References

Abernethy, D. R., and Schwartz, J. B.,Calcium-antagonist drugs, New England Journal of Medicine, 341(19), 1447-1457, 1999.

Bairoch, A., Apweiler, R., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., and Yeh, L. S. L., The universal protein resource (UniProt), Nucleic acids research, 33(suppl 1), D154-D159, 2005.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H.,and Bourne, P. E., The protein data bank, Nucleic acids research, 28(1), 235-242, 2000.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Brice, M. D., Rodgers, J. R. and Tasumi, M.,The protein data bank, European Journal of Biochemistry, 80(2), 319-324, 1977.

Bezanilla, F., Voltage-gated ion channels, NanoBioscience, IEEE Transactions on, 4(1), 34-48, 2005.

Bhasin, M., and Raghava, G. P. S., GPCRpred: an SVM-based method for prediction of families and subfamilies of G-protein coupled receptors, Nucleic Acids Research, 32(suppl 2), W383-W389, 2004.

Bhasin, M., and Raghava, G. P. S., GPCRsclass: a web tool for the classification of amine type of G-protein-coupled receptors, Nucleic acids research, 33(suppl 2), W143-W147, 2005.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., and Schneider, M.,The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003, Nucleic acids research, 31(1), 365-370, 2003.

Borro, L. C., Oliveira, S. R., Yamagishi, M. E., Mancini, A. L., Jardine, J. G., Mazoni, I. and Neshich, G., Predicting enzyme class from protein structure using Bayesian classification, Genet, Mol. Res, 5(1), 193-202, 2006.

Bradley, A. P., The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern recognition, 30(7), 1145-1159, 1997.

Breiman, L., Bagging predictors, Machine learning, 24(2), 123-140, 1996.

Breiman, L., Random forests, Machine learning, 45(1), 5-32, 2001.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A., Classification and regression trees, CRC press, 1984.

Bryll, R., Gutierrez-Osuna, R., and Quek, F., Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets, Pattern recognition, 36(6), 1291-1302, 2003.

Cai, C. Z., Han, L. Y., Ji, Z. L., Chen, X., and Chen, Y. Z., SVM-Prot: web-based support vector machine software for functional classification of a protein from its primary sequence, Nucleic acids research, 31(13), 3692-3697, 2003.

Cai, C. Z., Yuan, Q. F., Xiao, H. G., Liu, X. H., Han, L. Y., and Chen, Y. Z., Prediction of transmembrane proteins from their primary sequence by support vector machine approach, In Computational Intelligence and Bioinformatics (pp. 525-533). Springer Berlin Heidelberg, 2006.

Cai, Jun, Yanda Li, Classification of nuclear receptor subfamilies with RBF kernel in support vector machine, In Advances in Neural Networks–ISNN. pp. 680-685, 2005.

Cai, Y. D., and Chou, K. C., Using functional domain composition to predict enzyme family classes, Journal of proteome research, 4(1), 109-111, 2005.

Casari, G., Sander, C., and Valencia, A., A method to predict functional residues in proteins, Nature structural biology, (2), 171-8, 1995.

Chen, C, Tian, Y X, Zou, X Y, Cai, P X, Mo, J Y., Using pseudo-amino acid composition and support vector machine to predict protein structural class, Journal of Theoretical Biology, 243(3), 444-448, 2006.

Chen, W., and Lin, H., Identification of voltage-gated potassium channel subfamilies from sequence information using support vector machine. Computers in biology and medicine, 42(4), 504-507, 2012.

Chou, K. C., and Shen, H. B., MemType-2L: a web server for predicting membrane proteins and their types by incorporating evolution information

through Pse-PSSM. Biochemical and biophysical research communications, 360(2), 339-345, 2007.

Cortes, C., and Vapnik, V., Support-vector networks, Machine learning, 20(3), 273-297, 1995.

Cover, T., and Hart, P., Nearest neighbor pattern classification, Information Theory, IEEE Transactions on, 13(1), 21-27, 1967.

