

References

- [1] Abonyi, J., Feil, B.: Cluster analysis for data mining and system identification. Birkhäuser, Basel (2007)
- [2] Ackley, D.: A connectionist machine for genetic hillclimbing. Kluwer Academic Publishers, Dordrecht (1987)
- [3] Adamo, J.M.: Data Mining for association rules and sequential patterns: sequential and parallel algorithms. Springer, Heidelberg (2001)
- [4] Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proc. 1993 ACM SIGMOD Conference, Washington, pp. 207–216 (1993)
- [5] Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules in Large Databases. In: Proc. 20th International Conference on Very Large Data Bases, pp. 487–499. Morgan Kaufmann, San Francisco (1994)
- [6] Agrawal, R., Srikant, R.: Mining sequential patterns. In: Proc. 11th ICDE 1995, Taipei, Taiwan, pp. 3–14 (March 1995)
- [7] Agrawal, R., Srikant, R.: Mining sequential patterns: Generalizations and performance improvements. In: Apers, P.M.G., Bouzeghoub, M., Gardarin, G. (eds.) EDBT 1996. LNCS, vol. 1057, pp. 3–17. Springer, Heidelberg (1996)
- [8] Aldrich, J.H., Nelson, F.D.: Linear probability, logit and probit models. Sage University Papers, Thousand Oaks (1985)
- [9] Allen, D.M.: Mean square error of prediction as a criterion for selecting variables. *Technometrics* 13, 469–475 (1971)
- [10] Altman, D.G.: Practical statistics for medical research. Chapman and Hall, Boca Raton (1991)
- [11] Altman, D.G., Bland, J.M.: Diagnostic tests 1: Sensitivity and specificity. *BMJ* 308, 1552 (1994)
- [12] Altman, D.G., Bland, J.M.: Diagnostic tests 2: Predictive values. *BMJ* 309, 102 (1994)
- [13] Andersen, R.: Modern Methods for Robust Regression. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-152. Sage, Thousand Oaks (2008)
- [14] An, L., Tong, L.: Similarity, Boolean Reasoning and Rule Induction. In: Proc. Second International Symposium on Intelligent Information Technology Application, vol. 1, pp. 7–12 (2008)
- [15] Anderson, J.A.: An introduction to Neural Networks. MIT Press, Cambridge (1995)
- [16] Aris, R.: Mathematical modelling techniques. Pitman (1978)

- [17] Aspray, W., Burks, A.: Papers of John von Neumann on computing and computer theory. Charles Babbage Institute Reprint Series for the history of computing, vol. 12. MIT Press, Cambridge (1986)
- [18] Back, T.: Self-adaptation. In: Back, T., Fogel, D.B., Michalewicz, Z. (eds.) *Evolutionary computation 2: advanced algorithms and operators*, ch. 21, pp. 188–211. Institute of Physics Publishing (2000)
- [19] Back, T., Fogel, D.B., Michalewicz, Z. (eds.): *Evolutionary computation 1: basic algorithms and operators*. Institute of Physics Publishing, Bristol (2000)
- [20] Baker, J.E.: Reducing bias and inefficiency in the selection algorithm. In: Grefenstette, J.J. (ed.) *Proc. 2nd International Conference on Genetic Algorithms and Their Applications*, pp. 14–21. Lawrence Erlbaum, Hillsdale (1987)
- [21] Banzhaf, W., Nordin, P., Keller, R.E., Francone, F.D.: *Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications*. Morgan Kaufmann, San Francisco (1998)
- [22] Barricelli, N.A.: Esempi numerici di processi di evoluzione. *Methodos*, 45–68 (1954)
- [23] Barto, A.G., Sutton, R.S., Anderson, C.W.: Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics SMC-13*, 834–846 (1983)
- [24] Bates, D.M., Watts, D.G.: *Nonlinear Regression Analysis and Its Applications*. Wiley Series in Probability and Statistics. Wiley, Chichester (1988)
- [25] Bayes, T.: An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London* 53, 370–418 (1763)
- [26] Beachley, N.H., Harrison, H.L.: *Introduction to dynamic systems analysis*. Harper and Row, New York (1978)
- [27] Belciug, S., Gorunescu, F., Gorunescu, M., Salem, A.B.: Assessing Performances of Unsupervised and Supervised Neural Networks in Breast Cancer Detection. In: *Proc. 7th International Conference on INFormatics and Systems (INFOS 2010)*, ADEM, pp. 80–87 (2010)
- [28] Belciug, S., Gorunescu, F.: A hybrid neural network/genetic algorithm applied to the breast cancer detection and recurrence. *Expert Systems*. Wiley & Blackwell (2011) (in press)
- [29] Bender, E.A.: *An introduction to mathematical modeling*. John Wiley, Chichester (1978)
- [30] Berkhin, P.: *Survey of clustering data mining techniques*. Technical Report, Accrue Software, San Jose, California, <http://citeseer.nj.nec.com/berkhin02survey.html>
- [31] Berry, M., Linoff, G.: *Mastering Data Mining*. John Wiley and Sons, Chichester (2000)
- [32] Bezdek, J.C.: *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, New York (1981)
- [33] Beyer, H.G.: *The theory of evolution strategies*. Springer, Heidelberg (2010)
- [34] Bishop, C.: *Neural networks for pattern recognition*. Oxford University Press, Oxford (1995)
- [35] Blumer, A., Ehrenfeucht, A., Haussler, D., Warmuth, M.K.: Occam's razor. *Information Processing Letters* 24, 377–380 (1987)
- [36] Boring, S.: Sensitivity, specificity, and predictive value. In: Walker, H.K., Hall, W.D., Hurst, J.W. (eds.) *Clinical Methods: The History, Physical, and Laboratory Examinations*, ch. 6, 3rd edn. Butterworths, London (1990)
- [37] Borji, A., Hamidi, M.: Evolving a fuzzy rule-base for image segmentation. *World Academy of Science, Engineering and Technology* 28, 4–9 (2007)

- [38] Boser, B., Guyon, I., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proc. 5th Annual Workshop on Computational Learning Theory, San Mateo, California, USA, pp. 144–152 (1992)
- [39] Box, G.E.P.: Evolutionary operation: a method of increasing industrial productivity. *Applied Statistics* 6, 81–101 (1957)
- [40] Box, G.E.P., Jenkins, G.M.: *Time series analysis: Forecasting and control*. Holden-Day, San Francisco (1976)
- [41] Boyce, B.R., Meadow, C.T., Kraft, D.H.: *Measurement in information science*. Academic Press, London (1994)
- [42] Brause, R., Langsdorf, T., Hepp, M.: Neural Data Mining for Credit Card Fraud Detection. In: Proc. 11th IEEE International Conference on Tools with Artificial Intelligence, vol. 103, pp. 103–106 (1999)
- [43] Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: *Classification and regression trees*. Wadsworth and Brooks/Cole Advanced Books and Software, Monterey (1984)
- [44] Bremermann, H.J.: Optimization through evolution and recombination. In: Yovits, M.C., Jacobi, G.T., Goldstein, G.D. (eds.) *Self-Organizing Systems*, pp. 93–106. Spartan Books, Washington (1962)
- [45] Brown, M.P., Grundy, W.N., Lin, D., Cristianini, N., Sugnet, C.W., Furey, T.S., Ares Jr., M., Haussler, D.: Knowledge-based Analysis of Microarray Gene Expression Data by Using Support Vector Machines. *Proc. of the National Academy of Sciences USA* 97(1), 262–267 (2000)
- [46] Bull, L., Kovacs, T. (eds.): *Foundations of Learning Classifier Systems: An Introduction*. STUDEFUZZ, vol. 183. Springer, Heidelberg (2005)
- [47] Bull, L., Bernardo-Mansilla, E., Holmes, J. (eds.): *Learning Classifier Systems in Data Mining: An Introduction*. SCI, vol. 125. Springer, Heidelberg (2008)
- [48] Burbidge, R., Trotter, M., Buxton, B., Holden, S.: Drug Design by Machine Learning: Support Vector Machines for Pharmaceutical Data Analysis. *Computers and Chemistry* 26(1), 5–14 (2001)
- [49] Burges, C.: A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery* 2, 121–167 (1998)
- [50] Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., Zanasi, A.: *Discovering Data Mining from Concept to Implementation*. Prentice-Hall, Englewood Cliffs (1997)
- [51] Cacoulos, T.: Estimation of a multivariate density. *Ann. Inst. Stat. Math (Tokyo)* 18, 179–189 (1966)
- [52] Cazzaniga, M., Salerno, F., Borroni, G., Ceriani, R., Stucchi, G., Guerzoni, P., Casiraghi, M.A., Tommasini, M.: Prediction of asymptomatic cirrhosis in chronic hepatitis C patients: accuracy of artificial neural networks compared with logistic regression models. *Eur. J. Gastroenterol. Hepatol.* 21(6), 681–687 (2009)
- [53] Chaudhuri, S., Dayal, U.: Data Warehousing and OLAP for Decision Support. In: *ACM SIGMOD Int. Conf. on Management of Data*, vol. 26(2), pp. 507–508. ACM Press, Tucson (1997)
- [54] Chen, H., Wang, X.F., Ma, D.Q., Ma, B.R.: Neural network-based computer-aided diagnosis in distinguishing malignant from benign solitary pulmonary nodules by computed tomography. *Chin. Med. J. (Engl.)* 120(14), 1211–1215 (2007)
- [55] Cheng, C.-H., Chen Jr., S.: Diagnosing cardiovascular disease using an enhanced rough sets approach. *Applied Artificial Intelligence* 23(6), 487–499 (2009)
- [56] Cheng, J.-H., Chen, H.-P., Cheng, K.-L.: Business failure prediction model based on grey prediction and rough set theory. *WSEAS Transactions on Information Science and Applications* 6(2), 329–339 (2009)

