

Business Applications of Emulative Neural Networks

Yochanan Shachmurove

*Departments of Economics, The City College of the City University of New York and,
The University of Pennsylvania, yochanan@econ.sas.upenn.edu*

ABSTRACT

This paper surveys research on Emulative Neural Network (ENN) models as economic forecasters. ENNs are statistical methods that seek to mimic neural processing. They serve as trainable analytical tools that “learn” autonomously. ENNs are ideal for finding non-linear relationships and predicting seemingly unrecognized and unstructured behavioral phenomena. As computing power rapidly progresses, these models are increasingly desirable for economists who recognize that people act in dynamic ways with rational expectations. Unlike traditional regressions, ENNs work well with incomplete data and do not require normal distribution assumptions. ENNs can eliminate substantial uncertainty in forecasting, but never enough to completely overcome indeterminacy.

JEL: C3, C32, C45, C5, C63, F3, G15.

Keywords: Emulative neural networks; Dynamic interrelations; Nonlinear forecasting; Perceptron learning process; Multi-layer perceptron model; Learning; Observational indeterminability; Indeterminacies

I. INTRODUCTION

This paper surveys the significance of recent work on emulative neural networks (ENNs) by researchers across many disciplines in the light of issues of indeterminacy. Financial and economic forecasters have witnessed the recent development of a number of new forecasting models. Traditionally, popular forecasting techniques include regression analysis, time-series analysis, moving averages and smoothing methods, and numerous judgmental methods. However, all of these have the same drawback because they require assumptions about the form of population distribution. Regression models, for example, assume that the underlying population is normally distributed.

ENNs are members of a family of statistical techniques, as are flexible nonlinear regression models, discriminant models, data reduction models, and nonlinear dynamic systems (Sarle, 1994; Cheng and Tetterington, 1994). They are trainable analytic tools that attempt to mimic information processing patterns in the brain (Krishnaswamy, Gilbert, Pashley, 2000). Because they do not necessarily require assumptions about population distribution, economists, mathematicians and statisticians are increasingly using ENNs for data analysis. Not only do they not require assumptions about the underlying population but are also powerful forecasting tools that draw on the most recent developments in artificial intelligence research.

As Hardin (2002) observes in his essay, "Indeterminacy and Basic Rationality," statistical methods, such as neural networks, were developed partly as the product of the ordinal revolution in economics and choice theory. As he points out, because our choices have social and interactive contexts, it would be extremely difficult to construct a theoretical model that is capable of tracing all of these potential and actual responses and interactions. Such models are bound to exhibit fundamental indeterminacy. These indeterminacies are the inevitable product of strategic interactions among rational individuals who understand that their actions, or inactions, are going to be followed by reactions - those of the other participants in the strategic game and those of the environment. In such circumstances, one may find that responses are not similar, let alone unique.

This is especially true when we add time to the discussion. A player may react in ways very different to what was presupposed in response to an unexpected reaction by his opponent. Such models are inherently dependent upon, and sensitive to, initial conditions, which may not permit accurate predictions even for very near-future states (Brown and Chua, 1998; Smith, 1998; Stone, 1989; Bau and Shachmurove, 2002). As Domotor and Batitsky (2002) point out, a creeping amplification of error will eventually wipe out all predictive accuracy. Even the solar system, reported to be the oldest paradigm of a regular, predictable dynamical system, is unpredictable on the time scale of millions of years.

Supposing such a complex model could be constructed, once the model is subject to empirical testing, the problem of aggregation arises. Although ENNs perform well in the presence of missing data, large data sets are required to train them (Gilbert, Krishnaswamy, and Pashley, 2000). This, in turn, leads to Arrow's impossibility theorem. The central difficulty in focusing on individual preferences is the aggregation of those preferences into a general rule; thus, it may be impossible to determine an optimal allocation of resources in the face of disagreement.

For example, in the case of an election, Arrow's theorem starts with a finite set of outcomes (such as, candidates for office), a finite set of voters and their individual preferences for outcomes. The individual preferences are an unrestricted domain, satisfying the Pareto principle, and they are independent of irrelevant alternatives (independence from other influences). Thus, Arrow's Impossibility Theorem states that it is not possible to derive a complete and consistent social choice rule exclusively from individual preferences, except in dictatorships, which are characterized by the impossibility of determining the intensity of preference (Arrow, 1963; see also Hardin, 2002).

ENNs ignore that we are incapable of conceiving, designing, or constructing such a complicated interactive model of human behavior. The remaining option is to try to learn from past observations, and to deduce the main patterns of future events by tracing painstakingly similar interactions found in historical data. Rather than seeking, or imposing, some determinate principle, or even attributing uniqueness to the solutions generated for the complex makings of a system, we let historical data conveniently guide us through the indeterminable maze confronting us (Krippendorff, 2002).

In such cases, one proceeds under the assumption that it should be more helpful to imitate the past than estimate the proximity of the initial conditions to their putative later effects. This is because in economics and finance, as in any human behavior, initial conditions or past observations depend on future values owing to expectations; hence, the price of a stock today does not depend solely on its price yesterday or ten years ago but also on the expectations of future demand for, and supply of the product it represents.

ENNs are capable of processing data and performing calculations that humans cannot perform simply. Their answers, therefore, must be accepted on faith, yet ENNs themselves are at the mercy of the sample data and the quality of their features (Ripley, 1993; Sarle, 1994). They process data and make forecasts much the same way that statistical algorithms perform estimations.

A typical civil engineering situation in which ENNs are utilized is vehicle discharge headway at signalized intersections. In this situation, variables include driver behaviors, vehicle characteristics, and traffic environment. A three-layered ENN was created to predict this model. When the model was trained and validated with field data, it was proven to estimate fairly accurately headway for individual vehicles better than any other alternative model (Hung and Tong, 2002).

In applications specific to the financial and economic fields, the main focus for ENN technology so far has been with data involving variables in nonlinear relations. Granger (1991) maintains that in financial and economic data, nonlinear relationships are more likely to occur than linear relationships. This claim is supported by tests such as the regression test and the neural network test, which confirms this nonlinearity through the rejection of linearity and chaos where justifiable.

