References

D. H. Ackley, G. E. Hinton and T. J. Sejnowski (1985). "A learning algorithm for Boltzmann machines." *Cognitive Science* **9**: 147–169.

E. Alpaydin and C. Kaynak (1998). "Cascading Classifers." *Kybernetika* **34**(4): 369–374.

E. Anderson (1935). "The irises of the Gaspé Peninsula." *Bulletin of the American Iris Society* **59**: 2–5.

A. Bain (1873). *Mind and Body: The Theories of Their Relation*. London, Henry King.

P. V. Balakrishnan, M. C. Cooper, V. S. Jacob and P. A. Lewis. (1994). "A study of the classification capabilities of neural networks using unsupervised learning: A comparison with A-means clustering." *Psychometrika* **59**(4): 509–525.

E. B. Baum and D. Haussler (1989). "What size net gives valid generalization." *Neural Computation* **1**(1): 151–160.

W. G. Baxt and H. White (1995). "Bootstrapping confidence intervals for clinical input variable effects in a network trained to identify the presence of acute myocardial infarction." *Neural Computation* **7**(3): 624–638.

S. Becker and Y. LeCun (1989). Improving the convergence of back-propagation learning with second order methods. *Proceedings of the 1988 Connectionist Summer School*. D. Touretzky, G. E. Hinton and T. Sejnowski. San Mateo, CA, Morgan Kaufmann Publishers: 29–37.

R. E. Bellman (1961). *Adaptive Control Processes: A Guided Tour*. Princeton, NJ, Princeton University Press.

L. M. Belue and K. W. Bauer, Jr. (1995). "Determining input features for multilayer perceptrons." *Neurocomputing* **7**(2): 1A–121.

D. P. Bertsekas and J. Tsitsiklis (1996). *Neuro-Dynamic Programming*. Belmont, MA, Athena Scientific.

M. Bianchini and M. Gori (1996). "Optimal Learning in Artificial Neural Networks: A Review of Theoretical Results." *Neurocomputing* **13**(5): 313–346.

C. M. Bishop (1995). *Neural Networks for Pattern Recognition*. Oxford, Oxford University Press.

C. L. Blake and C. J. Merz (1998). UCI Repository of machine learning databases. Irvine, CA, University of California, Department of Information and Computer Science**:** http://www.ics.uci.edu/~mlearn/MLRepository.html.

L. Breiman (1996). "Heuristics of instability and stabilization in model selection." *Annals of Statistics* **24**(6): 2350–2383.

L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone (1984). *Classification and Regression Trees*, Kluwer Academic Publishers.

L. Breiman and P. Spector (1992). "Submodel selection and evaluation in regression: The X-random case." *International Statistical Review* **60**(3): 291–319.

J. S. Bridle (1990). Training Stochastic Model Recognition Algorithms as Networks can lead to Maximum Mutual Information Estimation of Parameters. *Advances in Neural Information Processing Systems*. D. S. Touretzky. San Mateo, CA, Morgan Kaufmann Publishers. **2**: 2A–217.

H. B. Burke (1996). The Importance of Artificial Neural Networks in Biomedicine. *Applications of Neural Networks in Environment, Energy, and Health*. P. E. Keller, S. Hashem, L. J. Kangas and R. T. Kouzes. Singapore, World Scientific Publishing**:** 145–153.

C. Cardaliaguet and E. Guillaume (1992). "Approximation of a function and its derivative with a neural network." *Neural Networks* **5**(2): 207–220.

G. A. Carpenter (1997). "Distributed learning, recognition, and prediction by ART and ARTMAP neural networks." *Neural Networks* **10**(8): 1473–1494.

G. A. Carpenter and S. Grossberg (1987a). "A massively parallel architecture for a self-organizing neural pattern recognition machine." *Computer Vision, Graphics, and Image Processing* **37**(1): 54–115.

G. A. Carpenter and S. Grossberg (1987b). "ART2: Stable self-organization of pattern recognition codes for analog input patterns." *Applied Optics* **26**(23): 4919–4930.

G. A. Carpenter and S. Grossberg (1990). ART3: Self-organization of Distributed Pattern Recognition Codes in Neural Network Hierarchies. *Proceedings of the International Conference on Neural Networks (INNC'90)*. Amsterdam, Kluwer Academic Publishers, North-Holland. **2**: 801–804.

