- [224] K. Karplus, C. Barrett, and R. Hughey. Hidden Markov Models for Detecting Remote Protein Homologies. *Bioinformatics*, 14(10):846–856, 1998.
- [225] H. Kashima, K. Tsuda, and A. Inokuchi. Marginalized Kernels between Labeled Graphs. In T. Faucett and N. Mishra, editors, *Proceedings of the Twentieth International Conference on Machine Learning*, pages 321–328, New York, NY, USA, 2003. AAAI Press.
- [226] H. Kashima, K. Tsuda, and A. Inokuchi. Kernels for graphs. In B. Schölkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 155–170. MIT Press, The MIT Press, Cambridge, Massachusetts, 2004.
- [227] E. Kellenberger, J. Rodrigo, P. Muller, and D. Rognan. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins*, 57(2):225–242, Nov 2004.
- [228] P. Kharchenko, D. Vitkup, and G. M. Church. Filling gaps in a metabolic network using expression information. *Bioinformatics*, 20 Suppl 1:I178–I185, Aug 2004.
- [229] H. Kim and H. Park. Protein secondary structure prediction based on an improved support vector machines approach. *Protein Eng.*, 16(8):553–560, Aug 2003.

- [230] H. Kim and H. Park. Prediction of protein relative solvent accessibility with support vector machines and long-range interaction 3D local descriptor. *Proteins*, 54(3):557–562, Feb 2004.
- [231] J. Kim, P.L. Krapivsky, B. Kahng, and S. Redner. Evolving protein interaction networks. E-print cond-mat/0203167, 2001.
- [232] J. H. Kim, J. Lee, B. Oh, K. Kimm, and I. Koh. Prediction of phosphorylation sites using SVMs. *Bioinformatics*, 20(17):3179–3184, 2004.
- [233] K. H. Kim, S. W. Bang, and S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals. *Med Biol Eng Comput*, 42(3):419–27, May 2004.
- [234] T. Kin, K. Tsuda, and K. Asai. Marginalized kernels for RNA sequence data analysis. In R.H. Lathrop, K. Nakai, S. Miyano, T. Takagi, and M. Kanehisa, editors, *Genome Informatics 2002*, pages 112–122. Universal Academic Press, 2002.
- [235] A. Kohlmann, C. Schoch, S. Schnittger, M. Dugas, W. Hiddemann, W. Kern, and T. Haferlach. Pediatric acute lymphoblastic leukemia (ALL) gene expression signatures classify an independent cohort of adult ALL patients. *Leukemia*, 18(1):63–71, 2004.
- [236] A. Koike and T. Takagi. Prediction of protein-protein interaction sites using support vector machines. *Protein Eng. Des. Sel.*, 17(2):165–173, Feb 2004.
- [237] D. Komura, H. Nakamura, S. Tsutsumi, H. Aburatani, and S. Ihara. Multidimensional support vector machines for visualization of gene expression data. *Bioinformatics*, 21(4):439–444, Feb 2005.
- [238] R. Kondor and J.-P. Vert. Diffusion kernels. In B. Schölkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 171–192. MIT Press, 2004.
- [239] R. I. Kondor and J. Lafferty. Diffusion Kernels on Graphs and Other Discrete Input. In *ICML 2002*, 2002.
- [240] Bette Korber, Montiano LaBute, and Karina Yusim. Immunoinformatics comes of age. *PLoS Comput Biol*, 2(6):e71, Jun 2006.
- [241] Zsafia Kote-Jarai, Richard D Williams, Nicola Cattini, Maria Copeland, Ian Giddings, Richard Wooster, Robert H tePoele, Paul Workman, Barry Gusterson, John Peacock, Gerald Gui, Colin Campbell, and Ros Eeles. Gene expression profiling after radiation-induced DNA damage is strongly predictive of BRCA1 mutation carrier status. *Clin. Cancer Res.*, 10(3):958–63, Feb 2004.

- [242] S. Kramer, E. Frank, and C. Helma. Fragment generation and support vector machines for inducing SARs. *SAR QSAR Environ Res*, 13(5):509–23, Jul 2002.
- [243] V. G. Krishnan and D. R. Westhead. A comparative study of machine-learning methods to predict the effects of single nucleotide polymorphisms on protein function. *Bioinformatics*, 19(17):2199–2209, 2003.
- [244] B. Krishnapuram, L. Carin, and A. Hartemink. Gene expression analysis: joint feature selection and classifier design. In B. Scholkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 299–317. MIT Press, 2004.
- [245] B. Krishnapuram, L. Carin, and A. Hartemink. Joint Classifier and Feature Optimization for Comprehensive Cancer Diagnosis Using Gene Expression Data. *J. Comput. Biol.*, 11(2-3):227–242, 2004.
- [246] Balaji Krishnapuram, Alexander J Hartemink, Lawrence Carin, and Mrio A T Figueiredo. A bayesian approach to joint feature selection and classifier design. *IEEE Trans Pattern Anal Mach Intell*, 26(9):1105–11, Sep 2004.
- [247] Romano T Kroemer. Structure-based drug design: docking and scoring. *Curr Protein Pept Sci*, 8(4):312–328, Aug 2007.
- [248] R. Kuang, E. Ie, K. Wang, K. Wang, M. Siddiqi, Y. Freund, and C. Leslie. Profile-based string kernels for remote homology detection and motif extraction. *Proc IEEE Comput Syst Bioinform Conf*, pages 152–160, 2004.
- [249] R. Kuang, E. Ie, K. Wang, K. Wang, M. Siddiqi, Y. Freund, and C. Leslie. Profile-based string kernels for remote homology detection and motif extraction. *J. Bioinform. Comput. Biol.*, 3(3):527–550, Jun 2005.
- [250] R. Kuang, C. S. Leslie, and A.-S. Yang. Protein backbone angle prediction with machine learning approaches. *Bioinformatics*, 20(10):1612–1621, 2004.
- [251] M. Kumar, M. Bhasin, N. K. Natt, and G. P. S. Raghava. BhairPred: prediction of beta-hairpins in a protein from multiple alignment information using ANN and SVM techniques. *Nucleic Acids Res*, 33(Web Server issue):W154–9, Jul 2005.
- [252] Stephen LaConte, Stephen Strother, Vladimir Cherkassky, Jon Anderson, and Xiaoping Hu. Support vector machines for temporal classification of block design fMRI data. *Neuroimage*, 26(2):317–29, Jun 2005.
- [253] Thomas Navin Lal, Michael Schrder, Thilo Hinterberger, Jason Weston, Martin Bogdan, Niels Birbaumer, and Bernhard Scholkopf. Support vector channel selection in BCI. *IEEE Trans Biomed Eng*, 51(6):1003–10, Jun 2004.

- [254] G. R. G. Lanckriet, T. De Bie, N. Cristianini, M. I. Jordan, and W. S. Noble. A statistical framework for genomic data fusion. *Bioinformatics*, 20(16):2626–2635, 2004.
- [255] G.R. Lanckriet, M. Deng, N. Cristianini, M.I. Jordan, and W.S. Noble. Kernel-based data fusion and its application to protein function prediction in yeast. In *Proceedings of the Pacific Symposium on Biocomputing*, pages 300–311, 2004.
- [256] G.R.G. Lanckriet, N. Cristianini, M.I. Jordan, and W.S. Noble. Kernel-based integration of genomic data using semidefinite programming. In B. Schlkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 231–259. MIT Press, 2004.
- [257] Zhiqiang Lao, Dinggang Shen, Zhong Xue, Bilge Karacali, Susan M Resnick, and Christos Davatzikos. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage*, 21(1):46–57, Jan 2004.
- [258] J. S. Lazo and P. Wipf. Combinatorial chemistry and contemporary pharmacology. *J Pharmacol Exp Ther*, 293(3):705–709, Jun 2000.
- [259] Dongkwon Lee, Sang Wook Choi, Myengsoo Kim, Jin Hyun Park, Moonkyu Kim, Jungchul Kim, and In-Beum Lee. Discovery of differentially expressed genes related to histological subtype of hepatocellular carcinoma. *Biotechnol Prog.*, 19(3):1011–5, 2003.
- [260] Y. Lee and C.-K. Lee. Classification of multiple cancer types by multicategory support vector machines using gene expression data. *Bioinformatics*, 19(9):1132–1139, 2003.
- [261] C. Leslie, E. Eskin, and W.S. Noble. The spectrum kernel: a string kernel for SVM protein classification. In Russ B. Altman, A. Keith Dunker, Lawrence Hunter, Kevin Lauerdale, and Teri E. Klein, editors, *Proceedings of the Pacific Symposium on Biocomputing 2002*, pages 564–575, Singapore, 2002. World Scientific.
- [262] C. Leslie, E. Eskin, J. Weston, and W.S. Noble. Mismatch String Kernels for SVM Protein Classification. In Suzanna Becker, Sebastian Thrun, and Klaus Obermayer, editors, *Advances in Neural Information Processing Systems 15*. MIT Press, 2003.
- [263] C. S. Leslie, E. Eskin, A. Cohen, J. Weston, and W. S. Noble. Mismatch string kernels for discriminative protein classification. *Bioinformatics*, 20(4):467–476, 2004.
- [264] D. Lett, M. Hsing, and F. Pio. Interaction profile-based protein classification of death domain. *BMC Bioinformatics*, 5(75), 2004.