- [57] Cherkassky, V., Mulier, F.: *Learning from data: concepts, theory and methods*. Wiley, Chichester (1998)
- [58] Chi, C.L., Street, W.N., Wolberg, W.H.: Application of artificial neural network-based survival analysis on two breast cancer datasets. In: *AMIA Annu. Symp. Proc.*, vol. 11, pp. 130–134 (2007)
- [59] Chiu, J.S., Wang, Y.F., Su, Y.C., Wei, L.H., Liao, J.G., Li, Y.C.: Artificial neural network to predict skeletal metastasis in patients with prostate cancer. *J. Med. Syst.* 33(2), 91–100 (2009)
- [60] Christensen, T., Neuberger, J., Crowe, J., et al.: Beneficial effect of azathioprine and prediction of prognosis in primary biliary cirrhosis: final results of an international trial. *Gastroenterology* 89, 1084–1091 (1985)
- [61] Clark, P., Niblett, T.: The CN2 Induction Algorithm. *Machine Learning* 3(4), 261–283 (1989)
- [62] Cliff, N., Krus, D.J.: Interpretation of canonical variate analysis: Rotated vs. unrotated solutions. *Psychometrika* 41(1), 35–42 (1976)
- [63] Cochran, W.G.: *Sampling Techniques*. John Wiley and Sons, Chichester (1977)
- [64] De Cock, M., Cornelis, C., Kerre, E.E.: Fuzzy Rough Sets: Beyond the Obvious. In: *Proc. FUZZIEEE 2004 (2004 IEEE International Conference on Fuzzy Systems)*, vol. 1, pp. 103–108 (2004)
- [65] Codd, E.F., Codd, S.B., Salley, C.T.: *Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate*. Codd and Date Inc. (1993)
- [66] Cohen, W.: Fast Effective Rule Induction. In: *Proc. 12th Intl. Conf. on Machine Learning, Tahoe City, CA, USA*, pp. 115–123 (1995)
- [67] Cooley, W.W., Lohnes, P.R.: *Multivariate data analysis*. Wiley, New York (1971)
- [68] Cortes, C., Vapnik, V.N.: Support vector networks. *Machine Learning* 20, 273–297 (1995)
- [69] Cost, S., Salzberg, S.: A Weighted Nearest Neighbor Algorithm for Learning with Symbolic Features. *Machine Learning* 10(1), 57–78 (1993)
- [70] Cover, T.M.: Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers EC-14*, 326–334 (1965)
- [71] Cox, D.R., Miller, D.: *The theory of stochastic processes*. Methuen (1965)
- [72] Cox, D.R.: Regression models and life tables. *J. Roy. Statist. Soc. [B]* 34, 187–202 (1972)
- [73] Cristianini, N., Shawe-Taylor, J.: *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, Cambridge (2000)
- [74] Darlington, R.B.: *Regression and linear models*. McGraw-Hill, New York (1990)
- [75] Darwiche, A.: *Modeling and reasoning with Bayesian Networks*. Cambridge University Press, Cambridge (2009)
- [76] Darwin, C.: *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 1st edn. John Murray, London (1859)
- [77] Dasarathy, B.V. (ed.): *Nearest Neighbour Norms: Nearest Neighbour Pattern Classification Techniques*. IEEE Computer Society Press, Los Alamitos (1991)
- [78] Davis, L.: *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, New York (1991)
- [79] De Jong, K.: *An analysis of the behaviour of a class of genetic adaptive systems*. PhD thesis, University of Michigan (1975)
- [80] De Jong, K.: *Evolutionary computation*. MIT Press, Cambridge (2002)

- [81] Delgado, M., Pegalajar, M.C., Cuellar, M.P.: Memetic evolutionary training for recurrent neural networks: an application to time-series prediction. *Expert Systems* 23(2), 99–115 (2006)
- [82] Domingos, P.: The Role of Occam's Razor in Knowledge Discovery. *Data Mining and Knowledge Discovery* 3, 409–425 (1999)
- [83] Dos Santos, W.P., de Souza, R.E., dos Santos Filho, P.B.: Evaluation of Alzheimer's disease by analysis of MR images using multilayer perceptrons and Kohonen SOM classifiers as an alternative to the ADC maps. In: *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 2118–2121 (2007)
- [84] Draper, N.R., Smith, H.: *Applied regression analysis*, 2nd edn. John Wiley, Chichester (1981)
- [85] Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V.: Support Vector Regression Machines. In: Mozer, M.C., Jordan, M.I., Petsche, T. (eds.) *Advances in Neural Information Processing Systems*, vol. 9, pp. 155–161. MIT Press, Cambridge (1997)
- [86] Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems* 17, 191–209 (1990)
- [87] Dubois, D., Prade, H.: Putting rough sets and fuzzy sets together. In: Slowinski, R. (ed.) *Intelligent Decision Support-Handbook of Applications and Advances of the Rough Sets Theory*, pp. 203–232. Kluwer Academic Publishers, Dordrecht (1992)
- [88] Duda, R.O., Hart, P.E.: *Pattern classification and scene analysis*. John Wiley, Chichester (1973)
- [89] Duda, R.O., Hart, P.E., Stork, D.H.: *Pattern Classification*, 2nd edn. Wiley Interscience, Hoboken (2001)
- [90] Dumitrescu, D.: *Genetic algorithms and evolutionary strategies applications in Artificial Intelligence and connex domains*. Blue Publishing House, Cluj-Napoca (2000)
- [91] Dunteman, G.H.: *Principal Components Analysis. Series: Quantitative applications in the social science*. Sage University Paper 69 (1989)
- [92] Dym, C.L., Ivey, E.S.: *Principles of mathematical modeling*. Academic Press, London (1980)
- [93] Dzubera, J., Whitley, D.: Advanced correlation analysis of operators for the Traveling Salesman Problem. In: Davidor, Y., Männer, R., Schwefel, H.-P. (eds.) *PPSN 1994. LNCS*, vol. 866, pp. 68–77. Springer, Heidelberg (1994)
- [94] Eiben, A.E., Schippers, C.A.: On evolutionary exploration and exploitation. *Fundamenta Informaticae* 35(1-4), 35–50 (1998)
- [95] Eiben, A.E., Hinterding, R., Michalewicz, Z.: Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 3(2), 124–141 (1999)
- [96] Eiben, A.E.: Multiparent recombination in evolutionary computing. In: Gosh, A., Tsutsui, S. (eds.) *Advances in Evolutionary Computation: Theory and Applications*, pp. 175–192. Springer, Heidelberg (2003)
- [97] Eiben, A.E., Smith, J.E.: *Introduction to Evolutionary Computing*. Springer, Heidelberg (2003)
- [98] Eiben, A.E., Michalewicz, Z., Schoenauer, M., Smith, J.E.: *Parameter Control in Evolutionary Algorithms. SCI*, vol. 54. Springer, Heidelberg (2007)
- [99] Egan, J.P.: *Signal detection theory and ROC analysis. Series in Cognition and Perception*. Academic Press, New York (1975)

- [100] El-Darzi, E., Gorunescu, M., Gorunescu, F.: A non-homogeneous kinetic model in the patients flow modeling. In: Gorunescu, F., El-Darzi, E., Gorunescu, M. (eds.) Proc. First East European Conference on Health Care Modelling and Computation-HCMC 2005, Craiova, Romania, pp. 99–106. Medical University Press (2005)
- [101] El-Darzi, E., Abbi, R., Vasilakis, C., Gorunescu, F., Gorunescu, M., Millard, P.: Length of stay-based clustering methods for patient grouping. In: McClean, S., Millard, P., El-Darzi, E., Nugent, C. (eds.) Intelligent Patient Management. Studies in Computational Intelligence, vol. 189, pp. 39–56. Springer, Heidelberg (2009)
- [102] El-Gamal, M.A., Mohamed, M.D.: Ensembles of Neural Networks for Fault Diagnosis in Analog Circuits. *Journal of Electronic Testing: Theory and Applications* 23(4), 323–339 (2007)
- [103] Elsner, U.: Static and Dynamic Graph Partitioning. A comparative study of existing algorithms. Logos (2002)
- [104] English, T.M.: Evaluation of evolutionary and genetic optimizers: no free lunch. In: Fogel, L.J., Angeline, P.J., Back, T. (eds.) Proc. 5th Annual Conference on Evolutionary Programming, pp. 163–169. MIT Press, Cambridge (1996)
- [105] Everitt, B., Landau, S., Leese, M.: Cluster analysis. Wiley, Chichester (2001)
- [106] Fausett, L.V.: Fundamentals of Neural Networks: architectures, algorithms and applications. Prentice-Hall, Englewood Cliffs (1993)
- [107] Fawcett, T.: An introduction to ROC analysis. *Pattern Recognition Letters* 27, 861–874 (2006)
- [108] Fayyad, U., Piasteky-Shapiro, G., Smyth, P., Uthurusamy, R.: Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press (1996)
- [109] Fibak, J., Pawlak, Z., Slowinski, K., Slowinski, R.: Rough sets based decision algorithm for treatment of duodenal ulcer by HSV. *Bull. Polish Acad. Sci. Biological Sci.* 34(10-12), 227–246 (1986)
- [110] Fisher, D.: Improving Inference through Conceptual Clustering. In: Proc. of 1987 AAAI Conferences, Seattle, Washington, pp. 461–465 (1987)
- [111] Fisher, D.: Knowledge acquisition via incremental conceptual clustering. *Machine Learning* 2, 139–172 (1987)
- [112] Fisher, R.A.: On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A* 222, 309–368 (1922)
- [113] Fisher, R.A.: The use of multiple measurements in taxonomic problems. *Annual Eugenics* 7(Part II), 179–188 (1936)
- [114] Foucart, T., Bensaber, A., Garnier, R.: *Methodes pratiques de la Statistique*. Masson (1987)
- [115] Fraser, A.S.: Simulation of Genetic Systems by Automatic Digital Computers. I. Introduction. *Australian Journal of Biological Sciences* 10, 484–491 (1957)
- [116] Frawley, W., Piasteky-Shapiro, G., Matheus, C.: Knowledge Discovery in Databases: An Overview. *AI Magazine* (1992)
- [117] French, M.N., Krajewski, W.F., Cuykendall, R.R.: Rainfall forecasting in space and time using a neural network. *Journal of Hydrology* 137(1-4), 1–31 (1992)
- [118] Freund, J.: Modern elementary statistics, 11th edn. Pearson/Prentice Hall (2004)
- [119] Fukunaga, K.: Introduction to statistical pattern recognition. Academic Press, London (1990)
- [120] Gao, X., Zhao, R., Qi, C., Shi, Z.: A rough set method to treat ECG signals for predicting coronary heart disease. *Journal of Biomedical Engineering* 25(5), 1025–1028 (2008)