G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds and D. B. Rosen (1992). "Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps." *IEEE Transactions on Neural Networks* **3**(5): 698–713.

G. A. Carpenter, S. Grossberg and J. H. Reynolds (1995). "A Fuzzy ARTMAP Nonparametric Probability Estimator for Nonstationary Pattern Recognition Problems." *IEEE Transactions on Neural Networks* **6**(6): 1330–1336.

G. A. Carpenter, S. Grossberg and D. B. Rosen (1991). "Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance Systems." *Neural Networks* **4**(6): 759–771.

G. A. Carpenter and W. D. Ross (1995). "ART-EMAP: A Neural Network Architecture for Object Recognition by Evidence Accumulation." *IEEE Transactions on Neural Networks* **6**(4): 805–818.

T. P. Caudell, S. D. G. Smith, R. Escobedo and M. Anderson (1994). "NIRS: Large Scale ART-1 Neural Architectures for Engineering Design Retrieval." *Neural Networks* **7**(9): 1339–1350.

L.-W. W. Chan and F. Fallside (1987). "An adaptive training algorithm for back-propagation networks." *Computer Speech and Language* **2**: 205–218.

T. Chen and H. Chen (1995). "Universal Approximation to Nonlinear Operators by Neural Networks with Arbitrary Activation Functions and Its Application to Dynamical Systems." *IEEE Transactions on Neural Networks* **6**(4): 9A–917.

G. Copson, R. Badcock, J. Boon and P. Britton (1997). "Articulating a systematic approach to clinical crime profiling." *Criminal Behaviour and Mental Health* **7**: 13–17.

T. M. Cover (1965). "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition." *IEEE Transactions on Electronic Computers* **E3-14**(3): 326–334.

G. Coward (1992). *Tree Book: Learning to Recognize Trees of British Columbia*. Victoria, BC, Canada, Forestry Canada.

N. Cristianini and J. Shawe-Taylor (2000). *An Introduction to Support Vector Machines*. Cambridge, Cambridge University Press.

G. V. Cybenko (1989). "Approximation by Superpositions of a Sigmoidal Function." *Mathematics of Control, Signals, and Systems* **2**(4): 303–314.

H. Dai and C. Macbeth (1997). "Effects of learning parameters on learning procedure and performance of a BPNN." *Neural Networks* **10**(8): 1505–1521.

J. S. Denker and Y. LeCun (1991). Transforming Neural-Net Output Levels to Probability Distributions. *Advances in Neural Information Processing Systems*. R. Lippmann, J. E. Moody and D. S. Touretzky. San Mateo, CA, Morgan Kaufmann Publishers. **3**: 853–859.

H. Drucker and Y. LeCun (1992). "Improving Generalization Performance Using Double Backpropagation." *IEEE Transactions on Neural Networks* **3**(6): 991–997.

H. Drucker, D. Wu and V. N. Vapnik (1999). "Support Vector Machines for Spam Categorization." *IEEE Transactions on Neural Networks* **10**(5): 1048–1054.

R. O. Duda and P. E. Hart (1973). *Pattern Analysis and Scene Classification*. New York, John Wiley & Sons.

R. O. Duda, P. E. Hart and D. G. Stork (2001). *Pattern Classification, 2nd Edition*. New York, John Wiley & Sons Inc.

R. Eckhorn, H. J. Reitboeck, M. Arndt and P. Dicke (1990). "Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex." *Neural Cooperativity* **2**(3): 293–307.

B. Efron (1979). "Bootstrap methods: Another look at the jackknife." *Annals of Statistics* **7**(1): 1–26.

B. Efron (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, Society for Industrial and Applied Mathematics.

B. Efron (1983). "Estimating the error rate of a prediction rule: Improvement on cross-validation." *Journal of the American Statistical Association* **78**: 316–331.

B. Efron and R. Tibshiran (1986). "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy." *Statistical Science* **1**: 54–77.

B. Efron and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. London, Chapman & Hall.

B. Efron and R. J. Tibshirani (1997). "Improvements on cross-validation: The .632+ bootstrap method." *Journal of the American Statistical Association* **92**: 548–560.

J. L. Elman (1991). "Distributed representations, simple recurrent networks, and grammatical structure." *Machine Learning* **7**: 195–225.

S. E. Fahlman (1989). Faster Learning Variations on Back Propagation: An Empirical Study. *Proceedings of the 1988 Connectionist Models Summer School*. D. Touretzky, G. E. Hinton and T. Sejnowski. San Mateo, CA, Morgan Kaufmann Publishers: 38–51.

