Proceedings of STRATEGICA International Conference October 24-25 2024, Bucharest, Romania

Finding an Architecture of Artificial Neural Networks for Determining the Hierarchy of the Influence of Macroeconomic Indicators on GDP

Constantin ILIE
OVIDIUS University of Constanta, Romania
124 Mamaia., 900527 Constanta, RO
constantin.ilie@365.univ-ovidius.ro

Margareta ILIE
OVIDIUS University of Constanta, Romania
124 Mamaia., 900527 Constanta, RO
margareta.ilie@365.univ-ovidius.ro

doi: 10.25019/STR/2024.018

Abstract
Artificial intelligence has become indispensable in making decisions to obtain and maintain a competitive advantage. In this sense, finding the most efficient architectures and forms of AI is essential. This paper aims to determine, with a view to the subsequent application, an architecture of the artificial neural network (ANN) based on the testing and validation results of the training process. Using feedforward, the artificial neural network, with ReLu and Adam functions on macroeconomics indicators, revealed the most fitted ANN for future use. Training and testing of all the chosen ANN's architecture were considered a success in terms of accuracy and results as feature importance. The influence of indices over the GDP's evolution results in the following hierarchical order: 1. Manufacture of computer, electronic, and optical products; 2. Manufacturing as a volume index of production; 3. Energy as volume index of production; 4. Business investment, and 5. Government investment.

Keywords
Artificial Neural Network; GDP (Gross Domestic Product); Investment; Manufacture; Simulation.

Introduction

The influence of macroeconomic indices on the evolution of GDP (Gross Domestic Product) is information that can be very useful for decision-making at the national level. For a better understanding and faster accommodation to the currently changed in the wild economic environment. Evolutions of these indices are nowadays more nonlinear. Thus, researchers and technicians should use tools that can adapt to these data types. One of these tools is AI, respectively the Artificial Neural Network (ANN) branch.

The use of the ANN to model and simulate GDP evolution has a history of research, especially for predictions according to time spans. Starting with the early ANN models, the prediction of the main economic indices, including GDP, has evolved through the application of ANNs in different structures and forms, but also through the application of analysis modules that improve the understanding of time series and relationships between various variables.

The search for the best structure options for training the best-trained ANNs is, therefore, still an emerging area of research in GDP forecasting. Since the field of ANN is still in the enhancement period, its application to economic index forecasting will be characterized by developing new ANNs and feature modules to improve the results.

The paper, in accordance with the research objective, is organized into three major parts. After the *Introduction* and *Literature review*, the third part (Methodology) reveals the data and the method used for the model and simulation of the influence of indices on the GDP evolution. The fourth part (Results and Discussion) is structured into two subchapters; the first subchapter (First database) shows the results of several ANN architectures and the feature importance that those architectures have found for the training process, while the second subchapter (Second database) uses the best architecture to model and simulate the influence for Manufacturing; Energy; Manufacture of electrical equipment; Business investment; Government investment indices over the GDP and to determine their future importance. Also, in part four, some conclusions are drawn.

The authors expected the ANN to consider the MAN, ENE, and ELE more important than the BI and GI when the ANN trains itself. Also, the authors expected to see an influence hierarchy of the MAN, ENE, and ELE, with the MAN first and ENE last.

The importance of the research emerges from the determination of a structure of an ANN that can correct the model and simulate influences between macroeconomic indices over the GDP, and offer a hierarchy of that influence through the feature importance.

Literature review

From Wu et al. (2021), we can understand that typical time series data has specific rules. Therefore, their paper analyzed and processed the GDP data from 1980 to 2020. Also, they used an ANN to predict the GDP value of China in the next 5 years (2021–2025). So, they can determine that the GDP value of China in the future is still in a rising stage, which is consistent with the historical trend. Therefore, using a neural network model to forecast the GDP of time series data is effective. Tkacz (2001) found that ANN produces statistically lower forecast errors for the year-over-year real GDP growth rate relative to linear and univariate models. Yet, such forecast improvements are less notable when forecasting quarterly real GDP growth. This should be considered in adding not only time data, but also indices that have some influence over the GDP. The same authors believe that more distinct non-linearities at the longer horizon are consistent with monetary policy's possible asymmetric effects on the real economy.