- [121] Gan, G., Ma, C., Wu, J.: Data clustering: Theory, algorithms, and applications. SIAM, Philadelphia (2007)
- [122] Garcia-Orellana, C.J., Gallardo-Caballero, R., Macias-Macias, M., Gonzalez-Velasco, H.: SVM and neural networks comparison in mammographic CAD. In: Conf. Proc. IEEE Eng. Med. Bio. Soc., pp. 3204–3207 (2007)
- [123] Garcia-Laencina, P.J., Sancho-Gomez, J.L., Figueiras-Vidal, A.R., Verleysen, M.: K-nearest neighbours with mutual information for simultaneous classification and missing data imputation. *Neurocomputing* 72(7-9), 1483–1493 (2009)
- [124] Gilchrist, W.: Statistical modeling. John Wiley, Chichester (1984)
- [125] Glasziou, P., Del Mar, C., Salisbury, J.: Evidence-based medicine workbook. BMJ Publishing Group (2003)
- [126] Glesner, M., Pochmuller, W.: Neurocomputers: An Overview of Neural Networks in VLSI. Chapman and Hall, Boca Raton (1994)
- [127] Gluck, R., Klimov, A.V.: Occam's Razor in Metacomputation: the Notion of a Perfect Process Tree. In: Cousot, P., Filé, G., Falaschi, M., Rauzy, A. (eds.) WSA 1993. LNCS, vol. 724, pp. 112–123. Springer, Heidelberg (1993)
- [128] Goh, C., Law, R., Mok, H.: Analyzing and forecasting tourism demand: A rough sets approach. *Journal of Travel Research* 46(3), 327–338 (2008)
- [129] Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading (1989)
- [130] Goldberg, D.E., Lingle, R.: Alleles, loci, and the travelling salesman problem. In: Grefenstette, J.J. (ed.) Proceedings 1st International Conference on Genetic Algorithms and Their Applications, pp. 154–159. Lawrence Erlbaum, Hillsdale (1985)
- [131] Goldberg, D.E.: Sizing Populations for Serial and Parallel Genetic Algorithms. Addison-Wesley, Reading (1989)
- [132] Goldstein, B.J., Mushlin, A.I.: Use of single thyroxine test to evaluate ambulatory medical patients for suspected hypothyroidism. *J. Gen. Intern. Med.* 2, 20–24 (1987)
- [133] Gordon, A.D.: Classification, 2nd edn. Chapman and Hall, Boca Raton (1999)
- [134] Gorsuch, R.L.: Factor analysis, 2nd edn. Lawrence Erlbaum Assoc. Inc., Hillsdale (1983)
- [135] Gorunescu, F.: k-means clustering: a heuristic approach to improve the treatments effectiveness. *Craiova Medical Journal* 5(3), 421–423 (2003)
- [136] Gorunescu, F.: Architecture of Probabilistic Neural Networks: estimating the adjustable smoothing factor. In: Proc. 4th International Conference on Artificial Intelligence and Digital Communications. Research Notes in Artificial Intelligence and Digital Communications (Subseries of Research Notes in Computer Science), vol. 104, pp. 56–62 (2004)
- [137] Gorunescu, F., Gorunescu, M., El-Darzi, E., Gorunescu, S.: An Evolutionary Computational Approach to Probabilistic Neural Network with Application to Hepatic Cancer Diagnosis. In: Tsymbal, A., Cunningham, P. (eds.) Proc. 18th IEEE International Symposium on Computer-Based Medical Systems -IEEE CBMS 2005, pp. 461–466. IEEE Computer Science Press, Los Alamitos (2005)
- [138] Gorunescu, F., Gorunescu, M., El-Darzi, E., Ene, M., Gorunescu, S.: Performance enhancement approach for Probabilistic Neural Networks. In: Gorunescu, F., El-Darzi, E., Gorunescu, M. (eds.) Proc. First East European Conference on Health Care Modelling and Computation-HCMC 2005, pp. 142–148. Medical University Press (2005)

- [139] Gorunescu, F., Gorunescu, M., El-Darzi, E., Ene, M., Gorunescu, S.: Statistical Comparison of a Probabilistic Neural Network Approach in Hepatic Cancer Diagnosis. In: Proc. Eurocon2005 -IEEE International Conference on "Computer as a tool", pp. 237–240. IEEE Press, Los Alamitos (2005)
- [140] Gorunescu, M., Gorunescu, F., Ene, M., El-Darzi, E.: A heuristic approach in hepatic cancer diagnosis using a probabilistic neural network-based model. In: Janssen, J., Lenca, P. (eds.) Proc. 11th Applied Stochastic Models and Data Analysis - ASMDA 2005, pp. 1016–1025 (2005)
- [141] Gorunescu, F., Gorunescu, M., El-Darzi, E., Gorunescu, S., Revett, K., Khan, A.: A Cancer Diagnosis System Based on Rough Sets and Probabilistic Neural Networks. In: Gorunescu, F., El-Darzi, E., Gorunescu, M. (eds.) Proc. First East European Conference on Health Care Modelling and Computation - HCCM 2005, pp. 149–159. Medical University Press (2005)
- [142] Gorunescu, F., Gorunescu, M., El-Darzi, E., Ene, M.: Comparing classifiers: Probabilistic Neural Networks vs. k-means clustering. In: Proc. 5th International Conference on Artificial Intelligence and Digital Communications. Research Notes in Artificial Intelligence and Digital Communications (Subseries of Research Notes in Computer Science), Annals University of Craiova, vol. 105, pp. 6–11 (2005)
- [143] Gorunescu, F., Gorunescu, M.: A heuristic approach in cancer research using evolutionary computation. International Journal of Computer Research (Nova Science) 12(4), 59–65 (2005)
- [144] Gorunescu, F.: Benchmarking Probabilistic Neural Network algorithms. In: Proc. 6th International Conference on Artificial Intelligence and Digital Communications. Research Notes in Artificial Intelligence and Digital Communications (Subseries of Research Notes in Computer Science), vol. 106, pp. 1–7 (2006)
- [145] Gorunescu, M., Gorunescu, F., Revett, K.: A Neural Computing-based approach for the Early Detection of Hepatocellular Carcinoma. International Journal of Biomedical Sciences, Transactions on Engineering, Computing and Technology (Enformatika) 17, 65–68 (2006)
- [146] Gorunescu, M., Revett, K.: A novel noninvasive cancer diagnosis using neural networks. In: Proc. 7th International Conference on Artificial Intelligence and Digital Communications. Research Notes in Artificial Intelligence and Digital Communications (Subseries of Research Notes in Computer Science), vol. 107, pp. 16–20 (2007)
- [147] Gorunescu, F., Gorunescu, M., Revett, K., Ene, M.: A hybrid incremental/Monte Carlo searching technique for the "smoothing" parameter of Probabilistic Neural Networks. In: Proc. International Conference on Knowledge Engineering, Principles and Techniques -KEPT 2007, pp. 107–114. Cluj University Press (2007)
- [148] Gorunescu, M., Gorunescu, F., Revett, K.: Investigating a Breast Cancer Dataset Using a Combined Approach: Probabilistic Neural Networks and Rough Sets. In: Proc. 3rd ACM International Conference on Intelligent Computing and Information Systems, ICICIS 2007, pp. 246–249. Police Press, Cairo (2007)
- [149] Gorunescu, F., Gorunescu, M., Gorunescu, S., Saftoiu, A., Vilmann, P.: Neural Computing: Application in Non-invasive Cancer Detection. Case Studies in Business, Industry and Government Statistics (CSBIGS) 2(1), 38–46 (2008)
- [150] Gorunescu, F., Saftoiu, A., Vilmann, P.: Neural Networks-based dynamic medical diagnosis. In: Proc. 4th ACM International Conference on Intelligent Computing and Information Systems, ICICIS 2009, pp. 30–35 (2009)

- [151] Gorunescu, F., Gorunescu, M., Saftoiu, A., Vilmann, P., Belciug, S.: Competitive/Collaborative Neural Computing System for Medical Diagnosis in Pancreatic Cancer Detection. *Expert Systems*. Wiley&Blackwell (2010), doi:10.1111/j.1468-0394.2010.00540.x
- [152] Gorunescu, F., El-Darzi, E., Belciug, S., Gorunescu, M.: Patient grouping optimization using a hybrid Self-Organizing Map and Gaussian Mixture Model for length of stay-based clustering system. In: *Proc. IEEE International Conference on Intelligent Systems, IS 2010*, pp. 173–178 (2010)
- [153] Gorunescu, F., Gorunescu, M., El-Darzi, E., Gorunescu, S.: A statistical framework for evaluating neural networks to predict recurrent events in breast cancer. *International Journal of General Systems* 39(5), 471–488 (2010)
- [154] Gourieroux, C., Monfort, A.: *Time series and dynamic models*. Cambridge University Press, Cambridge (1997)
- [155] Gower, J.C., Legendre, P.: Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification* 3, 5–48 (1986)
- [156] Green, D.M., Swets, J.M.: *Signal detection theory and psychophysics*. John Wiley and Sons Inc., New York (1966)
- [157] Groth, R.: *Data Mining. A hands-on approach for business professionals*. Prentice-Hall, Englewood Cliffs (1998)
- [158] Grubbs, F.: Procedures for Detecting Outlying Observations in Samples. *Technometrics* 11(1), 1–21 (1969)
- [159] Grunwald, P.: *The Minimum Description Length principle*. MIT Press, Cambridge (2007)
- [160] Grzymala, J., Siddhava, S.: Rough set Approach to Rule Induction from Incomplete Data. In: *Proc. IPMU 2004, 10th International Conference on information Processing and Management of Uncertainty in Knowledge-Based System (2004)*
- [161] Gurney, K.: *An introduction to Neural Networks*. CRC Press, Boca Raton (1997)
- [162] Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning* 46(1-3), 389–422 (2002)
- [163] Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *Journal of Machine Learning Research* 3, 1157–1182 (2003)
- [164] Halkidi, M., Batistakis, Y., Vazirgiannis, M.: On clustering validation techniques. *Journal of Intelligent Information Systems* 17(2/3), 107–145 (2001)
- [165] Hammer, B.: *Learning with Recurrent Neural Networks*. LNCIS, vol. 254. Springer, Heidelberg (2000)
- [166] Han, J., Kamber, M.: *Data Mining Concepts and Techniques*, 2nd edn. Morgan Kaufmann Publishers, San Francisco (2006)
- [167] Han, J., Pei, J., Yin, Y., Mao, R.: Mining frequent patterns without candidate generation. *Data Mining and Knowledge Discovery* 8, 53–87 (2004)
- [168] Hand, D., Mannila, H., Smyth, P.: *Principles of Data Mining*. MIT Press, Cambridge (2001)
- [169] Hand, D., Yu, K.: Idiot’s Bayes - not so stupid after all? *International Statistical Review* 69(3), 385–399 (2001)
- [170] Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36 (1982)
- [171] Hansen, L.P., Sargent, T.J.: *Robustness*. Princeton University Press, Princeton (2007)
- [172] Harris, S.: Total Information Awareness official responds to criticism. In: *Government Executive.com* (January 31, 2003),
<http://www.govexec.com/dailyfed/0103/013103h1.htm>