S. E. Fahlman and C. Liebiere (1990). The Cascade—Correlation Learning Architecture. *Advances in Neural Information Processing Systems*. D. S. Touretzky. San Mateo, CA, Morgan Kaufmann Publishers. **2**: 524–532.

B. G. Farley (1960). Self-Organizing Models for Learned Perception. *Self-Organizing Systems*. M. C. Yovits and S. Cameron. Oxford, UK, Pergamon Press.

B. G. Farley and W. A. Clark (1954). "Simulation of Self-Organizing Systems by Digital Computer." *IRE Transactions on Information Theory* **4**(4): 76–84.

B. G. Farley and W. A. Clark (1955). Generalization of pattern recognition in a self-organizing system. *Proceedings of the 1955 Western Joint Computer Conference*: 86–91.

R. A. Fisher (1936). "The use of multiple measurements in taxonomic problems." *Annual Eugenics* **7**(Part II): 179–188.

D. Fogel (1990). "An Information Criterion for Optimal Neural Network Selection." *IEEE Transactions on Neural Networks* **2**(5): 490–497.

D. H. Foley (1972). "Considerations of Sample and Feature Size." *IEEE Transactions on Information Theory* **IT-18**(5): 618–626.

N. Fraser (1998). Neural Network Follies. *http://neil.fraser.name/writing/tank/*.

B. R. Frieden (1983). *Probability, Statistical Optics, and Data Testing*. Berlin Heidelberg New York, Springer-Verlag.

K. Fukunaga (1990). *Introduction to Statistical Pattern Recognition*. San Diego, CA, Academic Press, Inc.

K. Funahashi (1989). "On the approximate realization of continuous mappings by neural networks." *Neural Networks* **2**(3): 183–192.

V. J. Geberth (1996). *Practical Homicide Investigation: Tactics, Procedures, and Forensic Techniques*. Boca Raton, Florida, CRC Publishing.

K. A. Gernoth and J. W. Clark (1995). "Neural Networks That Learn to Predict Probabilities: Global Models of Nuclear Stability and Decay." *Neural Networks* **8**(2): 291–311.

G. R. Gindi, A. F. Gmitro and K. Parthasarathy (1988). "Hopfield Model Associative Memory with Nonzero-Diagonal Terms in Memory Matrix." *Applied Optics* **27**(1): 129–134.

A. F. Gmitro, P. E. Keller and G. R. Gindi (1989). "Statistical performance of outer-product associative memory models." *Applied Optics* **28**(10): 1940–1948.

C. Goutte (1997). "Note on free lunches and cross-validation." *Neural Computation* **9**(6): 12A–1215.

S. Grossberg (1976a). "Adaptive pattern classification and universal recording: I. Parallel development and coding of neural detectors." *Biological Cybernetics* **23**: 121–134.

S. Grossberg (1976b). "On the Development of Feature Detectors in the Visual Cortex with Applications to Learning and Reaction-Diffusion Systems." *Biological Cybernetics* **21**(3): 145–159.

S. Grossberg (1976c). "Adaptive pattern classification and universal recording: II. Feedback, expectation, olfaction, illusions." *Biological Cybernetics* **23**: 187–202.

S. Grossberg (1987). *The Adaptive Brain I: Cognition, Learning, Reinforcement, and Rhythm*. Amsterdam, Elsevier/North-Holland.

M. Hagiwara (1992). Theoretical derivation of momentum term in back-propagation. *Proceedings of the International Joint Conference on Neural Networks (IJCNN'92)*. Piscataway, NJ, IEEE. **1**: 682–686.

C. L. Harris and J. L. Elman (1989). Representing variable information with simple recurrent networks. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ, Lawrence Erlbaum: 635–642.

D. Harrison, Jr. and D. L. Rubinfeld (1978). "Hedonic housing prices and the demand for clean air." *Journal of Environmental Economics and Management* **5**(1): 81–102.

S. Haykin (1994). *Neural Networks: A Comprehensive Foundation*. New York, Macmillan College Publishing Company.

M. A. Hearst, B. Scholkopf, S. Dumais, E. Osuna and J. Platt (1998). "Trends and Controversies: Support Vector Machines." *IEEE Intelligent Systems* **13**(4): 18–28.