Predicting the GDP can help better grasp future economic tendencies. The experimental results validate that ANN is reliable in predicting GDP and can be used for further applications in practice (Lai, 2022). Jafari-Samimi et al. (2007) said GDP is generally a significant element in macroeconomic analysis. Policymakers of a country use disparities in GDP for long-term planning. Considering the diverse economic conditions of a country, forecasting is a valuable tool to identify the variations of GDP for planning. They have shown that the ANN approach method is the best alternative to forecast the GDP for Iran. They also add that many factors can influence GDP and complicated nonlinear relations between them, so the results of traditional linear forecasting methods to forecast GDP were complicated to apply. Their research results show that using ANN for GDP forecasting can give a higher precision and has specific practical values (Bo, 2011).

The implementation of ANN studies applied in GDP forecasting has been realized since the potential of these tools was understood, as Tkacz and Hu (1999) made for the Bank of Canada. The paper sought to determine whether the forecasting performance of financial and monetary variables can be improved using neural network models. They obtained 15 to 19 percent lower root mean squared forecast errors than their linear model counterparts. In 2011, Saman studied the Romanian GDP evolution using two neural networks. One ANN considered the independent variables' short-term impact (one month) on GDP, and the second ANN evaluated the medium-term impact (one quarter). This different approach refers to the lag structure of the non-financial variables. The results show that the two neural models present suitable performance measures on the dataset. Also, the author concluded that improved forecast accuracy may capture more fundamental non-linearities between investment and financial variables and the actual output for a longer horizon. Jahn (2018) used an ANN regression model to demonstrate superior accuracy over the corresponding linear model. The work predicts the GDP growth rate of 15 industrialized economies from 1996 to 2016.

A comparative research of different types of ANN that can be applied to forecast GDP is done by Lai (2022). Thus, the results of the genetic algorithm – back-propagation neural network model, the particle swarm optimization (PSO) – Elman neural network (Elman NN) model, and the bat algorithm – long short-term memory model were analyzed. It was found that, based on the mean square error values, which were 0.0287, 0.0166, and 0.0465, respectively; the PSO-Elman NN model had the best performance in GDP prediction.

The forecast of other economic and/or financial variables can be added to the GDP forecast, as Kordanuli et al. (2016) did with the Hirschman–Herfindahl Index (HIHI). Using variables such as agriculture and industry-added value, final consumption expenditure of general government, gross fixed capital formation (investments), and fertility rate, the authors compare a back propagation (BP) and an extreme learning machine (ELM) learning algorithms for ANN and found that for those variables, the ELM shows better performances in applications of GDP and HHI forecasting.

The application of ANN for GDP forecasting can be achieved by implementing different forms and structures of networks. Longo et al. (2022) proposed a combination of a Recurrent Neural Network (RNN) and a Dynamic Factor model to forecast the aftermath of the 2008-09 global financial crisis by reducing the forecast error for the one-quarter horizon. Their research found that combining the two tools significantly improves the benchmark models for a one-quarter horizon. Hopp (2021) proposes a Long Short-Term Memory Artificial Neural Networks (LSTM), considering it well-suited to deal with economic time series. He showed that LSTMs produce superior results to dynamic factor model DFMs in nowcasting three separate variables: global merchandise export values and volumes, and global services exports. Also, more advantages include their ability to handle large numbers of input features in various time frequencies. However, their disadvantage is the inability to assign contributions of input features to model outputs, a problem common to all ANNs.

Methodology

Data

The choice of the analyzed data was based on the experience of the authors but also on the possibility of finding the correct information and values that can be used in the simulation with artificial intelligence. Thus, data used for training, validation, and implementation of the Artificial Neural Network (ANN) are as follows:

MAN = Manufacturing as production volume index, annual average, unadjusted data (i.e., neither seasonally adjusted nor calendar adjusted data), index, 2015=100 (Production in industry).

ENE = MIG - energy as volume index of production, annual average, unadjusted data (i.e., neither seasonally adjusted nor calendar adjusted data), index, 2015=100 (Production in industry).

ELE = Manufacture of computer, electronic, and optical products; manufacture of electrical equipment, annual average, unadjusted data (i.e., neither seasonally adjusted nor calendar adjusted data), index, 2015=100 (Production in industry).