Davies, M. N., Secker, A., Freitas, A. A., Mendao, M., Timmis, J., and Flower, D. R.., On the hierarchical classification of G protein-coupled receptors, Bioinformatics, 23(23), 3113-3118, 2007.

Dietterich, T. G., Ensemble methods in machine learning, In Multiple classifier systems (pp. 1-15). Springer Berlin Heidelberg, 2000.

Ding, C., and Peng, H., Minimum redundancy feature selection from microarray gene expression data, Journal of bioinformatics and computational biology, 3(02), 185-205, 2005.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., and Vapnik, V., Support vector regression machines, Advances in neural information processing systems, 9, 155-161, 1997.

Dudani, S. A., The distance-weighted k-nearest-neighbor rule, Systems, Man and Cybernetics, IEEE Transactions on, (4), 325-327, 1976.

Eijkelkamp, N., Linley, J. E., Baker, M. D., Minett, M. S., Cregg, R., Werdehausen, R., and Wood, J. N., Neurological perspectives on voltage-gated sodium channels, Brain, 135(9), 2585-2612, 2012.

Fang, Y, Guo, Y, Feng, Y, Li, M, Predicting DNA-binding proteins: approached from Chou's pseudo amino acid composition and other specific sequence features, Amino acids. 34(1), 103-109, 2008.

Frank, H. Y., and Catterall, W. A., Overview of the voltage-gated sodium channel family, Genome Biol, 4(207), 10-1186, 2003.

Freund, Y., and Schapire, R. E., Experiments with a new boosting algorithm, In ICML (Vol. 96, pp. 148-156), 1996.

Gao, Q B, Jin, Z C, Ye, X F, Wu, C, He, J., Prediction of nuclear receptors with optimal pseudo amino acid composition, Analytical biochemistry. 387(1), 54-59, 2009.

Gao, Q. B., and Wang, Z. Z., Classification of G-protein coupled receptors at four levels, Protein Engineering Design and Selection, 19(11), 511-516, 2006.

Garg, A, Bhasin M, Raghava G P., SVM-based method for subcellular localization of human proteins using amino acid compositions, their order and similarity search, J. Biol. Chem. 280, 14427–14432, 2005.

Garg, A., and Raghava, G. P., A machine learning based method for the prediction of secretory proteins using amino acid composition, their order and similarity-search, In silico biology, 8(2), 129-140, 2008.

Gu, Q, Ding, Y S, Zhang, T L., Prediction of G-Protein-Coupled Receptor Classes in Low Homology Using Chous Pseudo Amino Acid Composition with Approximate Entropy and Hydrophobicity Patterns, Protein and peptide letters. 17(5), 559-567, 2010

Guo, Y. Z., Li, M., Lu, M., Wen, Z., Wang, K., Li, G., and Wu, J., Classifying G protein-coupled receptors and nuclear receptors on the basis of protein power spectrum from fast Fourier transform, Amino Acids, 30(4), 397-402, 2006.

Gupta, R., Mittal, A., and Singh, K., A novel and efficient technique for identification and classification of GPCRs, Information Technology in Biomedicine, IEEE Transactions on, 12(4), 541-548, 2008.

Guyon, I., and Elisseeff, A., An introduction to variable and feature selection, The Journal of Machine Learning Research, 3, 1157-1182, 2003.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V., Gene selection for cancer classification using support vector machines, Machine learning, 46(1-3), 389-422, 2002.

Hagan, M. T., Demuth, H. B., and Beale, M. H., Neural network design (pp. 2-14), Boston: Pws Pub., 1996.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H., The WEKA data mining software: an update, ACM SIGKDD explorations newsletter, 11(1), 10-18, 2009.

Hanley, J. A., and McNeil, B. J., The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology, 143(1), 29-36, 1982.

Harmar, A. J., Hills, R. A., Rosser, E. M., Jones, M., Buneman, O. P., Dunbar, D. R., and Spedding, M., IUPHAR-DB: the IUPHAR database of G protein-

coupled receptors and ion channels, Nucleic acids research, 37(suppl 1), D680-D685, 2009.