D. O. Hebb (1949). *The Organization of Behavior*. New York, John Wiley & Sons.

R. Hecht-Nielsen (1987). "Counterpropation networks." *Applied Optics* **26**(23): 4979–4983.

J. S. U. Hjorth (1994). *Computer Intensive Statistical Methods Validation, Model Selection, and Bootstrap*. London, Chapman & Hall.

A. L. Hodgkin and A. F. Huxley (1952). "A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve." *Journal of Physiology* **117**: 500–544.

J. J. Hopfield (1982). "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the National Academy of Sciences USA* **79**: 2554–2558.

J. J. Hopfield (1984). "Neurons with graded response have collective computational properties like those of two-state neurons." *Proceedings of the National Academy of Sciences USA* **81**: 3088–3092.

J. J. Hopfield and D. W. Tank (1985). "'Neural' computation of decisions in optimization problems." *Biological Cybernetics* **52**: 141–152.

J. J. Hopfield and D. W. Tank (1986). "Computing with neural circuits: A model." *Science* **233**: 625–633.

K. Hornik, M. Stinchcombe and H. White (1989). "Multilayer feedforward networks are universal approximators." *Neural Networks* **2**: 359–366.

H. Hotelling (1933). "Analysis of a complex of statistical variables into principal components." *Journal of Educational Psychology* **24**: 417–441 and 498–520.

J. Huang, M. Georgiopoulos and G. L. Heileman (1995). "Fuzzy ART Properties." *Neural Networks* **8**(2): 203–213.

S.-C. Huang and Y.-F. Huang (1990). "Learning Algorithms for Perceptrons Using Back Propagation with Selective Updates." *IEEE Control Systems Magazine* **10**(3): 56–61.

C.-A. Hung and S.-F. Lin (1995). "Adaptive Hamming Net: A Fast-Learning ART 1 Model Without Searching." *Neural Networks* **8**(4): 605–618.

B. Hunt, M. S. Nadar, P. Keller, E. VonColln and A. Goyal (1993). "Synthesis of a Nonrecurrent Associative Memory Model Based on a Nonlinear Transformation in the Spectral Domain." *IEEE Transactions on Neural Networks* **4**(5): 873–878.

C. M. Hurvich and C.-L. Tsai (1989). "Regression and time series model selection in small samples." *Biometrika* **76**: 297–307.

D. R. Hush and B. G. Horne (1993). "Progress in Supervised Neural Networks: What's New Since Lippmann?" *IEEE Signal Processing Magazine* **10**(1): 8–39.

A. Hyvärinen and E. Oja (1999). Independent Component Analysis: A Tutorial. http://www.cis.hut.fi/∼ aapo/ps/NN00.pdf. Espoo, Finland, Helsinki University of Technology.

A. Hyvärinen and E. Oja (2000). "Independent Component Analysis: Algorithms and Applications." *Neural Networks* **13**(4-5): 4A–430.

J. E. Jackson (1991). *A User's Guide to Principal Components*. New York, John Wiley & Sons Inc.: 63–69.

R. A. Jacobs (1988). "Increased Rates of Convergence Through Learning Rate Adaptation." *Neural Networks* **1**(4): 295–307.

W. James (1890). *Principles of Psychology*. New York, Henry Holt.

D. S. Johnson and C. H. Papadimitriou (1985). Computational Complexity. *The Traveling Salesman Problem*. E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan and D. B. Shmoys. New York, John-Wiley & Sons: 37–85.

B. L. Kalman and S. C. Kwasny (1992). Why Tanh: Choosing a Sigmoidal Function. *Proceedings of the International Joint Conference on Neural Networks (IJCNN'92)*. Piscataway, NJ, IEEE. **4**: 578–581.

L. J. Kangas, K. M. Terrones, R. D. Keppel and R. D. La Moria (1998). Computer Aided Tracking and Characterization of Homicides & Sexual Assaults (CATCH). *Applications and Science of Computational Intelligence II—Proceedings of the SPIE*. K. L. Priddy, P. E. Keller, D. B. Fogel and J. C. Bezdek. Bellingham, WA, SPIE. **3722**: 250–260.

P. P. Kanjilal and D. N. Banerjee (1995). "On the Application of Orthogonal Transformation for the Design and Analysis of Feedforward Networks." *IEEE Transactions on Neural Networks* **6**(5): 1061–1070.

M. Kearns (1997). "A bound on the error of cross validation using the approximation and estimation rates, with consequences for the training-test split." *Neural Computation* **9**(5): 1143–1161.