BI = Business investment, annual, percentage.

GI = Government investment, annual, percentage.

GDP = Gross domestic product at market prices, current prices, million euros, annual average, Unadjusted data (i.e., neither seasonally adjusted nor calendar adjusted data).

Results

The data were recorded annually between 2000 and 2022 for Romania, extracted from the EUROSTAT website (Eurostat, 2023). From previous research, applying visual representation for the present data (Figure 1) reveals that between 2000 and 2022, the six indices had the following evolution:

- MAN has a slightly positive increase.
- ENE is almost constant with a slight decrease starting from 2009.
- ELE has the highest growth after the GDP.
- BI increased until 2009, then decreased almost slowly.
- GI has the most inflexion point of evolution, with a high rise in 2006 and abruptly decreased between 2014 and 2016.
- GDP has the highest growth.

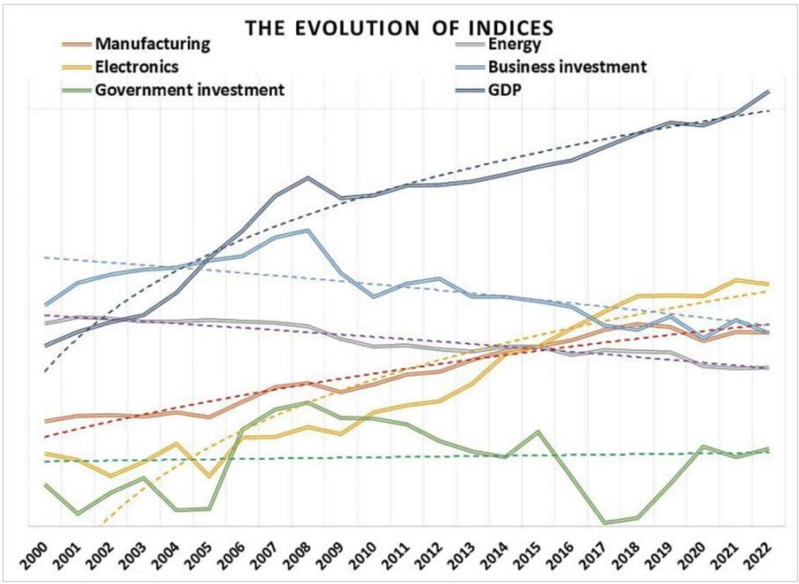
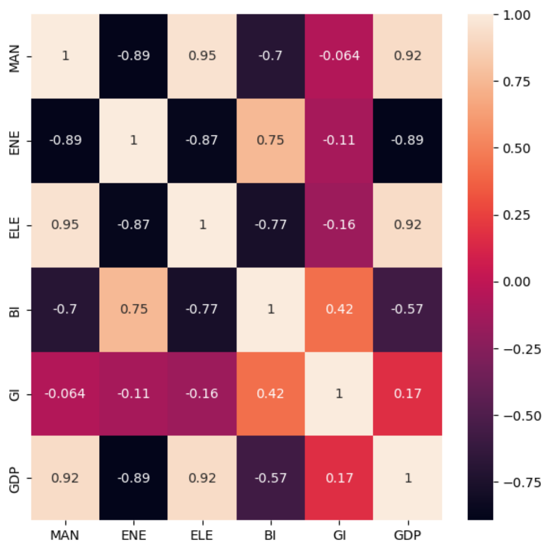
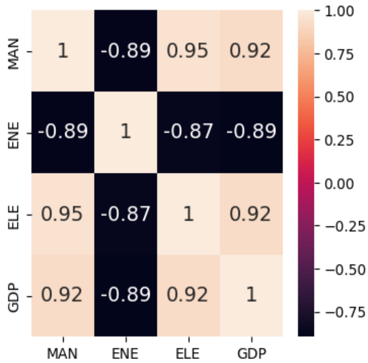


Figure 1. The evolution of indices
(Source: Authors' analysis using Excel)

Also, from Figure 1, you can see the trends (developed with linear functions). GDP, MAN, and ELE have positive trends, with the steepest increase in GDP. While GI is almost constant, ENE and BI have negative trends with almost the same decrease level.