- [173] Hassan, M.R., Hossain, M.M., Bailey, J., Ramamohanarao, K.: Improving k -Nearest Neighbour Classification with Distance Functions Based on Receiver Operating Characteristics. In: Daelemans, W., Goethals, B., Morik, K. (eds.) ECML PKDD 2008, Part I. LNCS (LNAI), vol. 5211, pp. 489–504. Springer, Heidelberg (2008)
- [174] Hassanien, A., Ali, J.: Image classification and retrieval algorithm based on rough set theory. *S. Afr. Comput. J.* 30, 9–16 (2003)
- [175] Hassanien, A.: Rough set approach for attribute reduction and rule generation: A case of patients with suspected breast cancer. *Journal of the American Society for Information Science and Technology* 55(11), 954–962 (2004)
- [176] Hassanien, A., Abraham, A., Peters, J., Schaefer, G., Henry, C.: Rough sets and near sets in medical imaging: a review. *IEEE Transactions on Information Technology in Biomedicine* 13(6), 955–968 (2009)
- [177] Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, Heidelberg (2001)
- [178] Hastie, T., Tibshirani, R.: Generalized additive models (with discussion). *Statistical Science* 1, 297–318 (1986)
- [179] Hastie, T., Tibshirani, R.: *Generalized additive models*. Chapman and Hall, Boca Raton (1990)
- [180] Haupt, R., Haupt, S.E.: *Practical genetic algorithms*. Wiley, Chichester (2004)
- [181] Haykin, S.: *Neural Networks, 2nd edn. A comprehensive foundation*. Prentice-Hall, Englewood Cliffs (1999)
- [182] Hertz, J., Krogh, A., Palmer, R.G.: *Introduction to the theory of neural computation*. Addison Wesley, Redwood City (1991)
- [183] Heyer, L.J., Kruglyak, S., Yoosheph, S.: Exploring expression data: identification and analysis of coexpressed genes. *Genome Res.* 9(11), 1106–1115 (1999)
- [184] Hill, T., Lewicki, P.: *Statistics, methods and applications. A comprehensive reference for science, industry, and data mining*. StatSoft (2006)
- [185] Hincapie, J.G., Kirsch, R.F.: Feasibility of EMG-based neural network controller for an upper extremity neuroprosthesis. *IEEE Trans. Neural. Syst. Rehabil. Eng.* 17(1), 80–90 (2009)
- [186] Hinton, G., Sejnowski, T.J. (eds.): *Unsupervised Learning: Foundations of Neural Computation*. MIT Press, Cambridge (1999)
- [187] Hoffmann, A.: *Paradigms of artificial intelligence*. Springer, Heidelberg (1998)
- [188] Holland, J.H.: *Adaptation*. In: Rosen, R., Snell, F.M. (eds.) *Progress in Theoretical Biology*, vol. 4, Plenum, New York (1976)
- [189] Holland, J.H.: *Adaptation in Natural and Artificial Systems, 1st edn*. University of Michigan Press, Ann Arbor (1975); MIT Press, Cambridge (1992)
- [190] Holte, R.C.: Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. *Machine Learning* 11, 63–90 (1993)
- [191] Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. *Proc. National Academy of Sciences of the USA* 79(8), 2554–2558 (1982)
- [192] Hoppner, F., Klawonn, F., Kruse, R., Runkler, T.: *Fuzzy Cluster Analysis*. Wiley, Chichester (1999)
- [193] Hoskinson, A.: Creating the ultimate research assistant. *Computer* 38(11), 97–99 (2005)
- [194] Hotelling, H.: Relations between two sets of variates. *Biometrika* 28, 321–377 (1936)
- [195] Hua, S., Sun, Z.: Support Vector Machine Approach for Protein Subcellular Localization Prediction. *Bioinformatics* 17(8), 721–728 (2001)

- [196] Huang, C., Dorsey, E., Boose, M.A.: Life Insurer Financial Distress Prediction: A Neural Network Model. *Journal of Insurance Regulation* 13(2), 131–167 (Winter 1994)
- [197] Huber, P.J.: *Robust Statistics*. John Wiley, Chichester (1981)
- [198] Huet, S., Bouvier, A., Poursat, M.-A., Jolivet, E.: *Statistical tools for nonlinear regression*, 2nd edn. Springer Series in Statistics. Springer, Heidelberg (2004)
- [199] Hunt, E.B., Marin, J., Stone, P.T.: *Experiments in Induction*. Academic Press, New York (1966)
- [200] Hwang, J.-N., Choi, J., Oh, S., Marks, R.: Query-based learning applied to partially trained multilayer perceptron. *IEEE Transactions on Neural Networks* 2(1), 131–136 (1991)
- [201] Ince, E.A., Ali, S.A.: Rule based segmentation and subject identification using fiducial features and subspace projection methods. *Journal of Computers* 2(4), 68–75 (2007)
- [202] Jackson, J.E.: *A User's Guide to Principal Components*. Wiley series in probability and mathematical statistics. Wiley-IEEE (2003)
- [203] Jain, A.K., Dubes, R.C.: *Algorithms for Clustering Data*. Prentice-Hall, New Jersey (1988)
- [204] Jambu, M.: *Exploratory and multivariate data analysis*. Academic Press, London (1991)
- [205] Jensen, F.: *Bayesian networks and decision graphs*. Statistics for Engineering and Information Science. Springer, Heidelberg (2001)
- [206] Jiang, Z., Yamauchi, K., Yoshioka, K., Aoki, K., Kuroyanagi, S., Iwata, A., Yang, J., Wang, K.: Support Vector Machine-Based Feature Selection for Classification of Liver Fibrosis Grade in Chronic Hepatitis C. *Journal of Medical Systems* 20, 389–394 (2006)
- [207] Jin, H., Lu, Y.: A non-inferiority test of areas under two parametric ROC curves. *Contemporary Clinical Trials* 30(4), 375–379 (2009)
- [208] Joachims, T.: *Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms*. Springer, Heidelberg (2002)
- [209] Jolliffe, I.T.: *Principal component analysis*, 2nd edn. Springer series in statistics. Springer, Heidelberg (2002)
- [210] Kahn, B., Strong, D., Wang, R.: Information Quality Benchmarks: Product and Service Performance. *Communications of the ACM*, 184–192 (April 2002)
- [211] Kalamatianos, D., Liatsis, P., Wellstead, P.E.: Near-infrared spectroscopic measurements of blood analytes using multi-layer perceptron neural networks. In: *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 3541–3544 (2006)
- [212] Kaufman, L., Rousseeuw, P.J.: *Finding groups in data: an introduction to cluster analysis*. John Wiley and Sons, Chichester (2005)
- [213] Khan, A.U., Bandopadhyaya, T.K., Sharma, S.: Classification of Stocks Using Self Organizing Map. *International Journal of Soft Computing Applications* 4, 19–24 (2009)
- [214] Kim, J.-O., Mueller, C.: *Introduction to factor analysis*. Series 07-013. Sage Publications, Beverly Hills (1978)
- [215] Kim, J.-O., Mueller, C.: *Factor Analysis. Statistical Methods and Practical Issues*. Series 07-014. Sage Publications, Beverly Hills (1978)
- [216] Kitajima, M., Hirai, T., Katsuragawa, S., Okuda, T., Fukuoka, H., Sasao, A., Akter, M., Awai, K., Nakayama, Y., Ikeda, R., Yamashita, Y., Yano, S., Kuratsu, J., Doi, K.: Differentiation of common large sellar-suprasellar masses effect of artificial neural network on radiologists' diagnosis performance. *Acad. Radiol.* 16(3), 313–320 (2009)
- [217] Kohonen, T.: *An introduction to neural computing*. *Neural Networks* 1, 3–16 (1988)

- [218] Kohonen, T.: *Self-Organizing Maps* (3rd extended edn.). Springer Series in Information Sciences, vol. 30. Springer, Heidelberg (2001)
- [219] Komorowski, J., Polkovski, L., Skowron, A.: *Rough Sets: A tutorial*. In: 11th European Summer School in Language, Logic and Information. Lecture Notes for ESSLLI 1999, Utrecht, Holland (1999)
- [220] Komorowski, J., Ohrn, A., Skowron, A.: *The ROSETTA Rough Set Software System*. In: Klossgen, W., Zytkow, J. (eds.) *Handbook of Data Mining and Knowledge Discovery*, ch. D.2.3. Oxford University Press, Oxford (2002)
- [221] Koski, T., Noble, J.: *Bayesian Networks: An Introduction*. Wiley Series in Probability and Statistics (2009)
- [222] Koza, J.: *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge (1992)
- [223] Kowalczyk, W., Piasta, Z.: *Rough sets inspired approach to knowledge discovery in business databases*. In: Wu, X., Kotagiri, R., Korb, K.B. (eds.) *PAKDD 1998*. LNCS, vol. 1394, Springer, Heidelberg (1998)
- [224] Kumar, A., Agrawal, D.P., Joshi, S.D.: *Multiscale rough set data analysis with application to stock performance modeling*. *Intelligent Data Analysis* 8(2), 197–209 (2004)
- [225] Kuo, S.J., Hsiao, Y.H., Huang, Y.L., Chen, D.R.: *Classification of benign and malignant breast tumors using neural networks and three-dimensional power Doppler ultrasound*. *Ultrasound Obstet. Gynecol.* 32(1), 97–102 (2008)
- [226] Lai, K.C., Chiang, H.C., Chen, W.C., Tsai, F.J., Jeng, L.B.: *Artificial neural network-based study can predict gastric cancer staging*. *Hepatogastroenterology* 55(86-87), 1859–1863 (2008)
- [227] Langdon, W., Poli, R.: *Foundations of genetic programming*. Springer, Heidelberg (2002)
- [228] Larose, D.: *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley, Chichester (2005)
- [229] Larose, D.: *Data mining methods and models*. Wiley Interscience, Hoboken (2006)
- [230] Lasko, T.A., Bhagwat, J.G., Zou, K.H., Ohno-Machado, L.: *The use of receiver operating characteristic curves in biomedical informatics*. *Journal of Biomedical Informatics* 38(5), 404–415 (2005)
- [231] Lebowitz, M.: *Experiments with incremental concept formation*. *Machine Learning* 2, 103–138 (1987)
- [232] Levine, M.: *Canonical analysis and factor comparison*. Sage Publications, Beverly Hills (1977)
- [233] Liang, K.H., Krus, D.J., Webb, J.M.: *K-fold crossvalidation in canonical analysis*. *Multivariate Behavioral Research* 30, 539–545 (1995)
- [234] Lin, T.Y., Cercone, N. (eds.): *Rough sets and data mining. Analysis of imprecise data*. Kluwer Academic Publishers, Dordrecht (1997)
- [235] Lin, T.Y., Tremba, J.: *Attribute Transformations on Numerical Databases*. In: Terano, T., Chen, A.L.P. (eds.) *PAKDD 2000*. LNCS, vol. 1805. Springer, Heidelberg (2000)
- [236] Lin, T.Y.: *Attribute transformations for data mining. I: Theoretical explorations*. *International Journal of Intelligent Systems* 17(2), 213–222 (2002)
- [237] Lin, T.Y., Yiyu, Y., Yao, Y.Y., Zadeh, L. (eds.): *Data Mining, Rough Sets and Granular Computing*. Physica-Verlag, Heidelberg (2002)
- [238] Lindeman, R.H., Merenda, P.F., Gold, R.: *Introduction to bivariate and multivariate analysis*. Scott, Foresman and Co., New York (1980)
- [239] Lippmann, R.P.: *An introduction to computing with neural nets*. *IEEE ASSP Magazine* 4, 4–22 (1987)