P. E. Keller and A. D. McKinnon (1999). Pulse-Coupled Neural Networks for Medical Image Analysis. *Applications and Science of Computational Intelligence II—Proceedings of the SPIE*. K. L. Priddy, P. E. Keller, D. B. Fogel and J. C. Bezdek. Bellingham, WA, SPIE. **3722**: 444–451.

P. E. Keller, D. L. McMakin, D. M. Sheen, A. D. McKinnon and J. W. Summet (2000). Privacy Algorithm for Airport Passenger Screening Portal. *Applications and Science of Computation Intelligence III – Proceedings of the SPIE*. K. L. Priddy, P. E. Keller and D. B. Fogel. Bellingham, WA, SPIE. **4055**: 476–483.

R. D. Keppel and R. Walter (1999). "Profiling Killers: A Revised Classification Model for Understanding Sexual Murder." *International Journal of Offender Therapy and Comparative Criminology* **43**(4): 417–437.

R. Kohavi (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'95)*. San Mateo, CA, Morgan Kaufmann Publishers: 1137–1143.

T. Kohonen (1972). "Correlation Matrix Memories." *IEEE Transactions on Computers* **3-21**: 353–359.

T. Kohonen (1982). "Self-organized formation of topologically correct feature maps." *Biological Cybernetics* **43**: 59–69.

T. Kohonen (1987). "Adaptive, associative, and self-organizing functions in neural computing." *Applied Optics* **26**(23): 4910–4918.

T. Kohonen (1988). "The 'Neural' Phonetic Typewriter." *IEEE Computer Magazine* **21**(3): A-22.

T. Kohonen (1989). *Self-Organization and Associative Memory*. Berlin Heidelberg London, Springer-Verlag.

B. Kosko (1987). "Constructing an associative memory." *Byte* **12**(10): 137–144.

B. Kosko (1987). "Adaptive bidirectional memories." *Applied Optics* **26**(23): 4947–4960.

B. Kosko (1988). "Bidirectional associative memory." *IEEE Transactions on Systems, Man and Cybernetics* **SM3-18**(1): 49–60.

B. Kosko (1992). *Neural Networks for Signal Processing*. Upper Saddle River, NJ, Prentice Hall.

A. Kowalczyk (1997). "Estimates of storage capacity of multilayer perceptron with threshold logic hidden units." *Neural Networks* **10**(8): 1417–1433.

M. A. Kraaijveld, J. Mao and A. K. Jain (1995). "A Nonlinear Projection Method Based on Kohonen's Topology Preserving Maps." *IEEE Transactions on Neural Networks* **6**(3): 548–559.

A. H. Kramer and A. L. Sangiovanni-Vincentelli (1989). Efficient parallel learning algorithms for neural networks. *Advances in Neural Information Processing Systems*. D. S. Touretzky. San Mateo, CA, Morgan Kaufmann Publishers. **1**: 40–48.

Y. LeCun (1985). Une procedure d'apprentissage pour reseau a seuil assymetrique. *Cognitiva '85: A la frontière de l 'intelligence Artificielle des Sciences de la Connaissance des Neuronsciences*: 599–604.

G. G. Lendaris, K. Mathia and R. E. Saeks (1999). "Linear Hopfield Networks and Constrained Optimization." *IEEE Transactions of Systems, Man & Cybernetics* **29**(1): 114–118.

K. Levenberg (1944). "A method for the solution of certain problems in least squares." *Quart. Applied Mathematics* **2**: 164–168.

T. Linblad and J. M. Kinser (1998). *Image Processing using Pulse-Coupled Neural Networks*. London, Springer.

R. Lippmann (1987). "An Introduction to Computing with Neural Networks." *IEEE ASSP Magazine* **4**(2): 4–22.

P. C. Mahalanobis (1936). "On the generalized distance in statistics." *Proceedings of the National Institute of Science of India* **2**: 49–53.

O. L. Mangasarian and D. R. Musicant (1999). "Successive Overrelaxation for Support Vector Machines." *IEEE Transactions on Neural Networks* **10**(5): 1032–1037.

J. Mao and A. K. Jain (1996). "A Self-Organizing Network for Hyperellipsoidal Clustering (HEC)." *IEEE Transactions on Neural Networks* **7**(1): 16–29.