- [265] H. Li, C. Ung, C. Yap, Y. Xue, Z. Li, Z. Cao, and Y. Chen. Prediction of genotoxicity of chemical compounds by statistical learning methods. *Chem. Res. Toxicol.*, 18(6):1071–1080, Jun 2005.
- [266] Jinyan Li, Huiqing Liu, James R Downing, Allen Eng-Juh Yeoh, and Limsoon Wong. Simple rules underlying gene expression profiles of more than six subtypes of acute lymphoblastic leukemia (ALL) patients. *Bioinformatics*, 19(1):71–8, Jan 2003.
- [267] L. Li, W. Jiang, X. Li, K.L. Moser, Z. Guo, L. Du, Q. Wang, E.J. Topol, Q. Wang, and S. Rao. A robust hybrid between genetic algorithm and support vector machine for extracting an optimal feature gene subset. *Genomics*, 85(1):16–23, 2005.
- [268] L. Li, H. Tang, Z. Wu, J. Gong, M. Gruidl, J. Zou, M. Tockman, and R.A. Clark. Data mining techniques for cancer detection using serum proteomic profiling. *Artif. Intell. Med.*, 32(2):71–83, 2004.
- [269] T. Li, C. Zhang, and M. Ogihara. A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression. *Bioinformatics*, 20(15):2429–2437, 2004.
- [270] H. Liang and Z. Lin. Detection of delayed gastric emptying from electrogastrograms with support vector machine. *IEEE Trans Biomed Eng*, 48(5):601–4, May 2001.
- [271] L. Liao and W. S. Noble. Combining pairwise sequence similarity and support vector machines for remote protein homology detection. In *Proceedings of the Sixth International Conference on Computational Molecular Biology*, 2002.
- [272] L. Liao and W.S. Noble. Combining Pairwise Sequence Similarity and Support Vector Machines for Detecting Remote Protein Evolutionary and Structural Relationships. *J. Comput. Biol.*, 10(6):857–868, 2003.
- [273] D. A. Liberles, A. Thorén, G. von Heijne, and A. Elofsson. The use of phylogenetic profiles for gene predictions. *Curr. Genom.*, 2002. To appear.
- [274] K. Lin, Y. Kuang, J. S. Joseph, and P. R. Kolatkar. Conserved codon composition of ribosomal protein coding genes in *Escherichia coli*, *Mycobacterium tuberculosis* and *Saccharomyces cerevisiae*: lessons from supervised machine learning in functional genomics. *Nucl. Acids Res.*, 30(11):2599–2607, 2002.
- [275] WuMei Lin, Xin Yuan, Powing Yuen, William I Wei, Jonathan Sham, PengCheng Shi, and Jianan Qu. Classification of in vivo autofluorescence spectra using support vector machines. *J Biomed Opt*, 9(1):180–6, 2004.
- [276] P. Lind and T. Maltseva. Support vector machines for the estimation of aqueous solubility. *J Chem Inf Comput Sci*, 43(6):1855–9, 2003.

- [277] J. Listgarten, S. Damaraju, B. Poulin, L. Cook, J. Dufour, A. Driga, J. Mackey, D. Wishart, R. Greiner, and B. Zanke. Predictive Models for Breast Cancer Susceptibility from Multiple Single Nucleotide Polymorphisms. *Clin. Cancer Res.*, 10(8):2725–2737, 2004.
- [278] H. X. Liu, C. X. Xue, R. S. Zhang, X. J. Yao, M. C. Liu, Z. D. Hu, and B. T. Fan. Quantitative prediction of logk of peptides in high-performance liquid chromatography based on molecular descriptors by using the heuristic method and support vector machine. *J Chem Inf Comput Sci*, 44(6):1979–86, 2004.
- [279] H. X. Liu, R. S. Zhang, X. J. Yao, M. C. Liu, Z. D. Hu, and B. T. Fan. QSAR study of ethyl 2-[(3-methyl-2,5-dioxo(3-pyrrolyl)amino]-4-(trifluoromethyl) pyrimidine-5-carboxylate: an inhibitor of AP-1 and NF-kappa B mediated gene expression based on support vector machines. *J Chem Inf Comput Sci*, 43(4):1288–96, 2003.
- [280] H. X. Liu, R. S. Zhang, X. J. Yao, M. C. Liu, Z. D. Hu, and B. T. Fan. Prediction of the isoelectric point of an amino acid based on GA-PLS and SVMs. *J Chem Inf Comput Sci*, 44(1):161–7, 2004.
- [281] H. X. Liu, R. S. Zhang, X. J. Yao, M. C. Liu, Z. D. Hu, and B. T. Fan. QSAR and classification models of a novel series of COX-2 selective inhibitors: 1,5-diarylimidazoles based on support vector machines. *J Comput Aided Mol Des*, 18(6):389–99, Jun 2004.
- [282] Huiqing Liu, Hao Han, Jinyan Li, and Limsoon Wong. An in-silico method for prediction of polyadenylation signals in human sequences. *Genome Inform Ser Workshop Genome Inform*, 14:84–93, 2003.
- [283] Huiqing Liu, Hao Han, Jinyan Li, and Limsoon Wong. Using amino acid patterns to accurately predict translation initiation sites. *In Silico Biol.*, 4(3):255–69, 2004.
- [284] Huiqing Liu, Jinyan Li, and Limsoon Wong. Use of extreme patient samples for outcome prediction from gene expression data. *Bioinformatics*, Jun 2005.
- [285] Jane Jijun Liu, Gene Cutler, Wuxiong Li, Zheng Pan, Sihua Peng, Tim Hoey, Liangbiao Chen, and Xuefeng Bruce Ling. Multiclass cancer classification and biomarker discovery using GA-based algorithms. *Bioinformatics*, 21(11):2691–7, Jun 2005.
- [286] Y. Liu. Active learning with support vector machine applied to gene expression data for cancer classification. *J. Chem. Inf. Comput. Sci.*, 44(6):1936–1941, 2004.
- [287] Y. Liu. A comparative study on feature selection methods for drug discovery. *J Chem Inf Comput Sci*, 44(5):1823–8, 2004.

- [288] Zhenqiu Liu, Dechang Chen, and Halima Bensmail. Gene expression data classification with kernel principal component analysis. *J Biomed Biotechnol*, 2005(2):155–9, 2005.
- [289] Siaw Ling Lo, Cong Zhong Cai, Yu Zong Chen, and Maxey C M Chung. Effect of training datasets on support vector machine prediction of protein-protein interactions. *Proteomics*, 5(4):876–84, Mar 2005.
- [290] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. J. C. H. Watkins. Text classification using string kernels. *J. Mach. Learn. Res.*, 2:419–444, 2002.
- [291] H. Lodhi, J. Shawe-Taylor, N. Cristianini, and C. J. C. H. Watkins. Text Classification using String Kernels. In *Adv. Neural Inform. Process. Syst.*, pages 563–569, 2000.
- [292] B. Logan, P. Moreno, B. Suzek, Z. Weng, and S. Kasif. A Study of Remote Homology Detection. Technical Report CRL 2001/05, Compaq Cambridge Research laboratory, June 2001.
- [293] Wei-Zhen Lu and Wen-Jian Wang. Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends. *Chemosphere*, 59(5):693–701, Apr 2005.
- [294] Y.J. Lu, D. Williamson, R. Wang, B. Summersgill, S. Rodriguez, S. Rogers, K. Pritchard-Jones, C. Campbell, and J. Shipley. Expression profiling targeting chromosomes for tumor classification and prediction of clinical behavior. *Genes Chromosomes Cancer*, 38(3):207–214, 2003.
- [295] Feng Luan, Ruisheng Zhang, Chunyan Zhao, Xiaojun Yao, Mancang Liu, Zhide Hu, and Botao Fan. Classification of the carcinogenicity of N-nitroso compounds based on support vector machines and linear discriminant analysis. *Chem Res Toxicol*, 18(2):198–203, Feb 2005.
- [296] P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret, and J.-P. Vert. Extensions of marginalized graph kernels. In R. Greiner and D. Schuurmans, editors, *Proceedings of the Twenty-First International Conference on Machine Learning (ICML 2004)*, pages 552–559. ACM Press, 2004.
- [297] P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret, and J.-P. Vert. Graph kernels for molecular structure-activity relationship analysis with support vector machines. *J. Chem. Inf. Model.*, 45(4):939–51, 2005.
- [298] W. H. Majoros, L. Pertea, and S. L. Salzberg. Efficient implementation of a generalized pair hidden Markov model for comparative gene finding. *Bioinformatics*, 21(9):1782–1788, May 2005.
- [299] S. K. Majumder, N. Ghosh, and P. K. Gupta. Support vector machine for optical diagnosis of cancer. *J Biomed Opt*, 10(2):024034, 2005.

- [300] M.Z. Man, G. Dyson, K. Johnson, and B. Liao. Evaluating methods for classifying expression data. *J. Biopharm. Stat.*, 14(4):1065–1084, 2004.
- [301] Yong Mao, Xiaobo Zhou, Daoying Pi, Youxian Sun, and Stephen T C Wong. Multiclass cancer classification by using fuzzy support vector machine and binary decision tree with gene selection. *J Biomed Biotechnol*, 2005(2):160–71, 2005.
- [302] E. M. Marcotte, M. Pellegrini, M. J. Thompson, T. O. Yeates, and D. Eisenberg. A combined algorithm for genome-wide prediction of protein function. *Nature*, 402:83–86, November 1999.
- [303] E.M. Marcotte, M. Pellegrini, H.-L. Ng, D.W. Rice, T.O. Yeates, and D. Eisenberg. Detecting Protein Function and Protein-Protein Interactions from Genome Sequences. *Science*, 285:751–753, 1999.
- [304] F. Markowetz, L. Edler, and M. Vingron. Support Vector Machines for Protein Fold Class Prediction. *Biometrical Journal*, 45(3):377–389, 2003.
- [305] Stephen Marsland, Jonathan Shapiro, and Ulrich Nehmzow. A self-organising network that grows when required. *Neural Netw*, 15(8-9):1041–58, 2002.
- [306] S. Martin, D. Roe, and J.-L. Faulon. Predicting protein-protein interactions using signature products. *Bioinformatics*, 21(2):218–226, Jan 2005.
- [307] Ann-Marie Martoglio, James W Miskin, Stephen K Smith, and David J C MacKay. A decomposition model to track gene expression signatures: preview on observer-independent classification of ovarian cancer. *Bioinformatics*, 18(12):1617–24, Dec 2002.
- [308] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296:910–913, 2002.
- [309] J. S. Mason, I. Morize, P. R. Menard, D. L. Cheney, C. Hulme, and R. F. Labaudiniere. New 4-point pharmacophore method for molecular similarity and diversity applications: overview of the method and applications, including a novel approach to the design of combinatorial libraries containing privileged substructures. *J Med Chem*, 42(17):3251–3264, Aug 1999.
- [310] Alvaro Mateos, Joaquin Dopazo, Ronald Jansen, Yuhai Tu, Mark Gerstein, and Gustavo Stolovitzky. Systematic learning of gene functional classes from DNA array expression data by using multilayer perceptrons. *Genome Res.*, 12(11):1703–15, Nov 2002.
- [311] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol*, 288(5):911–940, May 1999.

