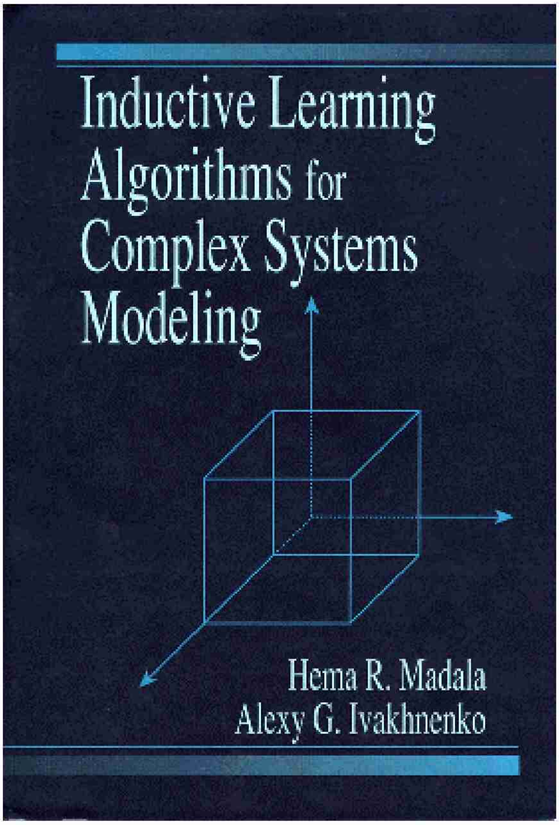
Inductive Learning Algorithms for Complex Systems Modeling

Hema R, Madala
Department of Mathematics and Computer Science
Clarkson University
Potsdam, New York

Alexy G, Ivakhnenko
Ukrainian Academy of Sciences
Institute of Cybernetics
Kiev, Ukraine

CRC Press
Boca Raton Ann Arbor London Tokyo



Madala H.R. and Ivakhnenko A.G.,

"Inductive Learning Algorithms for Complex System Modeling",

1994, CRC Press, ISBN: 0-8493-4438-7.

From abstract:

Inductive Learning Algorithms for Complex Systems Modeling is a professional monograph that surveys new types of learning algorithms for modeling complex scientific systems in science and engineering. The book features discussions of algorithm development, structure, and behavior; comprehensive coverage of all types of algorithms useful for this subject; and applications of various modeling activities (e.g., environmental systems, noise immunity, economic systems, clusterization, and neural networks). It presents recent studies on clusterization and recognition problems, and it includes listings of algorithms in FORTRAN that can be run directly on IBM-compatible PCs.

Inductive Learning Algorithms for Complex Systems Modeling will be a valuable reference for graduate students, research workers, and scientists in app lied mathematics, statistics, computer science, and systems science disciplines. The book will also benefit engineers and scientists from applied fields such as environmental studies, oceanographic modeling, weather forecasting, air and water pollution studies, economics, hydrology, agriculture, fisheries, and time series evaluations.

Features:

- Discusses algorithm development, structure, and behavior
- Presents comprehensive coverage of algorithms useful for complex systems modeling
- Includes recent studies on clusterization and recognition problems
- Provides listings of algorithms in FORTRAN that can be run directly on IBM-compatible PCs

Madala, Hema Rao.
Inductive learning algorithms for complex systems modeling / Hema Rao Madala and Alexey G. Ivakhnenko.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-4438-7
1. System analysis. 2. Algorithms. 3. Machine learning.
I. Ivakhnenko, Aleksei Grigo'evich. II. Title.
T57.6.M313 1993 93-24174
003--dc20 CIP

Direct all inquiries to CRC Press, Inc., 2000 Corporate Blvd., Boca Raton, Florida 33431.

International Standard Book Number 0-8493-4438-7
Library of Congress Card Number 93-24174
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

One can see the development of automatic control theory from single-cycled to the multicycled systems and to the development of feedback control systems that have brainlike network structures (Stafford Beer). The pattern recognition theory has a history of about fifty years—beginning with single-layered classificators, it developed into multi-layered neural networks and from there to connectionist networks. Analogical developments can be seen in the cognitive system theory starting with the simple classifications of the single-layered perceptrons and further extended to the system of perceptrons with the feedback links. The next step is the stage of "neuronets."

