An Econometric Model of GDP using Machine Learning through an Artificial Neural Network

Shivan Prasad
Student
Detroit Country Day School
Michigan, USA

Ashok K. Sinha
UST Software India Pvt. Ltd
New Dehli, India

ABSTRACT
Forecasting economic strength from several economic indicators is the primary concern of the area of econometrics. Gross domestic product measures the current value of all goods within a country and is one of the most prominent economic parameters used to evaluate economic development. With the understanding of how economic indicators affect the economy, planners can choose to allocate resources in certain industries to boost economic growth. This study builds an econometric model of US GDP with back-propagation artificial neural network architecture. Taking data from more than the past 70 years, this model will use machine learning to understand the behavior of three economic sectors of the United States. By analyzing the interconnecting behavior of the three sectors, it can form a mathematical relationship between them and produce an output of US GDP. This model accounts for nonlinearity and can be easily adapted to include more economic indicators than just the three sectors selected, sharpening the results and allowing the study of new relationships. After the machine learning process is complete, economists can adjust the input values to examine its effect on the resulting GDP from the derived mathematical relationship. The model's current implementation is found to be very satisfactory and can be useful for the future planning of economic activity.

Keywords
Artificial Neural Network (ANN), Gross Domestic Product(GDP)

1. INTRODUCTION
Ever since the end of World War II, the US economy has remained the most dominant on the global scale. However, the United States doesn't have the same stranglehold over the international economy as it had several decades ago From leading the world with a near 50 percent share of global GDP at the outset of the post-World War II era, the strength of the US economy has dwindled and allowed other developing nations to catch up. Although they still exist as one of the strongest economic powers by comprising 24.3 percent of global GDP, the country lags behind in many other socio-economic indicators[1]. For health, the US was ranked 36thinternationally for its health index, based on life expectancy at birth [2]. In terms of education, the US is ranked 10th for its education index, which is derived from expected and mean years of schooling . For income inequality, 14.5 percent of Americans are living below the poverty line and the nation has the fourth highest income inequality in the world[1]. Once leading in most of these areas, the US no longer meets its expectation of supremacy across several socio-economic parameters. At its current rate, many economists even predict that the US economy will be surpassed in GDP by China before 2030 and possibly India by 2050[3]. In order to continue steady economic growth, nations rely on economic researchers and planners to discover how to best allocate resources to progress the economy.

1.1 Economic Sectors
The function of any national economy begins by the harvesting of raw materials from natural resources. With these raw materials, manufacturers can add to their value by using them to produce useful goods for consumers. Then, the new profitable products can be marketed and sold domestically, or exported to other nations in demand of the products. This fundamental concept of a national economy can be broken down into three sectors.

The primary sector concerns the extraction of raw materials from natural resources. Natural resources serve as a foundation for a nation's economy, as any manufactured products are constructed through the use of harvested raw materials. This sector does not include any form of production or manufacturing other than the simple processing of raw materials to be sold directly to manufacturers for proper use. Some industries of the primary sector include agriculture, forestry, mining, timber, and commercial fishing[4].

Next, the secondary sector follows the primary sector by collecting raw materials and converting them into manufactured products. If consumers find a product useful, its value should be significantly higher than its raw material components, as it creates a demand from anyone that shows interest in its new functionality. Industries in the secondary sector include any form of construction, such as buildings or highways, or manufacturing, such as textiles or automobiles

The tertiary sector is also known as the service sector. After producing useful goods, the service sector involves marketing, retailing, import/export, transportation, and any other service or distribution of goods to finally return profits back to the producers. Services in the tertiary sector aren't limited to the distribution of manufactured products. This sector includes services such as administration, tourism, arts and recreation, and waste management, along with government and non-profit work.

Together, these three sectors envelop major aspects of a national economy. GDP, or gross domestic product, is one of the most prominent economic indicators used by economists and measures the values of goods and services newly produced in a country . For analyzing the growth of the US economy, examining the interconnecting relationship between industries in the three sectors can help with understanding the behavior of GDP growth in relation to these factors. With a defined relation between US GDP growth and changes in each economic sector, it would be possible to predict the effect of any economic changes or developments on the US GDP.