Consequently, new tests based on ENN systems have increased in popularity among economists. Researchers have examined the application of ENNs to financial markets where the nonlinear properties of financial data create too many difficulties for traditional methods of analysis (see Grudnitski and Osburn 1993; Altman, Marco, and Varetto 1994; Michie, Spiegelhalter and Taylor, 1994; Kaastra and Boyd 1995; and Swanson and White, 1995).

The remainder of this paper is organized as follows. Section II explains the

background and principles behind emulative neural networks in general. Section III more specifically looks into the multilayer perceptron model. Section IV discusses some of the advantages and disadvantages of using ENNs. Section V presents some of the recent applications of ENNs to business, economics, and finance. Section VI concludes.

II. BACKGROUND AND PRINCIPLES

A. Background

ENNs are information processing systems whose structure and function are modeled after the cognitive processes and organizational structure of neuro-biological systems. They are a class of models developed by cognitive scientists interested in understanding the biological nervous system and simulating the computational methods of the brain. ENNs are generally used as models for intelligence, learning, and data analysis. For data analysis, ENNs are used in numerous fields of inquiry, ranging from airport noise and property values (Collins and Evans, 1994) to student grade point average movements (Gorr, Nagin, and Szczypula, 1994) and agricultural crop yields (Joerding, Li, and Young, 1994; Wang, 1994).

ENNs have been applied in modeling market response (Dasgupta, Dispensa, and Ghose, 1994), collective behavior (Artyushkin, 1990), telecommunication flows (Fisher and Gopal, 1994), real estate valuation (Worzala, Lenk, and Silva, 1995), and even the determinants of military expenditure (Refenes, Kollias, and Zapranis, 1995). They are also used for analyzing relations among economic and financial phenomena, forecasting, data filtration, generating time-series, and optimization (Hawley, Johnson, and Raina, 1990; Refenes, 1995; Gately, 1996).

The most basic components of ENNs are highly interconnected processing elements, called neurons, which work independently in parallel. It is anticipated that self-organization of the neural interconnections will result in an artificial intelligence akin to human learning (Jain and Mao, 1996).

However, the intelligence and value of ENNs are still highly debated (Levine, 2000). On the one hand, neural networks are capable of processing vast amounts of data and making extremely accurate forecasts, but on the other hand, many authors argue that despite their accurate forecasting capacities, neural networks are not intelligent in the human sense, too much incapable of insightful imagination (Ripley, 1993; Sarle, 1994; Abdi and Edelman, 1999).

Furthermore, a difference in terminology raises a barrier to communication between economists and ENNs practitioners. What economists and statisticians call independent variables are called *inputs* in the neural network literature; dependent variables are known as *training values*; estimated values are called *outputs*; residuals are the same as *errors*; estimation is interchangeably represented by *training*, *learning*, *adaptation*, and *self-organization*; the neural network counterpart to an estimation criterion is called either an *error*, *cost*, or *Lyapunov function*; *training patterns* or *training pairs* are observations in the statistical literature; transformations are known as *functional links*; interpolation and extrapolation are called *generalization* (White, 1992; Sarle, 1994; Bishop, 1995; Devroye, Gyorfı and Lugosi, 1996; Ripley, 1996). In the neural network literature, *training sets* are

commonly used in a manner that corresponds to the sample in statistics.

B. Principles

The neuron is a basic functional element of the brain (see Chapter by Gur et al. 2002). An individual neuron consists of a cell body, dendrites, and an axon. Every individual neuron receives electrical stimuli from other neurons through the dendrites. The electrical stimuli is then amplified or de-amplified by the synapse and summated. If the sum of all stimuli is greater than the neuron's resistance threshold, provided by the cell body, then the neuron fires, producing a stimulus that passes through the axon to another neuron (Gleitman, 1991).

These synaptic connections are used to carry messages from one neuron to another. The strength of these connections varies. Neurons store information and learn meaningful patterns by strengthening their inter-connections. The synaptic process, modeled mathematically, serves as the theoretical basis for artificial neural networks. This section demonstrates the basic structure and behavior of artificial neurons. These neurons, on their own, are very simple structures and are generally more useful when connected together as part of a larger network.

A set of inputs x_i , a set of weights w_i , a threshold, u , an activation function, f , and a signal neuron output, y , where i is the degree (number of inputs) of the neuron. The weights represent the amplification or de-amplification of the process. Typically, the values of the inputs, x_i , are external and are stimuli from the environment. Alternatively, they may be generated from the outputs of other artificial neurons. They can be discrete values from a set, such as $\{0,1\}$, or real-valued numbers.

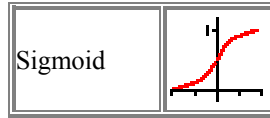
The weights, w_i are real-valued numbers that determine the contribution of each input to the neuron's weighted sum and eventually its output. The goal of neural network training algorithms is to determine the best possible set of weight values for the problem under consideration. Finding the optimal set is often a trade-off between computation time and minimizing the network error.

The threshold, u , is a real number that is subtracted from the weighted sum of the input values. Sometimes the threshold is referred to as a bias value. In this case, the real number is added to the weighted sum. For simplicity, the threshold can be regarded as another input / weight pair, where $w_0 = u$ and $x_0 = -1$.

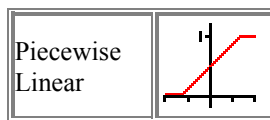
The activation function, f , can be very simple; for example, there is the unit step function (see below). However, the artificial neuron model has been expanded to include other functions such as the sigmoid, piecewise linear, and Gaussian functions. The activation functions that are often considered are shown below. A possible representation of the Unit Step equation is $f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$



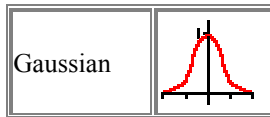
The Sigmoid equation is $f(x) = 1 / (1 + e^{-\beta x})$



The piecewise equation is $f(x) = \begin{cases} 0 & \text{if } x < x_{\min} \\ mx + b & \text{if } x_{\max} > x > x_{\min} \\ 1 & \text{if } x \geq x_{\max} \end{cases}$



The Gaussian equation is $f(x) = [1/(2\pi\sigma)^{0.5}] e^{[-(x-\mu)/(x-\mu)/2\sigma\sigma]}$



The artificial neuron output, y , computes its output according to the equation shown below. This is the output of the activation function for the current weighted sum. Once the output has been calculated, it can be passed to another neuron (or group of neurons) or sampled by the external environment. The interpretation of the neuron output depends upon the problem under consideration. For example, in pattern classification, an output of 1 would imply the input belongs to a certain class.

$$y = f(w_1 x_1 + w_2 x_2 + w_3 x_3 + \dots + w_n x_n + u)$$

$$y = f(w_0 x_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n)$$

Now, because usually learning is involved, one may want to include a learning formula for weight updating

$$W_i^{\text{new}} = W_i^{\text{old}} + k,$$

where k is a function involving a learning constant and errors at various (input, output and hidden) layers of the network (see more below). The output of the neuron, which depends on the neuron's threshold, is a non-linear transformation of the weighted sum of its total input, usually the asymmetric sigmoid or the hard limiter (Kollias and Refenes, 1996).