- [312] A. Matsuda, J.-P. Vert, H. Saigo, N. Ueda, H. Toh, and T. Akutsu. A novel representation of protein sequences for prediction of subcellular location using support vector machines. *Protein Sci.*, 14(11):2804–2813, 2005.
- [313] T. Mattfeldt, H.W. Gottfried, H. Wolter, V. Schmidt, H.A. Kestler, and J. Mayer. Classification of prostatic carcinoma with artificial neural networks using comparative genomic hybridization and quantitative stereological data. *Pathol. Res. Pract.*, 199(12):773–784, 2003.
- [314] Michael Mavroforakis, Harris Georgiou, Nikos Dimitropoulos, Dionisis Cavouras, and Sergios Theodoridis. Significance analysis of qualitative mammographic features, using linear classifiers, neural networks and support vector machines. *Eur J Radiol*, 54(1):80–9, Apr 2005.
- [315] Torsten Mayr, Christian Igel, Gregor Liebsch, Ingo Klimant, and Otto S Wolfbeis. Cross-reactive metal ion sensor array in a micro titer plate format. *Anal Chem*, 75(17):4389–96, Sep 2003.
- [316] J. D. McAuliffe, L. Pachter, and M. I. Jordan. Multiple-sequence functional annotation and the generalized hidden Markov phylogeny. *Bioinformatics*, 20(12):1850–1860, Aug 2004.
- [317] Larry McKnight and Padmini Srinivasan. Categorization of sentence types in medical abstracts. *AMIA Annu Symp Proc*, pages 440–4, 2003.
- [318] P. Meinicke, M. Tech, B. Morgenstern, and R. Merkl. Oligo kernels for datamining on biological sequences: a case study on prokaryotic translation initiation sites. *BMC Bioinformatics*, 5(169), 2004.
- [319] S.I. Meireles, A.F. Carvalho, R. Hirata, A.L. Montagnini, W.K. Martins, F.B. Runza, B.S. Stolf, L. Termini, C.E. Neto, R.L. Silva, F.A. Soares, E.J. Neves, and L.F. Reis. Differentially expressed genes in gastric tumors identified by cDNA array. *Cancer Lett.*, 190(2):199–211, Feb 2003.
- [320] Jordi Mestres. Computational chemogenomics approaches to systematic knowledge-based drug discovery. *Curr Opin Drug Discov Devel*, 7(3):304–313, May 2004.
- [321] M. Middendorf, E. Ziv, C. Adams, J. Hom, R. Koytcheff, C. Levovitz, G. Woods, L. Chen, and C. Wiggins. Discriminative topological features reveal biological network mechanisms. *BMC Bioinformatics*, 5(181), 2004.
- [322] Sven Mika and Burkhard Rost. NLProt: extracting protein names and sequences from papers. *Nucleic Acids Res*, 32(Web Server issue):W634–7, Jul 2004.
- [323] Sven Mika and Burkhard Rost. Protein names precisely peeled off free text. *Bioinformatics*, 20(Suppl. 1):i241–i247, 2004.

- [324] Tomohiro Mitsumori, Sevrani Fation, Masaki Murata, Kouichi Doi, and Hirohumi Doi. Gene/protein name recognition based on support vector machine using dictionary as features. *BMC Bioinformatics*, 6 Suppl 1:S8, 2005.
- [325] F. Miwakeichi, R. Ramirez-Padron, P. A. Valdes-Sosa, and T. Ozaki. A comparison of non-linear non-parametric models for epilepsy data. *Comput. Biol. Med.*, 31(1):41–57, Jan 2001.
- [326] F. Model, P. Adorjan, A. Olek, and C. Piepenbrock. Feature selection for DNA methylation based cancer classification. *Bioinformatics*, 17(Supp. 1):S157–S164, 2001.
- [327] E. J. Moler, M. L. Chow, and I. S. Mian. Analysis of molecular profile data using generative and discriminative methods. *Physiol. Genomics*, 4(2):109–126, Dec 2000.
- [328] S. Mukherjee, P. Tamayo, J. P. Mesirov, D. Slonim, A. Verri, and T. Poggio. Support vector machine classification of microarray data. Technical Report 182, C.B.L.C., 1998. A.I. Memo 1677.
- [329] K.-R. Müller, G. Rätsch, S. Sonnenburg, S. Mika, M. Grimm, and N. Heinrich. Classifying 'drug-likeness' with Kernel-based learning methods. *J Chem Inf Model*, 45(2):249–53, 2005.
- [330] K. Murphy and S. Mian. Modelling gene expression data using dynamic Bayesian networks. Technical report, Computer Science Division, University of California, Berkeley, CA., 1999.
- [331] E. Myasnikova, A. Samsonova, M. Samsonova, and J. Reinitz. Support vector regression applied to the determination of the developmental age of a Drosophila embryo from its segmentation gene expression patterns. *Bioinformatics*, 18(Suppl. 1):S87–S95, 2002.
- [332] Rajesh Nair and Burkhard Rost. Mimicking cellular sorting improves prediction of subcellular localization. *J Mol Biol*, 348(1):85–100, Apr 2005.
- [333] N.K. Natt, H. Kaur, and G.P. Raghava. Prediction of transmembrane regions of beta-barrel proteins using ANN- and SVM-based methods. *Proteins*, 56(1):11–18, 2004.
- [334] M. N. Nguyen and J. C. Rajapakse. Two-stage multi-class support vector machines to protein secondary structure prediction. *Pac Symp Biocomput*, pages 346–57, 2005.
- [335] Minh N Nguyen and Jagath C Rajapakse. Multi-class support vector machines for protein secondary structure prediction. *Genome Inform Ser Workshop Genome Inform*, 14:218–27, 2003.

- [336] Minh N Nguyen and Jagath C Rajapakse. Prediction of protein relative solvent accessibility with a two-stage SVM approach. *Proteins*, 59(1):30–7, Apr 2005.
- [337] H. Nielsen, J. Engelbrecht, S. Brunak, and G. von Heijne. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*, 10(1):1–6, 1997.
- [338] W. S. Noble. Support vector machine applications in computational biology. In B. Scholkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 71–92. MIT Press, 2004.
- [339] Rebekah K O’Donnell, Michael Kupferman, S. Jack Wei, Sunil Singhal, Randal Weber, Bert O’Malley, Yi Cheng, Mary Putt, Michael Feldman, Barry Ziober, and Ruth J Muschel. Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. *Oncogene*, 24(7):1244–51, Feb 2005.
- [340] R. A. O’Flanagan, G. Paillard, R. Lavery, and A. M. Sengupta. Non-additivity in protein-DNA binding. *Bioinformatics*, 21(10):2254–63, May 2005.
- [341] Scott Oloff, Shuxing Zhang, Nagamani Sukumar, Curt Breneman, and Alexander Tropsha. Chemometric analysis of ligand receptor complementarity: identifying complementary ligands based on receptor information (colibri). *J Chem Inf Model*, 46(2):844–851, 2006.
- [342] M. Opper and R. Urbanczik. Universal learning curves of support vector machines. *Phys Rev Lett*, 86(19):4410–3, May 2001.
- [343] M. Opper and O. Winther. Gaussian processes for classification: mean-field algorithms. *Neural Comput*, 12(11):2655–84, Nov 2000.
- [344] Tapio Pahikkala, Filip Ginter, Jorma Boberg, Jouni Jarvinen, and Tapio Salakoski. Contextual weighting for Support Vector Machines in literature mining: an application to gene versus protein name disambiguation. *BMC Bioinformatics*, 6(1):157, Jun 2005.
- [345] Fei Pan, Baoying Wang, Xin Hu, and William Perrizo. Comprehensive vertical sample-based KNN/LSVM classification for gene expression analysis. *J Biomed Inform*, 37(4):240–8, Aug 2004.
- [346] A. Pandey and M. Mann. Proteomics to study genes and genomes. *Nature*, 405:837–846, 2000.
- [347] K.-J. Park and M. Kanehisa. Prediction of protein subcellular locations by support vector machines using compositions of amino acids and amino acid pairs. *Bioinformatics*, 19(13):1656–1663, 2003.

