

References

- [1] G.G. Towell, J.W. Shavlik, Knowledge-Based Artificial Neural Networks, *Artificial Intelligence* 70(1-2), 119-165, 1994.
- [2] M. Golea, On the complexity of rule extraction from neural networks and network querying, In *Proceedings of the AISB'96 Workshop on Rule Extraction from Trained Neural Networks*, Brighton, UK, 1996, pp.51-59.
- [3] M. Craven, J. Shavlik, Learning Symbolic Rules Using Artificial Neural Networks, *Proceedings of the 10th International Conference on Machine Learning*, pp. 73-80, Amherst, MA. Morgan Kaufmann, 1993.
- [4] M. Craven, J. Shavlik, Using Sampling and Queries to Extract Rules from Trained Neural Networks, *Proceedings of the 11th International Conference on Machine Learning*, pp. 37-45, New Brunswick, NJ. Morgan Kaufmann, 1994.
- [5] R. Andrews, S. Geva, Inserting and extracting knowledge from constrained error back propagation networks , In *Proc. 6th Australian Conference on Neural Networks*, Sydney, NSW, 1995.
- [6] R. Andrews, S. Geva, On the effects of initialising a neural network with prior knowledge, In *Proceedings of the International Conference on Neural Information Processing (ICONIP'99)*, pages 251-256, Perth, Western Australia, 1999.
- [7] W. Duch, R. Adamczak, K. Grabczewski, A new methodology of extraction, optimization and application of crisp and fuzzy logical rules, *IEEE Transactions on Neural Networks* 12:277-306, 2001.
- [8] B. Hammer, A. Rechten, M. Strickert, T. Villmann, Rule extraction from self-organizing maps, in: J.R.Dorransoro (Ed.), *Artificial Neural Networks – ICANN 2002*, Springer, 877-882, 2002.
- [9] K. Saito, R. Nakano, Medical Diagnostic Expert Systems based on PDP model, in: *Proc. IEEE International Conference on Neural Networks*, San Diego, CA, vol.1, pp.255-2562, 1988.
- [10] S. Gallant, Connectionist Expert Systems, *Communications of the ACM*, 31(2):152-169, 1988.
- [11] S. Thrun, *Extracting Provably Correct Rules from Artificial Neural Networks*, Technical Report IAI-TR-93-5, Institut für Informatik, Universität Bonn, 1994.
- [12] A.B. Tickle, M. Orłowski, J. Diederich, *DEDEC: Decision Detection by Rule Extraction from Neural Networks*, QUT NRC, 1994.

- [13] U. Markowska-Kaczmar, W. Trelak, Fuzzy Logic and Evolutionary Algorithm - two Techniques in Rule Extraction from Neural Networks, *Neuro-computing* 63: 359-380, 2005.
- [14] D. Nauck, F. Klawonn, R. Kruse, Foundations of Neuro-Fuzzy Systems, Wiley, NY, USA, 1997.
- [15] J.-S.R. Jang, ANFIS: Adaptive-Network-Based Fuzzy Inference Systems, *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665-685, 1993.
- [16] A. Ultsch., Knowledge Extraction from Self-Organizing Neural Networks, In O. Opitz, B. Lausen, and R. Klar, editors, *Information and Classification*, pages 301–306, London, UK, 1993, Springer.
- [17] D. Haussler, *Convolutional Kernels on Discrete Structures*, Technical Report, Department of Computer Science, University of California at Santa Cruz, 1999.
- [18] C. Watkins, *Dynamic Alignment Kernels*. Technical Report, Department of Computer Science, Royal Holloway, University of London, 1999.
- [19] T. Gärtner, P.A. Flach, S. Wrobel, On Graph Kernels: Hardness Results and Efficient Alternatives. *Sixteenth Annual Conference on Computational Learning Theory and Seventh Kernel Workshop (COLT-2003)*, Springer, 2003.
- [20] T. Gärtner, A Survey of Kernels for Structured Data, *SIGKDD Explorations*, 2003.
- [21] S. Sonnenburg, G. Rätsch, A. Jagota, K.-R. Müller, New Methods for Splice Site Recognition, *ICANN* 2002.
- [22] T. Joachims, Text Categorization with Support Vector Machines: Learning with Many Relevant Features, *Proceedings of the European Conference on Machine Learning (ECML)*, Springer, 1998.
- [23] H. Lodhi, J. Shawe-Taylor, N. Cristianini, C. Watkins, Text Classification using String Kernels, *Journal of Machine Learning Research* 2:419-444, 2002.
- [24] C. Leslie, E. Eskin, J. Weston, W.S. Noble, Mismatch String Kernels for SVM Protein Classification, *NIPS*, 2002.
- [25] N. Cancedda, E. Gaussier, C. Goutte, J.-M. Renders, Word-sequence Kernels, *Journal of Machine Learning Research* 3:1059-1082, 2003.
- [26] M. Collins, N. Duffy, Convolutional Kernels for Natural Languages, *NIPS'2001*.