C. Perceptron Learning Process

This section presents a simple form of supervised learning entitled the perceptron-learning

algorithm. Using this procedure, one can train the perceptron to act as a binary logic unit. The simplest example is of a Single-Layer Perceptron Neural Networks (SLPNN). SLPNN consists of one or more artificial neurons in parallel. Each neuron provides one network output and is usually connected to all of the external inputs. Figure 1 is an example of a single-neuron, single-layer perceptron network.

The perceptron-learning algorithm can be described as follows:

1. Initialize the weights and threshold to small random numbers.
2. Present a vector to the neuron inputs and calculate the output.
3. Update the weights according to:
 $w_j(t+1) = w_j(t) + \zeta(d-y)x_j$
 where d is the desired output, t , is the iteration number, and ζ is the gain or step size, where $0.0 < \zeta < 1.0$
4. Repeat steps 2 and 3 until the iteration error is less than a user-specified error threshold or a predetermined number of iterations have been completed

Notice that learning only occurs when an error is made; otherwise, the weights are left unchanged. During training, it is often useful to measure the performance of the network as it attempts to find the optimal weight set. A common measure of error or cost function is sum-squared error. It is computed over the entire input vector / output vector pairs in the training set and is given by the equation below:

$$E = (1/2) \left(\left| y^{(1)} - d^{(1)} \right|^2 + \left| y^{(2)} - d^{(2)} \right|^2 + \left| y^{(p)} - d^{(p)} \right|^2 \right)$$

where p is the number of input/output vector pairs in the training set.

The power of neural computing comes from the threshold concept. It provides a way to transform complex interrelationships into simple yes-no situations. When the combination of several factors begins to become overly complex, the neuron model posits an intermediate yes-no node to retain simplicity.

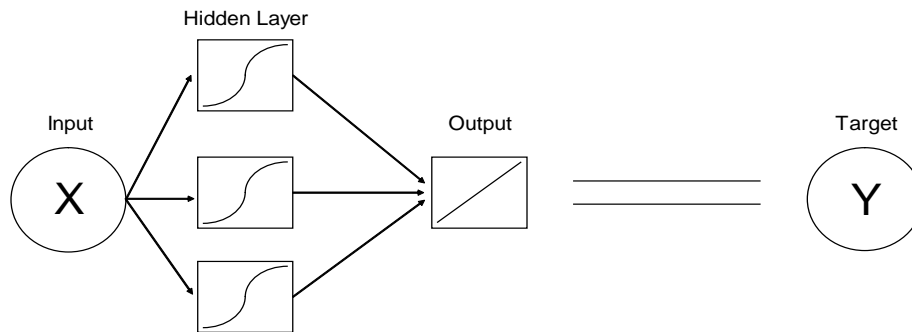
ENNs consist of interconnected layers that contain neuron-like units called processing elements. These processing elements allow an algorithm to “learn” and eventually recognize meaningful patterns in the data. As a given algorithm learns by synthesizing more training records, the weights between its interconnected processing elements strengthen and weaken dynamically. Determinacy may play a key role here, employing the rational choice of picking more rather than less value when given the weights of the preferences; thus, each of these processes yield determinate outcomes up to the limit of the device.

III. THE MULTILAYER PERCEPTRON MODEL

One of the most useful applications of neural networks to data analysis is the multilayer perceptron model (MLP). MLPs are nonlinear neural network models that can be used to approximate almost any function with a high degree of accuracy (White 1992). An MLP contains a hidden layer of neurons that uses nonlinear activation functions, such as a logistic function. Figure 1 offers a representation of an MLP with one hidden layer and a

single input and output. The MLP in figure 1 represents a simple nonlinear regression.

Figure 1
Multi-layer Perceptron with a Single Input and output

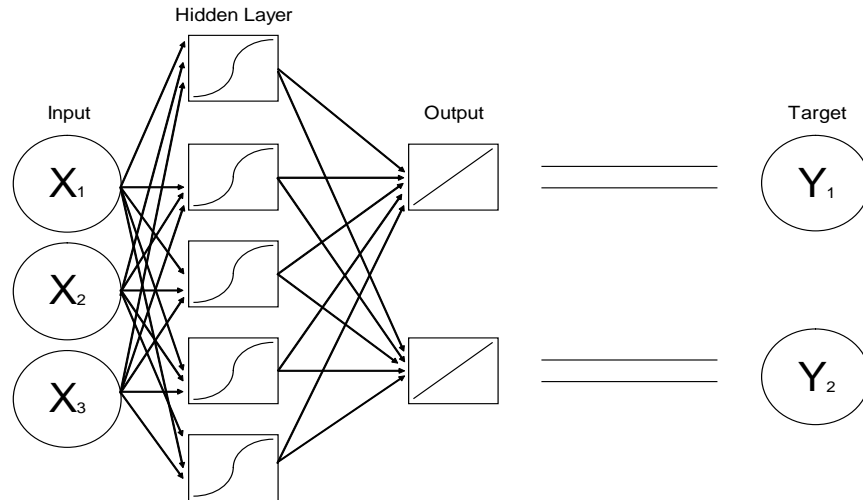


The number of inputs and outputs in the MLP, as well as the number of hidden layers, can be manipulated to analyze different types of data. Figure 2 presents a multilayer perceptron with multiple inputs and outputs. The MLP in figure 2 represents multivariate multiple nonlinear regression.