Hechenbichler, K., and Schliep, K., Weighted k-nearest-neighbor techniques and ordinal classification, 2004.

Ho, T. K., The random subspace method for constructing decision forests, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 20(8), 832-844, 1998.

Horn, F., Bettler, E., Oliveira, L., Campagne, F., Cohen, F. E., and Vriend, G. GPCRDB information system for G protein-coupled receptors, Nucleic acids research, 31(1), 294-297, 2003.

Hua, S., and Sun, Z., Support vector machine approach for protein subcellular localization prediction, Bioinformatics, 17(8), 721-728, 2001.

Huang, G., Zhang, Y., Chen, L., Zhang, N., Huang, T., and Cai, Y. D., Prediction of multi-type membrane proteins in human by an integrated approach. PloS one, 9(3), e93553.

Huang, W. L., Chen, H. M., Hwang, S. F., and Ho, S. Y., "Accurate prediction of enzyme subfamily class using an adaptive fuzzy k-nearest neighbor method, Biosystems, 90(2), 405-413, 2007, 2014.

Huang, Y, and Li, Y., Prediction of protein subcellular locations using fuzzy k-NN method. Bioinformatics, 20(1), 21-28, 2004.

Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W., CD-HIT Suite: a web server for clustering and comparing biological sequences, Bioinformatics, 26(5), 680-682, 2010.

Huang, Ying,Yanda Li, .Prediction of protein subcellular locations using fuzzy k-NN method, Bioinformatics .20.1 ,21-28, 2004.

Jensen, L. J., Skovgaard, M., and Brunak, S., Prediction of novel archaeal enzymes from sequence-derived features, Protein science, 11(12), 2894-2898, 2002.

Jentsch, T. J., Hübner, C. A., and Fuhrmann, J. C., Ion channels: function unravelled by dysfunction, Nature cell biology, 6(11), 1039-1047, 2004.

Kass, G. V., An exploratory technique for investigating large quantities of categorical data, Applied statistics, 119-127, 1980.

Kawashima, S., and Kanehisa, M., AAindex: amino acid index database, Nucleic acids research, 28(1), 374-374, 2000.

Keller, F., Naive Bayes Classifiers. Connectionist and Statistical Language Processing, Course at Universität des Saarlandes, 2002.

Keller, J. M., Gray, M. R., and Givens, J. A., A fuzzy k-nearest neighbor algorithm, Systems, Man and Cybernetics, IEEE Transactions on, (4), 580-585, 1985.

Kim, S Y, Sim, J, and Lee, J., Fuzzy k-nearest neighbor method for protein secondary structure prediction and its parallel implementation, In Computational Intelligence and Bioinformatics (pp. 444-453), Springer Berlin Heidelberg, 2006.

Kira, K., and Rendell, L. A., The feature selection problem: Traditional methods and a new algorithm. In AAAI (pp. 129-134), 1992.

Kolakowski Jr, L. F., GCRDb: a G-protein-coupled receptor database, Receptors and channels, 2(1), 1-7, 1993.

Kotsiantis, S., Combining bagging, boosting, rotation forest and random subspace methods, Artificial Intelligence Review, 35(3), 223-240, 2011.

Kumar, C., and Choudhary, A., A top-down approach to classify enzyme functional classes and sub-classes using random forest, EURASIP Journal on Bioinformatics and Systems Biology, 2012(1), 1-14, 2012.

Le Novère, N., and Changeux, J. P., The ligand gated ion channel database, Nucleic acids research, 27(1), 340-342, 1999.

Lee, B. J., Shin, M. S., Oh, Y. J., Oh, H. S., and Ryu, K. H., Identification of protein functions using a machine-learning approach based on sequence-derived properties, Proteome science, 7(1), 27, 2009.

Li, F M, Li, Q Z., Predicting Protein Subcellular Location Using Chou's Pseudo Amino Acid Composition and Improved Hybrid Approach, Protein Pept Lett, 15, 612-616, 2008.