- [27] H. Kashima, T. Koyanagi, Kernels for Semi-Structured Data, In *Proc. 19th International Conference on Machine Learning (ICML2002)*, Sydney, Australia, 2002.
- [28] J. Suzuki, Y. Sasaki, E. Maeda, Kernels for Structured Natural Language Data, *NIPS'2003*.
- [29] D. Zelenko, C. Aone, Kernel Methods for Relation Extraction, *Journal of Machine Learning Research*, submitted, 2003.
- [30] C. Cumby, D. Roth, On Kernel Methods for Relational Learning, *ICML'2003*.
- [31] T. Gärtner, K. Driessens, J. Ramon, Graph Kernels and Gaussian Processes for Relational Reinforcement Learning *Thirteenth International Conference on Inductive Logic Programming (ILP-2003)*, Springer, 2003.
- [32] T. Gärtner, Predictive Graph Mining with Kernel Methods, accepted for *Advanced Methods for Knowledge Discovery from Complex Data*.
- [33] H. Kashima, K. Tsuda, A. Inokuchi, Marginalized Kernels between Labeled Graphs, *ICML'2003*.
- [34] A. Passerini, P. Frasconi, L. De Raedt, Kernels on Prolog Proof Trees: Statistical Learning in the ILP Setting, In *ICML '05 Workshop on Approaches and Applications of Inductive Programming*, 2005.
- [35] T. Jaakkola, M. Diekhans, D. Haussler, A Discriminative Framework for Detecting Remote Protein Homologies, *Journal of Computational Biology* 7(1/2), 2000.
- [36] R. Karchin, *Classifying G-Protein Coupled Receptors with Support Vector Machines*, Master's Thesis, 2000.
- [37] Paul Pavlidis, Terrence S. Furey, Muriel Liberto, David Haussler, William Noble Grundy, Promoter region-based classification of genes, *Proceedings of the Pacific Symposium on Biocomputing*, January 3-7, 2001. pp. 151-163.
- [38] Nathan Smith, Mark Gales, Speech Recognition using SVMs , In T.G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems 14*. MIT Press, 2002.
- [39] George Siolas, Florence d'Alché-Buc: Mixtures of Probabilistic PCAs and Fisher Kernels for Word and Document Modeling. *ICANN 2002*: 769-776.
- [40] Schölkopf, B., K. Tsuda, J. Vert, *Kernel Methods in Computational Biology*, MIT Press, 2004.

- [41] K. Tsuda, M. Kawanabe, G. Rätsch, S. Sonnenburg, K.R. Müller, A New Discriminative Kernel from Probabilistic Models, *Neural Computation* 14(10):2397-2414, 2002.
- [42] K. Tsuda, T. Kin, K. Asai, Marginalized kernels for biological sequences, *Bioinformatics* 18, S1, 268-275, 2002.
- [43] Moreno, Pedro J.; Ho, Purdy P, *A New SVM Approach to Speaker Identification and Verification Using Probabilistic Distance Kernels*, Technical Report HPL-2004-7 , HP Labs.
- [44] T. Jebara, R. Kondor and A. Howard. Probability Product Kernels. *Journal of Machine Learning Research*, JMLR, Special Topic on Learning Theory, 5(Jul):819-844, 2004
- [45] Risi Imre Kondor, John D. Lafferty, Diffusion Kernels on Graphs and Other Discrete Input Spaces, *Proceedings of the Nineteenth International Conference on Machine Learning*, 315 - 322, 2002.
- [46] John D. Lafferty, Guy Lebanon, Information Diffusion Kernels. *NIPS* 2002: 375-382
- [47] J.-P. Vert and M. Kanehisa, Graph-driven features extraction from microarray data using diffusion kernels and kernel CCA, *Advances in Neural Information Processing Systems* 15 (NIPS 2002), Suzanna Becker, Sebastian Thrun and Klaus Obermayer (Eds), p. 1425-1432, MIT Press, Cambridge, MA, 2003.
- [48] J.-P. Vert, R. Thurman and W. S. Noble, Kernels for gene regulatory regions, *Advances in Neural Information Processing Systems* 18 (NIPS 2005), 2006.
- [49] L. Ralaivola, J. S. Swamidass, H. Saigo, and P. Baldi. Graph Kernels for Chemical Informatics. *Neural Networks*, special issue on Neural Networks and Kernel Methods for Structured Domains, in press, (2005).
- [50] S.C.Kremer, Spatiotemporal connectionist networks: a taxonomy and review, *Neural Computation* 13:249-306, 2001.
- [51] Barak A. Pearlmutter, Gradient Calculations for Dynamic Recurrent Neural Networks: A Survey, *IEEE Transactions on Neural Networks*, 6(5): 1212–1228, 1995.
- [52] Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- [53] Christian W. Omlin, C.L. Giles, Rule Revision with Recurrent Neural Networks, *Knowledge and Data Engineering* 8(1): 183-188, 1996.
- [54] Hava T. Siegelmann, Eduardo D. Sontag, Analog Computation Via Neural Networks, *Theoretical Computer Science* 131(2):331–360, 1994.

