THE COMPUTATIONAL INTELLIGENCE TECHNIQUES FOR PREDICTIONS - ARTIFICIAL NEURAL NETWORKS

Mary Violeta Bar Ph. D Student
University of Craiova
Faculty of Economics and Business Administration
Craiova, Romania

Abstract: : The computational intelligence techniques are used in problems which can not be solved by traditional techniques when there is insufficient data to develop a model problem or when they have errors.Computational intelligence, as he called Bezdek (Bezdek, 1992) aims at modeling of biological intelligence. Artificial Neural Networks(ANNs) have been applied to an increasing number of real world problems of considerable complexity. Their most important advantage is solving problems that are too complex for conventional technologies - problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found. In general, because of their abstraction from the biological brain, ANNs are well suited for problems that people are good at solving, but for which computers are not. The ability to accurately predict the future is fundamental to many decision activities in many functional areas of business.In this paper emphasize the advantages and disadvantages of using ANNs for predictions.

JEL classification: C45, C53,C63

Key words: Computational intelligence techniques, Artificial Neural Networks, prediction

1. INTRODUCTION

The ability to analyze massive data lags far behind the capability of gathering and storing it. This gives rise to new challenges for businesses and researchers in the extraction of useful information.

Increasing accuracy of forecasting can save millions for a company and is a major motivation for using formal methods of forecasting and systematic investigation of new methods and better prognosis.

Information technology in the past several years have created a lots of innovations in the area of business. More businesses and organizations are collecting high quality data on a large scale. The world is rather nonlinear and complex than linear because there are so many possible nonlinear relationships or structures. Most nonlinear models developed during the last two decades are likely parameters.

To use these models, the model must be specified first. Therefore, these models cannot be used if the data characteristics do not fit the model assumptions involved. The parametric approach is quite suitable for nonlinear problems with complex structures, but there is a lack of theories to suggest a specific form of the structure.

Artificial Neural Networks are algorithms and techniques that can be used to perform nonlinear statistical modeling and provide a new alternative to logistic regression, the most commonly used method for developing predictive models.

Artificial Neural Networks offer a number of advantages, including requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms.

2. OBJECTIVES

The main objective of this study is an analysis of the advantages and disadvantages using Artificial Neural Networks. An overview of the features of neural networks is presented, and the advantages and disadvantages of using this modeling technique are discussed.

3. METHODOLOGY

An Artificial Neural Network (ANN), is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system.

It is composed of a large number of highly interconnected processing elements working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well.

A Neural Network is a parallel computing system of several interconnected processor nodes. The input to individual network nodes is restricted to numeric values falling in the closed range [0,1]. Because of this, categorical data must be transformed prior to network training .

Another definition is given by Haykin : A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

Knowledge is acquired by the network through a learning process. Interneuron connection strengths known as synaptic weights are used to store the knowledge.

In 1982, John Hopefield published a paper showing how neural networks could be used for computational purposes. In 1984, Kohonen introduced a new algorithm he called an organizing feature map, which allowed for a process of using neural networks for unsupervised learning.

This opened a new branch of neural network research where no correct answer is required to learn or train a network. In 1986 Rumelhart, Hinton and Williams wrote a paper on the back-propagation method, which opened up a flurry of activity in the late 1980s and 1990s.

Neural networks are used extensively in the business world as predictive models. In particular, the financial services industry widely uses neural networks to model fraud in credit cards and monetary transactions.

Some of the well known types of neural networks are: Competitive Learning, the Boltzmann Machine, the Hopfield Network, the Kohonen network, the Adaptive Resonance Theory , and back propagation neural networks.

Although there are many other variations of neural networks, the back propagation network and its variants, as a subset of multilayer feed forward networks, are currently the most widely used networks in applications.

Neural networks attempt to mimic a neuron in a human brain, with each link described as a processing unit. Neural networks learn from experience and are useful in detecting unknown relationships between a set of input data and an outcome.

Like other approaches, neural networks detect patterns in data, generalize relationships found in the data, and predict outcomes. Neural networks have been especially noted for their ability to predict complex processes.

Processing elements, or processing units are linked to inputs and outputs. The process of training a network involves modifying the strength, or weight, of connections from the inputs to the output.

Increase or decreases in the strength of a connection is based on its importance for producing the proper outcome. A connection's strength depends on a weight it receives during a trial-and-error process.

This process uses a mathematical model for adjusting the weights, and is called a learning rule.

