

GDP and Unemployment Rate in Turkey: An Empirical Study using Neural Networks

¹Mounir Ben Mbarek, ²Rochdi Feki

¹ PhD, Department of Economics, University of Management and Economic Sciences of Sfax, Street of Airport km 4.5, LP 1088, 3018 Sfax, Tunisia
E-mail: mounir_benmbarek@yahoo.fr

² Professor, Department of Economics, University of Management and Economic Sciences of Sfax, Street of Airport km 4.5, LP 1088, 3018 Sfax, Tunisia
E-mail: rochdi.feki@escs.rnu.tn

ABSTRACT

In this paper, we analyze the forecasting performance of artificial neural networks (ANN) models to forecast the GDP and the unemployment rate (UR) in Turkey. Our object is to compare between two econometric techniques: the ANN and the classical econometric techniques particularly the vector autoregressive model (VAR) in terms of their performance. The results show that the ANN models are often more robust than the VAR model.

Keywords: *Artificial Neural Networks, VAR model, Forecasting.*

JEL Classification: C45, C53

1. INTRODUCTION

Finding an effective tool to predict the economic variables constitutes a major problem among the existing problems in the econometric studies. Indeed, recent researches suggest that artificial intelligence techniques, such as the ANN, are useful in predicting the uncertainties of financial and economic variables. It is difficult to predict some variable, such as the exchange rate, using the traditional statistical methods [11]. The applications of the ANN also applied successfully to the macroeconomic variables such as the economic growth [9]. The applications to macroeconomics are quite historic and still considered to be at the frontier of the economic empirical methods. [7] noted that the technique of ANN is an effective alternative compared to the existing methods as they can detect non-linear dependencies among the data set. [10] conducted a study in which he compared the time series models (ARIMA, exponential smoothing), linear models and models of ANN. His results showed that the error values of the ANN model were lower than those of other models especially in the GDP forecasts. In other words, making predictions of ANN is more efficient. [8] studied the nonlinear relationship between investment and GDP in the Roman economy. The results show that the technique presents a neuronal good performance on the data set. [4] studied the Malaysian GDP forecast using learning based on economic indicators. In this study, the author compared the neural networks and econometric approaches and showed that the ANN had better results in forecasting the GDP. [12] used the hybrid approach which is the combination of ARIMA and ANN models. He got better results with this approach. So, there are many studies which compared the ANN and the other econometric methods in terms of their performance.

A comparison the ANN model and the VAR model remains a topical issue since the approach, in the recent years, has become quite a common tool for the macroeconomic analysis. [3] examined the relationship between the GDP, investment in construction and equipment investment in the United Kingdom using cointegrated VAR without imposing restrictions. Cointegrated VAR model, also called the method of vector error correction (VECM), is used for the analysis of nonstationary time series. [1] studied the variation of inflation and the relative prices against remittances to Mexico starting from the generalized impulse responses and some estimates using the VAR model.

The use of the VAR approach, for various macroeconomic studies, has been criticized by [5] and [2]. These authors criticize the common use of the VAR approach, but note that the method has many uses. Leamer.E (1985) recognizes that the VAR model is useful for prediction or as a descriptive device without underlying theoretical framework. For the VAR analysis, he suggests that the economic significance of variables must be economically justified. Structural equation models and the VAR model were compared by [6] using data sets from the manufacturing sector of the Italian economy. Manera found that structural models have better results than the VAR with the substitutability of the factors of production in the long term.

Therefore, we will present the simplest types of ANN that are related to standard econometric techniques. In this article, parallels are drawn between the ANN and the econometric methods to facilitate the understanding of the readings made in econometrics. A better understanding of ANN can help economists to decide the appropriateness of using these models for economic and financial forecasts and the exceed other techniques such as the model of the autoregressive vector (VAR).

This article will focus on the comparison of the ANN with the other classical forecasting techniques (especially the VAR model) in terms of performance using the GDP and UR series Turkey.