- [240] Liu, H., Hussain, F., Tan, C.L., Dash, M.: Discretization: An Enabling Technique. *Data Mining and Knowledge Discovery* 6(4), 393–423 (2002)
- [241] Liu, H., Hiroshi, M.: *Feature Selection for Knowledge Discovery and Data Mining*. Kluwer Academic Publishers, Dordrecht (1998)
- [242] Liu, H., Hiroshi, M. (eds.): *Feature Extraction, Construction and Selection: A Data Mining Perspective*. Springer International Series in Engineering and Computer Science (1998)
- [243] Loeve, M.: *Probability Theory I*, 4th edn. Springer, Heidelberg (1977)
- [244] Luo, F.-L., Unbehauen, R.: *Applied Neural Networks for Signal Processing*. Cambridge University Press, Cambridge (1999)
- [245] Maglogiannis, I., Sarimveis, H., Kiranoudis, C.T., Chatzizoiannou, A.A., Oikonomou, N., Aidinis, V.: Radial basis function neural networks classification for the recognition of idiopathic pulmonary fibrosis in microscopic images. *IEEE Trans. Inf. Technol. Biomed.* 12(1), 42–54 (2008)
- [246] Maqsood, I., Abraham, A.: Weather analysis using ensemble of connectionist learning paradigms. *Applied Soft Computing* 7(3), 995–1004 (2007)
- [247] Mandic, D., Chambers, J.: *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. Wiley, Chichester (2001)
- [248] Mathias, K., Whitley, D.: Genetic Operators, the Fitness Landscape and the Traveling Salesman Problem. In: Manner, R., Manderick, B. (eds.) *Parallel Problem Solving from Nature*, pp. 219–228. Elsevier, Amsterdam (1992)
- [249] Maung, Z.M., Mikami, Y.: A rule-based syllable segmentation of Myanmar Text. In: *Proc. IJCNLP 2008 Workshop on NLP for Less Privileged Languages*, pp. 51–58 (2008)
- [250] McLachlan, G.J., Peel, D.: *Finite Mixture Models*. John Wiley and Sons, New York (2000)
- [251] Martinez, W.L., Martinez, A.L.: *Exploratory Data Analysis with MATLAB*. CRC Press, Boca Raton (2004)
- [252] Martinez, L.C., da Hora, D.N., Palotti, J.R., Meira, W., Pappa, G.: From an artificial neural network to a stock market day-trading system: a case study on the BM and F BOVESPA. In: *Proc. 2009 International Joint Conference on Neural Networks*, pp. 3251–3258 (2009)
- [253] McCulloch, W.S., Pitts, W.A.: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematics and Biophysics* 5, 115–133 (1943)
- [254] McLachlan, G., Krishnan, T.: *The EM Algorithm and extensions*, 2nd edn. Wiley, Chichester (2008)
- [255] McLaren, C.E., Chen, W.P., Nie, K., Su, M.Y.: Prediction of malignant breast lesions from MRI features: a comparison of artificial neural network and logistic regression techniques. *Acad. Radiol.* 16(7), 842–851 (2009)
- [256] Meinel, L.A., Stolpen, A.H., Berbaum, K.S., Fajardo, L.L., Reinhardt, J.M.: Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. *J. Magn. Reson. Imaging* 25(1), 89–95 (2007)
- [257] Menzies, T., Hu, Y.: Data Mining For Very Busy People. *IEEE Computer* 36(11), 22–29 (2003)
- [258] Metz, C.E.: Basic principles of ROC analysis. *Sem. Nuc. Med.* 8, 283–298 (1978)
- [259] Michalewicz, Z., Fogel, D.B.: *How to solve it: Modern Heuristics*. Springer, Heidelberg (2004)

- [260] Michalewicz, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*, 3rd edn. Springer, Heidelberg (1996)
- [261] Michalski, R.S.: Knowledge acquisition through conceptual clustering: A theoretical framework and an algorithm for partitioning data into conjunctive concepts. *International Journal of Policy Analysis and Information Systems* 4, 219–244 (1980)
- [262] Michalski, R.S., Stepp, R.E.: Learning from observation: Conceptual clustering. In: Michalski, R.S., Carbonell, J.G., Mitchell, T.M. (eds.) *Machine Learning: An Artificial Intelligence Approach*, pp. 331–363. Tioga, Palo Alto (1983)
- [263] Michel, A.N., Liu, D.: *Qualitative analysis and synthesis of recurrent neural networks*. Pure and Applied Mathematics. CRC Press, Boca Raton (2002)
- [264] Miller, B.L., Goldberg, D.E.: Genetic Algorithms, Tournament Selection and the Effects of Noise. *Complex Systems* 9, 193–212 (1996)
- [265] Minsky, M.L., Papert, S.A.: *Perceptrons*. MIT Press, Cambridge (1969) (expanded edn., 1987)
- [266] Mirzaaghadzadeh, A., Motameni, H., Karshenas, M., Nematzadeh, H.: Learning flexible neural networks for pattern recognition. *World Academy of Science, Engineering and Technology* 33, 88–91 (2007)
- [267] Mitchell, T.M.: *Machine learning*. McGraw-Hill, New York (1997)
- [268] Mitchell, M.: *An introduction to genetic algorithms (complex adaptive systems)*. MIT Press, Cambridge (1998)
- [269] Moore, D.S., McCabe, G.P.: *Introduction to the Practice of Statistics*, 3rd edn. W.H. Freeman, New York (1999)
- [270] Montana, D., Davis, L.: Training feedforward networks using genetic algorithms. In: *Proc. International Joint Conference on Artificial Intelligence*. Morgan Kaufmann, San Francisco (1989)
- [271] Mrozek, A., Plonka, L., Rough, L.: sets in image analysis. *Foundations of Computing Decision Sciences* 18(3-4), 259–273 (1993)
- [272] Mulaik, S.A.: *Foundations of Factor Analysis*, 2nd edn. Chapman and Hall/CRC Statistics in the Social and Behavioral Scie (2009)
- [273] Muller, K.R., Smola, A.J., Ratsch, G., Scholkopf, B., Kohlmorgen, J., Vapnik, V.: Predicting Time Series with Support Vector Machines. In: Gerstner, W., Hasler, M., Germond, A., Nicoud, J.-D. (eds.) *ICANN 1997*. LNCS, vol. 1327, pp. 999–1004. Springer, Heidelberg (1997)
- [274] Mustapha, N., Jalali, M., Jalali, M.: Expectation Maximization clustering algorithm for user modeling in Web usage mining systems. *European Journal of Scientific Research* 32(4), 467–476 (2009)
- [275] Neubauer, A.: Adaptive non-uniform mutation for genetic algorithms. In: *Computational Intelligence Theory and Applications*. LNCS, Springer, Heidelberg (1997)
- [276] Nguyen, H.S., Nguyen, S.H.: Pattern extraction from data. *Fundamenta Informaticae* 34, 129–144 (1998)
- [277] Nguyen, H.D., Yoshihara, I., Yasunaga, M.: Modified edge recombination operators of genetic algorithms for the traveling salesman problem. In: *26th Annual Conference of the IEEE Industrial Electronics Society, IECON 2000*, vol. 4, pp. 2815–2820 (2000)
- [278] Nigsch, F., Bender, A., van Buuren, B., Tissen, J., Nigsch, E., Mitchell, J.B.O.: Melting Point Prediction Employing k-nearest Neighbor Algorithms and Genetic Parameter Optimization. *Journal of Chemical Information and Modeling* 46(6), 2412–2422 (2006)
- [279] Norton, P.G., Dunn, E.V.: Snoring as a risk factor for disease: an epidemiological survey. *Br. Med. J.* 291, 630–632 (1985)

- [280] Obe, O.O., Shangodoyin, D.K.: Artificial neural network based model for forecasting sugar cane production. *Journal of Computer Science* 6(4), 439–445 (2010)
- [281] Ogulata, S.N., Sahin, C., Erol, R.: Neural network-based computer-aided diagnosis in classification of primary generalized epilepsy by EEG signals. *J. Med. Syst.* 33(2), 107–112 (2009)
- [282] Ohrn, A., Komorowski, J.: ROSETTA: A Rough Set Toolkit for Analysis of Data. In: *Proc. Third International Joint Conference on Information Sciences, Fifth International Workshop on Rough Sets and Soft Computing (RSSC 1997)*, vol. 3, pp. 403–407 (1997)
- [283] Ohrn, A.: ROSETTA Technical Reference Manual, Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway (2000)
- [284] Oliver, J., Wallace, C.S.: Inferring decision graphs. In: *Proc. of the 1992 Aust. Joint Conf. on Artificial Intelligence*, pp. 361–367 (1992)
- [285] Oliver, I.M., Smith, D.J., Holland, J.: A study of permutation crossover operators on the travelling salesman problem. In: Grefenstette, J.J. (ed.) *Proc. 2nd International Conference on Genetic Algorithms and Their Applications*, pp. 224–230. Lawrence Erlbaum, Hillsdale (1987)
- [286] O’Neill, S., Leahy, F., Pasterkamp, H., Tal, A.: The effects of chronic hyperinflation, nutritional status, and posture on respiratory muscle strength in cystic fibrosis. *Am. Rev. Respir. Dis.* 128, 1051–1054 (1983)
- [287] Osaki, S.: *Applied stochastic system modeling*. Springer, Heidelberg (1992)
- [288] Osuna, E., Freund, R., Girosi, F.: Training Support Vector Machines: An Application to Face Detection. In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 130–136 (1997)
- [289] Park, L.-J.: Learning of Neural Networks for Fraud Detection Based on a Partial Area Under Curve. In: Wang, J., Liao, X.-F., Yi, Z. (eds.) *ISNN 2005. LNCS*, vol. 3497, pp. 922–927. Springer, Heidelberg (2005)
- [290] Parzen, E.: On estimation of a probability density function and mode. *Ann. Math. Stat.* 33, 1065–1076 (1962)
- [291] Pawlak, Z.: *Rough Sets*. *International Journal of Computer and Information Sciences* 11, 341–356 (1982)
- [292] Pawlak, Z.: *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishing, Dordrecht (1991)
- [293] Pearson, K.: Contributions to the mathematical theory of evolution II: Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London* 186, 343–414 (1895)
- [294] Pearson, K.: Das Fehlgengesetz und seine Verallgemeinerungen durch Fechner und Pearson. A Rejoinder. *Biometrika* 4, 169–212 (1905)
- [295] Pearson, K.: On the generalized probable error in multiple normal correlation. *Biometrika* 6, 59–68 (1908)
- [296] Pedrycz, W., Skowron, A.: *Fuzzy and rough sets*. *Handbook of data mining and knowledge discovery*. Oxford University Press, Inc., New York (2002)
- [297] Pendaharkar, P.C.: A comparison of gradient ascent, gradient descent and genetic-algorithm-based artificial neural networks for the binary classification problem. *Expert Systems* 24(2), 65–86 (2007)
- [298] Pepe, M.S.: *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press, Oxford (2004)

