ON THE PREDICTION OF CHAOTIC DYNAMICS WITH ARTIFICIAL INTELLIGENCE TECHNIQUES

CIOBANU DUMITRU
PH.D. STUDENT, UNIVERSITY OF CRAIOVA,
e-mail:ciobanubebedumitru@yahoo.com

Abstract
Because of sensitive dependence on initial conditions (SDIC), characteristic to chaotic systems, the prediction of such system can be made with an accepted accuracy only for relatively small number of steps ahead. Using artificial techniques like neural networks and support vector machine to predict chaotic dynamics present advantages over traditional methods and usually they offers better results. In this paper, I highlight some of these advantages.

Keywords: Chaos Theory, Artificial Intelligence, Time Series, Exchange rate.

JEL Classification: C22, C32, G17.

1. Introduction

Chaotic systems are in fact complex deterministic systems with a large number of variables that influence the evolution of the process making it impossible for humans to simulate it and therefore making them seems unpredictable. This, also, makes it impossible to determine the initial state of the system knowing just the final state.

Most processes and systems that are found in nature involve the interaction of many factors, which allow us to catalog them as chaotic systems. Thus, chaos is met in: solar system dynamics, evolution of populations, the weather, chemical reactions, etc.

The economy can be seen as a chaotic system, a factor that brings a huge number of variables is direct involvement of people. The chaos from complex systems is known as deterministic chaos.

For the chaotic systems, the fact that they are deterministic does not make them predictable. However, the predictive power in the case of chaotic systems can be improved and can be illustrated by weather system for which predictions for short periods have reached to a very good accuracy.

We must emphasize the fact that the emergence and development of chaos theory could not have to take place before the invention of computers as simulation of complex systems with many variables could not have done without their help.

An important feature of chaotic systems is Sensitive Dependence on Initial Conditions (SDIC). This tells us that two initially close trajectories depart exponentially in a finite number of iterations, sometimes very quickly. In such a system, prediction is impossible except maybe the prediction for very short periods. The most used tool for identifying these processes from dynamical systems theory or experimental series is Lyapunov characteristic exponent (LCE).

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 [16].

There are a large number of linear forecasting models, such as moving average, exponential smoothing, time series regression and time series decomposition.

One of the most important and popular linear models is Autoregressive Integrated Moving Average - ARIMA, which was popularized by Box and Jenkins in the 1970s [2]. Often, ARIMA is also called Box-Jenkins model.

Although ARIMA models are quite flexible in modeling a wide range of time series models, their major limitation is given by the assumption of a linear form for the model. This means that a linear autocorrelation structure is assumed before the model to be according to the data. Therefore, an ARIMA model cannot capture nonlinear models that are normally encountered in the time series from economics and business. Approximation with linear models of complex real-world problems is not always satisfactory, as reflected in the well-known M-competition in the early 1980s [12].

Artificial Intelligence (AI) provides an alternative to conventional methods of prediction. Two of them we will present in this paper, namely Neural Networks (NN) and Support Vector Machines (SVM).

2. Strengths and weaknesses of Neural Networks

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be use to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A
A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions.

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 [1].

There are some downsides 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. Commercial-grade neural networks today have effectively eliminated overtraining through “bootstrapping holdout (test) samples”, and by monitoring test versus training errors [10].

Another issue with neural networks is training speed. Neural networks require many passes to build. This means that creating the most accurate models can be very time consuming [8]. However, the speed of most current machines is such that this is typically not much of an issue.

The major issues of concern today are the scalability problem, testing, verification, and integration of neural network systems into the modern environment. Neural network programs sometimes become unstable when applied to larger problems.

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. In addition, 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.

3. Strengths and weaknesses of Support Vector Machines

Support vector machines are simple concept, but very powerful, very well behaved in comparative tests with other popular classifiers [13], [14] and have been successfully applied for problems in many fields.

Some examples of applications in that support vector machines have proven their superiority are: identifying images, medical image classification, face recognition and visually speech recognition.

Besides solving some problems that many learning methods are facing, such as small samples, overtraining, large dimensions and local minimum, support vector machines have shown a power of generalization (in the case of support vector classification - SVC) or prediction (if support vector regression - SVR) better than artificial neural networks.