(a)



(b)

Figure 2. Correlation matrix of indices values: a) all indices; b) indices with absolute associations greater than 0.75 on GDP
(Source: Authors' analysis using Colab Notebooks and Python)

A correlation between indices values showed some possible direction of research. Correlation summarizes the strength and direction of the linear (straight-line) association between two quantitative variables. Denoted by r, it takes values between -1 and +1. A positive value for r indicates a positive association, and a negative value for r suggests a negative association. The closer r is to 1, the closer the data points fall to a straight line; thus, the linear association is stronger. The closer r is to 0, making the linear association weaker. The results of the correlation are shown in Figure 2, where Figure 2.b emphasizes the indices with absolute associations greater than 0.75 on GDP. The correlation results are used in the second part of the Results and Discussion part (in the Second database subchapter).

Method

As an already applied and demonstrated method, for the model and simulation of the nonlinear data values evolution, we used Artificial Neural Network (ANN) as a method. Based on the previous research (Ilie, 2021), the characteristics of the present problem and the present data type require a feed-forward ANN with a backpropagation training algorithm. Also, the training uses solver Adam (an algorithm for first-order gradient-based optimization of stochastic objective functions based on adaptive estimates of lower-order moments) and ReLU (an activation function that introduces the property of non-linearity to a deep learning model and solves the vanishing gradients issue). The other structural and training data are:

- the learning rate is constant, and it has a value of 0.01.
- the batch size is 5.
- 500 iterations for error change.

Starting from the above characteristics, the best architecture of the ANN (neurons and neuron layers) was searched by comparing several structures. From the first applied (coded ARH1), 5-11-1 (5 neurons in the input layer, 11 neurons in a single hidden layer, and one neuron in the output layer) with just one hidden layer, the architecture was modified by adding hidden layers and changing the number of hidden neurons. The results of the changed architectures are discussed further in the next chapter.

The results of feature importance are considered to determine the most influential input over the training of ANN. Feature importance refers to techniques that calculate a score for all input features for a given model - the scores simply represent the "importance" of each feature. A higher score means that the specific characteristic will have a more significant effect on the model used to predict a particular variable. Feature importance scores can provide insight into the dataset. The relative scores can highlight which features may be most relevant to the target, and the converse, which features are the least relevant. A domain expert may interpret this and could be used as the basis for gathering more or different data [https://machinelearningmastery.com/calculate-feature-importance-with-python/].

The Python sklearn’s MLPRegressor and Random Forest Regressor were used for the ANN's construction, application, and evaluation. As explained by scikit-learn.or,g “MLPRegressor implements a multi-layer perceptron (MLP) that trains using backpropagation with no activation function in the output layer, which can also be seen as using the identity function as activation function. Therefore, it uses the square error as the loss function, and the output is a set of continuous values. MLPRegressor also supports multi-output regression, in which a sample can have more than one target.”. It also explains that “A random forest is a meta estimator that fits several decision tree regressors on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. Trees in the forest use the best split strategy”.

Results and discussion

First database

The best architecture was searched using all five initial input indices (MAN, ENE, ELE, BI, GI). This helped ANN to understand better, in the training process, the influences of these indices over the GDP (as output data) and also showed cum ANN understood the feature importance correctly, in accordance with the initial information. After using different structures, the results of the best are shown in Table 1. Also, in Figure 3 are presented the evolution of the errors in the training process, throughout all the iterations.

Table 1. Structural features – architecture
(Source: Authors' analysis using Colab Notebooks and Python)

Structure			**Architecture code**
Input layer	Hidden layer(s)	Output layer	
5	11	1	ARH1
5	11-3	1	ARH2
5	9	1	ARH3
5	11-7-1	1	ARH4