Li, Z., Zhou, X., Dai, Z., and Zou, X., Classification of G-protein coupled receptors based on support vector machine with maximum relevance minimum redundancy and genetic algorithm, BMC bioinformatics, 11(1), 325, 2010.

Lin, H. H., Han, L. Y., Zhang, H. L., Zheng, C. J., Xie, B., and Chen, Y. Z., Prediction of the functional class of lipid binding proteins from sequence-derived properties irrespective of sequence similarity, Journal of lipid research, 47(4), 824-831, 2006.

Lin, H., and Ding, H., Predicting ion channels and their types by the dipeptide mode of pseudo amino acid composition, Journal of Theoretical Biology, 269(1), 64-69, 2011.

Lin, W Z, Fang, J A, Xiao, X, Chou, K C., iDNA-Prot: identification of DNA binding proteins using random forest with grey model, PLoS One. 6(9), e24756, 2011.

LIU, L. X., LI, M. L., TAN, F. Y., LU, M. C., WANG, K. L., GUO, Y. Z. and Jiang, L., Local Sequence Information-based Support Vector Machine to Classify Voltage-gated Potassium Channels, Acta biochimica et biophysica Sinica, 38(6), 363-371, 2006.

Lu, L, Qian, Z, Cai, Y D, Li, Y., ECS: An automatic enzyme classifier based on functional domain composition, Comput Biol Chem.. 31 (3), 226-232, 2007.

Ma, J. W., A novel method for predicting protein subcellular localization based on pseudo amino acid composition, Biochemistry and Molecular Biology Reports, 43(10), 670-676, 2010.

Mbah, A. N., Application of Hybrid Functional Groups to Predict ATP Binding Proteins, International Scholarly Research Notices, 2014.

Moore, J. T., Collins, J. L., and Pearce, K. H., The nuclear receptor superfamily and drug discovery, ChemMedChem, 1(5), 504-523, 2006.

Murphy, K. P,. Naive bayes classifiers, University of British Columbia, 2006.

Nagao, C., Nagano, N., and Mizuguchi, K., Prediction of detailed enzyme functions and identification of specificity determining residues by random forests, PloS one, 9(1), e84623, 2014.

Nasibov, E., and Kandemir-Cavas, C., Efficiency analysis of KNN and minimum distance-based classifiers in enzyme family prediction, Computational biology and chemistry, 33(6), 461-464, 2009.

Ozcift, A., and Gulten, A., Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning

algorithms, Computer methods and programs in biomedicine, 104(3), 443-45, 2011.

Papasaikas, P. K., Bagos, P. G., Litou, Z. I., Promponas, V. J., and Hamodrakas, S. J., PRED-GPCR: GPCR recognition and family classification server, Nucleic acids research, 32(suppl 2), W380-W382, 2004.

Peng, H, Long, F, and Ding, C., Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, Pattern Analysis and Machine Intelligence, IEEE Transactions on. 27(8), 1226-1238, 2005.

Peng, Z. L., Yang, J. Y., and Chen, X., An improved classification of G-protein-coupled receptors using sequence-derived features. BMC bioinformatics, 11(1), 420, 2010.

Petrova, N. V., and Wu, C. H., Prediction of catalytic residues using Support Vector Machine with selected protein sequence and structural properties, BMC bioinformatics, 7(1), 312, 2006.

Pourbasheer, E., Riahi, S., Ganjali, M. R., and Norouzi, P., Application of genetic algorithm-support vector machine (GA-SVM) for prediction of BK-channels activity. European journal of medicinal chemistry, 44(12), 5023-5028, 2009.

Qiu, J. D., Huang, J. H., Shi, S. P., and Liang, R. P., Using the concept of Chou's pseudo amino acid composition to predict enzyme family classes: an approach with support vector machine based on discrete wavelet transform, Protein and peptide letters, 17(6), 715-722, 2010.

Quinlan, J. R., Learning decision tree classifiers, ACM Computing Surveys (CSUR), 28(1), 71-72, 1996.