The complexity of the MLP can be adjusted by varying the amount of hidden layers, thus, transforming an MLP from a simple parametric model to a flexible non-parametric model (White, 1992; Kuan and White, 1994; Fine, 1999; Husmeier, 1999). An MLP consisting of only one layer is very close to the projection pursuit regression model, which allows computation of explicit formulae for bias and error about the mean in orientation estimates and curve estimates.

However, the projection pursuit model uses a flexible nonlinear smoother whereas the MLP uses *a priori* determined activation function in the hidden layer. The activation function determines the activation or the internal state of a neuron. It is applied to the weighted inputs, and it results in an output signal. As the number of hidden layers is allowed to increase, an MLP becomes a useful alternative to kernel regression analysis, a nonparametric approach that makes no assumptions regarding the functional form of the unknown function (Green, 2000), and smoothing splines. These splines are parametrically linear, and they are easier to fit to a nonlinear regression curve.

Figure 2
Multi-layer Perceptron with Multiple Inputs and outputs



Spline regressions are estimation methods that investigate both shifts of slopes and intercepts by approximating the shifts and changes in intercepts by separate linear functions with dummy variables. The process is similar to estimating two or more separate regressions at once without losing degree of freedom by using linear restrictions on the estimated coefficients. If the variances differ across the different regression lines, however, ambiguity arises in the estimation of the disturbance variance. With a nonlinear activation function, MLP is genuinely nonlinear in the parameters and therefore takes more computer time to fit than kernels or splines (White, 1992; Eubank, 1994).

The Kernel method performs a local averaging of the observations when estimating the regression function. Crucial for the kernel regression estimation is the choice of global or local bandwidths. Too small bandwidths will lead to a wiggly curve while too large ones will smooth away important details.

There are different uses for MLPs with more than one hidden layer. In principal component analysis, a linear model for dimensionality reduction in which inputs and targets are the same variables, one hidden layer is insufficient for improving upon principal components. A nonlinear generalization of principal components can be obtained with the addition of two hidden layers, where the second layer serves as a bottleneck and the first and third layers provide the second nonlinearity (Sarle, 1994).

Cascading a group of single layer neurons forms multilayer-perceptron models. In a three-layer network, for example, there is an input layer, an output layer, and a hidden layer wherever intermediate nonlinear transforms occur. The complexity of the model can be varied according to how many layers are incorporated. The nodes of different layers are densely interconnected through direct links. At the input layers, the nodes receive the values of input variables and multiply them through the network, layer by layer.

The middle layer nodes are often characterized as feature-detectors. The number of

hidden layers and the number of nodes in each hidden layer can be selected arbitrarily, but too many nodes in the middle layer produce a neural network that merely memorizes the input data and lacks the ability to generalize. Most commonly, the hidden layer includes at least 75% of the number of input nodes. The initial weights of the connections can be chosen randomly. In thousands of iterations, the neural net's weights shift subtly, allowing it to learn (memorize) the patterns and recall them in the future (Baets, 1994). The random choosing of the initial conditions generates a state of indeterminacy in terms of possible patterns or outcomes. Specifically, from an observer's perspective, the process is not observationally determinable (Krippendorff, 2002).

The computed output is compared to the known output. If the computed output is "correct" (identical), then nothing more is necessary. If the computed output is "incorrect", then the weights are adjusted to make the computed output closer to the known output. This process is continued for a large number of cases, or time-series, until the net gives the correct output for a given input. The entire collection of cases learned is called a "training sample" (Anderson, 1995; Anderson and Rosenfeld, 1998; Kecman, 2001).

ENNs are programmed to learn up to a given threshold of error. After the neural network learns up to the error threshold, the weight adaptation mechanism is turned off, and the neural network is tested on known cases it has not seen before, generating the rate of true error. In a well-trained neural network, the error threshold and the true error should be identical (Baets, 1994).

Networks themselves are not correct or incorrect; it is their outputs that are such. But networks can be analyzed for component or structural redundancy. Although well-trained neural networks have determinate error rates, the outputs of each ENN follow the indeterminacy principal. The most likely solution is never 100% accurate; thus the solutions are never exactly the same. The primary cause is the indeterminacy of the adjustment of the weight during the "training" process, which may differ each time for each ENN.

IV. ADVANTAGES AND DISADVANTAGES TO USING ARTIFICIAL NEURAL NETWORKS

Econometric models can be used as forecasting tools, but each tool requires a different method using different procedures, each of which offers specific advantages and disadvantages. ENNs have many advantages over conventional methods of analysis. First, they have the ability to analyze complex patterns quickly and with a high degree of accuracy. Second, artificial neural networks make no assumptions about the nature of the distribution of the data. They are not, therefore, biased in their analysis. Instead of making assumptions about the underlying population, ENNs with at least one middle layer use the data to develop an internal representation of the relationship between the variables. Consequently, better results can be expected with neural networks when the relationship between the variables does not fit an assumed model.

Third, since time-series data is dynamic in nature, it is necessary to have nonlinear tools in order to discern relationships among time-series (Maasoumi, Khotanzad, and Abaye, 1994). ENNs are best at discovering these types of relationships. Fourth, neural networks perform well with missing or incomplete data. Whereas traditional regression

analysis is not adaptive, indiscriminately processing older data together with new data, ENNs readjust their weights as new input data becomes available (Gilbert, Krishnaswamy, and Pashley, 2000). Fifth, compared with an econometric model, it is easier to use ENNs when a forecast needs to be obtained in a shorter period of time.

Since economic, financial and social systems are complex and subject to human reactions and counter-reactions by different agents or players, it is difficult, if not impossible, to write down a complete model with all the potential reactions and counter-reactions determined. In such complex systems, it is natural to turn to models, which emulate and simulate the economy or the society in question. This is what the neural network methodology is capable of delivering. Models are difficult to understand and are impractical to handle if one tries to control all of the possible variables and potential outcomes in the systems and account for all their dynamic interactions.

One treatment of this indeterminacy involves probability and statistics; the other is the fuzziness of input and output variables and the fact that economic and finance data arriving in different time intervals are always subject to major revisions. Using the neural network vocabulary, the transfer function, described in Section 2B, can often be multi-valued, and the weights and other (e.g., learning) parameters can at best be only interval-valued. Traditional econometrics techniques are not able to handle such models due to these indeterminacies. ENNs suggest an alternative not by solving or reducing the indeterminacy but by being able nevertheless to forecast with some degree of accuracy.