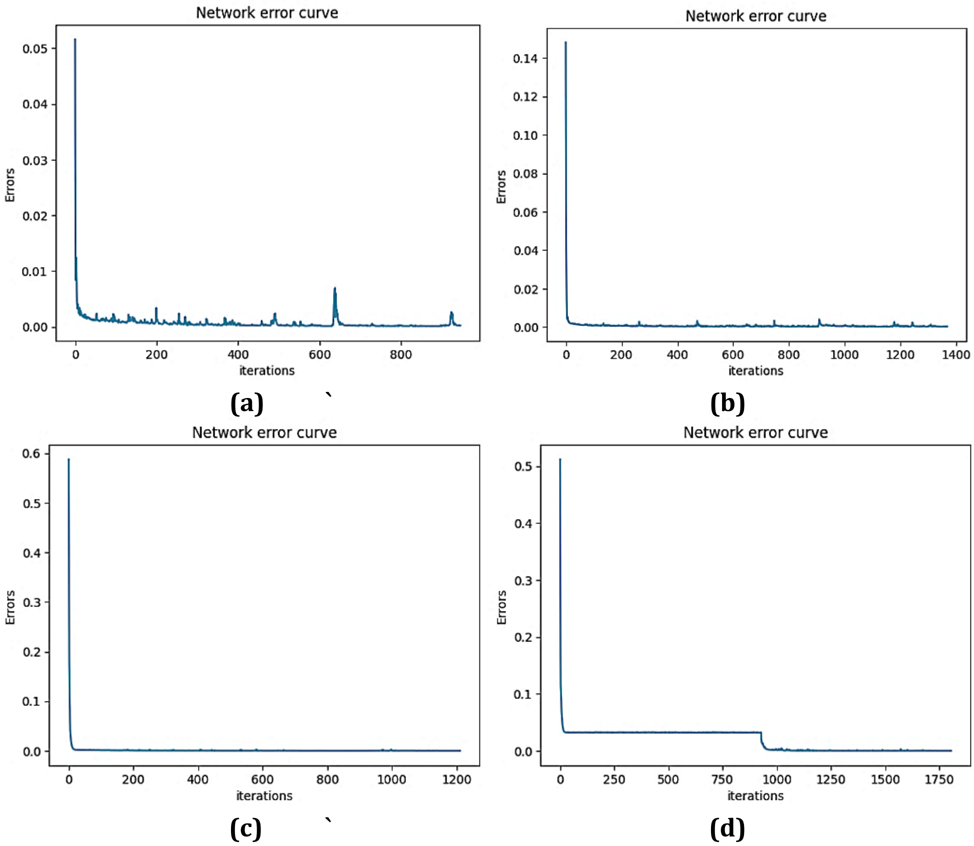


Figure 3. Network loss curve: a) ARH1; b) ARH2; c) ARH3; d) ARH4
(Source: Authors' analysis using Colab Notebooks and Python)

Analyzing the loss curve images from Figure 3, we can say that all the training have more than enough iterations for the lowest loss finding.

The comparison between the results of the training and the results of the testing for each of the four architectures are shown in Table 2, as the accuracy of the training and testing. For each of those evaluations, the mean absolute error and the mean squared error were used. The smaller the value, the better the training and testing.

Table 2. Accuracy analysis
(Source: Authors' analysis using Colab Notebooks and Python)

Architecture code	**Test**				**Train**			
	Mean absolute error		Mean squared error		Mean absolute error		Mean squared error	
ARH1	0.15370	IV	0.03672	IV	0.00998	III	0.00017	III
ARH2	*0.01786*	*I*	*0.00033*	*I*	0.01391	IV	0.00035	IV
ARH3	0.08545	II	0.00812	II	0.00733	II	*0,000084*	*I*
ARH4	0.09958	III	0.01723	III	*0.00640*	*I*	0,000089	II
		Hierarchy		Hierarchy		Hierarchy		Hierarchy

The feature importance was determined for each of them using the same four architectures, see Figure 4.