Quinlan, J. R., Bagging, boosting, and C4. 5, In AAAI/IAAI, Vol. 1 (pp. 725-730), 1996.

Quinlan, J. R., C4. 5: programs for machine learning, Elsevier. 2014.

Quinlan, R., Data mining tools See5 and C5. 0, 2004.

Rao, H. B., Zhu, F., Yang, G. B., Li, Z. R., and Chen, Y. Z., Update of PROFEAT: a web server for computing structural and physicochemical features of proteins and peptides from amino acid sequence, Nucleic acids research, 39(suppl 2), W385-W390, 2011.

Robinson-Rechavi, M., Garcia, H. E., and Laudet, V., The nuclear receptor superfamily, Journal of cell science, 116(4), 585-586, 2003.

Rodriguez, J. J., Kuncheva, L. I., and Alonso, C. J., Rotation forest: A new classifier ensemble method, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(10), 1619-1630, 2006.

Saeys, Y., Inza, I., and Larrañaga, P., A review of feature selection techniques in bioinformatics, bioinformatics, 23(19), 2507-2517, 2007.

Saha, S., Zack, J., Singh, B., and Raghava, G. P. S., VGIchan: prediction and classification of voltage-gated ion channels, Genomics, proteomics and bioinformatics, 4(4), 253-258, 2006.

Sandhiya, S., and Dkhar, S. A., Potassium channels in health, disease and development of channel modulators, 2009.

Sands, Z., Grottesi, A., and Sansom, M. S., Voltage-gated ion channels, Current Biology, 15(2), R44-R47, 2005.

Schalkoff, R. J., Artificial neural networks (pp. 162-164). New York: McGraw-Hill, 1997.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C., SWISS-MODEL: an automated protein homology-modeling server, Nucleic acids research, 31(13), 3381-3385, 2003.

Shen, H B, Yang, J, and Chou, K C., Fuzzy KNN for predicting membrane protein types from pseudo-amino acid composition, Journal of theoretical biology, 240(1), 9-13, 2006.

Shen, H. B., and Chou, K. C., PseAAC: a flexible web server for generating various kinds of protein pseudo amino acid composition, Analytical biochemistry, 373(2), 386-388, 2008.

Shi, J. Y., Zhang, S. W., Pan, Q., Cheng, Y. M., and Xie, J., Prediction of protein subcellular localization by support vector machines using multi-scale energy and pseudo amino acid composition, Amino Acids, 33(1), 69-74, 2007.

Sim, J, Kim, S Y, and Lee, J., Prediction of protein solvent accessibility using fuzzy k-nearest neighbor method, Bioinformatics, 21(12), 2844-2849, 2005.

Skolnick, J., and Fetrow, J. S., From genes to protein structure and function: novel applications of computational approaches in the genomic era, Trends in biotechnology, 18(1), 34-39, 2000.

Söding, J., Protein homology detection by HMM–HMM comparison, Bioinformatics, 21(7), 951-960, 2005.

UniProt Consortium, The universal protein resource (UniProt), Nucleic acids research, 36(suppl 1), D190-D195, 2008.

Volpato, V., Adelfio, A., and Pollastri, G., Accurate prediction of protein enzymatic class by N-to-1 Neural Networks, BMC bioinformatics, 14(Suppl 1), S11, 2013.

Vroling, B, Thorne, D, McDermott, P, Joosten, H J, Attwood, T K, Pettifer, S, Vriend, G., NucleaRDB: information system for nuclear receptors, Nucleic acids research, gkr960, 2011.

Wang, J, Sung, W K, Krishnan, A, Li, K B., Protein subcellular localization prediction for Gram-negative bacteria using amino acid subalphabets and a combination of multiple support vector machines, BMC Bioinformatics, 6, 174, 2005.

Wang, M., Yang, J., Liu, G. P., Xu, Z. J., and Chou, K. C., Weighted-support vector machines for predicting membrane protein types based on pseudo-amino acid composition, Protein Engineering Design and Selection, 17(6), 509-516, 2004.