Training continues until a neural network produces outcome values that match the known outcome values within a specified accuracy level, or until it satisfies some other stopping criteria.

Figure no.1 demonstrates a neural network. Each of the processing units takes many inputs and generates an output that is a nonlinear function of the weighted sum of the inputs.

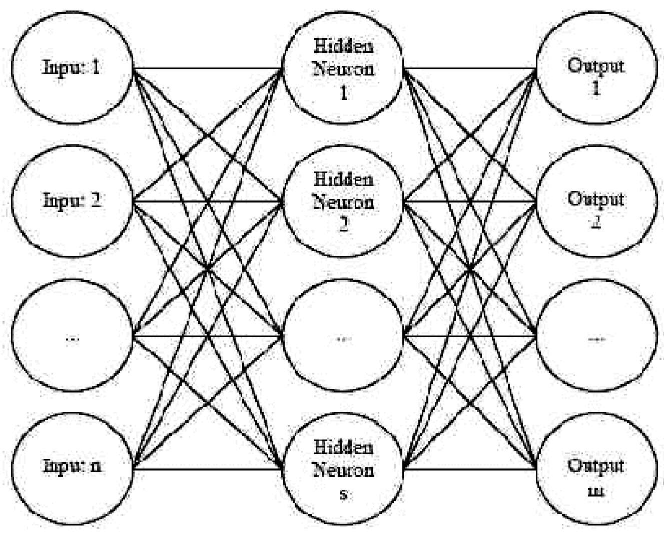

Figure no.1 An ANN model

The weights assigned to each of the inputs are obtained during a training process in which outputs generated by the nets are compared with target outputs. The answers you want the network to produce are compared with generated outputs, and the deviation between them is used as feedback to adjust weights.

The process of readjusting weights is important to increasing a model's accuracy. Notice there are also hidden nodes, or middle layer nodes, in Figure 1. These hidden nodes are associated with the weighting process. The number of hidden nodes can be adjusted and there can be multiple levels of hidden nodes. The number of inputs, hidden nodes, outputs, and the weighting algorithms for the connections between nodes determine the complexity of a neural network, its accuracy, and the time it takes to create the neural network model.

Because the configuration of hidden nodes and weights is so critical to neural networks, there are many approaches for finding the right number of hidden nodes and readjusting weights.

Network Design Parameters: Employing a neural network requires an understanding of a number of network design options. Be advised that there are no definate rules for choosing the settings of these parameters a priori. Since the solution space associated with each problem is not known, an number of different network runs must be undertaken before the user can determine with relative confidence a suitable combination.

Number of Input Nodes: These are the independent variables which must be adjusted to fall into a range of 0 to 1. The number of nodes is fixed by the number of inputs. Inputs must not be nominal scale, but can be binary or better ordinal. Such inputs can be accommodated by providing a separate input node for each category which is associated with a binary (0 or 1) input.

Number of Output Nodes: The number of output nodes depends on the purposes of the research and they are also adjusted to fall within the range of 0-1.

Number of middle or hidden layers: The hidden layers allow a number of potentially different combinations of inputs that might results in high (or low) outputs. Each successive hidden layer represents the possibility of recognizing the importance of combinations of combinations

Number of Hidden Layers: The more nodes there are the greater the number of different input combinations that the network is able to recognize.

Number of Nodes Per Hidden Layer: Generally all nodes of any one layer are connected to all nodes of the previous and the following layers. This can be modified at the discretion of the user however.

Initial Connection Weights: The weights on the input links are initialized to some random potential solution. Because the training of the network depends on the initial starting solution, it can be important to train the network several times using different starting points.

Some users may have reason to start the training with some particular set of link weights. It is possible, for example to find a particularly promising starting point using a genetic algorithm approach to weight initialization.

Initial Node Biases: Node bias values impart a significance of the input combinations feeding into that node. In general node biases are allowed to be modified during training, but can be set to particular values at network initialization time. Modification of the node biases can be also allowed or disallowed.

Learning Rate: At each training step the network computes the direction in which each bias and link value can be changed to calculate a more correct output. The rate of improvement at that solution state is also known. A learning rate is user-designated in order to determine how much the link weights and node biases can be modified based on the change direction and change rate. The higher the learning rate (max. of 1.0) the faster the network is trained. However, the network has a better chance of being trained to a local minimum solution. A local minimum is a point at which the network stabilizes on a solution which is not the most optimal global solution.