- [55] H. T. Siegelmann, E. D. Sontag, Turing Computability with Neural Nets, *Applied Mathematics Letters* 4(6): 77-80, 1991.
- [56] W. Maass, and P. Orponen (1998) On the effect of analog noise in discrete-time analog computation. *Neural Computation* 10(5):1071-1095.
- [57] Wolfgang Maass, Eduardo D. Sontag, Analog neural nets with Gaussian or other common noise distributions cannot recognize arbitrary regular languages, *Neural Computation* 11(3): 771 - 782, 1999.
- [58] B. Hammer, P. Tino, Recurrent neural networks with small weights implement definite memory machines, *Neural Computation* 15(8), 1897-1929, 2003.
- [59] Yoshua Bengio, Patrice Simard, Paolo Frasconi, Learning Long-Term Dependencies with Gradient Descent is Difficult, *IEEE Transactions on Neural Networks*, 5(2):157-166, 1994.
- [60] B. Hammer, Generalization ability of folding networks, *IEEE Transactions on Knowledge and Data Engineering*, 13(2): 196-206, 2001.
- [61] J.J. Steil, H. Ritter, Maximization of stability ranges for RNNs subject to on-line adaptation, *ESANN'99*, 73-79, 1999.
- [62] J.Suykens, B. de Moor, J. Vandewalle, Robust local stability of multilayer neural networks, *IEEE Transactions on Neural Networks* 11(1): 222-229, 2000.
- [63] Sepp Hochreiter, Jurgen Schmidhuber, Long Short-Term Memory, *Neural Computation*, Vol 9, 1735-1780, 1997.
- [64] M. Boden, Generalization by symbolic abstraction in cascaded recurrent networks, *Neurocomputing*, 57, pp. 87-104. 2004.
- [65] Nigel Crook, Tjeerd Olde Scheper: Adaptation Based on Memory Dynamics in a Chaotic Neural Network. *Cybernetics and Systems* 33(4): 341-378 (2002).
- [66] Martin Hülse, Steffen Wischmann, Frank Pasemann: Structure and function of evolved neuro-controllers for autonomous robots. *Connect. Sci.* 16(4): 249-266 (2004).
- [67] Robert Haschke, Jochen J. Steil, Input Space Bifurcation Manifolds of Recurrent Neural Networks, *Neurocomputing*, 64C, 25-38, 2005.
- [68] Rodriguez, P., Wiles, J., Elman, J. A Recurrent Neural Network That Learns to Count, *Connection Science*, Vol. 11, No. 1, 5-40, 1999.
- [69] P. Tino, B.G. Horne, C.L. Giles: Attractive Periodic Sets in Discrete Time Recurrent Networks (with Emphasis on Fixed Point Stability and Bifurcations in Two-Neuron Networks). *Neural Computation*, 13(6), pp. 1379-1414, 2001.