However, there are some drawbacks connected with the use of ENNs. For one, ENNs are not all-purpose problem solvers. Thus far, there is no structured methodology available for choosing, developing, training, and verifying an ENN. The output quality of ENNs may be unpredictable regardless of the design and implementation schedule. Some researchers maintain that no estimation or prediction errors are calculable when using ENNs (Caporaletti, Dorsey, Johnson, and Powell, 1994) due to constant "learning" by the process. Also, ENNs are "black boxes" because it is impossible to figure out how relations in their hidden layers are estimated (Li, 1994, Gilbert, Krishnaswamy, and Pashley, 2000). This is a perfect example of observational indeterminability. The main factor in this indeterminacy is that ENNs are autopoietic systems, producing their own patterns from a set of inputs that will be needed to operate the very network of production in the future (Krippendorff, 2002).

One may claim that regression models are generally based on the first principles and the law of motions, which has theoretical foundations (typical in physics) whereas ENN modeling is based on emulation or simulation. Simulation is much weaker from the standpoint of identification of structure and explanatory power because they accept the given structure whereas emulation tries to learn and update itself in the process. Therefore, it is important to understand what network methodology offers and fails to offer when, e.g. compared with the traditional derivation of equations of motion for a given dynamical system, based on fundamental laws.

Another drawback is that neural networks have long training times. Excessive iterations are required to train neural networks (Gilbert, Krishnaswamy, and Pashley, 2000). Reducing training time is crucial because building a neural network forecasting system is a process of trial and error; hence, the more experiments a researcher can run in a finite time period, the more confident he can be of the result. The network also tends to

base its predictions of future events on “memories” of similar situations from the past (Ruggiero 1994).

In the case of financial markets, neural networks quantify the influence of major financial variables and the impact that these relationships have on the future price movement of the target market (Gorr, 1994). The networks learn from carefully crafted training data containing such variables as interest rates, currency prices, commodity prices, the slope of the yield curve, the movement in major commodity prices, the movement in major financial averages, internal market data, technical indicators, and so forth. Neural networks are data-dependent, so the algorithms are only as good as the data shown to them (Meade, 1995). In this light, ENNs may be thought as weakly deterministic systems that converge to a predictable eigen-behavior (Krippendorff's, 2002).

After the learning process, when given specific input data, it may be possible to predict a general solution that would be produced by an ENN. Since the future is often different from the past in financial markets, the user needs to test the predictions on out-of-sample ranges to ensure the reliability of the network. The results are generally indeterministic. Kanas (2001) used out-of-sample data of monthly returns forecasts for Dow Jones and the Financial Times stock indices, using both a linear and neural network model. Neither model performed well in predicting directional changes in the two indices, but they did help to support the conclusion that the underlying relationship between stock prices is not linear (Kanas, 2001).

ENNs may become a bit overzealous and undertake to fit a curve to some data even when there is no relationship. They tend to under- or over- fit data (Gilbert, Krishnaswamy, and Pashley, 2000). It is always possible to build a neural net or a mathematical function that exactly fits all the historical data such as a time series, but the predictive capability of such a system is relatively nonexistent. This over-fitting is because the noise and anomalies in the data do not allow the network to predict with any accuracy.

An ENN's rigor thus suffers from poor generalization capability. Enough repetitions will help to produce results with extremely high R-squared values, but they will have no relevance in reality. If a user relies on the results of an ENN in the belief that an ENN has high predictive power, a series of decisions will ensue that may prove disastrous. The user should always be aware that an ENN is not a strongly deterministic system that contains explicit cause-and-effect relationships (Ruggiero, 1995).

However, work with hidden layer feed-forward networks is progressing towards limiting this problem (Corradi and White, 1995). In feed-forward networks, the weight flows from the inputs in the forward direction, never cycling back to an input node of an output node of a previous layer. Although restricted to one direction each, the choices of the weight flows are interactive and often indeterminate within each individual network.

Using already smoothed variables, which give the system less incentive to fit its own curve to the data, can reduce problems. It is also extremely important to use only economically significant variables for inputs. Every user will have to define what he considers the most economically significant variables. Studies show that using a few well-chosen variables will give significantly better results than trying to use every economic indicator as a viable predictor (Schwartz, 1995).

Studies show that networks produce the best results when used in conjunction with an expert. The expert prompts the neural network at key decision nodes, and he or she

allows the user to enter their opinion as to the weight, or importance, of a specific variable. By removing highly correlated signals that have a fairly direct relationship between two inputs, the user can assess more correctly the predictive power of the inputs still present and thereby construct a better model. With such checked interaction, expert-guided networks command greater predictive power. These interactive choices present a form of indeterminacy, which results from strategic interaction, differing each time for each expert.

V. APPLICATIONS OF ENNs TO BUSINESS, ECONOMICS, AND FINANCE

Emulative neural networks are used in a widening range of applications, including airline security control, investment management and risk control (Brockett, Cooper, Golden, and Pitakong, 1994), industrial management and production (Davies, Thomas, and Shaw, 1994; Satake, Morikawa, and Nakamura, 1994; Eberts and Habibi, 1995), as well as in forecasting stock price indexes and derivative securities (Hutchinson, Poggio, and Lo, 1994; Li, 1994; Fish, Barnes, and Milam, 1995; Shachmurove and Witkowska, 2001), and predicting exchange rates (Kuan and Liu, 1995) and thrift failures.

ENNs have found ardent supporters among various avant-garde portfolio managers, investment banks, and trading firms. Most of the major investment banks, such as Goldman Sachs and Morgan Stanley, have dedicated departments to the implementation of neural networks. Fidelity Investments has set up a mutual fund whose portfolio allocation is based solely on recommendations produced by an artificial neural network. The fact that major companies in the financial industry are investing resources in emulative neural networks indicates that artificial neural networks may serve as an important method of financial forecasting in the future.