2. THE DATA

In this article we will focus on the GDP series and the annual unemployment rates in Turkey. The data used in this article were taken from "the International Monetary Fund (IMF)" in which we studied a series of annual GDP and the unemployment rate in Turkey (UR). Our data from the year 1980 until the year 2010, GDP is measured in national currency (Billions), and the unemployment rate (Percent of total Labor Force). As shown in the following table 1:

Table 1. Descriptive statistics

	GDP	UR
Mean	62.82484	8.663419
Median	59.18300	8.198000
Maximum	105.6800	14.02800
Minimum	30.48700	6.497000
Std. Dev.	22.82943	1.756544
Skewness	0.382082	1.119404
Kurtosis	2.055204	3.989358
Jarque-Bera Probability	1.907256 0.385340	7.738489 0.020874
Sum	1947.570	268.5660
Sum Sq. Dev.	15635.48	92.56338
Observations	31	31

3. THE MODELS

3.1. The ANN model

The design of an ANN model is to choose a number of parameters before any application of the model. First, we have to choose a number of layers, usually for forecasting problems. We create a model of three layers: input, hidden and output. Then, we define the number of neurons in each layer, which will be optimized during model selection.

Thereafter, we have to choose the appropriate learning algorithm and the parameters related to the learning rate.

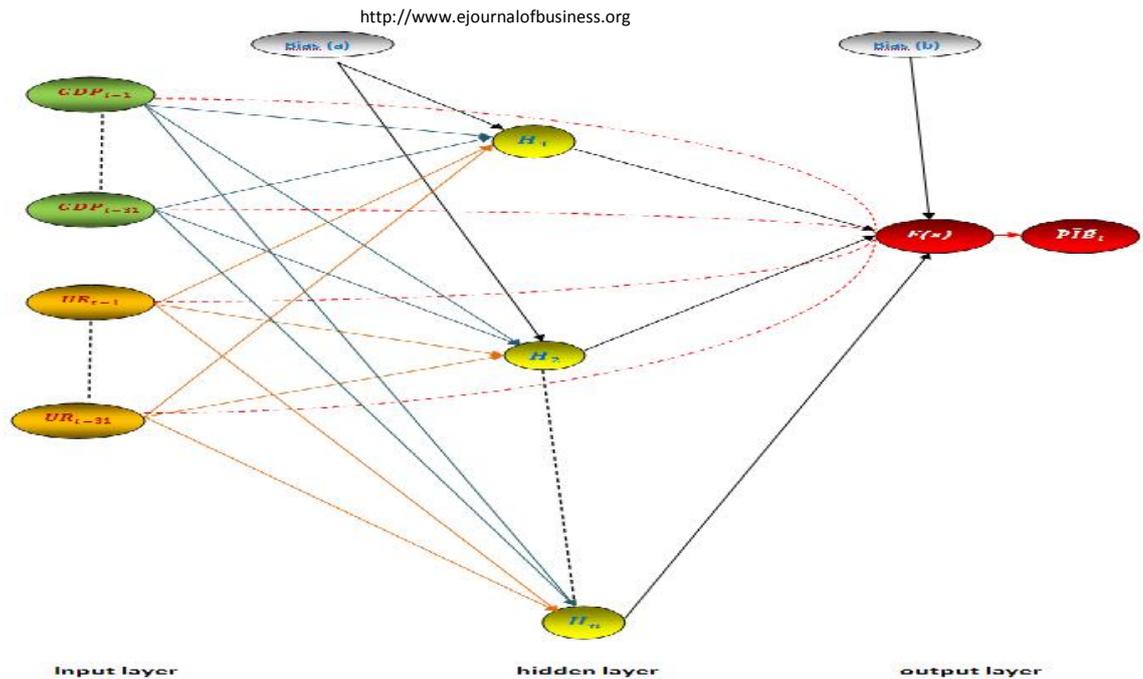
The choice of the transfer or activation functions of the hidden and output neurons is also of great importance for the model architecture. Generally, we use a sigmoid non-linear function at the hidden layer, and we proceed to linear functions at the output layer. The number of neurons in the output layer has a very important meaning either for the (GDP) and the (UR) in our case. It is about to identify the provisional horizon by selecting the number of neurons in the output layer.

The main purpose of optimizing the architecture of an ANN is to de the number of weights and neurons in order to reduce the network complexity and improve the computation time. Indeed, if we uses for example fewer neurons, we will be obtain poor performance in learning and generalization, and on the other hand, if we use a very large number of neurons, it will support a limit of the generalization capacity, what we call learning by heart.

Determining the number of inputs and outputs of the network in this example is a simple process. Our neural network will contain 62 entries (31 observations for each explanatory variable (or, 2 variables)).