- [299] Piatetsky-Shapiro, G.: Discovery, analysis, and presentation of strong rules. In: Piatetsky-Shapiro, G., Frawley, W. (eds.) *Knowledge Discovery in Databases*, AAAI/MIT Press, Cambridge/MA (1991)
- [300] Polkowski, L., Skowron, A. (eds.): *Rough Sets in Knowledge Discovery 1: Methodology and Applications*. Physica-Verlag, Heidelberg (1998)
- [301] Polkowski, L., Skowron, A. (eds.): *Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems*. Physica-Verlag, Heidelberg (1998)
- [302] Polkowski, L., Tsumoto, S., Lin, T.Y. (eds.): *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems*. STUDEFUZZ. Physica-Verlag, Heidelberg (2000)
- [303] Polkowski, L.: *Rough Sets*. Physica-Verlag, Heidelberg (2002)
- [304] Pontil, M., Verri, A.: Support Vector Machines for 3D Object Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(6), 637–646 (1998)
- [305] Price, R., Shanks, G.: A Semiotic Information Quality Framework. In: *Proc. IFIP International Conference on Decision Support Systems (DSS2004): Decision Support in an Uncertain and Complex World*, Prato (2004)
- [306] Provost, F., Fawcett, T.: Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions. In: *Proc. 3rd International Conference on Knowledge Discovery and Data Mining (KDD 1997)*, pp. 43–48. AAAI Press, Menlo Park (1997)
- [307] Provost, F., Fawcett, T., Kohavi, R.: The case against accuracy estimation for comparing induction algorithms. In: Shavlik, J. (ed.) *Proc. ICML 1998*, pp. 445–453. Morgan Kaufmann, San Francisco (1998)
- [308] Quinlan, J.R.: Induction of decision trees. *Machine Learning* 1, 81–106 (1986)
- [309] Quinlan, J.R., Rivest, R.L.: Inferring decision trees using the minimum description length principle. *Information and Computation* 80(3), 227–248 (1989)
- [310] Quinlan, J.R.: *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Francisco (1993)
- [311] Radzikowska, A.M., Kerre, E.E.: A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems* 126, 137–156 (2002)
- [312] Raj, V.S.: Better performance of neural networks using functional graph for weather forecasting. In: *Proc. 12th WSEAS International Conference on Computers*, pp. 826–831 (2008)
- [313] Rechenberg, I.: *Evolutionsstrategie Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. PhD thesis (1971); reprinted by Fromman-Holzboog (1973)
- [314] Revett, K., Gorunescu, F., Gorunescu, M., El-Darzi, E., Ene, M.: A Breast Cancer Diagnosis System: A Combined Approach Using Rough Sets and Probabilistic Neural Networks. In: *Proc. Eurocon 2005-IEEE International Conference on Computer as a tool*, Belgrade, Serbia, pp. 1124–1127. IEEE Computer Society Press, Los Alamitos (2005)
- [315] Revett, K., Gorunescu, F., Gorunescu, M., Ene, M.: Mining A Primary Biliary Cirrhosis Dataset Using Rough Sets and a Probabilistic Neural Network. In: *Proc. of the IEEE Intelligent Systems 2006 -IS06*, Westminster, London, pp. 284–289. IEEE Press, Los Alamitos (2006)
- [316] Revett, K., Gorunescu, F., Salem, A.B.: Feature Selection in Parkinson’s Disease: A Rough Sets Approach. In: *Proc. International Multiconference on Computer Science and Information Technology-IMCSIT 2009*, Mrgowo, Poland, vol. 4, pp. 425–428 (2009)

- [317] Ripley, B.D.: *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge (1996)
- [318] Rissanen, J.: Modeling by the shortest data description. *Automatica* 14, 465–471 (1978)
- [319] Rissanen, J.: *Information and Complexity in Statistical Modeling*. Springer, Heidelberg (2007)
- [320] *** ROC graphs: Notes and practical considerations for researchers. HP Labs, Tech. Rep. HPL-2003-4 (2003)
- [321] Roemer, V.M., Walden, R.: Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios for electronic foetal heart rate monitoring using new evaluation techniques. *Z. Geburtshilfe Neonatol.* 214(3), 108–118 (2010)
- [322] Rokach, L., Maimon, O.: *Data mining with decision trees*. Series in Machine Perception and Artificial Intelligence, vol. 69. World Scientific, Singapore (2007)
- [323] Romesburg, H.C.: *Cluster Analysis for Researchers*, 2nd edn. Krieger Pub. Co. (2004)
- [324] Rosenblatt, F.: The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychoanalytic Review* 65, 386–408 (1958)
- [325] Rosenblatt, F.: *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Spartan, Washington (1962)
- [326] Ryan, T.: *Modern Regression Methods, Solutions Manual*, 2nd edn. Wiley Series in Probability and Statistics. Wiley, Chichester (2009)
- [327] Saad, E.W., Choi, J.J., Vian, J.L., Wunsch, D.C.: Query-based learning for aerospace applications. *IEEE Trans. Neural Netw.* 14(6), 1437–1448 (2003)
- [328] Sandage, A.: EDWIN HUBBLE. *JRASC* 83(6), 1889–1953 (1989)
- [329] Saftoiu, A., Gorunescu, F., Rogoveanu, I., Ciurea, T., Ciurea, P.: A Dynamic Forecasting Model to Predict the Long-Term Response During Interferon Therapy in Chronic Hepatitis C. *World Congresses of Gastroenterology* 59(3), 349 (1998)
- [330] Saftoiu, A., Ciurea, T., Gorunescu, F., Rogoveanu, I., Gorunescu, M., Georgescu, C.: Stochastic modeling of the tumor volume assessment and growth patterns in hepatocellular carcinoma. *Journal of Oncology/Bulletin du Cancer* 91(6), 162–166 (2004)
- [331] Saftoiu, A., Vilmann, P., Gorunescu, F., Gheonea, D.I., Gorunescu, M., Ciurea, T., Popescu, G.L., Iordache, A., Hassan, H., Iordache, S.: Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. *Gastrointestinal Endoscopy* 68(6), 1086–1094 (2008)
- [332] Sanzogni, L., Kerr, D.: Milk production estimates using feed forward artificial neural networks. *Computer and Electronics in Agriculture* 32(1), 21–30 (2001)
- [333] Schaffer, J.D.: Some Effects of Selection Procedures on Hyperplane Sampling by Genetic Algorithms. In: Davis, L. (ed.) *Genetic Algorithms and Simulated Annealing*. Pitman (1987)
- [334] Schaffer, J.D., Whitley, D., Eshelman, L.: Combination of Genetic Algorithms and Neural Networks: The state of the art. In: *Combination of Genetic Algorithms and Neural Networks*. IEEE Computer Society, Los Alamitos (1992)
- [335] Schervish, M.J.: P-Values: What They Are and What They Are Not. *The American Statistician* 50(3), 203–206 (1996)
- [336] Schlkopf, B., Smola, A.: *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. MIT Press, Cambridge (2001)
- [337] Schwefel, H.P.: *Numerische Optimierung von Computer-Modellen*. PhD thesis (1974); reprinted by Birkhäuser (1977)

- [338] Seber, G.A.F., Wild, C.J.: *Nonlinear Regression*. Wiley Series in Probability and Statistics. Wiley, Chichester (1989)
- [339] Senthilkumaran, N., Rajesh, R.: A Study on Rough Set Theory for Medical Image Segmentation. *International Journal of Recent Trends in Engineering* 2(2), 236–238 (2009)
- [340] Shafer, J., Agrawal, R., Mehta, M.: SPRINT: A scalable parallel classifier for Data Mining. In: *Proc. 22nd VLDB Conference Mumbai (Bombay), India*, pp. 544–555 (1996)
- [341] Shakhnarovich, G., Darrell, T., Indyk, P. (eds.): *Nearest-Neighbor Methods in Learning and Vision*. MIT Press, Cambridge (2005)
- [342] Shank, D.B., McClendon, R.W., Paz, J., Hoogenboom, G.: Ensemble artificial neural networks for prediction of dew point temperature. *Applied Artificial Intelligence* 22(6), 523–542 (2008)
- [343] Shvaytser, H., Peleg, S.: Fuzzy and probability vectors as elements of vector space. *Information Sciences* 36, 231–247 (1985)
- [344] Shu, C., Burn, D.H.: Artificial neural network ensembles and their application in pooled flood frequency analysis. *Water Resources Research* 40(9), W09301, 1–10 (Abstract) (2004), doi: 10.1029/2003WR002816
- [345] Shuai, J.-J., Li, H.-L.: Using rough set and worst practice DEA in business failure prediction. In: Ślezak, D., Yao, J., Peters, J.F., Ziarko, W.P., Hu, X. (eds.) *RSFDGrC 2005*. LNCS (LNAI), vol. 3642, pp. 503–510. Springer, Heidelberg (2005)
- [346] Skowron, A.: Boolean reasoning for decision rules generation. In: Komorowski, J., Raś, Z.W. (eds.) *ISMIS 1993*. LNCS, vol. 689, pp. 295–305. Springer, Heidelberg (1993)
- [347] Slezak, D.: Approximate Entropy Reducts. *Fundamenta Informaticae* 53(3,4), 365–390 (2002)
- [348] Slowinski, R., Zopounidis, C.: Applications of the rough set approach to evaluation of bankruptcy risk. *International J. Intelligent Systems in Accounting, Finance and Management* 4(1), 27–41 (1995)
- [349] Smith, T.F.: Occam’s razor. *Nature* 285(5767), 620 (1980)
- [350] Smith, J.E., Fogarty, T.C.: Operator and parameter adaptation in genetic algorithms. *Soft Computing* 1(2), 81–87 (1997)
- [351] Song, T., Jamshidi, M., Lee, R.R., Hung, M.: A modified probabilistic neural network for partial volume segmentation in brain MR image. *IEEE Trans. Neural Networks* 18(5), 1424–1432 (2007)
- [352] Song, J.H., Venkatesh, S.S., Conant, E.A., Arger, P.H., Sehgal, C.M.: Comparative analysis of logistic regression and artificial neural network for computer-aided diagnosis of breast masses. *Acad. Radiol.* 12(4), 487–495 (2005)
- [353] Sorjamaa, A., Hao, J., Lendasse, A.: Mutual Information and k -Nearest Neighbors Approximator for Time Series Prediction. In: Duch, W., Kacprzyk, J., Oja, E., Zadrozny, S. (eds.) *ICANN 2005*. LNCS, vol. 3697, pp. 553–558. Springer, Heidelberg (2005)
- [354] Spackman, K.A.: Signal detection theory: Valuable tools for evaluating inductive learning. In: *Proc. 6th International Workshop on Machine Learning*, pp. 160–163. Morgan Kaufman, San Mateo (1989)
- [355] Spearman, C.: General intelligence objectively determined and measured. *American Journal of Psychology* 15, 201–293 (1904)
- [356] Spearman, C.: *The nature of intelligence and the principles of cognition*. Macmillan, London (1923)
- [357] Spearman, C.: *The abilities of man*. Macmillan, London (1927)

