Inductive Modelling World Wide the State of the Art

Miroslav Šnorek, Pavel Kordík

Dept. of Computer Science and Engineering, Karlovo nam. 13, 121 35 Praha 2, Czech Rep.

snorek@fel.cvut.cz, kordikp@fel.cvut.cz

Abstract. *In this contribution, we summarize the state of the art of the inductive modelling world wide. Recently, there is a trend to utilize evolutionary algorithms for optimization of inductive models. Also some new approaches in feature selection using inductive modelling, model validation, ensembles of inductive models, etc. are to be described.*

Keywords

Inductive modeling, state of the art.

1 Introduction

The capability of induction is fundamental for human thinking. It is the next human ability that can be utilized in soft-computing, besides that of learning and generalization. The induction means gathering small pieces of information, combining it, using already collected information in the higher abstraction level to get complex overview of the studied object or process. Inductive modeling methods utilize the process of induction to construct models of studied systems. The construction process is highly efficient, it starts from the minimal form and the model grows according to system complexity. It also works well for systems with many inputs. Where the traditional modeling methods fail, due to the "curse of dimensionality" phenomenon, the inductive methods are capable to build reliable models. The problem is decomposed into small subtasks. At first, the information from most important inputs is analyzed in the subspace of low dimensionality, later the abstracted information is combined to get a global knowledge of the system variables relationship.

There are several methods for inductive models construction commonly known as Group Method of Data Handling (GMDH) introduced by Ukrainian scientist Ivachknenko in 1966 [1]. The GMDH theory or polynomial networks are called Statistical Learning Networks [2] in the United States of America. They were developed more or less independently. In general, in inductive modelling and more specifically the GMDH theory models are generated from data.

This contribution maps the state of the art of the inductive modeling world wide.

2 The state of the art in the GMDH related research

Polynomial Neural Networks (PNN) [4] are also GMDH type networks. The units called partial descriptions having different transfer function of polynomial type [4] are evolved by a genetic algorithm (GA).

In [5] the structural optimization of the fuzzy polynomial neural network (FPNN) is realized via standard GA whereas in the case of the parametric optimization a standard least square method based learning is used. The article [6] uses GA to optimize the structure of original MIA GMDH neural network whereas the coefficients are solved by the Singular Value Decomposition (SVD) method.

The hybrid architecture of the network is employed in the polynomial harmonic GMDH (phGMDH) [7], where the harmonic inputs are passed to a polynomial network whose architecture is built using the MIA GMDH algorithm.

The Group of Adaptive Models Evolution (GAME) [3] uses neurons (units) with several possible types of transfer function (linear, polynomial, sigmoid, harmonic, preceptron net, etc.). The structure of models is optimized layer by layer by niching genetic algorithm [24]. Ensemble [26] of models is generated to be able to estimate the bounds of output confidence.

A novel algorithm based on GMDH for designing MLP neural networks can be found in [8]. This idea is similar to the one presented in [9], where Cascade Correlation Networks are enhanced by the GMDH.

In [7], the iterative gradient descent training algorithm is offered for improving the performance of polynomial neural networks. The Back-Propagation algorithm is derived for multilayered networks with polynomial activation functions. We believe that if "powerful enough" optimization techniques are used during the construction stage, it is not necessary to readjust parameters after the polynomial network is built. This readjustment of parameters might suggest they were not set optimally.

The AIC and PSS criterion are used in revised GMDH-type algorithm [10] to find the optimal number of neurons and layers of the GMDH networks. Such regularization takes into account just the complexity of GMDH network. Outputs from neurons in a layer can be highly correlated resulting to a redundant GMDH network.

The recent article [11] introduces a GMDH-based feature ranking and selection algorithm. This algorithm builds GMDH networks of gradual complexity, rewarding features selected by smaller networks. In this thesis we propose three different algorithms of feature ranking that can also supply the proportional significance of features.

The idea of Twice Multilayered GMDH networks with active neurons (neuronet) was originally published in [29]. It is shown that any learning forecasting or pattern recognition algorithm, having self-organizing abilities can be used as an active neuron in twice-multilayered neural network, which has self-organizing abilities too [27, 28]. This idea is very close to ensembling, where diverse models are combined to get better performance. A neuronet will be efficient only if active neurons are weak learners and if their diversity is ensured (they demonstrate diverse errors).