In the current example, an architecture with sometimes only two hidden cells was selected and some other times more cellules. Although this architecture is probably too simplistic to capture the complexity of the present problem, this simplicity assists the understanding of the dynamics of the model. The purpose of this paper is thus to give a concrete example of a network, rather than develop a high-performance tool for forecasting. Therefore, the network presented here should not be considered as the best model the neural networks could offer.

(Fig A) illustrates the network architecture, It is a fully connected network, ie. all the inputs are connected to all of the hidden units. Polarized limits were included for the hidden units and the output unit. In addition, direct connections were added between the inputs and outputs, which show an augmented ANN. As explained below, and to implement this method, the sample was divided into three separate phases: a training set, a set of validation and a test set.



(Figure 1: A neural network to predict the increased GDP)

It is therefore an ANN with a single hidden layer having a varied number of cells, if we consider an activation function of sigmoid type, then we can write:

$$\widehat{GDP}_t = b_0 + a_{1,GDP}GDP_{t-1} + \dots + a_{n,GDP}GDP_{t-n} + \frac{b_1}{1 + e^{-(a_{01,GDP} + a_{11,GDP}GDP_{t-1} + \dots + a_{n1,GDP}GDP_{t-n})}} + \frac{b_2}{1 + e^{-(a_{02,GDP} + a_{12,GDP}UR_{t-1} + \dots + a_{n2,GDP}UR_{t-n})}}$$

$$\widehat{UR}_t = b_0 + a_{1,UR}UR_{t-1} + \dots + a_{n,UR}UR_{t-n} + \frac{b_1}{1 + e^{-(a_{01,UR} + a_{11,UR}GDP_{t-1} + \dots + a_{n1,UR}GDP_{t-n})}} + \frac{b_2}{1 + e^{-(a_{02,UR} + a_{12,UR}UR_{t-1} + \dots + a_{n2,UR}UR_{t-n})}}$$

3.2. The VAR model:

In this section we give an overview of more statistical procedures that are necessary for a VAR approach. The first step of this approach is to determine if the data series are stationary or not, If they are not stationary and integrated of the same order therefore the cointegrated (VAR) model must be adapted (also called the method of vector error correction (VECM)). The determination of the appropriate number of lags based on the Akaike and Schwarz criteria. To investigate the exogeneity of the variables we use the Granger causality test. In the case of VAR model, each equation can be

estimated by OLS, independently of each other (or the likelihood method). Since there are a lot of factors to estimate a VAR model, it is preferable to perform a test of causation before attempting to estimate the VAR model. This will eliminate the model to estimate the variables that do not intervene in the dependent variable.

Either the VAR (p).

$$GDP_t = a_{GDP} + \sum_{i=1}^p b_{GDP,i}GDP_{t-i} + \sum_{i=1}^p c_{GDP,i}GDP_{i-t} + \epsilon_{GDP,t}$$

$$UR_t = a_{UR} + \sum_{i=1}^p b_{UR,i}GDP_{t-i} + \sum_{i=1}^p c_{UR,i}UR_{i-t} + \epsilon_{UR,t}$$

4. EMPIRICAL RESULTS

4.1. The estimation of an ANN model

We used two sets of the GDP and UR to find the ANN can reproduce well the two series. We used the Matlab version 7.11.

Various tests have been done by 70% for training, 15% for validation and 15% for the estimate, every time we varied the number of delays and the number of cells in the hidden layer, the MSE varies, we obtain different results as shown in the table 2 below:

Table 2. choice of neural network architecture

Number of hidden neurons	Number of delays	Mean squared error: MSE			Number of delays	Mean squared error: MSE		
		Training	validation	testing		training	Validation	testing
(2)	1	4.55317	2.39037	4.13468	6	6.38015	85.97206	207.15141
	2	3.10766	4.43487	7.25906	7	2.05635	1.76235	28.53869
	3	1.34035	8.72208	0.807128	8	92.5644	21.62707	92.17826
	4	0.609698	17.42576	0.0109885	9	2.21177	16.24088	22.37271
	5	5.06511	6.68640	3.59894	10	0.743187	3.90970	84.83808
(3)	2	1.14728	4.16102	2.02565	-	-	-	-
(4)	2	1.01858	1.38924	4.81816	-	-	-	-
(5)	2	1.71266	8.93524	4.22430	-	-	-	-
(6)	2	42.57558	8.00583	63.46563	-	-	-	-
(7)	2	0.349688	2.32030	7.10429	-	-	-	-
(8)	2	150.2622	2455.72445	13574820	-	-	-	-

We chose this model from a set of models based on the mean square error (MSE) (MSE). This is an ANN model of order (4) with (2) cells in the hidden layer, as shown in the following output, so this is a ANN(4,2).