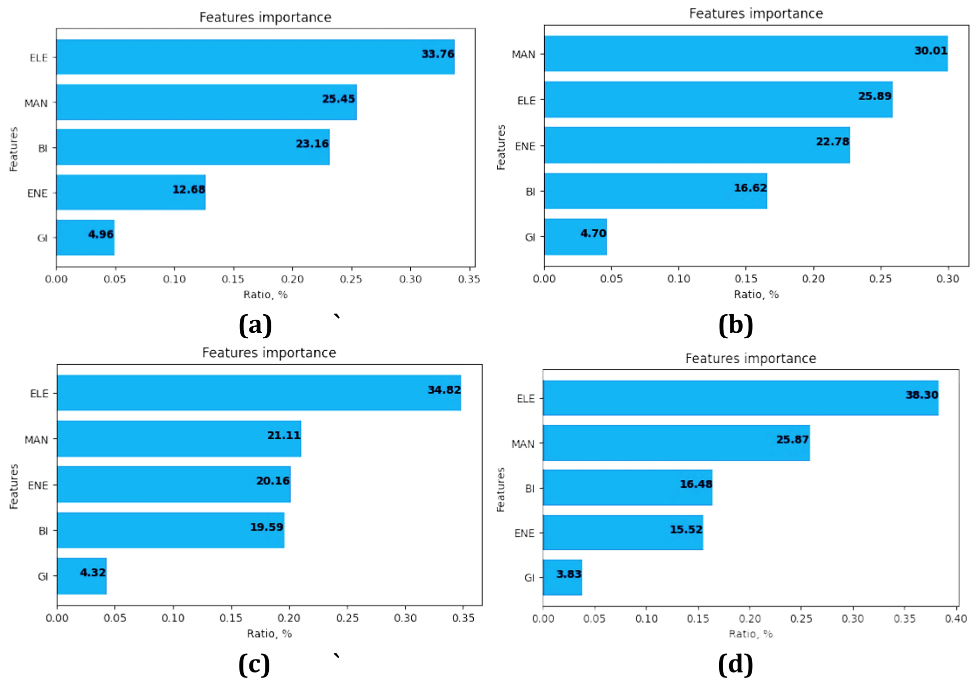


Figure 4. Feature importance: a) ARH1; b ARH2; c) ARH3; d) ARH4
(Source: Authors' analysis using Colab Notebooks and Python)

Table 3. Feature importance
(Source: Authors' analysis using Colab Notebooks and Python)

Architecture code	MAN	ENE	ELE	BI	GI
ARH1	24.54	12.68	*33.76*	23.16	4.96
ARH2	*30.01*	22.78	25.89	16.62	4.70
ARH3	21.11	20.16	*34.82*	19.59	4.32
ARH4	25.87	15.52	*38.30*	16.48	3.85
Average	25.38	17.79	*33.19*	18.96	4.46

The feature importance values shown in Figure 4 and Table 3 can be concluded by the fact that in almost all architecture (except ARH2, which considers MAN the most influential), the index with the most importance is ELE. Calculating a ratio for all the indices on all the architectures, the hierarchy is: 1. ELE; 2. MAN; 3. BI and ENE; 4. GI.

In accordance with Table 2, the best architecture for testing is ARH2; as for the training, the smallest error belongs to ARH3 and ARH4 (as absolute and squared error, respectively).

Further, we compare the Random Forest Regressor (RFR) score. Random forest is an ensemble learning algorithm based on decision tree learners. The estimator fits multiple decision trees on randomly extracted subsets from the dataset and averages their prediction. With RFR, we can check the accuracy of predicted data for each architecture (Table 4).

Table 4. Architecture score with Random Forest Regressor
(Source: Authors' analysis using Colab Notebooks and Python)

Architecture code	**Model score on testing data**		**Model score on training data**	
ARH1	0.963619	III	*0.986663*	*I*
ARH2	0.821638	IV	0.983406	III
ARH3	*0.986697*	*I*	0.985347	II
ARH4	0.980224	II	0.982159	IV

According to RFR's score, the best architecture belongs to ARH3 for testing and ARH1 for training. Considering the values from Tables 2 and 3, we chose the best architecture for the present problem – the ARH3 (5-9-1).

Second database

For further research, we considered the ARH3 to be applied for the model and simulation of only the best 3 "influencers" - MAN, ENE, ELE, named "second database." Another reason for this choice is that GDP has more direct connections with GI and indirect with BI through policies. The application of this architecture followed the same steps as in the case of all five indices used. The logic behind this was to compare the results of applying the same architecture for the two databases. Thus, in Figure in Figure 5 is the evolution of the network loss curve using the second database.