3 Connection to neural networks and genetic algorithms

Neural networks are closely connected to the inductive modelling, although they have different background. The GMDH evolved from the mathematical description of a system by means of Kolmogorov-Gabor polynomial [12]. Neural networks were at the beginning biologically oriented. Later, powerful optimization methods for neural networks (Back-Propagation of error) were invented, allowing building multilayered networks of neurons (MLP) capable of solving nonlinear problems. It has been shown [13] that MLPs are equivalent to mathematical description of a system by means of the Koglomorov theorem (although inner functions are very complex and they have almost fractal character).

In recent time both neural networks and GMDH algorithms are optimized by genetic algorithms and it is even harder to distinguish the boundary between both theories. Of course, some neural networks are very different from GMDH (recurrent, modular, spiking neural networks, etc.) [14].

The article [15] shows that the problem of two intertwined spirals can be successfully solved by the MultiLayered Perceptron (MLP), where weights are evolved by the Genetic Algorithm. Number of function evaluation is in this case much higher than when using standard BackPropagation, because the information about gradient of error is not utilized in the GA. On the other hand, in some applications, Genetic approach gives better results than BackPropagation [16].

The Cascade Correlation algorithm [17] is capable of solving extremely difficult problems. It performs optimally on "spiral" benchmarking problem (a network consisting of less than 20 neurons is generated). According to experiments on real-world data performed in [18], the algorithm has difficulties with avoiding premature convergence to complex topological structures. The main advantage of the Cascade Correlation algorithm is also its main disadvantage. It easily solves extremely difficult problems therefore it is likely to overfit.

The recent article [19] proposes an improvement of the Cascade Correlation Algorithm [17]. The original algorithm assumes fully connected network. Each neuron is connected to all features and all previously built neurons. The improvement called Evolving Cascade Correlation Networks (ECCN) [19] uses techniques from GMDH theory [1] to choose just relevant inputs for each neuron. Cascade networks evolved by ECCN overfit data less than fully connected cascade networks.

Recently, very interesting algorithm for designing recurrent neural networks was proposed in [20, 21]. The NeuroEvolution Through Augmenting Topologies (NEAT) is designed for solving reinforcement learning tasks [22], but can be applied also to supervised learning problems.
Similarly to the GMDH, NEAT networks grow from a minimal structure up to optimal complexity. The topology and also weights of the NEAT networks are evolved using niching genetic algorithm [23].

The NEAT is primary designed for reinforcement learning. It can also solve standard supervised problems. We have applied the NEAT to Two intertwined spiral problem [24], but it failed. The reason why NEAT is unable to evolve successful networks solving the spiral problem is probably that a) the chromosome is too big when evolving architecture and weights of network simultaneously, b) niching alone is unable to protect complex structures.

4 Conclusion

In this paper we presented the state of the art of the GMDH related research. With recent application of evolutionary computation methods the boundary between GMDH and modern neural networks almost disappears. However some differences still exist.

Most of the GMDH algorithms tend to avoid the black box approach with the ability to extract comprehensible math formulae describing modeled systems.

The advantage of inductive modeling is the ability to deal with irrelevant inputs allowing building credible models even for high dimensional short data samples.

5 Acknowledgements

This research is partially supported by the grant Automated Knowledge Extraction (KJB201210701) of the Grant Agency of the Academy of Science of the Czech Republic and the research program "Transdisciplinary Research in the Area of Biomedical Engineering II" (MSM6840770012) sponsored by the Ministry of Education, Youth and Sports of the Czech Republic.

References

[1] Ivakhnenko A. G.: Polynomial theory of complex systems. IEEE Transactions on Systems, Man, and Cybernetics, SMC-1(1):364378, 1971.

[2] A. Barron and R. Baron, “Statistical learning networks: A unifying. view,” in Proc. 20th Symp. Interface, 1988, pp. 192-203.

[3] P. Kordik. Fully Automated Knowledge Extraction using Group of Adaptive Models Evolution. PhD thesis, Dep. of Comp. Sci. and Computers, FEE, CTU Prague, Czech Republic, September 2006.

[4] Oh, S.; Pedrycz, W. & Park, B. (2003), 'Polynomial neural networks architecture: analysis and design', Computers and Electrical Engineering 29, 703-725.

[5] Oh, S.K. & Pedrycz, W. (2002), 'The design of self-organizing Polynomial Neural Networks', Inf. Sci. 141, 237-258.

[6] Oh, S. & Pedrycz, W. (2005), 'A new approach to self-organizing fuzzy polynomial neural networks guided by genetic optimization', Physics Letters A(345), 88-100.