- [348] A. Passerini and P. Frasconi. Learning to discriminate between ligand-bound and disulfide-bound cysteines. *Protein Eng. Des. Sel.*, 17(4):367–373, 2004.
- [349] R. Pastor-Satorras, E. D. Smith, and R. V. Solé. Evolving protein interaction networks through gene duplication. Technical report, Santa Fe Institute, 2002. Working paper 02-02-008.
- [350] D.J. Patterson, K. Yasuhara, and W.L. Ruzzo. Pre-mRNA secondary structure prediction aids splice site prediction. In Russ B. Altman, A. Keith Dunker, Lawrence Hunter, Kevin Lauerdale, and Teri E. Klein, editors, *Proceedings of the Pacific Symposium on Biocomputing 2002*, pages 223–234. World Scientific, 2002.
- [351] S. Pavey, P. Johansson, L. Packer, J. Taylor, M. Stark, P.M. Pollock, G.J. Walker, G.M. Boyle, U. Harper, S.J. Cozzi, K. Hansen, L. Yudt, C. Schmidt, P. Hersey, K.A. Ellem, M.G. O’Rourke, P.G. Parsons, P. Meltzer, M. Ringner, and N.K. Hayward. Microarray expression profiling in melanoma reveals a BRAF mutation signature. *Oncogene*, 23(23):4060–4067, May 2004.
- [352] P. Pavlidis, T. S. Furey, M. Liberto, D. Haussler, and W. N. Grundy. Promoter Region-Based Classification of Genes. In *Pacific Symposium on Biocomputing*, pages 139–150, 2001.
- [353] P. Pavlidis, J. Weston, J. Cai, and W.N. Grundy. Gene functional classification from heterogeneous data. In *Proceedings of the Fifth Annual International Conference on Computational Biology*, pages 249–255, 2001.
- [354] P. Pavlidis, J. Weston, J. Cai, and W.S. Noble. Learning gene functional classifications from multiple data types. *J. Comput. Biol.*, 9(2):401–411, 2002.
- [355] Paul Pavlidis, Ilan Wapinski, and William Stafford Noble. Support vector machine classification on the web. *Bioinformatics*, 20(4):586–7, Mar 2004.
- [356] M. Pellegrini, E. M. Marcotte, M. J. Thompson, D. Eisenberg, and T. O. Yeates. Assigning protein functions by comparative genome analysis: Protein phylogenetic profiles. *Proc. Natl. Acad. Sci. USA*, 96:4285–4288, April 1999.
- [357] S. Peng, Q. Xu, X.B. Ling, X. Peng, W. Du, and L. Chen. Molecular classification of cancer types from microarray data using the combination of genetic algorithms and support vector machines. *FEBS Lett.*, 555(2):358–362, 2003.
- [358] Bjoern Peters and Alessandro Sette. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics*, 6:132, 2005.

- [359] Tho Hoan Pham, Kenji Satou, and Tu Bao Ho. Prediction and analysis of beta-turns in proteins by support vector machine. *Genome Inform Ser Workshop Genome Inform*, 14:196–205, 2003.
- [360] Tho Hoan Pham, Kenji Satou, and Tu Bao Ho. Support vector machines for prediction and analysis of beta and gamma-turns in proteins. *J. Bioinform. Comput. Biol.*, 3(2):343–58, Apr 2005.
- [361] Dariusz Plewczynski, Adrian Tkacz, Adam Godzik, and Leszek Rychlewski. A support vector machine approach to the identification of phosphorylation sites. *Cell Mol Biol Lett*, 10(1):73–89, 2005.
- [362] Dariusz Plewczynski, Adrian Tkacz, Lucjan Stanislaw Wyrwicz, and Leszek Rychlewski. AutoMotif server: prediction of single residue post-translational modifications in proteins. *Bioinformatics*, 21(10):2525–7, May 2005.
- [363] N. Pochet, F. De Smet, J. A. K. Suykens, and B. L. R. De Moor. Systematic benchmarking of microarray data classification: assessing the role of non-linearity and dimensionality reduction. *Bioinformatics*, 20(17):3185–3195, Nov 2004.
- [364] J. Podani, Z.N. Oltvai, H. Jeong, B. Tombor, A.-L. Barabási, and E. Szathmáry. Comparable system-level organization of Archaea and Eukaryotes. *Nat. Genet.*, 29:54–56, 2001.
- [365] Poggio and Girosi. A Sparse Representation for Function Approximation. *Neural Comput*, 10(6):1445–54, Jul 1998.
- [366] M. Pontil and A. Verri. Properties of support vector machines. *Neural Comput*, 10(4):955–74, May 1998.
- [367] J. Prados, A. Kalousis, J.C. Sanchez, L. Allard, O. Carrette, and M. Hilario. Mining mass spectra for diagnosis and biomarker discovery of cerebral accidents. *Proteomics*, 4(8):2320–2332, 2004.
- [368] K. N Bhanu Prakash, A. G. Ramakrishnan, S. Suresh, and Teresa W P Chow. Fetal lung maturity analysis using ultrasound image features. *IEEE Trans Inf Technol Biomed*, 6(1):38–45, Mar 2002.
- [369] J. Qian, J. Lin, N. M. Luscombe, H. Yu, and M. Gerstein. Prediction of regulatory networks: genome-wide identification of transcription factor targets from gene expression data. *Bioinformatics*, 19(15):1917–1926, 2003.
- [370] J. Qian, N. M. Luscombe, and M. Gerstein. Protein Fold and Family Occurrence in Genomes: Power-Law Behaviour and Evolutionary Model. *J. Mol. Biol.*, 313:673–681, 2001.

- [371] J. Qin, D. P. Lewis, and W. S. Noble. Kernel hierarchical gene clustering from microarray expression data. *Bioinformatics*, 19(16):2097–2104, 2003.
- [372] Zhong Qin, Qiang Yu, Jun Li, Zhi-Yi Wu, and Bing-Min Hu. Application of least squares vector machines in modelling water vapor and carbon dioxide fluxes over a cropland. *J Zhejiang Univ Sci*, 6(6):491–5, Jun 2005.
- [373] Gajendra P S Raghava and Joon H Han. Correlation and prediction of gene expression level from amino acid and dipeptide composition of its protein. *BMC Bioinformatics*, 6(1):59, Mar 2005.
- [374] J.-C. Rain, L. Selig, H. De Reuse, V. Battaglia, C. Reverdy, S. Simon, G. Lenzen, F. Petel, J. Wojcik, V. Schächter, Y. Chemama, A. Labigne, and P. Legrain. The protein-protein interaction map of *Helicobacter pylori*. *Nature*, 409:211–215, 2001.
- [375] S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, C.H. Yeang, M. Angelo, C. Ladd, M. Reich, E. Latulippe, J.P. Mesirov, T. Poggio, W. Gerald, M. Loda, E.S. Lander, and T.R. Golub. Multiclass cancer diagnosis using tumor gene expression signatures. *Proc. Natl. Acad. Sci. USA*, 98(26):15149–15154, Dec 2001.
- [376] H. Rangwala and G. Karypis. Profile-based direct kernels for remote homology detection and fold recognition. *Bioinformatics*, 21(23):4239–4247, Dec 2005.
- [377] M. Rarey, B. Kramer, T. Lengauer, and G. Klebe. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol*, 261(3):470–489, Aug 1996.
- [378] G. Rättsch and S. Sonnenburg. Accurate splice site detection for *Caenorhabditis elegans*. In B. Schölkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 277–298. MIT Press, 2004.
- [379] G. Rättsch, S. Sonnenburg, and B. Schölkopf. RASE: recognition of alternatively spliced exons in *C.elegans*. *Bioinformatics*, 21(Suppl. 1):i369–i377, Jun 2005.
- [380] Stefan A Rensing, Dana Fritzowsky, Daniel Lang, and Ralf Reski. Protein encoding genes in an ancient plant: analysis of codon usage, retained genes and splice sites in a moss, *Physcomitrella patens*. *BMC Genomics*, 6(1):43, Mar 2005.
- [381] I. Res, I. Mihalek, and O. Lichtarge. An evolution based classifier for prediction of protein interfaces without using protein structures. *Bioinformatics*, 21(10):2496–501, May 2005.
- [382] P. Rice, I. Longden, and A. Bleasby. Emboss: the european molecular biology open software suite. *Trends Genet*, 16(6):276–277, Jun 2000.

- [383] Simon B Rice, Goran Nenadic, and Benjamin J Stapley. Mining protein function from text using term-based support vector machines. *BMC Bioinformatics*, 6 Suppl 1:S22, 2005.
- [384] Henning Riedesel, Bjrn Kolbeck, Oliver Schmetzer, and Ernst-Walter Knapp. Peptide binding at class I major histocompatibility complex scored with linear functions and support vector machines. *Genome Inform Ser Workshop Genome Inform*, 15(1):198–212, 2004.
- [385] Risau-Gusman and Gordon. Generalization properties of finite-size polynomial support vector machines. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 62(5 Pt B):7092–9, Nov 2000.
- [386] S. Risau-Gusman and M. B. Gordon. Statistical mechanics of learning with soft margin classifiers. *Phys Rev E Stat Nonlin Soft Matter Phys*, 64(3 Pt 1):031907, Sep 2001.
- [387] Thorsteinn Rögnvaldsson and Liwen You. Why neural networks should not be used for HIV-1 protease cleavage site prediction. *Bioinformatics*, 20(11):1702–9, Jul 2004.
- [388] J. R. Rose, Jr. Turkett, W. H., I. C. Oroian, W. W. Laegreid, and J. Keele. Correlation of amino acid preference and mammalian viral genome type. *Bioinformatics*, 2005.
- [389] F. P. Roth, J. D. Hughes, P. W. Estep, and G. M. Church. Finding dna regulatory motifs within unaligned noncoding sequences clustered by whole-genome mrna quantitation. *Nat Biotechnol*, 16(10):939–945, October 1998.
- [390] S. Rudd and I. V. Tetko. Eclair—a web service for unravelling species origin of sequences sampled from mixed host interfaces. *Nucleic Acids Res*, 33(Web Server issue):W724–7, Jul 2005.
- [391] S. Ruepp, F. Boess, L. Suter, M. C. de Vera, G. Steiner, T. Steele, T. Weiser, and S. Albertini. Assessment of hepatotoxic liabilities by transcript profiling. *Toxicol Appl Pharmacol*, Jun 2005.
- [392] J. Saeh, P. Lyne, B. Takasaki, and D. Cosgrove. Lead hopping using SVM and 3D pharmacophore fingerprints. *J Chem Inf Model*, 45(4):1122–1133, Jul 2005.
- [393] P. Saetrom. Predicting the efficacy of short oligonucleotides in antisense and RNAi experiments with boosted genetic programming. *Bioinformatics*, 20(17):3055–3063, 2004.
- [394] Y. Saeys, S. Degroeve, D. Aeyels, P. Rouz, and Y. Van de Peer. Feature selection for splice site prediction: A new method using EDA-based feature ranking. *BMC Bioinformatics*, 5(64), 2004.