D. W. Marquardt (1963). "An algorithm for least-squares estimation of nonlinear parameters." *Journal of the Society of Industrial Applied Mathematics* **11**: 431–441.

S. Marriott and R. F. Harrison (1995). "A Modified Fuzzy ARTMAP Architecture for the Approximation of Noisy Mappings." *Neural Networks* **8**(4): 619–641.

T. Masters (1995). *Advanced Algorithms for Neural Networks: A C++ Sourcebook*. New York, John Wiley & Sons.

W. S. McCulloch and W. H. Pitts (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics* **5**: 115–133.

R. J. McEliece, E. C. Posner, E. R. Rodemich and S. S. Venkates (1987). "The capacity of the Hopfield associative memory." *IEEE Transactions on Information Theory* **33**(4): 461–482.

R. G. Miller (1974). "The jackknife, a review." *Biometrika* **61**: 1–15.

M. L. Minsky and S. A. Papert (1969). *Perceptrons: An introduction to Computational Geometry*. Cambridge, MA, MIT Press.

C. Z. Mooney and R. D. Duval (1993). *Bootstrapping: A Nonparametric Approach to Statistical Inference*, Sage Publications.

B. Moore (1988). ART 1 and Pattern Clustering. *Proceedings of the 1988 Connectionist Summer School*. D. Touretzky, G. E. Hinton and T. Sejnowski. San Mateo, CA, Morgan Kaufmann Publishers: 174–185.

F. M. Mulier and V. S. Cherkassky (1995). "Statistical Analysis of Self-organization." *Neural Networks* **8**(5): 717–727.

I. Nabney (2002). *Netlab: Algorithms for Pattern Recognition*. Berlin Heidelberg London, Springer-Verlag.

T. Nitta (1997). "An extension of the back-propagation algorithm to complex numbers." *Neural Networks* **10**(8): 1391–1415.

E. Oja (1991). Data compression, feature extraction, and auto-association in feedforward neural networks. *Proceedings of the 1991 International Conference on Artificial Neural Networks (ICANN'91)*. T. Kohonen, K. Makisara, O. Simula and J. Kangas. Amsterdam, Elsevier Science Publishers B. V. **1**: 737–746.

E. Oja (1991). "A simplified neuron model as a principal component analyzer." *Journal of Mathematical Biology* **15**: 267–273.

Y.-H. Pao (1989). *Adaptive Pattern Recognition and Neural Networks*. Reading, MA, Addison-Wesley Publishing Company, Inc.

D. Parker (1982). Learning-logic. *Invention Report S81-64, File 1*. Palo Alto, CA, Stanford University, Office of Technology Licensing.

M. Plutowski, S. Sakata and H. White (1994). Cross-validation estimates IMSE. *Advances in Neural Information Processing Systems*. J. D. Cowan, G. Tesauro and J. Alspector. San Mateo, CA, Morgan Kaufmann Publishers. **6**: 391–398.

T. Poggio (1975). "On optimal nonlinear associative recall." *Biological Cybernetics* **19**: 201–209.

B. T. Poljak (1964a). О некоторых способах ускорения сходимости итерационных методов." *Журн. выч. мат. и мат. физ. — Zhurnal Vychislitel'noi Matematiki I Matematicheskoi Fiziki* **4**(5): 791–803.

B. T. Poljak (1964b). "Some methods of speeding up the convergence of iteration methods." *USSR Computational Mathematics and Mathematical Physics* **4**(5): 1–17.

K. L. Priddy, S. K. Rogers, D. W. Ruck, G. L. Tarr and M. Kabrisky (1993). "Bayesian selection of important features for feedforward neural networks." *Neurocomputing* **5**(2): 91–103.

K.L. Priddy (2004). "A comparative analysis of machine classifiers." *Intelligent Computing: Theory and Applications II, Proceedings of SPIE*. K. L. Priddy. Bellingham, WA, SPIE. **5421**: 142–148.

M. H. Quenouille (1949). "Approximate tests of correlation in time series." *Journal of the Royal Statistical Society B* **11**: 18–84.

M. H. Quenouille (1956). "Notes on bias reduction." *Biometrika* **43**: 353–360.

S. Raudys (2000). "How Good are Support Vector Machines?" *Neural Networks* **13**(1): 17–19.

S. Raudys (2001). *Statistical and Neural Classifiers: An Integrated Approach to Design*. Berlin Heidelberg London, Springer-Verlag.