1.2 Artificial Neural Networks
Artificial neural networks [ANN] are computational models of how the human brain sends and receives signals throughout

the body. Using a system, an ANN establishes the relative importance, or weightage, of each input in producing the output]. The first step to creating an ANN is to determine the most significant sectors that could have an effect on the output. These sectors are variables that will be used as inputs for the neural network. Next, data must be collected that provides several instances of each input with their corresponding output(s). The objective of using an ANN is to develop the relationship between each input's behavior on the output as a function. The discovery and development of this function through computational methods is known as machine learning. The machine learning algorithm will use the simple architecture of an artificial neuron known as a perceptron [5].

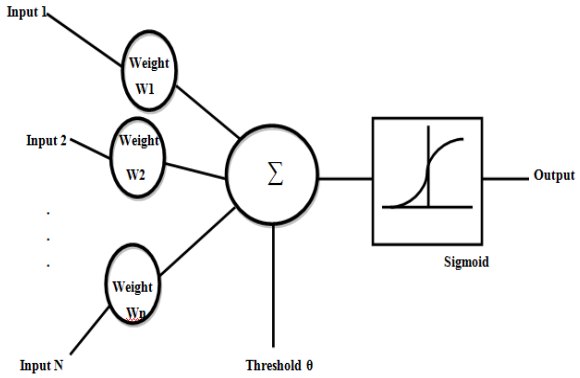

Fig. 1: Model of single layer perceptron structure [13]

2. LITERATURE SURVEY

There have been several proposed econometric models of GDP in different nations. Obeng et. al[6] in their paper have established a positive correlation between GDP growth rate and exchange rate in Ghana and suggested a stabilization of monetary and fiscal policies by policy makers. However, the model proposed by the authors used linear regression analysis to establish the relationship. While a linear approximation may be viable in developing countries, this experiment uses sigmoid activation to account for nonlinearity in the inputs because developed countries have reached a saturation point in many economic indicators.

Mahmud [7] has built a model relating Indian population growth and economic development. His objective was to support either a positive, negative, and neutral relation between the two parameters. He used a time series data from 1980 – 2013 in a Solow growth model to test the relationship and established a positive relationship between economic and population growth. Although his econometric model provides a relationship between the two, the model can only be used in developing countries and has actually been proven to be inaccurate in many cases[8].

Camacho and Martinez-Martin [9] have used a single-index dynamic factor model that was previously used to frequently measure overall economic activity from GDP, industrial production, sales, income, and employment. They have developed this model to utilize survey data and financial indicators for the forecasting of US GDP. Although their model has produced satisfactory immediate results, it only has the ability to accurately produce short-term results.

3. PROPOSED MODEL

The objective of this paper is to establish a relationship between the different economic sectors of the United States. The model aims to account for nonlinear relationships and remain accurate over stable economic periods.

This econometric model will use artificial neural network architecture along with the process of machine learning to derive an adequate relationship between the several inputs. The neural network model used is known as a multilayer perceptron supervised network. This network attempts to create a model that can correctly relate the inputs to their outputs using past data by comparing computed outputs to the desired outputs, to then be able to produce reasonable outputs in the future when the desired output is unknown. It uses three-layer architecture: an input layer, hidden layer, and output layer. Here, the X's are the outputs of input layer neurons and b is the bias, W's are the weights between input layer i and hidden layer j, and V's are the weights between hidden layers k and the output layers. At each of the hidden layer neurons the net input is

$$Z_j = \sum_{i=1}^{n} W_{ij}X_i + b_j. \; ; \; Ho = 1/(1+e^{-\sum_j KZ_j}); \quad y = \sum_{k=1}^{m} VkHo$$

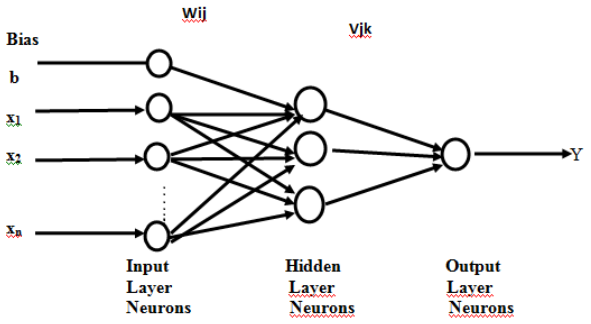

Fig. 2: Example of back propagation artificial neural network

The assumed activation function is chosen to be a sigmoid function, in which the output of the hidden layer $H_o = 1/(1+e^{\Sigma-kz_j})$ where k is a constant[10]. The above diagram is for informational purposes and doesn't reflect the structure of the neural network in this study. For this study, there are three inputs, the three sectors, and two hidden layers of six and four neurons that result in one output, the US GDP. The datasets for each parameter must be normalized to allow the neural network to observe the behavior of each parameter and output on the same scale.