- [395] Y. Saeys, S. Degroeve, D. Aeyels, Y. Van de Peer, and P. Rouze. Fast feature selection using a simple estimation of distribution algorithm: a case study on splice site prediction. *Bioinformatics*, 19(Suppl. 1):ii179–ii188, 2003.
- [396] H. Saigo, J.-P. Vert, N. Ueda, and T. Akutsu. Protein homology detection using string alignment kernels. *Bioinformatics*, 20(11):1682–1689, 2004.
- [397] N. Salim, J. Holliday, and P. Willett. Combination of fingerprint-based similarity coefficients using data fusion. *J Chem Inf Comput Sci*, 43(2):435–442, 2003.
- [398] J. Salomon and D. R. Flower. Predicting Class II MHC-Peptide binding: a kernel based approach using similarity scores. *BMC Bioinformatics*, 7:501, 2006.
- [399] Pamela A Sample, Michael H Goldbaum, Kwokleung Chan, Catherine Boden, Te-Won Lee, Christiana Vasile, Andreas G Boehm, Terrence Sejnowski, Chris A Johnson, and Robert N Weinreb. Using machine learning classifiers to identify glaucomatous change earlier in standard visual fields. *Invest Ophthalmol Vis Sci*, 43(8):2660–5, Aug 2002.
- [400] Marta Sanchez-Carbayo, Nicholas D Socci, Juan Jose Lozano, Wentian Li, Elizabeth Charytonowicz, Thomas J Belbin, Michael B Prystowsky, Angel R Ortiz, Geoffrey Childs, and Carlos Cordon-Cardo. Gene discovery in bladder cancer progression using cDNA microarrays. *Am. J. Pathol.*, 163(2):505–16, Aug 2003.
- [401] Deepak Sarda, Gek Huey Chua, Kuo-Bin Li, and Arun Krishnan. pSLIP: SVM based protein subcellular localization prediction using multiple physicochemical properties. *BMC Bioinformatics*, 6(1):152, Jun 2005.
- [402] G. Schneider and P. Wrede. Artificial neural networks for computer-based molecular design. *Prog Biophys Mol Biol*, 70(3):175–222, 1998.
- [403] B. Schölkopf, K. Tsuda, and J.-P. Vert. *Kernel Methods in Computational Biology*. MIT Press, The MIT Press, Cambridge, Massachusetts, 2004.
- [404] B. Schölkopf, J. Weston, E. Eskin, C. Leslie, and W.S. Noble. A Kernel Approach for Learning from Almost Orthogonal Patterns. In *Proceedings of ECML 2002*, 2002.
- [405] Holger Schwender, Manuela Zucknick, Katja Ickstadt, Hermann M Bolt, and G. E. N. I. C. A. network. A pilot study on the application of statistical classification procedures to molecular epidemiological data. *Toxicol Lett*, 151(1):291–9, Jun 2004.
- [406] M. Seeger. Covariance Kernels from Bayesian Generative Models. In *Adv. Neural Inform. Process. Syst.*, volume 14, pages 905–912, 2002.

- [407] Matthias Seeger. Gaussian processes for machine learning. *Int J Neural Syst*, 14(2):69–106, Apr 2004.
- [408] E. Segal, N. Friedman, D. Koller, and A. Regev. A module map showing conditional activity of expression modules in cancer. *Nat. Genet.*, 36(10):1090–1098, Oct 2004.
- [409] E. Segal, M. Shapira, A. Regev, D. Pe’er, D. Botstein, D. Koller, and N. Friedman. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat. Genet.*, 34(2):166–176, Jun 2003.
- [410] M. R. Segal, K. D. Dahlquist, and B. R. Conklin. Regression approaches for microarray data analysis. *J. Comput. Biol.*, 10(6):961–980, 2003.
- [411] N. H. Segal, P. Pavlidis, C. R. Antonescu, R. G. Maki, W. S. Noble, D. DeSantis, J. M. Woodruff, J. J. Lewis, M. F. Brennan, A. N. Houghton, and C. Cordon-Cardo. Classification and Subtype Prediction of Adult Soft Tissue Sarcoma by Functional Genomics. *Am. J. Pathol.*, 163(2):691–700, Aug 2003.
- [412] N. H. Segal, P. Pavlidis, W. S. Noble, C. R. Antonescu, A. Viale, U. V. Wesley, K. Busam, H. Gallardo, D. DeSantis, M. F. Brennan, C. Cordon-Cardo, J. D. Wolchok, and A. N. Houghton. Classification of Clear-Cell Sarcoma as a Subtype of Melanoma by Genomic Profiling. *J. Clin. Oncol.*, 21(9):1775–1781, May 2003.
- [413] M. Seike, T. Kondo, K. Fujii, T. Okano, T. Yamada, Y. Matsuno, A. Gemma, S. Kudoh, and S. Hirohashi. Proteomic signatures for histological types of lung cancer. *Proteomics*, Jul 2005.
- [414] T.Z. Sen, A. Kloczkowski, R.L. Jernigan, C. Yan, V. Honavar, K.M. Ho, C.Z. Wang, Y. Ihm, H. Cao, X. Gu, and D. Dobbs. Predicting binding sites of hydrolase-inhibitor complexes by combining several methods. *BMC Bioinformatics*, 5(205), 2004.
- [415] Pasak Senawongse, Andrew R Dalby, and Zheng Rong Yang. Predicting the phosphorylation sites using hidden markov models and machine learning methods. *J Chem Inf Model*, 45(4):1147–52, 2005.
- [416] J.R. Serra, E.D. Thompson, and P.C. Jurs. Development of binary classification of structural chromosome aberrations for a diverse set of organic compounds from molecular structure. *Chem. Res. Toxicol.*, 16(2):153–163, 2003.
- [417] William Shannon, Robert Culverhouse, and Jill Duncan. Analyzing microarray data using cluster analysis. *Pharmacogenomics*, 4(1):41–52, Jan 2003.

- [418] Roded Sharan and Eugene W Myers. A motif-based framework for recognizing sequence families. *Bioinformatics*, 21 Suppl 1:i387–i393, Jun 2005.
- [419] R. She, F. Chen, K. Wang, M. Ester, J.L. Gardy, and F.S.L. Brinkman. Frequent-subsequence-based prediction of outer membrane proteins. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 436–445. ACM Press, 2003.
- [420] Lei Shi and Fabien Campagne. Building a protein name dictionary from full text: a machine learning term extraction approach. *BMC Bioinformatics*, 6(1):88, Apr 2005.
- [421] M. A. Shipp, K. N. Ross, P. Tamayo, A. P. Weng, J. L. Kutok, R. C. T. Aguiar, M. Gaasenbeek, M. Angelo, M. Reich, G. A. Pinkus, T. S. Ray, M. A. Koval, K. W. Last, A. Norton, T. A. Lister, J. Mesirov, D. S. Neuberg, E. S. Lander, J. C. Aster, and T. R. Golub. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat. Med.*, 8(1):68–74, 2002.
- [422] Ali Shoeb, Herman Edwards, Jack Connolly, Blaise Bourgeois, S. Ted Treves, and John Guttag. Patient-specific seizure onset detection. *Epilepsy Behav*, 5(4):483–98, Aug 2004.
- [423] J. A. Siepen, S. E. Radford, and D. R. Westhead. Beta Edge strands in protein structure prediction and aggregation. *Protein Sci.*, 12(10):2348–2359, 2003.
- [424] F. Slanina and M. Kotrla. Random networks created by biological evolution. *Phys. Rev. E*, 62(5):6170–6177, 2000.
- [425] P. A. Smith, M. J. Sorich, L. S C Low, R. A. McKinnon, and J. O. Miners. Towards integrated ADME prediction: past, present and future directions for modelling metabolism by UDP-glucuronosyltransferases. *J Mol Graph Model*, 22(6):507–17, Jul 2004.
- [426] R. V. Solé, R. Pastor-Satorras, E. D. Smith, and T. Kepler. A Model of Large-Scale Proteome Evolution. Technical report, Santa Fe Institute, 2001. Working paper 01-08-041.
- [427] Minghu Song, Curt M Breneman, Jinbo Bi, N. Sukumar, Kristin P Bennett, Steven Cramer, and Nihal Tugcu. Prediction of protein retention times in anion-exchange chromatography systems using support vector regression. *J Chem Inf Comput Sci*, 42(6):1347–57, 2002.
- [428] S. Sonnenburg, G. Rätsch, A. Jagota, and K.-R. Müller. New methods for splice-site recognition. In JR. Dorronsoro, editor, *Proc. International conference on artificial Neural Networks ? ICANN?02*, number 2415 in LNCS, pages 329–336. Springer Berlin, 2002.

