ARMENIAN NEURAL NETWORK CENTER

Business Proposition

Prof. Ashot Chilingarian,

Head of Cosmic Ray Division, Alikhanyan Physics Institute

2 Alikhanyan Brothers, Yerevan 36, Armenia,
Phone/Fax, 3741-344377, e-mail chili@crdlx5.yerphi.am

Proposition:

This is a business proposition to create the Armenian Neural Network Center (ANNC). A Group of highly experienced and highly motivated Armenian physicists and mathematicians who are internationally recognized for their expertise in this field are available for this venture as a private business. Proceeding from the problem the customer needs to investigate, we will work with them to help them gather relevant data, then we will provide the solution in the form of ”trained” network software, specific to the problem. The trained network for solving the problem can be easily tested and certified and easily implemented as a decision-making tool at the customer's site. As new data becomes available, the network can be "re-trained" and sent as an upgrade to the user in a very short time. -

Taking into account the positive feedback from current users of our packages in the scientific community, we feel we can apply our expertise in Neural Network solutions to industrial, medical, defense, management, and other users. Problems dealing with stochastic processes in complex nonlinear systems require robust algorithms for effective analysis of large amounts of nonparametric stochastic data. With our experience in writing successful algorithms and successfully implementing them in software, we can offer our expertise to various categories of customers. The solution will be available in user-friendly, ready-to-use format to the customer.

An example of implementing our methods for a practical, and important problem solution is the patent ***"Multivariate random search method with multiple starts and early stop for identification of differentially expressed genes based on microarray data"***, described in a provisional application filed in the United States Patent and Trademark Office on March 1, 2002, and accorded Application No. 60/361,068.In 2004 patent was authorized in US, Canada and Europe under N 03713675/1-24-02-US0305730.

Another our Internet information product – ***"Data Visualisation Interactive Network for the Aragats Space-environmental Center" –*** *DVIN for ASEC*, was officially announced as the world's best project in the category of e-science at the first World Summit on the Information Society (WSIS) 2003 in Geneva, December 9-13.

Brief Explanation:

A large number of processes in science, industry and medicine can be modeled by means of Bayesian statistical paradigm and Artificial Neural Networks. These techniques are employed when dealing with multidimensional and analytically difficult-to-formulate systems and data.

The analysis of the bulk of multidimensional data with high levels of stochastic variability requires development of new, powerful, reliable, and robust analysis and decision-making tools to meet the challenge of understanding and utilizing the abundant amount of information available from numerous sources.

The main goal of the statistical inference in astro-particle physics is the development of high-speed and effective algorithms for the analysis of multidimensional experimental data in several different areas of interest: classification of elemental particles and ions from the complicated images or points in N-dimensional feature space, enhancement of signal-to-noise ratio in searching for new physics, estimation of particle energy, particle arrival direction and other continuous parameters, deciding on the physical hypothesis supported by the experimentally available data, and others.

In industry the problems to solve include plant process controls such as the optimization of transportation rout planning to the determination of chip topology, design processes of new types of products, multidimensional comparison of product characteristics, quality control, and many-many others.

In finance we can outline decision making on “good” and “bad” credit, optimization of money investment strategy, analysis of marketing studies, estimation of numerous micro and macroeconomic parameters, etc.

In medicine the issues include diagnostics and drug design. The amount of information to process in medical research reaches intractable amounts particularly in the field of genomics. To outline and study “biological pathways” one needs techniques of optimal feature subset selection and multivariate comparisons.

In the area of national defense and intelligence information processing efficient and accurate analysis of the mountains of information available, which must be categorized according to severity of potential danger, has become a very high priority in the aftermath of September 11, 2001.

In all of the above-mentioned areas and more, optimal solutions are now sought with the help of intensive simulation and/or experimentation trials. For each of the alternative types of activity, huge amounts of multidimensional correlated information are collected to form a basis for scientific, process control, financial, or medical decisions. Usually it is impossible to find analytical family of functions to describe these distributions, therefore, we are obliged to use nonparametric methods of analysis. Thus, we have to deal with multivariate distributions with high level of statistical variability.

The key points in the required approach are feature selection, that is, the reduction of the initial multidimensional metric space of measured, or simulated parameters, multivariate density estimation modes which provide fastest convergence to theoretically achievable limits, and fast, reliable neural network training (learning) methods.

The collection of multivariate nonparametric methods, developed in united methodology was implemented in the original ANI software package. The first version of ANI software was developed in the early 80’s and has been continually enhanced and used extensively in the cosmic ray physics community since then.

The traditional nonparametric estimates, such as Parzen window or K Nearest Neighbor (KNN) models, are dependent on the window size or particular number of neighbors, which present difficult uniform choices over the whole distribution support range. The Bayesian decision rules implemented in ANI are based on original nonparametric methods of multivariate probability density estimation and exploit the inherent robustness of the distribution median. The Median Estimate method for ANI has proven to be very robust and provides better estimates than any of the fixed value methods, such as Parzen window or KNN.

The ANI Neural Network toolbox provides different training scenarios using the stochastic learning paradigm. It uses evolutionary algorithms consisting of random variations and survival-of-the fittest principles, for solving optimization problems. The search for "good" solutions is performed in the Feed Forward Neural Networks weights space, implementing different modifications of evolutionary algorithms. The network performance is continuously monitored during training, thus making it possible to formulate stopping rules, and therefore avoiding overtraining.

The methodologies implemented in the ANI package and User's Guide is available online at http://crdlx5.yerphi.am/ani/index.html.