B. D. Ripley (1996). *Pattern Recognition and Neural Networks*. Cambridge, Cambridge University Press.

N. Rochester, J. H. Holland, H. L. H. and W. L. Duda (1956). "Tests on a Cell Assembly Theory of the Action of the Brain Using a Large Digital Computer." *IRE Transactions on Information Theory* **IT-2**(3): 80–93.

S. K. Rogers (1997). Tools for Pattern Recognition. *EENG 617 Class Handout*. Wright-Patterson AFB, OH, Air Force Institute of Technology.

F. Rosenblatt (1958). "The Perceptron: a Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review* **65**: 386–408.

F. Rosenblatt (1959). *Principles of Neurodynamics*. New York, Spartan Books.

W. A. Rosenblith and H. B. Barlow (1990). "Sensory communications." *Scientific American*.

D. W. Ruck, S. K. Rogers and M. Kabrisky (1990a). "Feature selection using a multilayer perceptron." *Journal of Neural Network Computing* **2**(2): 40–48.

D. W. Ruck, S. K. Rogers, M. Kabrisky, M. E. Oxley and B. S. Suter (1990b). "The Multilayer Perceptron as an Approximation to a Bayes Optimal Discriminant Function." *IEEE Transactions on Neural Networks* **1**(4): 296–298.

D. E. Rumelhart, G. E. Hinton and R. J. Williams (1986). Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructures of Cognition. 1: Foundations*. D. E. Rumelhart and J. L. McClelland. Cambridge, MA, MIT Press. **1**: 318–362.

T. D. Sanger (1989a). An optimality principle for unsupervised learning. *Advances in Neural Information Processing Systems*. D. S. Touretzky. San Mateo, CA, Morgan Kaufmann Publishers. **1**: A–19.

T. D. Sanger (1989b). "Optimal unsupervised learning in a single-layer linear feedforward neural network." *Neural Networks* **2**(6-7): 459–473.

T. J. Sejnowski and C. R. Rosenberg (1987). "Parallel Networks that Learn to Pronounce English Text." *Complex Systems* **1**: 145–168.

J. Shao (1993). "Linear model selection by cross-validation." *Journal of the American Statistical Association* **88**: 486–494.

J. Shao (1997). "An asymptotic theory for linear model selection." *Statistica Sinica* **7**: 221–264.

J. Shao and D. Tu (1995). *The Jackknife and Bootstrap*. Berlin Heidelberg London, Springer-Verlag.

T. A. B. Snijders (1988). On cross-validation for predictor evaluation in time series. *On Model Uncertainty and Its Statistical Implications*. T. K. Dijkstra. Berlin, Springer-Verlag: 56–69.

F. Stäger and M. Agarwal (1997). "Three methods to speed up the training of feedforward and feedback perceptrons." *Neural Networks* **10**(8): 1435–1443.

M. Stone (1977). "Asymptotics for and against cross-validation." *Biometrika* **64**: 29–35.

M. Stone (1979). "Comments on model selection criteria of Akaike and Schwarz." *Journal of the Royal Statistical Society, Series B* **41**: 276–278.

W. S. Stornetta and B. A. Huberman (1987). An Improved Three-Layer, Back Propagation Algorithm. *Proceedings of the IEEE Conference on Neural Networks (ICNN'87)*. Piscataway, NJ, IEEE. **2**: 637–644.

Y. Suzuki (1995). "Self-Organizing QRS-Wave Recognition in ECG Using Neural Networks." *IEEE Transactions on Neural Networks* **6**(6): 1469–1477.

D. W. Tank and J. J. Hopfield (1986). "Simple Neural Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit." *IEEE Transactions on Circuits and Systems* **33**(5): 533–541.

G. Tarr, K. Priddy and S. Rogers (1992). "NeuralGraphics: A general purpose environment for neural network simulation," *Applications of Artificial Neural Networks III, Proceedings of SPIE.* S. K. Rogers Bellingham, WA, SPIE. **1709**: 1047–1056.

R. Tibshirani (1996). "A comparison of some error estimates for neural network models." *Neural Computation* **8**: 152–163.