- [429] M. J. Sorich, J. O. Miners, R. A. McKinnon, D. A. Winkler, F. R. Burden, and P. A. Smith. Comparison of linear and nonlinear classification algorithms for the prediction of drug and chemical metabolism by human UDP-glucuronosyltransferase isoforms. *J Chem Inf Comput Sci*, 43(6):2019–24, 2003.
- [430] Michael J Sorich, Ross A McKinnon, John O Miners, David A Winkler, and Paul A Smith. Rapid prediction of chemical metabolism by human UDP-glucuronosyltransferase isoforms using quantum chemical descriptors derived with the electronegativity equalization method. *J Med Chem*, 47(21):5311–7, Oct 2004.
- [431] Florence L Stahura and Jrgen Bajorath. Virtual screening methods that complement HTS. *Comb Chem High Throughput Screen*, 7(4):259–69, Jun 2004.
- [432] B.J. Stapley, L.A. Kelley, and M.J. Sternberg. Predicting the sub-cellular location of proteins from text using support vector machines. In Russ B. Altman, A. Keith Dunker, Lawrence Hunter, Kevin Lauerdale, and Teri E. Klein, editors, *Proceedings of the Pacific Symposium on Biocomputing 2002*, pages 374–385. World Scientific, 2002.
- [433] A. Statnikov, C. F. Aliferis, I. Tsamardinos, D. Hardin, and S. Levy. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. *Bioinformatics*, 2005. To appear.
- [434] Alexander Statnikov, Constantin F Aliferis, and Ioannis Tsamardinos. Methods for multi-category cancer diagnosis from gene expression data: a comprehensive evaluation to inform decision support system development. *Medinfo*, 11(Pt 2):813–7, 2004.
- [435] Guido Steiner, Laura Suter, Franziska Boess, Rodolfo Gasser, Maria Cristina de Vera, Silvio Albertini, and Stefan Ruepp. Discriminating different classes of toxicants by transcript profiling. *Environ. Health Perspect.*, 112(12):1236–48, Aug 2004.
- [436] Alexander Sturn, John Quackenbush, and Zlatko Trajanoski. Genesis: cluster analysis of microarray data. *Bioinformatics*, 18(1):207–8, Jan 2002.
- [437] A. I. Su, J. B. Welsh, L. M. Sapinoso, S. G. Kern, P. Dimitrov, H. Lapp, P. G. Schultz, S. M. Powell, C. A. Moskaluk, H. F.Jr. Frierson, and G. M. Hampton. Molecular Classification of Human Carcinomas by Use of Gene Expression Signatures. *Cancer Res.*, 61(20):7388–7393, 2001.
- [438] Yang Su, T.M. Murali, Vladimir Pavlovic, Michael Schaffer, and Simon Kasif. RankGene: identification of diagnostic genes based on expression data. *Bioinformatics*, 19(12):1578–1579, 2003.

- [439] Y.F. Sun, X.D. Fan, and Y.D. Li. Identifying splicing sites in eukaryotic RNA: support vector machine approach. *Comput. Biol. Med.*, 33(1):17–29, 2003.
- [440] Zhenghong Sun, Xiaoli Fu, Lu Zhang, Xiaoli Yang, Feizhou Liu, and Gengxi Hu. A protein chip system for parallel analysis of multi-tumor markers and its application in cancer detection. *Anticancer Res*, 24(2C):1159–65, 2004.
- [441] J. A. Suykens, J. Vandewalle, and B. De Moor. Optimal control by least squares support vector machines. *Neural Netw*, 14(1):23–35, Jan 2001.
- [442] S. J. Swamidass, J. Chen, J. Bruand, P. Phung, L. Ralaivola, and P. Baldi. Kernels for small molecules and the prediction of mutagenicity, toxicity and anti-cancer activity. *Bioinformatics*, 21(Suppl. 1):i359–i368, Jun 2005.
- [443] Y. Takaoka, Y. Endo, S. Yamanobe, H. Kakinuma, T. Okubo, Y. Shimazaki, T. Ota, S. Sumiya, and K. Yoshikawa. Development of a method for evaluating drug-likeness and ease of synthesis using a data set in which compounds are assigned scores based on chemists’ intuition. *J Chem Inf Comput Sci*, 43(4):1269–75, 2003.
- [444] Koichi Takeuchi and Nigel Collier. Bio-medical entity extraction using support vector machines. *Artif. Intell. Med.*, 33(2):125–37, Feb 2005.
- [445] A. Talukder and D. Casasent. A closed-form neural network for discriminatory feature extraction from high-dimensional data. *Neural Netw*, 14(9):1201–18, Nov 2001.
- [446] Thomas Tang, Jinbo Xu, and Ming Li. Discovering sequence-structure motifs from protein segments and two applications. *Pac Symp Biocomput*, pages 370–81, 2005.
- [447] R. D. Teixeira, A. P. Braga, R. H. Takahashi, and R. R. Saldanha. Recent advances in the MOBJ algorithm for training artificial neural networks. *Int J Neural Syst*, 11(3):265–70, Jun 2001.
- [448] Reiji Teramoto, Mikio Aoki, Toru Kimura, and Masaharu Kanaoka. Prediction of siRNA functionality using generalized string kernel and support vector machine. *FEBS Lett.*, 579(13):2878–82, May 2005.
- [449] Sushil K Thukral, Paul J Nordone, Rong Hu, Leah Sullivan, Eric Galambos, Vincent D Fitzpatrick, Laura Healy, Michael B Bass, Mary E Cosenza, and Cynthia A Afshari. Prediction of nephrotoxicant action and identification of candidate toxicity-related biomarkers. *Toxicol Pathol*, 33(3):343–55, 2005.
- [450] Liang Tian and Afzel Noore. A novel approach for short-term load forecasting using support vector machines. *Int J Neural Syst*, 14(5):329–35, Oct 2004.

- [451] M. Tobita, T. Nishikawa, and R. Nagashima. A discriminant model constructed by the support vector machine method for HERG potassium channel inhibitors. *Bioorg. Med. Chem. Lett.*, 15(11):2886–90, Jun 2005.
- [452] Richard W Tothill, Adam Kowalczyk, Danny Rischin, Alex Bousioutas, Izhak Haviv, Ryan K van Laar, Paul M Waring, John Zalberg, Robyn Ward, Andrew V Biankin, Robert L Sutherland, Susan M Henshall, Kwun Fong, Jonathan R Pollack, David D L Bowtell, and Andrew J Holloway. An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. *Cancer Res.*, 65(10):4031–40, May 2005.
- [453] C.A. Tsai, C.H. Chen, T.C. Lee, I.C. Ho, U.C. Yang, and J.J. Chen. Gene selection for sample classifications in microarray experiments. *DNA Cell Biol.*, 23(10):607–614, 2004.
- [454] A. Tsirigos and I. Rigoutsos. A sensitive, support-vector-machine method for the detection of horizontal gene transfers in viral, archaeal and bacterial genomes. *Nucleic Acids Res.*, 33(12):3699–707, 2005.
- [455] K. Tsuda, S. Akaho, and K. Asai. The em Algorithm for Kernel Matrix Completion with Auxiliary Data. *J. Mach. Learn. Res.*, 4:67–81, 2003.
- [456] K. Tsuda, M. Kawanabe, G. Rätsch, S. Sonnenburg, and K.-R. Müller. A new discriminative kernel from probabilistic models. *Neural Computation*, 14(10):2397–2414, 2002.
- [457] K. Tsuda, T. Kin, and K. Asai. Marginalized Kernels for Biological Sequences. *Bioinformatics*, 18:S268–S275, 2002.
- [458] K. Tsuda and W.S. Noble. Learning kernels from biological networks by maximizing entropy. *Bioinformatics*, 20:i326–i333, 2004.
- [459] Nihal Tugcu, Minghu Song, Curt M Breneman, N. Sukumar, Kristin P Bennett, and Steven M Cramer. Prediction of the effect of mobile-phase salt type on protein retention and selectivity in anion exchange systems. *Anal Chem*, 75(14):3563–72, Jul 2003.
- [460] W. L. Tung and C. Quek. GenSo-FDSS: a neural-fuzzy decision support system for pediatric ALL cancer subtype identification using gene expression data. *Artif. Intell. Med.*, 33(1):61–88, Jan 2005.
- [461] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleish, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, and J. M. Rothberg. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, 403:623–627, 2000.

- [462] G. Valentini. Gene expression data analysis of human lymphoma using support vector machines and output coding ensembles. *Artif. Intell. Med.*, 26(3):281–304, Nov 2002.
- [463] Anirudh Vallabhaneni and Bin He. Motor imagery task classification for brain computer interface applications using spatiotemporal principle component analysis. *Neurol Res*, 26(3):282–7, Apr 2004.
- [464] A. Vazquez, A. Flammini, A. Maritan, and A. Vespignani. Modeling of protein interaction networks. E-print cond-mat/0108043, Aug 2001.
- [465] J. C. et al. Venter. The Sequence of the Human Genome. *Science*, 291(5507):1304–1351, 2001.
- [466] W. Vercoutere, S. Winters-Hilt, H. Olsen, D. Deamer, D. Haussler, and M. Akeson. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat Biotechnol*, 19(3):248–52, Mar 2001.
- [467] J.-P. Vert. Support vector machine prediction of signal peptide cleavage site using a new class of kernels for strings. In R. B. Altman, A. K. Dunker, L. Hunter, K. Lauerdale, and T. E. Klein, editors, *Proceedings of the Pacific Symposium on Biocomputing 2002*, pages 649–660. World Scientific, 2002.
- [468] J.-P. Vert. A tree kernel to analyze phylogenetic profiles. *Bioinformatics*, 18:S276–S284, 2002.
- [469] J.-P. Vert. Kernel methods in computational biology. Technical Report ccscd-00012124, CNRS-HAL, Oct 2005.
- [470] J.-P. Vert and M. Kanehisa. Graph-driven features extraction from microarray data. Technical Report 0206055, Arxiv physics, 2002.
- [471] J.-P. Vert and M. Kanehisa. Extracting active pathways from gene expression data. *Bioinformatics*, 19:238ii–234ii, 2003.
- [472] J.-P. Vert and M. Kanehisa. Graph-driven features extraction from microarray data using diffusion kernels and kernel CCA. In S. Becker, S. Thrun, and K. Obermayer, editors, *Adv. Neural Inform. Process. Syst.*, pages 1449–1456. MIT Press, 2003.
- [473] J.-P. Vert, H. Saigo, and T. Akutsu. Local alignment kernels for biological sequences. In B. Schölkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 131–154. MIT Press, The MIT Press, Cambridge, Massachusetts, 2004.
- [474] J.-P. Vert, R. Thurman, and W. S. Noble. Kernels for gene regulatory regions. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Adv. Neural Inform. Process Syst.*, volume 18, pages 1401–1408, Cambridge, MA, 2006. MIT Press.