[7] Nariman-Zadeh, N.; Darvizeh, A.; Jamali, A. & Moeini, A. (2005), 'Evolutionary design of generalized polynomial neural networks for modelling and prediction of explosive forming process', Journal of Materials Processing Technology(165), 1561-1571.

[8] Nikolaev, N.Y. & Iba, H. (2003), 'Polynomial harmonic GMDH learning networks for time series modeling', Neural Networks(16), 1527-1540.
[9] Gilbar & Thomas, C. (2002),'A new GMDH type algorithm for the development of neural networks for pattern recognition', PhD thesis, FLORIDA ATLANTIC UNIVERSITY.
[10] Schetinin, V. (2003), 'A Learning Algorithm for Evolving Cascade Neural Networks', Neural Processing Letters(17), 21-31.
[11] Kondo, T.; Ueno, J. & Kondo, K. (2005), 'Revised GMDH-Type Neural Networks Using AIC or PSS Criterion and Their Application to Medical Image Recognition', Journal of Advanced Computational Intelligence and Intelligent Informatics 9(3), 257-267.
[12] Abdel-Aal, R. (2005), 'GMDH-based feature ranking and selection for improved classification of medical data', Journal of Biomedical Informatics 38, 456-468.
[13] Madala, H.R. & Ivakhnenko, A.Ration, B., ed. (1994), Inductive Learning Algorithm for Complex System Modelling, CRC Press.
[14] Kurkova, V. (1991), 'Kolmogorov's Theorem Is Relevant', Neural Computation 3, 617-622.
[15] Koutník, J. (2004), 'Modular Neural networks for Analysis and Recognition of Real Data'(DC-PSR-2004-08), Technical report, Czech Technical University in Prague, FEE, CTU Prague, Czech Republic.
[16] Mandischer, M. (2002), 'A comparison of evolution strategies and backpropagation for neural network training', Neurocomputing(42), 87-117.
[17] Sexton, R.S. & Gupta, J.N.D. (2000), 'Comparative evaluation of genetic algorithm and backpropagation for training neural networks', Information Sciences(129), 45-59.
[18] Fahlman, S.E. & Lebiere, C. (1991),'The Cascade-Correlation Learning Architecture'(CMU-CS-90-100), Technical report, Carnegie Mellon University Pittsburgh, USA.
[19] Wickera, D.; Rizkib, M.M. & Tamburinoa, L.A. (2002), 'E-Net:Evolutionary neural network synthesis', Neurocomputing 42, 171-196.
[20] Schetinin, V. (2003), 'A Learning Algorithm for Evolving Cascade Neural Networks', Neural Processing Letters(17), 21-31.
[21] Stanley, K.O. & Miikkulainen, R. (2002),Continual Coevolution Through Complexification., in 'GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference, New York, USA, 9-13 July 2002', pp. 113-120.
[22] Stanley, K.O. & Miikkulainen, R. (2002), 'Evolving Neural Networks through Augmenting Topologies', Evolutionary Computation 10(2), 99-127.
[23] Stanley, K.; Bryant, B. & Miikkulainen, R. (2005), 'Real-time neuroevolution in the NERO video game', Evolutionary Computation, IEEE Transactions on 9(6), 653-668.
[24] Mahfoud, S.W. (1995),'Niching Methods for Genetic Algorithms'(95001), Technical report, Illinois Genetic Algorithms Laboratory (IlliGaL), University of Ilinios at Urbana-Champaign.
[25] Juille, H. & Pollack, J.B. (1996),Co-evolving intertwined spirals, in Lawrence J. Fogel, Peter J. Angeline & T. Baeck, ed.,'Proceedings of the Fifth Annual Conference on Evolutionary Programming', MIT Press, , pp. 461-467.
[26] Brown, G.: Diversity in Neural Network Ensembles, PhD thesis, The University of Birmingham, 2004
[27] Ivakhnenko, A.G., Müller, J.-A.: Self-organisation of nets of active neurons. SAMS, vol.20, no.1-2, 1995, pp.93-106.
[28] Ivakhnenko, A.G.; Savchenko, E.A. & Ivakhnenko, G.A.: Problems of future GMDH algorithms development, Systems Analysis Modelling Simulation 43, p. 1301 – 1309, 2003.
[29] Ivakhnenko,A.G., Ivakhnenko,G.A. and Muller,J.A., Self-Organization of Neural Networks with Active Neurons. Pattern Recognition and Image Analysis, 1994, vol.4, no.2, pp.185-196.