- [358] Specht, D.F.: Probabilistic neural networks for classification mapping or associative memory. In: Proc. IEEE International Conference on Neural Networks, vol. 1, pp. 525–532 (1988)
- [359] Specht, D.F.: Probabilistic neural networks. *Neural Networks* 3, 110–118 (1990)
- [360] Specht, D.F.: Probabilistic neural networks and the polynomial adaline as complementary techniques for classification. *IEEE Trans. on Neural Networks* 1(1), 111–121 (1990)
- [361] Specht, D.F.: A Generalized Regression Neural Network. *IEEE Transactions on Neural Networks* 2(6), 568–576 (1991)
- [362] Srisawat, A., Phientrakul, T., Kijssirikul, B.: SV-kNNC: An Algorithm for Improving the Efficiency of k-Nearest Neighbor. In: Yang, Q., Webb, G. (eds.) *PRICAI 2006. LNCS (LNAI)*, vol. 4099, pp. 975–979. Springer, Heidelberg (2006)
- [363] Stefansky, W.: Rejecting Outliers in Factorial Designs. *Technometrics* 14, 469–479 (1972)
- [364] Steinwart, I., Christmann, A.: *Support Vector Machines. Information Science and Statistics.* Springer, Heidelberg (2008)
- [365] Stevens, J.: *Applied multivariate statistics for the social sciences.* Erlbaum, Hillsdale (1986)
- [366] Stoean, C., Preuss, M., Gorunescu, R., Dumitrescu, D.: Elitist generational genetic chromodynamics - a new radii-based evolutionary algorithm for multimodal optimization. In: 2005 IEEE Congress on Evolutionary Computation (CEC 2005), pp. 1839–1846. IEEE Computer Society Press, Los Alamitos (2005)
- [367] Stoean, C., Preuss, M., Dumitrescu, D., Stoean, R.: Cooperative Evolution of Rules for Classification. In: *IEEE Post-proceedings Symbolic and Numeric Algorithms for Scientific Computing (SYNASC 2006)*, pp. 317–322. IEEE Press, Los Alamitos (2006)
- [368] Stoean, R., Preuss, M., Stoean, C., El-Darzi, E., Dumitrescu, D.: An Evolutionary Approximation for the Coefficients of Decision Functions within a Support Vector Machine Learning Strategy. *Foundations on Computational Intelligence* 1, 83–114 (2009)
- [369] Stoean, R., Preuss, M., Stoean, C., El-Darzi, E., Dumitrescu, D.: Support Vector Machine Learning with an Evolutionary Engine. *Journal of the Operational Research Society, Special Issue: Data Mining and Operational Research: Techniques and Applications* 60(8), 1116–1122 (2009)
- [370] Stoean, R., Stoean, C., Lupsor, M., Stefanescu, H., Badea, R.: Evolutionary-Driven Support Vector Machines for Determining the Degree of Liver Fibrosis in Chronic Hepatitis C. *Artificial Intelligence in Medicine.* Elsevier (2010) (in press)
- [371] Sunay, A.S., Cunedioğlu, U., Yılmaz, B.: Feasibility of probabilistic neural networks, Kohonen self-organizing maps and fuzzy clustering for source localization of ventricular focal arrhythmias from intravenous catheter measurements. *Expert System* 26(1), 70–81 (2009)
- [372] Sutton, R.S., Barto, A.G.: *Reinforcement learning.* MIT Press, Cambridge (1999)
- [373] Suykens, J.A.K., Vandewalle, J.P.L., De Moor, B.L.R.: *Artificial neural networks for modelling and control of non-linear systems.* Kluwer Academic Publishers, Dordrecht (1996)
- [374] Swets, J.A.: The relative operating characteristic in psychology. A technique for isolating effects of response bias finds wide use in the study of perception and cognition. *Science* 182(4116), 990–1000 (1973)
- [375] Swets, J.A., Dawes, R.M., Monahan, J.: Better decisions through science. *Scientific American* 283, 82–87 (2000)

- [376] Świniarski, R.W.: Rough Sets and Bayesian Methods Applied to Cancer Detection. In: Polkowski, L., Skowron, A. (eds.) RSCTC 1998. LNCS (LNAI), vol. 1424, pp. 609–616. Springer, Heidelberg (1998)
- [377] Syswerda, G.: Uniform crossover in genetic algorithms. In: Saffer, J.D. (ed.) Proc. 3rd International Conference on Genetic Algorithms, pp. 2–9. Morgan Kaufmann, San Francisco (1989)
- [378] Tan, P.-N., Steinbach, M., Kumar, V.: Introduction to Data Mining. Addison-Wesley, Reading (2005)
- [379] Tay, F., Shen, L.: Economic and financial prediction using rough sets model. *European Journal of Operational Research* 141(3), 641–659 (2002)
- [380] Thomson, S.K.: Sampling, 2nd edn. Series in Probability and Statistics. Wiley, Chichester (2002)
- [381] Thuraisingham, B.: Data Mining: Technologies, Techniques, Tools, and Trends. CRC Press, Boca Raton (1999)
- [382] Thurstone, L.L.: Multiple factor analysis. *Psychological Review* 38, 406–427 (1931)
- [383] Titterton, D.M., Smith, A.F.M., Makov, U.E.: Statistical Analysis of Finite Mixture Distributions. John Wiley and Sons, New York (1985)
- [384] Tong, S., Chang, E.: Support Vector Machine Active Learning for Image Retrieval. In: Proc. of the 9th ACM international conference on Multimedia, vol. 9, pp. 107–118 (2001)
- [385] Toussaint, G.T.: Geometric proximity graphs for improving nearest neighbor methods in instance-based learning and data mining. *International Journal of Computational Geometry and Applications* 15(2), 101–150 (2005)
- [386] Tukey, J.: Exploratory Data Analysis. Addison-Wesley, Reading (1977)
- [387] Unnikrishnan, N., Mahajan, A., Chu, T.: Intelligent system modeling of a three-dimensional ultrasonic positioning system using neural networks and genetic algorithms. *Proc. Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 217(5), 367–377 (2003)
- [388] Vajda, P., Eiben, A.E., Hordijk, W.: Parameter control methods for selection operators in genetic algorithms. In: Rudolph, G., Jansen, T., Lucas, S., Poloni, C., Beume, N. (eds.) PPSN 2008. LNCS, vol. 5199, pp. 620–630. Springer, Heidelberg (2008)
- [389] Vapnik, V.N., Chervonenkis, Y.: On the uniform convergence of relative frequencies of events to their probabilities. *Theoretical Probability and its applications* 17, 264–280 (1971)
- [390] Vapnik, V.N.: Estimation of dependencies based on empirical data (translated by S. Kotz). Springer, Heidelberg (1982)
- [391] Vapnik, V.N.: The nature of statistical learning theory. Springer, Heidelberg (1995)
- [392] Vapnik, V.N., Golowich, S.E., Smola, A.J.: Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. In: NIPS 1996, Denver, USA, pp. 281–287 (1996)
- [393] Vapnik, V.N.: Statistical learning theory. Wiley, Chichester (1998)
- [394] Ward, J.H.: Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58(301), 236–244 (1963)
- [395] Velleman, P., Hoaglin, D.: The ABC's of EDA: Applications, Basics, and Computing of Exploratory Data Analysis. Duxbury, Boston (1981)
- [396] Wand, Y., Wang, R.: Anchoring Data Quality Dimensions in Ontological Foundations. *Communications of the*, pp. 86–95. ACM Press, New York (1996)
- [397] Wang, J., Neskovic, P., Cooper, L.N.: Improving nearest neighbor rule with a simple adaptive distance measure. *Pattern Recognition Letters* 28, 207–213 (2007)

- [398] Westphal, C., Blaxton, T.: *Data Mining Solutions*. John Wiley, Chichester (1998)
- [399] Whitley, D., Kauth, J.: Genitor: A different genetic algorithm. In: Proc. of the Rocky Mountain Conference on Artificial Intelligence, pp. 118–130 (1988)
- [400] Whitley, D.: The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. In: Proc. of the Third International Conference on Genetic Algorithms, pp. 116–123. Morgan Kaufmann, San Francisco (1989)
- [401] Whitley, D., Hanson, T.: Optimizing neural networks using faster more accurate genetic search. In: Proc. 3rd International Conference on Genetic Algorithms, pp. 391–396. Morgan Kaufmann, San Francisco (1989)
- [402] Whitley, D., Starkweather, T., Fuquay, D.'.A.: Scheduling problems and traveling salesman: The genetic edge recombination operator. In: Proc. 3rd International Conference on Genetic Algorithms, pp. 133–140. Morgan Kaufmann, San Francisco (1989)
- [403] Whitley, D., Starkweather, T., Shaner, D.: *Traveling Salesman and Sequence Scheduling: Quality Solutions Using Genetic Edge Recombination*. Handbook of Genetic Algorithms. Van Nostrand (1990)
- [404] Whitley, D.: *Genetic Algorithms in Engineering and Computer Science*. In: Periaux, J., Winter, G. (eds.) John Wiley and Sons Ltd., Chichester (1995)
- [405] Whitley, D.: Permutations. In: Back, T., Fogel, D.B., Michalewicz, Z. (eds.) *Evolutionary computation 1: basic algorithms and operators*, ch. 33.3, pp. 274–284. Institute of Physics Publishing, Bristol (2000)
- [406] Widrow, B., Hoff, M.E.: Adaptive switching circuits. IRE WESCON Convention Record 4, 96–104 (1960); reprinted in Anderson and Rosenfeld (1988)
- [407] Widrow, B., Lehr, M.A.: 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. Proc. of the IEEE 78(9), 1415–1442 (1990)
- [408] Wikinews: U.S. Army intelligence had detected 9/11 terrorists year before (August 18, 2005),
http://en.wikinews.org/w/index.php?title=U.S._Army_intelligence_had_detected_9/11_terrorists_year_before%2C_says_officer&oldid=130741
- [409] Wishart, G.C., Warwick, J., Pitsinis, V., Duffy, S., Britton, P.D.: Measuring performance in clinical breast examination. *Br. J. Surg.* 7(8), 1246–1252 (2010)
- [410] Witten, I.H., Eibe, F.: *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edn. Morgan Kaufmann, San Francisco (2005)
- [411] Wood, S.: *Generalized additive models. An introduction with R*. Chapman and Hall/CRC (2006)
- [412] Yao, J., Herbert, J.: Financial time-series analysis with rough sets. *Applied Soft Computing* 9(3), 1000–1007 (2009)
- [413] Ye, N., Farley, T.: A scientific approach to cyberattack detection. *Computer* 38(11), 55–61 (2005)
- [414] Ye, N., Chen, Q.: Computer intrusion detection through EWMA for autocorrelated and uncorrelated data. *IEEE Trans. Reliability* 52(1), 73–82 (2003)
- [415] Ye, N.: Mining computer and network security data. In: Ye, N. (ed.) *The Handbook of Data Mining*, pp. 617–636. Lawrence Erlbaum Assoc., Mahwah (2003)
- [416] Yeh, W., Huang, S.W., Li, P.C.: Liver Fibrosis Grade Classification with B-mode Ultrasound. *Ultrasound in Medicine and Biology* 29(9), 1229–1235 (2003)
- [417] Zadeh, L.: Fuzzy sets. *Information and Control* 8, 338–353 (1965)
- [418] Zaki, M.J., Parthasarathy, S., Ogihara, M., Li, W.: New algorithms for fast discovery of association rules. Technical Report TR651 (1997)