- [475] J.-P. Vert, K. Tsuda, and B. Schölkopf. A primer on kernel methods. In B. Schölkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 35–70. MIT Press, 2004.
- [476] J.-P. Vert and Y. Yamanishi. Supervised graph inference. In L. K. Saul, Y. Weiss, and L. Bottou, editors, *Adv. Neural Inform. Process. Syst.*, volume 17, pages 1433–1440. MIT Press, Cambridge, MA, 2005.
- [477] A. Vinayagam, R. Knig, J. Moormann, F. Schubert, R. Eils, K.-H. Glatting, and S. Suhai. Applying Support Vector Machines for Gene Ontology based gene function prediction. *BMC Bioinformatics*, 5(1):116, Aug 2004.
- [478] Kristian Vlahovicek, Lszl Kajn, Vilmos Agoston, and Sndor Pongor. The SBASE domain sequence resource, release 12: prediction of protein domain-architecture using support vector machines. *Nucleic Acids Res*, 33(Database issue):D223–5, Jan 2005.
- [479] A. Wagner. The Yeast Protein Interaction Network Evolves Rapidly and Contains Few Redundant Duplicate Genes. *Mol. Biol. Evol.*, 18:1283–1292, 2001.
- [480] M. Wagner, D. Naik, and A. Pothen. Protocols for disease classification from mass spectrometry data. *Proteomics*, 3(9):1692–1698, 2003.
- [481] M. Wagner, D.N. Naik, A. Pothen, S. Kasukurti, R.R. Devineni, B.L. Adam, O.J. Semmes, and G.L. Wright Jr. Computational protein biomarker prediction: a case study for prostate cancer. *BMC Bioinformatics*, 5(26), 2004.
- [482] Grace Wahba. Soft and hard classification by reproducing kernel Hilbert space methods. *Proc Natl Acad Sci U S A*, 99(26):16524–30, Dec 2002.
- [483] J. Wang, W.-K. Sung, A. Krishnan, and K.-B. Li. Protein subcellular localization prediction for Gram-negative bacteria using amino acid sub-alphabets and a combination of multiple support vector machines. *BMC Bioinformatics*, 6(1):174, Jul 2005.
- [484] Kai Wang, Ekachai Jenwitheesuk, Ram Samudrala, and John E Mittler. Simple linear model provides highly accurate genotypic predictions of HIV-1 drug resistance. *Antivir Ther*, 9(3):343–52, Jun 2004.
- [485] Long-Hui Wang, Juan Liu, Yan-Fu Li, and Huai-Bei Zhou. Predicting protein secondary structure by a support vector machine based on a new coding scheme. *Genome Inform Ser Workshop Genome Inform*, 15(2):181–90, 2004.
- [486] M. Wang, J. Yang, and K-C. Chou. Using string kernel to predict signal peptide cleavage site based on subsite coupling model. *Amino Acids*, 28(4):395–402, Jun 2005.

- [487] M. Wang, J. Yang, G.-P. Liu, Z.-J. Xu, and K.-C. Chou. Weighted-support vector machines for predicting membrane protein types based on pseudo-amino acid composition. *Protein Eng. Des. Sel.*, 17(6):509–516, 2004.
- [488] M-L. Wang, W-J. Li, M-L. Wang, and W-B. Xu. Support vector machines for prediction of peptidyl prolyl cis/trans isomerization. *J Pept Res*, 63(1):23–8, Jan 2004.
- [489] Ming-Lei Wang, Hui Yao, and Wen-Bo Xu. Prediction by support vector machines and analysis by Z-score of poly-L-proline type II conformation based on local sequence. *Comput. Biol. Chem.*, 29(2):95–100, Apr 2005.
- [490] Yu Wang, Igor V Tetko, Mark A Hall, Eibe Frank, Axel Facius, Klaus F X Mayer, and Hans W Mewes. Gene selection from microarray data for cancer classification—a machine learning approach. *Comput. Biol. Chem.*, 29(1):37–46, Feb 2005.
- [491] J. J. Ward, L. J. McGuffin, B. F. Buxton, and D. T. Jones. Secondary structure prediction with support vector machines. *Bioinformatics*, 19(13):1650–1655, 2003.
- [492] Christopher A Waring and Xiuwen Liu. Face detection using spectral histograms and SVMs. *IEEE Trans Syst Man Cybern B Cybern*, 35(3):467–76, Jun 2005.
- [493] Jeffrey F Waring, Roger G Ulrich, Nick Flint, David Morfitt, Arno Kalkuhl, Frank Staedtler, Michael Lawton, Johanna M Beekman, and Laura Suter. Interlaboratory evaluation of rat hepatic gene expression changes induced by methapyrilene. *Environ Health Perspect*, 112(4):439–48, Mar 2004.
- [494] M. K. Warmuth, J. Liao, G. Rätsch, M. Mathieson, S. Putta, and C. Lemmen. Active learning with support vector machines in the drug discovery process. *J Chem Inf Comput Sci*, 43(2):667–673, 2003.
- [495] M. K. Warmuth, G. Rätsch, M. Mathieson, L. Liao, and C. Lemmen. Active learning in the drug discovery process. In T.G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Adv. Neural Inform. Process. Syst.*, volume 14, pages 1449–1456. MIT Press, 2002.
- [496] C. Watkins. Dynamic alignment kernels. In A.J. Smola, P.L. Bartlett, B. Schölkopf, and D. Schuurmans, editors, *Advances in Large Margin Classifiers*, pages 39–50. MIT Press, Cambridge, MA, 2000.
- [497] J. D. Watson and F. H. C. Crick. A Structure for Deoxyribose Nucleic Acid. *Nature*, 171:737, 1953.

- [498] E. A. Weathers, M. E. Paulaitis, T. B. Woolf, and J. H. Hoh. Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS Lett.*, 576(3):348–352, 2004.
- [499] Griffin Weber, Staal Vinterbo, and Lucila Ohno-Machado. Building an asynchronous web-based tool for machine learning classification. *Proc AMIA Symp*, pages 869–73, 2002.
- [500] J. Weston, F. Pérez-Cruz, O. Bousquet, O. Chapelle, A. Elisseeff, and B. Schölkopf. Feature selection and transduction for prediction of molecular bioactivity for drug design. *Bioinformatics*, 19(6):764–771, 2003.
- [501] W. J. Wilbur. Boosting naive Bayesian learning on a large subset of MEDLINE. *Proc AMIA Symp*, pages 918–22, 2000.
- [502] R.D. Williams, S.N. Hing, B.T. Greer, C.C. Whiteford, J.S. Wei, R. Natrajan, A. Kelsey, S. Rogers, C. Campbell, K. Pritchard-Jones, and J. Khan. Prognostic classification of relapsing favorable histology Wilms tumor using cDNA microarray expression profiling and support vector machines. *Genes Chromosomes Cancer*, 41(1):65–79, Sep 2004.
- [503] D. Wilton, P. Willett, K. Lawson, and G. Mullier. Comparison of ranking methods for virtual screening in lead-discovery programs. *J Chem Inf Comput Sci*, 43(2):469–74, 2003.
- [504] S. Winters-Hilt, W. Vercoutere, V.S. DeGuzman, D. Deamer, M. Akeson, and D. Haussler. Highly accurate classification of Watson-Crick basepairs on termini of single DNA molecules. *Biophys. J.*, 84(2):967–976, 2003.
- [505] B. Wu, T. Abbott, D. Fishman, W. McMurray, G. Mor, K. Stone, D. Ward, K. Williams, and H. Zhao. Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data. *Bioinformatics*, 19(13):1636–1643, 2003.
- [506] Dan Xie, Ao Li, Minghui Wang, Zhewen Fan, and Huanqing Feng. LOCSVMPSI: a web server for subcellular localization of eukaryotic proteins using SVM and profile of PSI-BLAST. *Nucleic Acids Res.*, 33(Web Server issue):W105–10, Jul 2005.
- [507] Lei Xie, Li Xie, and Philip E Bourne. A unified statistical model to support local sequence order independent similarity searching for ligand-binding sites and its application to genome-based drug discovery. *Bioinformatics*, 25(12):i305–i312, Jun 2009.
- [508] E. P. Xing, W. Wu, M. I. Jordan, and R. M. Karp. LOGOS: A modular Bayesian model for de novo motif detection. *J. Bioinform. Comput. Biol.*, 2:127–154, 2004.
- [509] M. Xiong, X. Fang, and J. Zhao. Biomarker Identification by Feature Wrappers. *Genome Res.*, 11(11):1878–1887, 2001.

- [510] Jian xiong Dong, Adam Krzyzak, and Ching Y Suen. Fast SVM training algorithm with decomposition on very large data sets. *IEEE Trans Pattern Anal Mach Intell*, 27(4):603–18, Apr 2005.
- [511] Xiu-Qin Xu, Chon K Leow, Xin Lu, Xuegong Zhang, Jun S Liu, Wing-Hung Wong, Arndt Asperger, Sren Deininger, and Hon-Chiu Eastwood Leung. Molecular classification of liver cirrhosis in a rat model by proteomics and bioinformatics. *Proteomics*, 4(10):3235–45, Oct 2004.
- [512] C. X. Xue, R. S. Zhang, H. X. Liu, M. C. Liu, Z. D. Hu, and B. T. Fan. Support vector machines-based quantitative structure-property relationship for the prediction of heat capacity. *J Chem Inf Comput Sci*, 44(4):1267–74, 2004.
- [513] C. X. Xue, R. S. Zhang, H. X. Liu, X. J. Yao, M. C. Liu, Z. D. Hu, and B. T. Fan. An accurate QSPR study of O-H bond dissociation energy in substituted phenols based on support vector machines. *J Chem Inf Comput Sci*, 44(2):669–77, 2004.
- [514] C. X. Xue, R. S. Zhang, H. X. Liu, X. J. Yao, M. C. Liu, Z. D. Hu, and B. T. Fan. QSAR models for the prediction of binding affinities to human serum albumin using the heuristic method and a support vector machine. *J Chem Inf Comput Sci*, 44(5):1693–700, 2004.
- [515] C. X. Xue, R. S. Zhang, M. C. Liu, Z. D. Hu, and B. T. Fan. Study of the quantitative structure-mobility relationship of carboxylic acids in capillary electrophoresis based on support vector machines. *J Chem Inf Comput Sci*, 44(3):950–7, 2004.
- [516] L. Xue, F. L. Stahura, J. W. Godden, and J. Bajorath. Fingerprint scaling increases the probability of identifying molecules with similar activity in virtual screening calculations. *J Chem Inf Comput Sci*, 41(3):746–753, 2001.
- [517] Y. Xue, Z. R. Li, C. W. Yap, L. Z. Sun, X. Chen, and Y. Z. Chen. Effect of molecular descriptor feature selection in support vector machine classification of pharmacokinetic and toxicological properties of chemical agents. *J Chem Inf Comput Sci*, 44(5):1630–8, 2004.
- [518] Y. Xue, C. W. Yap, L. Z. Sun, Z. W. Cao, J. F. Wang, and Y. Z. Chen. Prediction of P-glycoprotein substrates by a support vector machine approach. *J Chem Inf Comput Sci*, 44(4):1497–505, 2004.
- [519] Y. Yabuki, T. Muramatsu, T. Hirokawa, H. Mukai, and M. Suwa. GRIF-FIN: a system for predicting GPCR-G-protein coupling selectivity using a support vector machine and a hidden Markov model. *Nucleic Acids Res.*, 33(Web Server issue):W148–53, Jul 2005.