J. W. Tukey (1958). "Bias and Confidence in Not-Quite Large Samples." *Annals of Mathematical Statistics* **29**: 614.

V. N. Vapnik (1995). *The Nature of Statistical Learning Theory*. Berlin Heidelberg London, Springer-Verlag.

B. Verma (1997). "Fast Training of Multilayer Perceptrons." *IEEE Transactions on Neural Networks* **8**(6): 1314–1320.

R. Vitthal, P. Sunthar and C. D. Rao (1995). "The Generalized Proportional-Integral-Derivative (PID) Gradient Descent Back Propagation Algorithm." *Neural Networks* **8**(4): 563–569.

S. Watanabe and K. Fukumizu (1995). "Probabilistic Design of Layered Neural Networks Based on Their Unified Framework." *IEEE Transactions on Neural Networks* **6**(3): 691–702.

S. M. Weiss and C. A. Kulikowski (1991). *Computer Systems That Learn*. San Mateo, CA, Morgan Kaufmann Publishers.

P. J. Werbos (1974). "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences." Ph.D. Thesis, Cambridge, MA.

P. J. Werbos (1994). *The Roots of Backpropagation*. New York, John Wiley & Sons.

B. Widrow (1962). Generalization and information Storage in Networks of Adaline Neurons. *Self-Organizing Systems*. M. C. Yovits, G. T. Jacobi and G. D. Goldstein. Washington, D.C., Spartan Books.

B. Widrow and M. E. Hoff, Jr. (1960). Adaptive Switching Circuits. *1960 IRE WESCON Convention Record, Part 4*. New York, Institute of Radio Engineers: 96–104.

B. Widrow and M. A. Lehr (1990). "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation." *Proceedings of the IEEE* **78**: 1415–1441.

A. L. Wilkes and N. J. Wade (1997). "Bain on Neural Networks." *Brain and Cognition* **33**: 295–305.

R. J. Williams and D. Zipser (1989). "A learning algorithm for continually running fully recurrent neural networks." *Neural Computation* **1**(2): 270–280.

R. J. Williams and D. Zipser (1990). Gradient-based learning algorithms for recurrent connectionist networks. *Technical Report NU-CCS-90-9*. Boston, Northeastern University, College of Computer Science.

R. J. Williams and D. Zipser (1995). Gradient-Based Learning Algorithms for Recurrent Networks and Their Computational Complexity. *Backpropagation:*

Theory, Architectures, and Applications. Y. Chauvin and D. E. Rumelhart. Hillsdale, NJ, Lawrence Erlbaum Publishers: 433–486.

J. R. Williamson (1996). "Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps." *Neural Networks* **9**(5): 881–897.

G. V. Wilson and G. S. Pawley (1988). "On the stability of the traveling salesman problem algorithm of Hopfield and Tank." *Biological Cybernetics* **58**(1): 63–70.

Y. Zheng and J. F. Greenleaf (1996). "The Effect of Concave and Convex Weight Adjustments on Self-Organizing Maps." *IEEE Transactions on Neural Networks* **7**(1): 87–96.

H. Zhu and R. Rohwer (1996). "No free lunch for cross-validation." *Neural Computation* **8**(7): 1421–1426.

Index

About the Authors

Paul E. Keller received a B.S. degree in physics from Boise State University in 1985, an M.S. degree in electro-optics from the University of Dayton in 1987, and a Ph.D. in optical sciences from the University of Arizona in 1991. He has been a scientist with Battelle at the Pacific Northwest National Laboratory since 1992. His research areas include optics, photonics, neural networks, sensor data analysis, image processing, and computational physics. He has worked on applications in homeland security, nuclear nonproliferation, national defense, medicine, environmental technologies, telecommunications, computing, and energy distribution. He has taught several short courses on neural networks and their applications and also served as an adjunct faculty at Washington State University in the area of neural networks. He has authored and co-authored more than 60 conference papers, journal articles, and book chapters on the topics of neural networks, medical technologies, optical computing, and Internet technologies. He holds three U.S. patents, with several pending.

Kevin L. Priddy received a B.S. degree in electrical engineering from Brigham Young University in 1982, an M.S. degree in electro-optics from the Air Force Institute of Technology in 1985, and a Ph.D. in electrical engineering from the Air Force Institute of Technology in 1992. Dr. Priddy is currently a team leader in the ATR & Fusion Algorithms Branch, Sensors Directorate, of the Air Force Research Laboratory. Dr. Priddy has over 40 conference papers and journal articles in the fields of artificial neural networks and electrical engineering. He has been awarded two patents and has others pending in the computational intelligence field. His research interests include computational intelligence, automated recognition systems, intelligent sensors, image processing and signal processing.