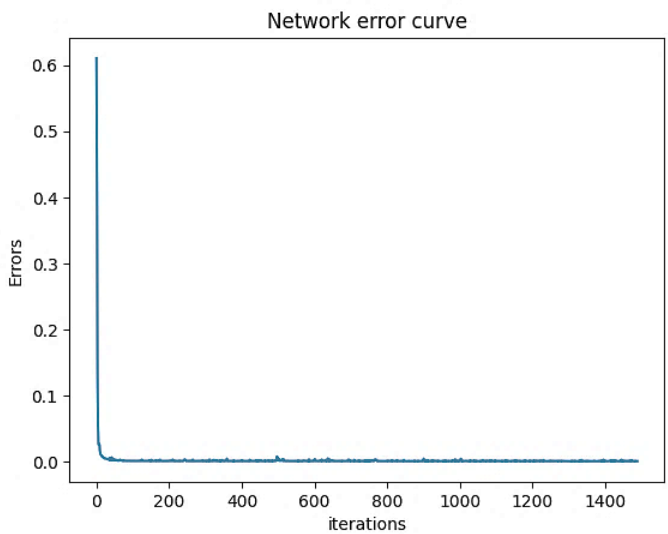


Figure 5. Network loss curve. Second database
(Source: Authors' analysis using Colab Notebooks and Python)

In Table 5, there is a comparison between the results of training and testing. The results show better training for the first database than the second, which is explained by a larger number of data, thus a more accurate possibility to understand the problem. The testing results are better for the second database model due to the smaller data sets that have been compared in Table 5.

Table 5. Accuracy. Second database
(Source: Authors' analysis using Colab Notebooks and Python)

Architecture code	Test		Train	
	Mean absolute error	Mean squared error	Mean absolute error	Mean squared error
ARH3	0.08545	0.00812	*0.00733*	*0,000084*
ARH3 – 2nd	*0.03563*	*0.00156*	0.03231	0.00211

The RFR score is better for the second base model, as can be seen in Table 6.

Table 6. Architecture score with Random Forest Regressor. Second database
(Source: Authors' analysis using Colab Notebooks and Python)

Architecture code	Model score on testing data	Model score on training data
ARH3	0.986697	0.985347
ARH3 – 2nd	*-0.254311*	*0.986479*

The feature importance was determined for each of them, see Figure 6 and Table 7.

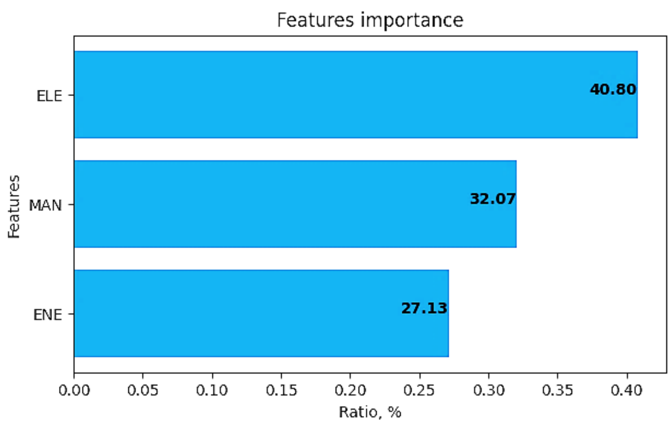


Figure 6. Feature importance. Second database
(Source: Authors' analysis using Colab Notebooks and Python)

Table 7. Architecture score with Random Forest Regressor. Second database
(Source: Authors' analysis using Colab Notebooks and Python)

Architecture code	MAN	ENE	ELE	BI	GI
ARH3	21.11	20.16	34.82	19.59	4.32
ARH3 – 2nd	32.07	27.13	40.80		

The feature importance, which in the case of the second base model doesn't incorporate BI and GI, is the same for both database models; the hierarchy is: 1. ELE; 2. MAN; 3. ENE.

Conclusions

The research objective was to determine the best architecture for the model and simulation of the influence of MAN, ENE, ELE, BI, and GI macroeconomic indices over the GDP, considering recorded data for Romania between 2000 and 2022. Based on the training and testing accuracy and the Random Forest Regressor score, the best ANN architecture was determined to be 5-9-1. For this architecture, the feature importance of input data was in hierarchical order: 1. Manufacture of computer, electronic, and optical products; 2. Manufacturing as volume index of production; 3. Energy as volume index of production; 4. Business investment, and 5. Government investment.

The ANN was trained and tested using the same architecture with only three input indices (MAN, ENE, ELE). The feature importance gave the same hierarchy as were using all five input indices: 1. Manufacture of computer, electronic, and optical products; 2. Manufacturing as volume index of production; 3. Energy as volume index of production.