- [70] P.Baldi, S.Brunak, P.Frasconi, G.Pollastri, and G.Soda. Bidirectional Dynamics for Protein Secondary Structure Prediction, in *Sequence Learning: Paradigms, Algorithms, and Applications*, R. Sun and L. Giles Editors, Springer Verlag, (2000).
- [71] A. Micheli, D. Sona, A. Sperduti, Contextual Processing of Structured Data by Recursive Cascade Correlation. *IEEE Transactions on Neural Networks*. Vol. 15, n. 6, Pages 1396-1410, November 2004.
- [72] B. Hammer, A. Micheli, A. Sperduti, M. Strickert. Recursive Self-organizing Network Models. *Neural Networks*, Elsevier Science. Volume 17, Issues 8-9 , October-November 2004, Pages 1061-1085.
- [73] M.Strickert, B.Hammer, Merge SOM for temporal data, *Neurocomputing* 64:39-72, 2005.
- [74] T. Voegtlin. Recursive self-organizing maps. *Neural Networks*, 15(8/9):979-991, 2002.
- [75] M. Hagenbuchner, A. Sperduti, and A. Tsoi. A self-organizing map for adaptive processing of structured data. *IEEE Transactions on Neural Networks*, 14(3):491-505, 2003.
- [76] G. Chappell and J. Taylor. The temporal Kohonen map. *Neural Networks*, 6:441-445, 1993.
- [77] Hinton, G. (1988). Representing Part-Whole hierarchies in connectionist networks. In *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*. Montreal, 48-54.
- [78] Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence*, 46:159-216.
- [79] Pollack, J. B. (1988). *Recursive Auto-Associative Memory: Devising Compositional Distributed Representations*. MCCS-88-124, Las Cruces: Computing Research Laboratory, New Mexico State University.
- [80] Pentti Kanerva, Binary Spatter-Coding of Ordered K-Tuples, *ICANN*, 869-873, 1996.
- [81] A. Sperduti. Labeling RAAM. *Connection Science*, 6(4), 1994.
- [82] B. Hammer, Recurrent networks for structured data - a unifying approach and its properties, *Cognitive Systems Research* 3(2), 145-165, 2002.
- [83] Andreas Kuchler and Christoph Goller. Inductive Learning in Symbolic Domains Using Structure-Driven Recurrent Neural Networks. In Günther Görz and Steffen Hölldobler, editors, *KI-96: Advances in Artificial Intelligence*, Lecture Notes in Computer Science (LNCS 1137), pages 183-197, Berlin, 1996. Springer

- [84] Marco Gori, Andreas Küchler, and Alessandro Sperduti. On the Implementation of Frontier-to-Root Tree Automata in Recursive Neural Networks. *IEEE Transactions on Neural Networks*, 10(6):1305-1314, 1999.
- [85] Stephan Schulz, Andreas Küchler, and Christoph Goller. Some Experiments on the Applicability of Folding Architecture Networks to Guide Theorem Proving. In D.D. Dankel II, editor, *Proc. of the 10th FLAIRS Conference*, Daytona Beach, pages 377-381. Florida AI Research Society, 1997.
- [86] S. Schulz, A. Küchler, C. Goller, Some Experiments on the Applicability of Folding Architecture Networks to Guide Theorem Proving, *Proc. of the 10th FLAIRS, Daytona Beach*, Florida AI Research Society, 377-381, 1997.
- [87] Sturt, P., Costa, F., Lombardo, V., and Frasconi, P. (2003). Learning first-pass attachment preferences with dynamic grammars and recursive neural networks. *Cognition*, 88(2):133-169
- [88] Costa, F., Frasconi, P., Lombardo, V., Sturt, P., and Soda, G. (2005). Ambiguity resolution in incremental parsing of natural language. *IEEE Transactions on Neural Networks*.
- [89] Vullo, A. and Frasconi, P. (2004). Disulfide Connectivity Prediction Using Recursive Neural Networks and Multiple Alignments. *Bioinformatics*, 20(5):653-9
- [90] Diligenti, M., Frasconi, P., and Gori, M. (2003). Hidden Tree Markov Models for Image Document Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4):519-523
- [91] Vullo, A. and Frasconi, P. (2003). A Recursive Connectionist Approach for Predicting Disulfide Connectivity in Proteins. In *Eighteenth Annual ACM Symposium on Applied Computing (SAC '03)*, pages 67-71. Special Track on Bioinformatics.
- [92] Vullo, A. and Frasconi, P. (2003). Prediction of Protein Coarse Contact Maps. *Journal of Bioinformatics and Computational Biology*, 1(2):411-431
- [93] Diligenti, M., Frasconi, P., and Gori, M. (2001). Image Document Categorization using Hidden Tree-Markov Models and Structured Representations. In *Second International Conference on Advances in Pattern Recognition*, volume 2013 of Lecture Notes in Computer Science, pages 147-156. Springer.
- [94] Ceroni, A., Frasconi, P., and Pollastri, G. (2005). Learning Protein Secondary Structure from Sequential and Relational Data. *Neural Networks*. to appear.