Momentum Rate: To help avoid settling into a local minimum, a momentum rate allows the network to potentially skip through local minima. A history of change rate and direction are maintained and used, in part, to push the solution past local minima. A momentum rate set at the maximum of 1.0 may result in training which is highly unstable and thus may not achieve even a local minimum, or the network may take an inordinate amount of training time.

If set at a low of 0.0, momentum is not considered and the network is more likely to settle into a local minimum. A process of "simulated annealing" is performed if the momentum rate starts high and is slowly shifted to 0 over a training session.

Like other statistical and mathematical solutions, back propagation networks can be over- parameterized. This leads to the ability of the statistics to find parameters which can accurately compute the desired output at the expense of the system's ability to interpolate and compute appropriate output for different inputs.

To ensure that a back propagation neural network is not over parameterized, the training data must be split into a training and a testing set. It is the performance of the trained network on the data reserved for testing that is the most important measure of training success.

4. ANALYSES

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

Other advantages include:

- Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
- Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.
- Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
- Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

Neural networks are universal approximators, and they work best if the system you are using them to model has a high tolerance to error. However they work very well for:

- capturing associations or discovering regularities within a set of patterns;
- where the volume, number of variables or diversity of the data is very great;
- the relationships between variables are vaguely understood;
- the relationships are difficult to describe adequately with conventional approaches.

The greatest strength of neural networks is their ability to accurately predict outcomes of complex problems. In accuracy tests against other approaches, neural networks are always able to score very high.

There are some downfalls to neural networks.

First, they have been criticized as being useful for prediction, but not always in understanding a model. It is true that early implementations of neural networks were criticized as “black box” prediction engines; however, with the new tools on the market today, this criticism is debatable.

Secondly, neural networks are susceptible to over-training. If a network with a large capacity for learning is trained using too few data examples to support that capacity, the network first sets about learning the general trends of the data. This is desirable, but then the network continues to learn very specific features of the training data, which is usually undesirable.

Such networks are said to have memorized their training data, and lack the ability to generalize.

The mathematical theories used to guarantee the performance of an applied neural network are still under development.

The solution for the time being may be to train and test these intelligent systems much as we do for humans. Also there are some more practical problems like:

- the operational problem encountered when attempting to simulate the parallelism of neural networks. Since the majority of neural networks are simulated on sequential machines, giving rise to a very rapid increase in processing time requirements as size of the problem expands. One solution to this problem is to implement neural networks directly in hardware, but these need a lot of development still.
- instability to explain any results that they obtain. Networks function as "black boxes" whose rules of operation are completely unknown.

5. CONCLUSIONS

ANNs provide an analytical alternative to conventional techniques which are often limited by strict assumptions of normality, linearity, variable independence etc.

Because an ANN can capture many kinds of relationships it allows the user to quickly and relatively easily model phenomena which may have been very difficult or impossible to explain otherwise.

Depending on the nature of the application and the strength of the internal data patterns you can generally expect a network to train quite well. This applies to problems where the relationships may be quite dynamic or non-linear.

Combining multiple models to improve forecast accuracy has been extensively studied in the literature.

The idea of association models is the assumption that the basic structure of real data is difficult or impossible to model by an exact model and the fact that different models can play a complementary role in capturing different data models.

REFERENCES

1. Grossberg, S — Adaptive pattern classification and universal recoding, ii: Feedback, expectation, olfaction, and illusions., Biological Cybernetics, 23, 1976
2. Haykin, S — Neural Networks: A Comprehensive Foundation, Macmillan, New York, p. 2, 1994
3. Hinton,G.E., Sejnowski,T.J., Ackley,D.H — Boltzmann Machines: Constraint satisfaction networks that learn, Technical Report CMU-CS-84-119, Carnegie-Mellon University, 1984
4. Hopfield, J — Neural networks and physical systems with emergent collective computational abilities, Proceedings of the National Academy of Sciences of the USA, vol. 79, no. 8 (April 1982), 1982
5. Kohonen, T — An introduction to neural computing, Neural Networks, Volume 1, Issue 1, 1988
6. Kohonen, T — Self-Organization and Associative Memory, Heidelberg New York, 1984
7. Rumelhart, D.E.,Zipser,D — Feature discovery by competitive learning". Cognitive Science, 9, 1985
8. Wang,M., Rees, S.J., Liao, S.Y — Building an online purchasing behavior analytical system with neural network, Edited by Zanasi, Brebbia and Melli, DataMining III., WIT Press, 2002.