ENN processing has many managerial applications. The use of ENNs for production planning and controls (PPC) has potential to grow significantly because more than one function of a PPC system can be supported by an ENN. Such functions include forecasting in the framework of master-production scheduling and consumption-driven material requirement planning; determination of lot sizes for production and procurement; capacity adjustment; scheduling; and short-term intervention into running processes (Corsten and May, 1996).

Yoon and Swales (1990) compare ENNs to discriminant analysis. The technique of discriminant analysis is generally used to build a procedure that not only considers the number of correct and incorrect classifications of the data but also takes into account the cost of each type of classification. Yoon and Swales (1990) show that the prediction of stock price performance based on an ENN model is superior to prediction based on a discriminant analysis.

Surkan and Singleton (1990) also find that ENN models perform better than discriminant analysis in predicting future assignments of ratings to bonds. This may lead to an inaccurate assumption of determinacy, where - given a set of initial conditions - it may be presupposed that future prices of stocks and bonds may be predicted. However, the actual values (or ENN outcomes) fluctuate unpredictably, indicating a noticeable behavior of indeterminacy.

Trippi and DeSieno (1992) apply an ENN system to the modeling of trades in Standard and Poor's 500 index futures. They find that the dynamics of the ENN system

helps to outperform a passive approach to investment (a buy-and-hold strategy) in the index; thus, they favor the implementation of ENNs to the financial decision making process.

Donaldson, Kamstra, and Kim (1993) contribute to the study of the capital asset pricing model as well as ENNs by examining the fat tails and heteroskedasticity in stock return data. Using data from the stock indexes of London, New York, Tokyo, and Toronto, they find that ENN models outperformed many traditional models, including the autoregressive conditioned heteroskedasticity (ARCH) model, in removing leptokurtosis and symmetric and asymmetric heteroskedasticity from the stock index data. This superior capability allows the ENN model to be utilized (perhaps a little too confidently) as a deterministic tool.

Hill, Marquez, O'Connor, and Remus (1994) offer a balanced assessment of the potential of ENNs. They find emulative neural networks to be comparable to traditional statistical methods. In regards to the forecasting potential of artificial neural networks, they find that such networks perform as well as classical statistical models for forecasting yearly time-series, but that for monthly and quarterly time-series, ENNs may actually outperform statistical models. ENNs do not necessarily outperform regression in modeling human decision-making, except when nonlinear elements are involved. This is not a case of explicit determinacy as it is more likely an indication of a weakly deterministic system (Krippendorff, 2002).

Kuo and Reitsch (1996) test the accuracy of forecasts produced by both multiple regression and neural network models. They test their models on ten different out-of-sample data sets and analyze the forecasting errors of each of the models. The results indicate that ENNs outperform conventional methods in all cases.

VI. CONCLUSION

The technology of ENNs has existed theoretically for decades. However, the explosion in computational processing power in the last ten years has revived research using such networks. Developing even a simple ENN requires millions of calculations (Perry, 1994). ENNs are revolutionizing statistical computing in many fields as they not only "learn" autonomously but are also ideal for finding nonlinear relationships in data (Hoptroff, 1993). A perfect example of an autopoietic system, an ENN is the best available tool today for analyzing and even forecasting indeterministic systems.

ENN models are becoming highly desirable statistical tools for statisticians and economists who have internalized the hallmark of rational expectations, and now believe that people react to one another and to policy rules in strategic and dynamic ways. The construction of an explicitly detailed dynamic model, however, is extremely difficult, for such a model must be capable of tracing all of the potential and actual responses and interactions. Thus, models are bound to exhibit indeterminacy in the sense that it may be impossible to develop a unique solution to a given situation.

The basis of such indeterminacy is the fuzziness of input and output variables as well as possible multi-valued transfer functions, where the weights and other (e.g., learning) parameters can only be interval-valued. A model's optimal performance will result in an approximation of a solution or a most likely outcome, which is the inevitable

product of strategic interactions among variables and/or rational individuals.

The ability to deal with many processing elements makes neural computing faster than conventional computing. In addition, parallelism makes it robust and fault-tolerant in the sense that performance does not degrade significantly even if one of the nodes fails. Researchers are concluding that most economic and financial problems are nonlinear; that simple cause-and-effect relationships rarely exist; that, instead, most problems encountered are fuzzy patterns, which relate to multiple variables.

Nonetheless, as many have argued (see, for example Sarle, 1994), statistics and ENNs are not competing methodologies for data analysis. Rather, the paths of statistics and ENNs often cross. There are many useful neural network models for nonlinear data analysis, such as the MLP model, and there is room for many more applications of statistics to neural networks, especially in regard to estimation criteria, optimization algorithms, confidence intervals, diagnostics, and graphical methods.

As they do not require an exact specification of the functional equations, emulative neural systems can be applied to predict economic phenomena - especially unrecognized, unstructured, and non-stationary processes. Thus, ENNs are highly suitable for analyzing economic systems. ENNs have proven themselves to be adequate also for searching out and identifying nonlinear relationships and for pinpointing those variables that hold the highest predictive value. After extensive training, ENNs are able to eliminate substantial amounts of uncertainty in economic forecasts, although never completely overcoming indeterminacy.

The unique qualities and potential of this technology assure it a leading place in the search for a better forecasting technique. ENNs will be utilized as analytical tools in an attempt to resolve residual questions of indeterminacy. As computer power increases, processing time decreases, and personal computers become more available, the use of emulative neural networks will continue to increase.