The sudden decrease in the evolution of the network errors determined in the training process (Figure 3 a-d and Figure 5) demonstrates the speed with which different ANN architectures learn the influence links between independent variables and GDP and self-organize to reproduce them. It is evident that after 250 iterations the network error values have very small decreases, in some cases even difficult to determine (ARH2; ARH3; second database). This shows the efficiency of the chosen structures, and even if the problem of blocking in local minima arises, these situations were minimized by selecting the training using solver Adam and the activation function ReLu. The training and testing of all the chosen ANN's architecture were considered a success in terms of accuracy and results as feature importance.

The limits of the present work arise from the small number of data sets used (data between 2000 and 2022), only 23 sets, as found on the Eurostat website.

For future research, the number of inputs and outputs can be increased, and the algorithm can be changed. The research from the present work can be applied in its current form, even if incipient, to all countries for which the necessary data can be found. The methodology can be applied to similar Eastern European countries and compare the results for further analyses.

References

Wu, J.; Yue, S.H. (2021). Prediction of GDP in Time Series Data Based on Neural Network Model. 2021 *IEEE International Conference on Artificial Intelligence and Industrial Design (AIID)*, 20-23. 10.1109/AIID51893.2021.9456509

Tkacz, G. (2001). Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, 17, 57-69. https://doi.org/10.1016/S0169-2070(00)00063-7

Lai, H. A. (2022). Comparative study of different neural networks in predicting gross domestic product. *Journal of Intelligent Systems*, 31, 601-610. https://doi.org/10.1515/jisys-2022-0042

Jafari-Samimi, A. et al. (2007). A Comparison Between Time Series, Exponential Smoothing, and Neural Network Methods to Forecast GDP of Iran. *Iranian economic review*, 12, 19-35.

Bo, C. (2011). Application of neural networks in the GDP forecasting. *Information Technology*.

Tkacz, G., and Hu, S. (1999). Forecasting GDP Growth Using Artificial Neural Networks, *Bank of Canada Working Paper 99-3*, Department of Monetary and Financial Analysis, ISSN 1192-5434, ISBN 0-662-27537-3, https://www.bankofcanada.ca/wp-content/uploads/2010/05/wp99-3.pdf

Saman, C. (2011). Scenarios of the Romanian GDP Evolution with Neural Models, *Romanian Journal of Economic Forecasting*, 14(4), 141-162. https://ipe.ro/rjef/rjef4_11/rjef4_2011p129-140.pdf

Longo, L., Riccaboni, M., Rungi, A., (2022). A neural network ensemble approach for GDP forecasting, *Journal of Economic Dynamics and Control*, 134, ISSN 0165-1889, https://doi.org/10.1016/j.jedc.2021.104278.

Jahn, M. (2018). Artificial neural network regression models: Predicting GDP growth, *Hamburg Institute of International Economics (HWWI)*, ISSN 1861-504X, https://www.econstor.eu/bitstream/10419/182108/1/1030440859.pdf

Lai, H. (2022). A comparative study of different neural networks in predicting gross domestic product. *Journal of Intelligent Systems*, 31(1), 601-610. https://doi.org/10.1515/jisys-2022-0042

Kordanuli, B., Barjaktarović, L., Jeremić, L., Meysam, A. (2016). Appraisal of artificial neural network for forecasting of economic parameters. *Physica A: Statistical Mechanics and its Applications*. 465. http://dx.doi.org/10.1016/j.physa.2016.08.062

Hopp, D. (2021). Economic Nowcasting with Long Short-Term Memory Artificial Neural Networks (LSTM), *United Nations Conference on Trade and Development*, UNCTAD Research Paper No. 62, UNCTAD/SER.RP/2021/5, 1-21, https://unctad.org/system/files/official-document/ser-rp-2021d5_en.pdf

Ilie, M., et al. (2021). Artificial Neural Networks in Management and Business Administration, THE BEST ROMANIAN MANAGE-MENT STUDIES 2019-2020, 3, 51-67, Paperback and eBook, DOI: 10.22618/TP.LIB.BRMS2021.
https://machinelearningmastery.com/calculate-feature-importance-with-python/

https://scikit-learn.org/stable/modules/neural_networks_supervised.html