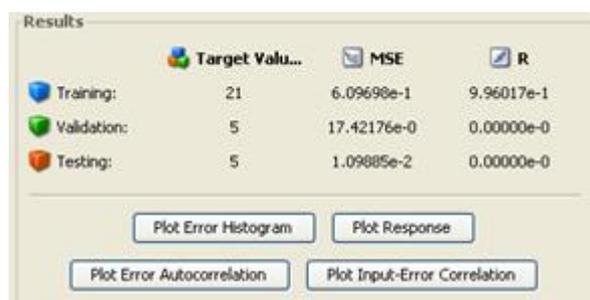


Figure 2. An attempt to estimate a neural network with two hidden layers and four delays.

Among the criteria that allow us to compare the quality of different models, we find the AIC criterion (Akaike information criterion).

$$AIC = n * \log (MSE) + 2 K$$

(K is the number of variables and n the number of observations)

$$AIC = -11.23993216$$

4.2. Results of estimating a VAR model

To estimate our model, the (Eviwes 6) software was used. We first study the stationarity of the series used. The results of the Dickey-Fuller show that the two sets of the GDP and the UR are integrated with order one, so they are stationary in one differentiation.

Table 3. Unit root test (ADF and PP)

ADF						
variables	level			1 st Difference		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
GDP	4.229386	0.673356	-2.039747	-1.882498***	-5.491931***	-5.573970***
UR	0.649099	-1.090120	-2.116783	-5.377479***	-5.468634***	-5.431271***
Phillips-Perron						
variables	level			1 st Difference		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
GDP	7.657201	1.321764	-2.039747	-3.599352***	-5.494170***	-5.827484***
UR	1.361245	-0.947137	-2.116561	-5.356683***	-5.496258	-5.708760

i) : Without intercept, (ii) : with an intercept, and (iii) : with an intercept and trend.
 *, **and***: asterisks mean a p-value less than 1%, 5% and 10%.

We then treat the new differentiated series, GDP and UR.

$$\begin{aligned}
 DGDP_{2011} &= a_{GDP} + \sum_{1980}^{2010} b_{GDP,t} DGDP_t \\
 &+ \sum_{1980}^{2010} c_{GDP,t} DUR_t + \varepsilon_{GDP,t} \\
 DUR_{2011} &= a_{UR} + \sum_{1980}^{2010} b_{UR,t} DGDP_t \\
 &+ \sum_{1980}^{2010} c_{UR,t} DUR_t + \varepsilon_{UR,t}
 \end{aligned}$$

We precisely determine the number of lags attributed to the VAR model by looking at the AIC, Schwartz and log likelihood given below the table of the output estimate. We add further delays (an even number) in the model and choose the VAR model with a number of delays which check these 3 criteria (minimum for AIC and Schwartz, and maximum for the log likelihood criterion).

Table 4. Estimation of the VAR(1) model

	D(GDP)	D(UR)
D(GDP(-1))	0.032913 (0.23514) [0.13997]	0.038007 (0.07296) [0.52091]
D(UR (-1))	1.046548 (0.80774) [1.29566]	-0.094492 (0.25064) [-0.37701]
C	2.225067 (0.89566) [2.48427]	0.096693 (0.27792) [0.34791]
Determinant resid covariance (dof adj.)		7.281713
Determinant resid covariance		5.853077
Log likelihood		-107.9195
Akaike information criterion		7.856515
Schwarz criterion		8.139404

The results of the (Eviews 6) software, presented below, lead us to retain lag length (1) which provides the best estimates of the model.