REFERENCES

- Abdi, H., Valentin, D., and Edelman, B., 1999, *Neural Networks*, Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-124, Thousand Oaks, CA: Sage.
- Altman, E.I., G. Marco, and F. Varetto, 1994, "Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience)," *Journal of Banking and Finance*, 18(3), May, 505-29.
- Anderson, J.A., 1995, *An Introduction to Neural Networks*, Cambridge, MA: The MIT Press.
- Anderson, J.A., and Rosenfeld, E., (editors), 1998, *Talking Nets: An Oral History of Neural Networks*, Cambridge, MA: The MIT Press.
- Arrow, Kenneth J., 1951 and 1963, *Social Choice and Individual Values*, New Haven, CT: Yale University Press, first and second Editions.
- Artyushkin, V.F., 1990, "Neural Network Ensembles as Models of Interdependence in Collective Behavior," *Mathematical Social Sciences*, 19(2), April, 167-77.
- Baets, W., Venugopal V., 1994, "Neural Networks and Their Applications in Marketing Management." *Cutting Edge*, 16-20.
- Bau, Haim and Shachmurove, Yochanan, 2002, "Lack of Predictability in Deterministic Systems" in Ciprut, Jose V., Editor, *Indeterminacy: The Mapped, The Navigable, and the Uncharted*, forthcoming.
- Bishop, C.M., 1995, *Neural Networks for pattern Recognition*, Oxford: Oxford University Press.
- Brockett, P., W. Cooper, L. Golden, U. Pitaktong, 1994, "A Neural Network Method for Obtaining an Early Warning of Insurer Insolvency," *Journal of Risk and Insurance*, 61(3), September, 402-24.
- Brown, R. and Chua, L.O., 1998, "Clarifying Chaos II: Bernoulli Chaos, Zero Lyapunov Exponents and Strange Attractors", *International Journal of Bifurcation and Chaos*, Vol. 8, 1-32.
- Caporaletti, L.E., R.E. Dorsey, J.D. Johnson, and W.A. Powell, 1994, "A Decision Support System for In-Sample Simultaneous Equation System Forecasting Using Artificial Neural Systems," *Decision Support Systems*, 11, 481-495.
- Cheng, B. and D.M. Titterington, 1994, "Neural Networks: A Review from a Statistical Perspective," *Statistical Sciences*, 9(1), 2-54.
- Collins, A. and A. Evans, 1994, "Aircraft Noise and Residential Property Values: An Artificial Neural Network Approach," *Journal of Transport Economics and Policy*, 28(2), May, 175-97.
- Corradi, V. and H. White, 1995, "Regularized Neural Networks: Some Convergence Rate Results," *Neural Computation*, 7, 1225-1244.
- Dasgupta, C.G., G.S. Dispensa, and S. Ghose, 1994, "Comparing the Predictive Performance of a Neural Network Model with Some Traditional Market Response Models," *International Journal of Forecasting*, 10(2), September, 235-44.
- Davies, A., P.V. Thomas, and M.W. Shaw, 1994, "The Utilization of Artificial Intelligence to Achieve Availability Improvement in Automated Manufacture," *International Journal of Production Economics*, 37(2-3), December, 259-74.

- Domotor, Zoltan and Batitsky, Vadim, 2002, "Chaos, Complexity and Indeterminism" in Ciprut, Jose V., Editor, *Indeterminacy: The Mapped, The Navigable, and the Uncharted*, forthcoming.
- Donaldson, R.G., M. Kamstra, and H.Y. Kim, 1993, "Evaluating Alternative Models for Conditional Stock Volatility: Evidence from International Data," University of British Columbia.
- Devroye, L., Györfi, L., and Lugosi, G., 1996, *A Probabilistic Theory of Pattern Recognition*, NY: Springer.
- Eberts, R. and S. Habibi, 1995, "Neural Network-Based Agents for Integrating Information for Production Systems," *International Journal of Production Economics*, 38(1), March, 73-84.
- Eubank, R.L., 1994, "A Simple Smoothing Spline," *The American Statistician*, 48, 103-106.
- Fine, T.L., 1999, *Feed-forward Neural Network Methodology*, NY: Springer.
- Fish, K.E., J.H. Barnes, and M.W. Aiken, 1990, "Artificial Neural Networks: A New Methodology for Industrial Market Segmentation," *Industrial Marketing Management*, Oct.
- Fischer, M.M. and S. Gopal, 1994, "Artificial Neural Networks: A New Approach to Modeling Interregional Telecommunication Flows," *Journal of Regional Science*, 34(4), November, 503-27.
- Gately, E., 1996, *Neural Networks for Financial Forecasting*. New York: John Wiley and Sons, Inc.
- Gleitman, H., 1991, *Psychology*, W.W. Norton and Company, New York.
- Gilbert, E.W., C.R. Krishnaswamy, and M.M. Pashley, 2000, "Neural Network Applications in Finance: A Practical Introduction,".
- Gorr, W.L., 1994, "Neural Networks in Forecasting: Special Section: Editorial: Research Prospective on Neural Network Forecasting," *International Journal of Forecasting*, 10(1), June, 1-4.
- Gorr, W.L., D. Nagin, and J. Szczypula, 1994, "Comparative Study of Artificial Neural Network and Statistical Models for Predicting Student Grade Point Averages," *International Journal of Forecasting*, 10(1), June, 17-34.
- Granger, C.W.J., 1991, "Developments in the Nonlinear Analysis of Economic Series," *Scandinavian Journal of Economics*, 93(2), 263-76.
- Greene, William H., 2000, *Econometric Analysis*, Fourth Edition, Prentice Hall, Upper Saddle River, NJ.
- Grudnitski, G. and L. Osburn, 1993, "Forecasting S&P and Gold Futures Prices: An Application of Neural Networks," *Journal of Futures Markets*, 13(6), September, 631-43.
- Gur, Ruben C., Contreras, Diego, and Gur, Raquel E., 2002, "Function and Indeterminacy: Brain and Behavior" in Ciprut, Jose V., Editor, *Indeterminacy: The Mapped, The Navigable, and the Uncharted*, forthcoming.
- Hardin, Russell, 2002, "Indeterminacy and Basic Rationality," in Ciprut, Jose V., Editor, *Indeterminacy: The Mapped, The Navigable, and the Uncharted*, forthcoming.
- Hawley, D.D., J.D. Johnson, and D. Raina, 1990, "Artificial Neural Systems: A New Tool for Financial Decision-Making," *Financial Analysis Journal*, Nov/Dec, 63-72.