One of the great open frontiers in the study of systems science, cybernetics, and engineering is the understanding of the complex nonlinear phenomena which arise naturally in the world we live in. Historically, most achievements were based on the deductive approach. But with the advent of significant theoretical breakthroughs, layered inductive networks, and associated modern high-speed digital computing facilities, we have witnessed progress in understanding more realistic and complicated underlying nonlinear systems. Recollect, for example, the story of Rosenblatt's perceptron theory. Until recently, the absence of good mathematical description with the demonstration by Minsky and Papert (1969) that only linear descrimination could be represented by two-layered perceptron, led to a waning of interest in multilayered networks. Still Rosenblatt's terminology has not been recovered; for example, we say "hidden units" instead of Rosenblatt's "association units" and so on.

Moving in the direction of unification we consider the inductive learning technique called Group Method of Data Handling (GMDH), the theory originated from the theory of perceptron and is based on the principle of self-organization. It was developed to solve the problems of pattern recognition, modeling, and predictions of the random processes. The new algorithms that are based on the inductive approach are very similar to the processes in our brain. Scientists who took part in the development have accepted "this science" as a unification of pattern recognition theory, cybernetics, informatics, systems science, and various other fields. Inspite of this, "this science" is quickly developing, and everybody feels comfortable in using "this science" for complex problem-solving. This means that this new scientific venture unifies the theories of pattern recognition and automatic control into one metascience. Applications include the studies on environmental systems, economical systems, agricultural systems, and time-series evaluations. The combined Control Systems (CCS) group of the Institute of Cybernetics, Kiev (Ukraine) has been a pioneering leader in many of these developments. Contributions to the field have come from many research areas of different disciplines. This indicates a healthy breadth and depth of interest in the field and a vigor in associated research. Developments could be more effective if we become more attentive to one another.

Since 1968 layered perceptron-like networks have been used in inductive learning algorithms, particularly in the training mode. The algebraic and the finite-difference type of polynomial equations which are linear in coefficients and nonlinear in parameters are used for the process predictions. In the network, many arbitrary links of connection-weights are obtained, several partial equations are generated, and the links are selected by our choice. The approach was originally suggested by Frank Rosenblatt to choose the coefficients of the first layer of links in a random way.

The polynomials of a discrete type of Volterra series (finite-difference and algebraic forms) are used in the inductive approach for several purposes:

First—for the estimation of coefficients by the least-squares method using explicit or implicit patterns of data. When the eigenvalues of characteristic equation are too small, this method leads to very biased estimates and the quality of predictions is decreased. This problem is avoided with the developments of objective systems analysis and cluster analysis.

Second—the polynomial equations are used for the investigation of selection characteristic by using the consistency (Shannon's displacement) criterion of minimum according to Shannon's second-limit theorem (analogical law is known in communication theory). The structure of optimal model is simplified when the noise dispersion in the data is increased. When Shannon's displacement is present, selection of two-dimensional model structures is used. When the displacement is absent, the selection of two one-dimensional model structures are used—first, the optimal set of variables, then the optimal structure of the model are found. The use of objective criteria in canonical form simplifies this procedure further.

Third—the use of polynomial equations are organized "by groups" in the selection procedure to get a smooth characteristic with single minimum. Selection "by groups" allows one to apply the simple stopping rule "by minimum" or "by the left corner rule." In multilevel algorithms, for example, each group includes a model candidate of similar structure of an equal number of members; and

Fourth—the equations are used to prove the convergence of iteration processes in multilayered algorithms. The convergence exists for some criteria in a mean-square sense called internal convergence; for others it is called external convergence. In the latter case, there is a necessity for certain "regularization" means.

This book covers almost last twenty years of research—from basic concepts to the recent developments in inductive learning algorithms conducted by the CCS group.

Chapter 1 is concerned with the basic approach of induction and the principle of self-organization. We also describe the selection criteria and general features of the algorithms.

Chapter 2 considers various inductive learning algorithms: multilayer, single-layered combinatorial, multi-layered aspects of combinatorial, multi-layered with propagating residuals, harmonical algorithms, and some new algorithms like correlational and orthogonalized partial descriptions. We also describe the scope of long-range quantitative predictions and levels of dialogue language generalization with subjective versus multilevel objective analysis.

Chapter 3 covers noise immunity of algorithms in analogy with the information theory. We also describe various selection criteria, their classification and analysis, the aspects of the asymptotic properties of external criteria, and the convergence of algorithms.

Chapter 4 concentrates on the description of physical fields and their representation in the finite-difference schemes, as these are important in complex systems modeling. We also explain the model formulations of cyclic processes.

Chapter 5 coverage is on how unsupervised learning or clustering might be carried out with the inductive type of learning technique. The development of new algorithms like objective computerized clustering (OCC) is presented in detail.