- [95] Marcialis, G. L., Roli, F., and Frasconi, P. (2001). Fingerprint Classification by Combination of Flat and Structural Approaches. In *3rd Int. Conf. on Audio- and Video-Based Biometric Person Authentication*, volume 2091 of Lecture Notes in Computer Science, pages 241-246. Springer
- [96] Vullo, A. and Frasconi, P. (2003). A Recursive Connectionist Approach for Predicting Disulfide Connectivity in Proteins. In *Eighteenth Annual ACM Symposium on Applied Computing (SAC '03)*, pages 67-71
- [97] M. Bianchini, P. Mazzoni, L. Sarti, and F. Scarselli, "Face Spotting in Color Images using Recursive Neural Networks". Accepted to *IAPR - TC3 International Workshop on Artificial Neural Networks in Pattern Recognition*, Firenze (ITALY), September, 12-13, 2003.
- [98] M. Bianchini, M. Gori, L. Sarti, F. Scarselli, "Recursive Processing of Cyclic Graphs", *IEEE Transactions on Neural Networks*, 2005, to appear.
- [99] M. Diligenti, M. Gori, M. Maggini, and E. Martinelli, "Adaptive graphical pattern recognition for the classification of company logos", *Pattern Recognition*, vol. 34, pp. 2049-2061, 2001
- [100] B. Hammer, A. Micheli, A. Sperduti. Universal Approximation Capability of Cascade Correlation for Structures. *Neural Computation*. Vol. 17, No. 5, Pages 1109-1159, (C) 2005 MIT press. ISSN 0899-7667
- [101] A. Micheli, A. Sperduti, A. Starita, A.M. Bianucci. A Novel Approach to QSPR/QSAR Based on Neural Networks for Structures, Chapter in Book : *Soft Computing Approaches in Chemistry*, pp. 265-296, H. Cartwright, L. M. Sztandera, Eds., Springer-Verlag, Heidelberg, March 2003. ISBN 3-540-00245-6.
- [102] A.M. Bianucci, A. Micheli, A. Sperduti, A. Starita. Application of Cascade Correlation Networks for Structures to Chemistry, *Applied Intelligence Journal* (Kluwer Academic Publishers), Special Issue on "Neural Networks and Structured Knowledge" Vol. 12 (1/2): 117-146, 2000
- [103] A. Micheli, D. Sona, A. Sperduti, Contextual Processing of Structured Data by Recursive Cascade Correlation. *IEEE Transactions on Neural Networks*. Vol. 15, n. 6, Pages 1396- 1410, November 2004
- [104] B. Hammer, *Learning with Recurrent Neural Networks*, Springer Lecture Notes in Control and Information Sciences 254, Springer-Verlag, 2000 (see <http://www.in.tu-clausthal.de/hammer/papers/> for an online version)
- [105] Duch, Setiono, Zurada, Computational intelligence methods for understanding of data, *Proc. of the IEEE* 92(5):771- 805, 2004
- [106] Gärtner, Lloyd, Flach, Kernels and distances for structured data, *Machine Learning*, 57, 2004.

- [107] Hammer, Steil, Perspectives on learning with recurrent networks, in: Verleysen, *ESANN'2002*, D-side publications, 357-368, 2002
- [108] Jacobsson, Rule extraction from recurrent neural networks: a taxonomy and review, *Neural Computation*, 17:1223-1263, 2005
- [109] Hammer, Perspectives on Learning Symbolic Data with Connectionistic Systems, in: Kühn, Menzel, Menzel, Ratsch, Richter, Stamatescu, *Adaptivity and Learning*, 141-160, Springer, 2003.
- [110] Frasconi, Gori, Sperduti, A General Framework for Adaptive Processing of Data Structures, *IEEE Transactions on Neural Networks*, 9(5):768-786,1998
- [111] Hammer, Jain, Neural methods for non-standard data, in: Verleysen, *ESANN'2004*, D-side publications, 281-292, 2004