- Hill, T., L. Marquez, M. O'Connor, and W. Remus, 1994, "Artificial Neural Network Models for Forecasting and Decision Making," *International Journal of Forecasting*, 10(1), June, 5-15.
- Hoptroff, R.G., 1993, "The Principles and Practice of Time Series Forecasting and Business Modeling Using Neural Nets," *Neural Computing and Applications*, 1, 59-66.
- Hung, W.T., and Tong, H.Y., 2002, "Neural Network Modeling of Vehicle Discharge Headway at Signalized Intersection: Model Descriptions and Results," January.
- Husmeier, D., 1999, *Neural Networks for Conditional Probability Estimation: Forecasting Beyond Point Predictions*, Berlin: Springer Verlag.
- Hutchinson, J. M., T. Poggio, and A.W. Lo, 1994, "A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks," April.
- Jain, Anil K. and Mao, Jiamchang, 1996, "Artificial Neural Networks: A Tutorial," *IEEE Computer*, March, 31-44.
- Joerding, W.H., Y. Li, and D.L. Young, 1994, "Feed-forward Neural Network Estimation of a Crop Yield Response Function," *Journal of Agricultural and Applied Economics*, 26(1), July, 252-63.
- Kaastra, I. and M.S. Boyd, 1995, "Forecasting Futures Trading Volume Using Neural Networks," *Journal of Futures Markets*, 15(8), December, 953-70.
- Kanas, Angelos, 2001, "Neural Network Linear Forecasts for Stock Returns," July.
- Kecman, V., 2001, *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*, Cambridge, MA: The MIT Press.
- Kollias, C. and A. Refenes, 1996, "Modeling the Effects of Defense Spending Reductions on Investment Using Neural Networks in the Case of Greece," *Center of Planning and Economic Research*, Athens.
- Krippendorff, Klaus, 2002, "Four (In)Determinabilities--not One" in Ciprut, Jose V., Editor, *Indeterminacy: The Mapped, The Navigable, and the Uncharted*, forthcoming.
- Kuan, C.M. and H. White, 1994, "Artificial Neural Networks: An Econometric Perspective," *Econometric Reviews*, 13(1), 1-91.
- Kuan, C.M. and T. Liu, 1995, "Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks," *Journal of Applied Econometrics*, 10(4), Oct.-Dec., 347-64.
- Levine, D. S., 2000, *Introduction to Neural and Cognitive Modeling*, 2nd ed., Lawrence Erlbaum: Hillsdale, N.J., 2000.
- Li, E.Y., 1994, "Artificial Neural Networks and Their Business Applications," *Information and Management*, Nov.
- Maasoumi, E., A. Khotanzad, and A. Abaye, 1994, "Artificial Neural Networks for Some Macroeconomic Series: A First Report," *Econometric Reviews*, 13(1), 105-22.
- Meade, N., 1995, "Neural Network: Time Series Forecasting of Financial Markets," *International Journal of Forecasting*, December, v11n4, 601-602.
- Michie, D., Spiegelhalter, D.J., and Taylor, C.C., 1994, *Machine Learning, Neural and Statistical Classification*, Ellis Horwood.
- Ormerod, P., J.C. Taylor, and T. Walker, 1991, "Neural Networks in Economics," in Taylor, M.P., ed. *Money and Financial Markets*, Cambridge, Mass. and Oxford: Blackwell, 341-53.

- Perry, William G., 1994, "What Is Neural Network Software?" *Cutting Edge*, September, 12-15.
- Ripley, B.D., 1993, "Statistical Aspects of Neural Networks," in Barndorff-Nielsen, O.E., J.L. Jensen, and W.S. Kendall eds., *Networks and Chaos: Statistical and Probabilistic Aspects*, London: Chapman and Hall.
- Ripley, B.D., 1996, *Pattern Recognition and Neural Networks*, Cambridge: Cambridge University Press.
- Refenes, A., 1995, *Neural Networks in the Capital Markets (Ed.)* Chichester, England: John Wiley and Sons, Inc.
- Refenes, A.N., C. Kollias, and A. Zaprani, 1995, "External Security Determinants of Greek Military Expenditure: An Empirical Investigation Using Neural Networks," *Defense and Peace Economics*, 6(1), 27-41.
- Ruggiero, M.A., 1994, "Training Neural Nets for Inter-market Analysis," *Futures: The Magazine of Commodities and Options*, August, v23n9, 42-44.
- Ruggiero, M.A., 1995, "Building a Real Neural Net," *Futures*, June, v24n6, 44-46.
- Satake, T., K. Morikawa, and N. Nakamura, 1994, "Neural Network Approach for Minimizing the Makespan of the General Job-Shop," *International Journal of Production Economics*, 33(1-3), January, 67-74.
- Sarle, W.S., 1994, "Neural Networks and Statistical Models," *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute, April, 1538-1550.
- Schwartz, T.J., 1995, "Mr. Spock Has Arrived," *Futures: The magazine of Commodities and Options*, January 1995, v24n1, 46-48.
- Shachmurove, Y. and D. Witkowska, 2001, "Dynamic interrelation Among Major World Stock Market: A Neural Network Analysis", *International Journal of Business*, Volume 6, Number 1, 1-22.
- Smith, P., 1998, *Explaining Chaos*, Cambridge University Press.
- Sohl, Jeffrey E., and A.R. Venkatachalam, 1995, "A Neural Network Approach to Forecasting Model Selection", *Information and Management*, 29, 297-303.
- Stone, M., 1989, "Chaos, Prediction and Laplacean Determinism", *American Philosophical Quarterly*, Vol. 26, 123-131.
- Surkan, A. and J. Singleton, 1990, "Neural Networks for Bond Rating Improved by Multiple Hidden Layers," *Proceedings of the 1990 International Joint Conference on Neural Networks*, 4, 157-162.
- Swanson, N.R. and H. White, 1995, "A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks," *Journal of Business and Economic Statistics*, 13(3), July, 265-75.
- Trippi, R.R. and D. DeSieno, 1992, "Trading Equity Index Futures with a Neural Network," *Journal of Portfolio Management*, 19(1), Fall, 27-33.
- Wang, F., 1994, "The Use of Artificial Neural Networks in a Geographical Information System for Agricultural Land-Suitability Assessment," *Environment and Planning*, 26(2), February, 265-84.
- White, H., 1992, *Artificial Neural Networks: Approximation and Learning Theory*, With A. R. Gallant, Cambridge and Oxford: Blackwell.
- Worzala, E., M. Lenk, and A. Silva, 1995, "An Exploration of Neural Networks and Its

- Application to Real Estate Valuation,” *Journal of Real Estate Research*, 10(2), 185-201.
- Yoon, Y. and G. Swales, 1990, “Predicting Stock Price Performance,” *Proceeding of the 24th Hawaii International Conference on System Sciences*, 4, 156-162.