4.3. Comparisons between neural network method and the VAR model

A direct method, in comparing the effectiveness of both approaches is to compare their mean square error or the lowest AIC before calculating the predictions. The AIC of the ANN is lower than of the VAR model. Then, our

prediction based on an ANN model is more efficient than make of the VAR model, as besides, the R^2 of the ANN model is superior to R^2 of the VAR model which shows that there is a good adjustment of the Explanatory variables for the ANN than the VAR model. Therefore, it can be concluded that the ANN model is more performing than the VAR.

Table 5. The AIC criterion

Model	AIC
ANN(4,2)	-11.23993216
VAR(1)	7.856515

After analyzing the estimates obtained, we can say that the ANN models have the advantage to approximating any functional dependency. The network learns and models the dependency itself without needing to be told the assumptions or restrictions. In addition, comparisons with classical statistical methods have shown that neural networks provide better resistance to noisy data (in our case the observations from 1980 to 2010).

5. CONCLUSION

In this work, the objective is to compare the performance of the ANN model with the conventional techniques, more precisely the VAR model. We exposed the essential elements to understand why, and in what cases, it is advantageous to implement the ANN. Presenting a typical application, we have tried to show, in a concrete way, what the economist can expect from this technique. If understood and used properly, the ANNs are statistical tools that adjust nonlinear functions to very general sets of points. As with any statistical method, using the ANN requires that the available data is should be sufficiently numerous and representative and the ANN are sparse approximations. The ANNs are used to model static phenomena networks (not curly) and dynamic networks (curly). It is always desirable, whenever possible, to use for the network design, the mathematical knowledge available to model the phenomenon: the ANNs are not necessarily "black boxes", while the main objective is to address the lack of an overview of the ANN with an emphasis on their relevance in economics.

An application that focuses on the prediction of the Turkish GDP series confirmed the theoretical superiority of the ANN. Indeed, the ANN enable to provide a model of which exceeds the one obtained with the VAR model.

Finally, the use of the ANN in the economic forecasting is still very rich. In addition, the basic skills of the economist may be implemented in order to obtain new

<http://www.ejournalofbusiness.org>

applications (hybrid model) beyond the technical power of the ANN, so as to use them as the basis of a heuristic rich problems and economic behavior.

REFERENCES

- [1] Balderas, J. U. and H. Nath. 2008: Inflation and Relative Price Variability in Mexico: The Role of Remittances. *Applied Economics Letters*. 15(1–3). pp. 181–185.
- [2] Cooley, T. F. and Leroy, S. F. 1985. .Atheoretical Macroeconometrics: A Critique. *Journal of Monetary Economics*. Vol 16(3), pp 283-308.
- [3] Jumah, A., and Kunst, R. M. 2008, Modeling Macroeconomic Subaggregates: An Application of Nonlinear Cointegration. *Macroeconomic Dynamics*, 12(2), 151-171.
- [4] Junoh, M. Z. H. M., 2004, Predicting GDP Growth in Malaysia Using Knowledge-based Economy Indicators: A Comparison between Neural Network and Econometric Approach, *Sunway College Journal*, vol. 1, 2004, pp. 39-50.
- [5] Leamer, E. 1985, Vector Autoregressions for Causal Inference?, in K. Brunner - A.H. Meltzer (a cura di), *Understanding Monetary Regime*, Carnegie-Rochester Conference Series on Public Policy, 22, pp. 255-304.
- [6] Manera, M. 2006, Modelling Factor Demands with SEM and VAR: An Empirical Comparison. *Journal of Productivity Analysis*, Vol. 26(2), pp. 121-146.
- [7] Morariu, N, Iancu, E, and Vlad, S, 2009. A Neural Network Model for Time-Series Forecasting. *Romanian journal of economic forecasting*. Vol. (2009) Pages: 213-223
- [8] Sman, C 2011. scenario of the romanian GDP evolution with neural model, *Romanian journal of economic forecasting*. Vol. (2011) Pages: 129-140
- [9] Tkacz, G, 1999. Neural Network Forecasts of Canadian GDP Growth Using Financial Variables. Bank of Canada, mimeo, April 1999.
- [10] Tkacz, G, 2001. Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting* vol. 17 (2001) page: 57–69
- [11] Verkooijen, W, 1996. A Neural Network Approach to Long-Run Exchange Rate Prediction. *Computational Economics*, 9, p.51-65.
- [12] Zhang, G. P., 2003. Time Series Forecasting using a Hybrid ARIMA and Neural Network Model, *Neurocomputing* 50, pp. 159-175.