- [520] Y. Yamanishi, J.-P. Vert, and M. Kanehisa. Heterogeneous data comparison and gene selection with kernel canonical correlation analysis. In B. Scholkopf, K. Tsuda, and J.P. Vert, editors, *Kernel Methods in Computational Biology*, pages 209–230. MIT Press, 2004.
- [521] Y. Yamanishi, J.-P. Vert, and M. Kanehisa. Protein network inference from multiple genomic data: a supervised approach. *Bioinformatics*, 20:i363–i370, 2004.
- [522] Y. Yamanishi, J.-P. Vert, and M. Kanehisa. Supervised enzyme network inference from the integration of genomic data and chemical information. *Bioinformatics*, 21:i468–i477, 2005.
- [523] Y. Yamanishi, J.-P. Vert, A. Nakaya, and M. Kanehisa. Extraction of correlated gene clusters from multiple genomic data by generalized kernel canonical correlation analysis. *Bioinformatics*, 19(Suppl. 1):i323–i330, 2003.
- [524] C. Yan, D. Dobbs, and V. Honavar. A two-stage classifier for identification of protein-protein interface residues. *Bioinformatics*, 20(Suppl. 1):i371–i378, 2004.
- [525] C. Yan, V. Honavar, and D. Dobbs. Identification of interface residues in protease-inhibitor and antigen-antibody complexes: a support vector machine. *Neural Comput. & Applic.*, 13:123–129, 2004.
- [526] Z. R. Yang and K.-C. Chou. Bio-support vector machines for computational proteomics. *Bioinformatics*, 20(5):735–741, 2004.
- [527] Zheng Rong Yang. Biological applications of support vector machines. *Brief Bioinform*, 5(4):328–38, Dec 2004.
- [528] X. J. Yao, A. Panaye, J. P. Doucet, R. S. Zhang, H. F. Chen, M. C. Liu, Z. D. Hu, and B. T. Fan. Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks, and multiple linear regression. *J Chem Inf Comput Sci*, 44(4):1257–66, 2004.
- [529] C. W. Yap, C. Z. Cai, Y. Xue, and Y. Z. Chen. Prediction of torsade-causing potential of drugs by support vector machine approach. *Toxicol Sci*, 79(1):170–7, May 2004.
- [530] C. W. Yap and Y. Z. Chen. Prediction of Cytochrome P450 3A4, 2D6, and 2C9 Inhibitors and Substrates by Using Support Vector Machines. *J Chem Inf Model*, 45(4):982–92, 2005.
- [531] C.H. Yeang, S. Ramaswamy, P. Tamayo, S. Mukherjee, R.M. Rifkin, M. Angelo, M. Reich, E. Lander, J. Mesirov, and T. Golub. Molecular classification of multiple tumor types. *Bioinformatics*, 17(Suppl. 1):S316–S322, 2001.

- [532] S. M. Yiu, Prudence W. H. Wong, T.W. Lam, Y.C. Mui, H. F. Kung, Marie Lin, and Y. T. Cheung. Filtering of Ineffective siRNAs and Improved siRNA Design Tool. *Bioinformatics*, 21(2):144–151, Jan 2005. To appear.
- [533] Y. Yoon, J. Song, S.H. Hong, and J.Q. Kim. Analysis of multiple single nucleotide polymorphisms of candidate genes related to coronary heart disease susceptibility by using support vector machines. *Clin. Chem. Lab. Med.*, 41(4):529–534, 2003.
- [534] C. Yu, N. Zavaljevski, F. J. Stevens, K. Yackovich, and J. Reifman. Classifying noisy protein sequence data: a case study of immunoglobulin light chains. *Bioinformatics*, 21(Supp 1):i495–i501, Jun 2005.
- [535] C.-S. Yu, C.-J. Lin, and J.-K. Hwang. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.*, 13(5):1402–1406, 2004.
- [536] C.S. Yu, J.Y. Wang, J.M. Yang, P.C. Lyu, C.J. Lin, and J.K. Hwang. Fine-grained protein fold assignment by support vector machines using generalized npeptide coding schemes and jury voting from multiple-parameter sets. *Proteins*, 50(4):531, 6 2003.
- [537] J.-k. Yu, S. Zheng, Y. Tang, and L. Li. An integrated approach utilizing proteomics and bioinformatics to detect ovarian cancer. *J Zhejiang Univ Sci B*, 6(4):227–31, Apr 2005.
- [538] J. S. Yu, S. Ongarello, R. Fiedler, X. W. Chen, G. Toffolo, C. Cobelli, and Z. Trajanoski. Ovarian cancer identification based on dimensionality reduction for high-throughput mass spectrometry data. *Bioinformatics*, 21(10):2200–9, May 2005.
- [539] J.K. Yu, Y.D. Chen, and S. Zheng. An integrated approach to the detection of colorectal cancer utilizing proteomics and bioinformatics. *World J. Gastroenterol.*, 10(21):3127–3131, 2004.
- [540] Kun Yu, Nikolai Petrovsky, Christian Schnbach, Judice Y L Koh, and Vladimir Brusic. Methods for prediction of peptide binding to MHC molecules: a comparative study. *Mol Med*, 8(3):137–148, Mar 2002.
- [541] Z. Yuan, K. Burrage, and J.S. Mattick. Prediction of protein solvent accessibility using support vector machines. *Proteins*, 48(3):566–570, 2002.
- [542] Z. Yuan, J.S. Mattick, and R.D. Teasdale. SVMtm: support vector machines to predict transmembrane segments. *J. Comput. Chem.*, 25(5):632, 6 2004.
- [543] N. M. Zaki, S. Deris, and R. Illias. Application of string kernels in protein sequence classification. *Appl. Bioinformatics*, 4(1):45–52, 2005.

- [544] N. Zavaljevski, F.J. Stevens, and J. Reifman. Support vector machines with selective kernel scaling for protein classification and identification of key amino acid positions. *Bioinformatics*, 18(5):689–696, 2002.
- [545] V. V. Zernov, K. V. Balakin, A. A. Ivaschenko, N. P. Savchuk, and I. V. Pletnev. Drug discovery using support vector machines. The case studies of drug-likeness, agrochemical-likeness, and enzyme inhibition predictions. *J Chem Inf Comput Sci*, 43(6):2048–56, 2003.
- [546] G. L. Zhang, A. M. Khan, K. N. Srinivasan, J. T. August, and V. Brusic. MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides. *Nucleic Acids Res*/, 33(Web Server issue):W172–W179, Jul 2005.
- [547] Qidong Zhang, Sukjoon Yoon, and William J Welsh. Improved method for predicting beta-turn using support vector machine. *Bioinformatics*, 21(10):2370–4, May 2005.
- [548] S.-W. Zhang, Q. Pan, H.-C. Zhang, Y.-L. Zhang, and H.-Y. Wang. Classification of protein quaternary structure with support vector machine. *Bioinformatics*, 19(18):2390–2396, 2003.
- [549] X. H-F. Zhang, K. A. Heller, I. Hefter, C. S. Leslie, and L. A. Chasin. Sequence Information for the Splicing of Human Pre-mRNA Identified by Support Vector Machine Classification. *Genome Res.*, 13(12):2637–2650, 2003.
- [550] Z. Zhang, S. Kochhar, and M. G. Grigorov. Descriptor-based protein remote homology identification. *Protein Sci.*, 42(2):431–444, 2005.
- [551] Y. Zhao, C. Pinilla, D. Valmori, R. Martin, and R. Simon. Application of support vector machines for T-cell epitopes prediction. *Bioinformatics*, 19(15):1978–1984, 2003.
- [552] GuoDong Zhou, Dan Shen, Jie Zhang, Jian Su, and SoonHeng Tan. Recognition of protein/gene names from text using an ensemble of classifiers. *BMC Bioinformatics*, 6 Suppl 1:S7, 2005.
- [553] GuoDong Zhou, Jie Zhang, Jian Su, Dan Shen, and ChewLim Tan. Recognizing names in biomedical texts: a machine learning approach. *Bioinformatics*, 20(7):1178–90, May 2004.
- [554] Xin Zhou and K. Z. Mao. LS Bound based gene selection for DNA microarray data. *Bioinformatics*, 21(8):1559–64, Apr 2005.
- [555] Lingyun Zhu, Baoming Wu, and Changxiu Cao. Introduction to medical data mining. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*, 20(3):559–62, Sep 2003.

- [556] A. Zien, G. Rätsch, S. Mika, B. Schölkopf, T. Lengauer, and K.-R. Müller. Engineering support vector machine kernels that recognize translation initiation sites. *Bioinformatics*, 16(9):799–807, 2000.