Chapter 6 takes up some of the applications related to complex systems modeling such as weather modeling, ecological, economical, agricultural system studies, and modeling of solar activity. The main emphasis of the chapter is on how to use specific inductive learning algorithms in a practical situation.

Chapter 7 addresses application of inductive learning networks in comparison with the artificial neural networks that work on the basis of averaged output error. The least mean-square (LMS) algorithm (adaline), backpropagation, and self-organization boolean-logic techniques are considered. Various simulation results are presented. One notes that the backpropagation technique which is encouraged by many scientists, is only one of several possible ways to solve the systems of equations to estimate the connection coefficients of a feed-forward network.

Chapter 8 presents the computational aspects of basic inductive learning algorithms. Although an interactive software package for inductive learning algorithms which includes multilayer and combinatorial algorithms was recently released as a commercial package (see *Soviet Journal of Automation and Information Sciences* N6, 1991), the basic source of these algorithms along with the harmonical algorithm are given in chapter 8.

The book should be useful to scientists and engineers who have experience in the scientific aspects of information processing and who wish to be introduced to the field of inductive learning algorithms for complex systems modeling and predictions, clustering, and neural-net computing, especially these applications.

This book should be of interest to researchers in environmental sciences, macro-economical studies, system sciences, and cybernetics in behavioural and biological sciences because it shows how existing knowledge in several interrelated computer science areas intermesh to provide a base for practice and further progress in matters germane to their research.

This book can serve as a text for senior undergraduate or for students in their first year of a graduate course on complex systems modeling. It approaches the matter of information processing with a broad perspective, so the student should learn to understand and follow important developments in several research areas that affect the advanced dynamical systems modeling. Finally, this book can also be used by applied statisticians and computer scientists who are seeking new approaches.

The scope of these algorithms is quite wide. There is a wide perspective in which to use these algorithms; for example, multilayered theory of statistical decisions (particularly in case of short-data samples) and algorithm of rebinarization (continued values recovery of input data). The "neuronet," that is realized as a set of more developed component-perceptrons in the near future, will be similar to the House of Commons, in which decisions are accepted by the voting procedure. Such voting networks solve problems related to pattern recognition, clustering, and automatic control. There are other ideas of binary features applied in the application of "neuronets," especially when every neuron unit is realized by two-layered Rosenblatt's perceptron.

The authors hope that these new ideas will be accepted as tools of investigation and practical use - the start of which took place twenty years ago for original multilayered algorithms. We invite readers to join us in beginning "this science" which has fascinating perspectives.

H. R. Madala and A. G. Ivakhnenko

Acknowledgments

We take this opportunity to express our thanks to many people who have, in various ways, helped us to attain our objective of preparing this manuscript.

Going back into the mists of history, we thank colleagues, graduate students, and co-workers at the Combined Control Systems Division of the Institute of Cybernetics, Kiev (Ukraine). Particularly, one of the authors (HM) started working on these algorithms in 1977 for his doctoral studies under the Government of India fellowship program. He is thankful to Dr. V.S. Stepashko for his guidance in the program. Along the way, he received important understanding and encouragement from a large group of people. They include Drs. B.A. Akishin, V.I. Cheberkus, D. Havel, M.A. Ivakhnenko, N.A. Ivakhnenko, Yu.V. Koppa, Yu.V. Kostenko, P.I. Kovalchuk, S.F. Kozubovskiy, G.I. Krotov, P.Yu. Peka, S.A. Petukhova, B.K. Svetalskiy, V.N. Vysotskiy, N.I. Yunusov and Yu.P. Yurachkovskiy. He is grateful for those enjoyable and strengthening interactions.

We also want to express our gratitude and affection to Prof. A.S. Fokas, and high regard for a small group of people who worked closely with us in preparing the manuscript. Particularly, our heartfelt thanks go to Cindy Smith and Frank Michielsen who mastered the TeX package. We appreciate the help we received from the staff of the library and computing services of the Educational Resources Center, Clarkson University. We thank other colleagues and graduate students at the Department of Mathematics and Computer Science of Clarkson University for their interest in this endeavor.

We are grateful for the help we received from CRC Press, particularly, Dr. Wayne Yuhasz, Executive Editor. We also would like to acknowledge the people who reviewed the drafts of the book - in particular, Prof. N. Bourbakis, Prof. S.J. Farlow and Dr. H. Rice - and also Dr. I. Havel for fruitful discussions during the manuscript preparation.

We thank our families. Without their patience and encouragement, this book would not have come to this form.