Unlike neural networks, support vector machines have far fewer parameters to be set which makes it easier to determine a suitable structure for a studied problem.

Support vector regression has emerged as an adaptation of support vector classification to forecasting. Support vector regression has the same advantages as classification and finds increasingly more applications.
Chen, Jeong and Hardie have conducted a comparative study for predicting financial gains in which the support vector regression obtained better results in estimating an ARIMA model over the MLE (Maximum Likelihood Estimation) and recurrent MLP (Multi Linear Perceptron) [3]. Although the results of support vector machines, in both classification and regression, are very good, we cannot say that exceed in any circumstances other methods.

Support vector machines (SVMs) are promising methods for classification and regression, due to solid mathematical foundation provides several important properties that other methods do not possess. These aspects allow us to consider that the combined methods are the key to achieving superior results. Development of algorithms that perform grouping before determining support vector is an example of such approaches and the work of Yu, Yang & Han highlights this [15].

4. Case study. Predicting the exchange rate EUR-LEU.

Time series used is the euro-leu exchange rates and was downloaded from the website of National Bank of Romania at [http://www.bnr.ro/Raport-statistic-606.aspx](http://www.bnr.ro/Raport-statistic-606.aspx). The program used to perform simulations was Matlab version 7.12.0 (R2011a).

![Fig. 1. The evolution of exchange rate euro-leu, 1845 observations.](image)

In (Fig. 1) is shown the evolution of exchange rate between euro and leu for 1845 observations. From this, I kept 100 observations for tests and I used the other 1745 for training the networks.

I trained a NAR (nonlinear autoregressive) Neural Network that use nine past values of the exchange rate and predict the next value of the time series. The structure of the network is presented in (Fig. 2) and consist in 100 hidden neurons with sigmoidal activation function and an output neuron with linear activation function.

![Fig. 2. The structure of one network used for prediction.](image)

I used the neural network to predict the next 100 values. First, I used as inputs the observed values, introducing nine measured values; the neural network provides our next prediction for the exchange rate. The results were satisfactory, see (Fig. 3) and (Fig. 4), with \( \text{MSE} = 5.2317 \times 10^{-5} \) for predictions and \( 5.0690 \times 10^{-5} \) for mean predictions.

This was not the only case in which the mean predictions have returned a mean squared error better than the predictions.
Using neural networks to predict the exchange rate is a good alternative to traditional predictive methods. The fact that predictions for a longer period not working is not a minus of neural networks over other methods but underlines the chaotic nature of time series euro-leu. Due to the chaotic nature of exchange rate time series, prediction for several steps, theoretically, would be possible only with a very complicated model. There is not such a model yet so other methods are also unable to make such predictions.
I built a support vector machine with time delay (autoregressive). From the dataset of 1845 observations we used for training 1745 and kept 100 observations for tests.

I made 100 predictions one step each time that is I used 8 known values and support vector machine was used to predict the next value.

In (Fig. 6.) the observed values of exchange rate are represented in black and with dotted red are the predictions obtained with support vector machine.

(Fig. 7.) and (Fig. 8.) represent the errors obtained for the predictions made on 100 observations that were not used in training process. These figures plot the same values but using a different scale on the Oy axis, the second representation shows that the size of errors varies in a certain band.
I measured performance of support vector machine using MSE (Mean Squared Error). For the training data used in the considered example I obtained $\text{MSE}=2.5059 \times 10^{-4}$. If the case of 100 observations kept for testing the returned mean square error was $\text{MSE}=5.7754 \times 10^{-5}$.

5. Conclusions

Approach to modeling nonlinear time series is probably more adequate for most real-world problems. 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 models 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.

In economy, the majority of historical data are available as time series. Detecting chaotic nature of the processes that have provided such data is important from this regard to establish a correct prediction horizon.

Using artificial intelligence techniques for predicting chaotic dynamics has several advantages over traditional approaches:

- Not need to specify a model.
- The model is derived from existing data.
- The model may change over time.
- Predict outcomes with accuracy for complex problems.

6. Bibliography