Wang, P, Xiao, X., NRPred-FS: A Feature Selection based Two-level Predictor for Nuclear Receptors, J Proteomics Bioinform S, 9, 2, 2014.

Wang, P, Xiao, X, Chou, K C., NR-2L: a two-level predictor for identifying nuclear receptor subfamilies based on sequence-derived features., PLoS One, 6(8), e23505, 2011.

Wang, T., Yang, J., Dimensionality Reduction Method for Predicting Membrane Proteins and Their Types, 2010.

Wang, Y C, Wang, X B, Yang, Z X, Deng, N Y., Prediction of enzyme subfamily class via pseudo amino acid composition by incorporating the conjoint triad feature, Protein Pept. Lett. 17, 1441–1449, 2010.

Wang, Y. C., Wang, Y., Yang, Z. X., and Deng, N. Y., Support vector machine prediction of enzyme function with conjoint triad feature and hierarchical context, BMC systems biology, 5(Suppl 1), S6, 2011.

Watson, J. D., Laskowski, R. A., and Thornton, J. M., Predicting protein function from sequence and structural data, Current opinion in structural biology, 15(3), 275-284, 2005.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., and Yaschenko, E., Database resources of the national center for biotechnology information, Nucleic acids research, 35(suppl 1), D5-D12, 2007.

Willett, P., Wilton, D., Hartzoulakis, B., Tang, R., Ford, J., and Madge, D., Prediction of ion channel activity using binary kernel discrimination, Journal of chemical information and modeling, 47(5), 1961-1966, 2007.

Wold, S., Esbensen, K., and Geladi, P., Principal component analysis, Chemometrics and intelligent laboratory systems, 2(1), 37-52, 1987.

Worster, A., Fan, J., and Upadhye, S., Understanding receiver operating characteristic (ROC) curves, Cjem, 8(1), 19-20, 2006.

Xiao, X., Wang, P., and Chou, K. C., iNR-PhysChem: a sequence-based predictor for identifying nuclear receptors and their subfamilies via physical-chemical property matrix, PloS one, 7(2), e30869, 2012.

Yadav, A., and Jayaraman, V. K., Structure based function prediction of proteins using fragment library frequency vectors, Bioinformation, 8(19), 953. 2012.

Yang, X. G., Luo, R. Y., and Feng, Z. P., Using amino acid and peptide composition to predict membrane protein types, Biochemical and biophysical research communications, 353(1), 164-169, 2007.

Yogeeswari, P., Ragavendran, J. V., Thirumurugan, R., Saxena, A., and Sriram, D., Ion channels as important targets for antiepileptic drug design, Current drug targets, 5(7), 589-602, 2004.

Yu, L., and Liu, H., Feature selection for high-dimensional data: A fast correlation-based filter solution, In ICML (Vol. 3, pp. 856-863), 2003.
Yu, Y., SVM-RFE Algorithm for Gene Feature Selection, Computer Engineering, 2008.

Zhang, C. X., and Zhang, J. S., RotBoost: A technique for combining Rotation Forest and AdaBoost, Pattern Recognition Letters, 29(10), 1524-1536, 2008.
Zhang, T. L., Ding, Y. S., and Chou, K. C., Prediction protein structural classes with pseudo-amino acid composition: approximate entropy and hydrophobicity pattern, Journal of theoretical biology, 250(1), 186-193, 2008.

Zhao, P. Y., and Ding, Y. S., Prediction of membrane protein types by an ensemble classifier based on pseudo amino acid composition and approximate entropy, In BioMedical Engineering and Informatics, 2008. BMEI 2008. International Conference on (Vol. 1, pp. 164-168). IEEE, 2008.

Zhou, X. B., Chen, C., Li, Z. C., and Zou, X. Y., Using Chou's amphiphilic pseudo-amino acid composition and support vector machine for prediction of enzyme subfamily classes, Journal of theoretical biology, 248(3), 546-551, 2007.