- [419] Zaknich, A.: Neural networks for intelligent signal processing. Series in Innovative Intelligence, vol. 4. World Scientific, Singapore (2003)
- [420] Zhang, C., Zhang, S.: Association Rule Mining. Models and Algorithms. LNCS (LNAI), vol. 2307. Springer, Heidelberg (2002)
- [421] Zhong, N.: Using Rough Sets with Heuristics for Feature Selection. *Journal of Intelligent Information Systems* 16, 199–214 (2001)
- [422] Ziarko, W.: *Rough Sets, Fuzzy Sets and Knowledge Discovery*. Springer, Heidelberg (1994)
- [423] Zur, R.M., Pesce, L.L., Jiang, Y.: The effect of two priors on Bayesian estimation of "Proper" binormal ROC curves from common and degenerate datasets. *Acad. Radiol.* 17(8), 969–979 (2010)
- [424] Zweig, M.H., Campbell, G.: Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39(8), 561–577 (1993)

Index

- accuracy
 - classification, 320
 - empirical, 329
 - model, 12
 - predictive, 16
 - true, 329
- activation function, 195
 - Gaussian, 196
 - Heaviside, 196
 - linear, 196
 - logistic, 196
 - piecewise-linear/ramp function, 196
 - sigmoid, 196
 - step/hard limiter, 196
- ADALINE, 214
- additive model, 127
- adjustment model, 127
- algorithm
 - back-propagation, 213
 - greedy, 164
 - Hunt, 161
 - k-means, 284
 - k-nearest(neighbor), 256
 - tree building/SPRINT, 164
- anomaly, 148
- ARIMA model, 128
- artificial neuron, 194
- association rule discovery, 23
- autopredictive model, 128
- average absolute deviation, 68

- B-boundary region, 265
- B-lower approximation, 265
- B-upper approximation, 265

- Bayes decision rule, 222
- Bayesian decision rule, 188
- Bayesian selection, 32
- bias, 195
- bivariate joint distribution, 69
- boundary region, 261

- chromosome, 290
- classification, 15
- clustering, 19
 - agglomerative, 272
 - conceptual, 272
 - divisive, 272
 - hierarchical, 272
 - k-means, 284
 - method, 155
- conditional distribution, 69
- confidence interval (95%), 67, 330
- confusion matrix, 319
- contingency/correlation table, 80
- correlation
 - coefficient, 74
 - method, 72
 - Pearson's, 73
- correlation matrix, 85
- cost matrix, 319
- covariance, 73
- covariance matrix, 133
- Cox proportional hazard model, 121
- Cox proportional hazards regression analysis, 121
- crisp set, 265
- cross-validation, 258, 284
- crossover

- cycle, 306
- edge, 304
- edge-3, 304
- N-point, 302
- one-point, 302
- order, 306
- partially mapped, 303
- uniform, 302
- crossover rate, 301
- cube
 - multidimensional, 138
 - OLAP, 138
- cumulative frequency/empirical distribution, 61
- data, 45
 - categorical data, 54
 - censored, 55
 - document data, 47
 - duplicate, 52
 - missing data, 51
 - nominal data, 54
 - numerical data, 53
 - ordered data, 49
 - percentage, 54
 - qualitative data, 54
 - quantitative data, 53
 - rank, 54
 - rate and ratio, 54
 - record data, 47
 - score, 55
 - transaction data, 47
- data/character, 56
- decile, 62
- decision system, 262
- density, 69
- dependent/outcome variable, 75
- descriptive methods, 14
- discernibility matrix, 261, 263
- discriminant function, 215, 239
- distance
 - Chebychev, 278
 - cosine, 279
 - Euclidean, 278
 - fuzzy, 280
 - Hamming, 278
 - Jaccard (index), 279
 - Mahalanobis, 279
 - Manhattan/city block/taxicab, 278
 - Minkowski, 278
 - Pearson's r , 279
 - power, 278
 - Tanimoto, 279
- distribution
 - exponential, 102
 - Gamma, 102
 - log-normal, 104
 - marginal, 69
 - Weibull, 103
- distribution (probability), 61
- duration analysis, 120
- dynamic series, 124
- elbow criterion, 22
- equivalence classes, 261
- estimation, 28
- evolution strategy, 291
- evolutionary algorithm, 289
- evolutionary programming, 291
- expectation, 70
 - conditional, 70
- explanatory model, 128
- factor analysis, 130
- feedback, 206
- fitness measure, 290
- fitness proportional selection, 297
- fitness/evaluation function, 294
- forecasting, 126
- formula
 - Bayes, 186
 - conditional probability, 187
 - total probability, 186
- free model, 31
- Gaussian mixture model, 288
- gene, 290, 293
- generalized additive model, 123, 124
- generalized linear model, 123
- genetic algorithm, 291
 - exploitation, 313
 - exploration, 313
 - fundamental/canonical, 312
 - parameter control, 310
 - parameter tuning, 310
 - parameters, 309
- genetic programming, 291
- genotype, 290
- geometric mean, 65

- goodness-of-fit, 110, 111
- gradient, 213
- harmonic mean, 65
- hazard function, 122
- Henry's line, 112
- hierarchical model, 31
- histogram, 82
- HOLAP, 139
- hypercube (OLAP), 138
- independent variables, 69
- independent/predictor variable, 75
- index
 - entropy, 166, 169
 - Gini, 165, 166
 - misclassification, 166
 - misclassification/error, 171
- indiscernibility relation, 262, 263
- indiscernible objects, 261
- information gain, 169
- information system, 261
 - core, 267
- information table, 261
- initialization, 294
- inner-product kernel, 237
- intercept, 75, 105
- Iris
 - Setosa, 141
 - Versicolour, 141
 - Virginica, 141
- kurtosis, 101
- learning
 - association rule, 249
 - Boltzmann, 209
 - competitive, 209
 - error-correction, 209
 - Hebbian, 209
 - memory-based, 209
 - polynomial, 243
 - reinforcement, 209, 210
 - supervised, 9, 209
 - unsupervised, 9, 211
 - unsupervised/self-organized, 209
 - with a teacher, 209
 - without a teacher, 209
- learning classifier system, 291
- learning-rate parameter, 199
- least squares method, 75
- likelihood, 154, 186
- likelihood ratio negative, 323
- likelihood ratio positive, 323
- logit, 117
- lower approximation, 261
- MADALINE, 214
- margin of separation, 235, 238
- mean/average, 63
- median, 63
- median absolute deviation, 68
- minimum description length, 181
- mode, 64
- model, 27
 - McCulloch-Pitts, 197
 - Rosenblatt's perceptron, 197
- MOLAP, 139
- Monte Carlo roulette wheel, 297
- moving average, 125
- multiplicative model, 127
- mutation
 - binary representation, 307
 - bitwise, 307
 - creep, 308
 - fixed distribution, 309
 - floating-point representation, 308
 - insert, 309
 - integer representation, 308
 - inversion, 309
 - non-uniform, 308
 - permutation representation, 309
 - random resetting, 308
 - scramble, 309
 - swap, 309
 - uniform, 308
- mutation rate, 307
- naive Bayes/Idiot's Bayes, 188
- nearest-neighbor method, 155
- negative predictive value, 322
- neural network, 191
 - feedforward, 206
 - Hopfield, 216
 - Kohonen self-organizing, 215
 - multilayer feedforward, 207
 - probabilistic, 218
 - radial-basis function, 214, 243
 - recurrent, 207, 208

- single-layer feedforward, 207
- noise, 51
- Occam's razor, 181
- odds ratio, 120
- OLAP
 - dicing, 145
 - drill-down, 145
 - pivot, 145
 - roll-up, 145
 - slicing, 145
- optimal hyperplane, 238
- outlier, 51, 148
- parents selection, 297
- Parzen's estimates, 218
- Parzen-Cacoulos classifier, 219
- percentile, 62
- perceptron, 197
 - multi-layer, 212
 - single-layer, 212
 - two-layer, 243
- perceptron convergence algorithm, 202
- phenotype, 290
- population, 296
- positive predictive value, 322
- posterior probability, 32, 186
- predictive methods, 14
- predictor, 15
- principal components, 134
- principal components analysis, 133
- prior probability, 32, 186
- probability density function, 69
- probability mass function, 69
- prognostic index, 120
- quantile, 62
- quartile, 62
- random variable, 56, 61
- range, 68
 - interquartile range, 68
- ranking selection, 299
- rate
 - false negative, 323
 - false positive, 323
 - true positive, 322
- recombination
 - arithmetic/intermediate, 303
 - floating-point representation, 303
 - multi-parent, 307
 - permutation representation, 303
 - simple arithmetic, 303
 - single arithmetic, 303
 - uniform arithmetic, 303
 - whole arithmetic, 303
- recombination operator
 - binary representation, 302
 - integer representation, 303
- recombination/crossover, 301
- reduct, 261, 267
- regression, 25
 - backward stepwise, 106
 - coefficient, 75, 105
 - equation, 75
 - forward stepwise, 106
 - line, 75
 - linear, 75
 - multiple linear, 105
 - multiple linear logistic, 117
 - multiple linear regression equation, 105
 - nonlinear, 76
- reliability theory, 120
- representation
 - binary, 293
 - floating point/real-coded, 293
 - integer, 293
 - permutation, 293
- representation (of individuals), 293
- residual, 75
- robustness, 16
- ROC curve, 323
 - area under (AUC), 325
- ROLAP, 139
- rough set, 265
- rule
 - confidence, 250
 - lift, 250
 - support, 250
- rule-based classification, 252
- scalability, 16
- scatterplot matrix, 87
- seasonal decomposition, 127
- sensitivity, 322
- sequential patterns, 25
- similarity measure, 273, 277
 - mixed-weighted, 280
 - weighted, 280

- skewness, 99
- slack variables, 242, 247
- smoothing parameter, 219
- specificity, 322
- standard deviation, 67
- statistical series, 59
- statistical variable, 61
- stochastic universal sampling, 298
- strategy
 - age-based, 300
 - fitness-based, 300
- support vector machine, 234
 - evolutionary, 247
- support vector regression, 246
- support vectors, 235, 240
- survival analysis, 120
- survival function, 122
- survival time, 121
- synaptic weight, 195
- taxonomy, 15
- test
 - Grubbs-outliers, 153
 - likelihood-outliers, 154
- testing dataset, 15
- theorem
 - Bayes' formula, 186
 - perceptron convergence, 199
 - Rosenblatt, 199
 - Total probability formula, 186
- time series, 124
- total fitness, 295
- total probability formula, 32
- tournament selection, 299
- training dataset, 15
- transform
 - Hotelling, 136
 - Karhunen-Loeve, 136
- tree
 - classification, 161
 - minimum n, 180
 - misclassification costs, 179
 - overfitting, 180
 - prior probabilities, 179
 - proportion of objects, 180
 - pruning, 180
 - regression, 161
 - stopping rules, 180
 - underfitting, 180
- trend analysis, 127
- universe, 262
- upper approximation, 261
- variance, 66
- VC dimension, 241
- XOR problem, 243
- z-test, 328