Here, the normalized x = (x – min x)/(max x – min x) - 0.5

Since the hyperbolic tangent sigmoid activation function being used has a range from -1 to 1, this form of normalization properly centers the data. The model is implemented using the Neurolab library for Neural Network tools in Python .

3.1 Hypothesis

The United States' gross domestic product per year is a non-linear function of the input vector of the three economic sectors.

GDP= *f[P, S, T]*, where the output GDP is the United States Gross Domestic Product

P is the gross output of the primary sector, involving natural resources and agriculture.

S is the gross output of the secondary sector, involving construction and manufacturing.
T is the gross output of the tertiary sector, involving product distribution and other services.

4. IMPLEMENTATION & RESULTS

4.1 Normalized Data

Fig. 3 Normalized plot of GDP and Econ. Sectors from 1947-1982

Fig. 3 presents the normalized inputs and outputs from the testing data set to show that they display similar growth patterns. This validates basis of the hypothesis, since the growth of each industry clearly resembles the GDP growth pattern to varying degrees.

4.2 Error Progress and Iterations

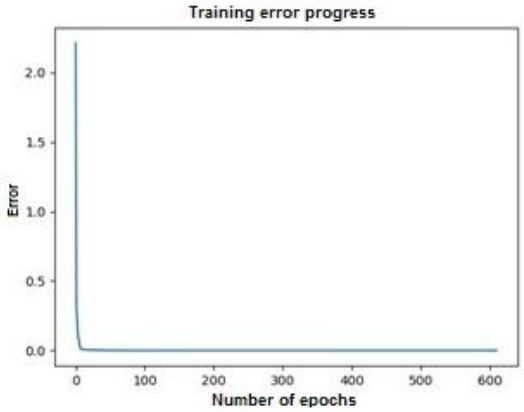

Fig. 4 Training Error per Epoch

The goal of the neural network was to decrease the calculated error to a value of 0.00001 within 2000 epochs. This figure shows that the goal of learning was reached after over 600 epochs through the dataset, with each epoch consisting of 36 iterations of analysis (one for each recorded timeframe within the testing dataset).

4.3 Computed Model vs. Testing Data

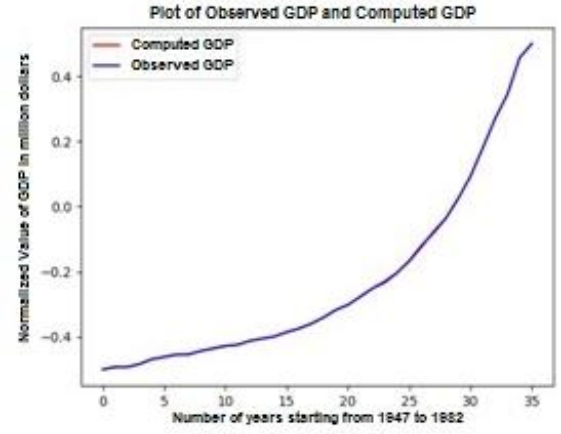

Fig. 5 Comparison between Computed Model and Observed (Actual) GDP from Testing Data

The computed GDP comes from the final model after either the goal of learning is reached or the program has exceeded the maximum amount of allowed epochs. The derived model then computes an output for each of the 36 training examples in the testing .csv file. Here, the computed and observed GDP are extremely difficult to distinguish, meaning that the computed model has produced extremely accurate results for the given inputs.

4.4 Computed Model vs. Validation Data

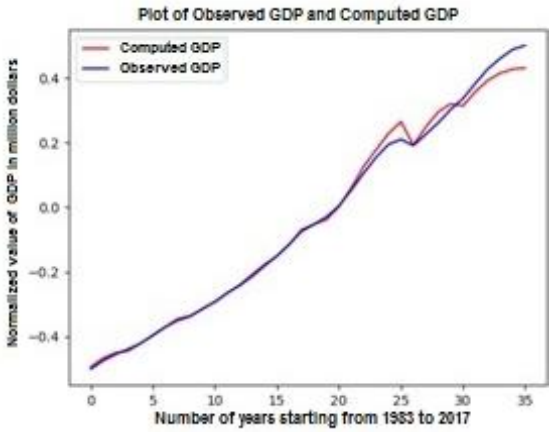

Fig. 6: Comparison between Computed Model and Observed (Actual) GDP from Validation Data

Here, the computed GDP model uses its derived weights and mathematical relationship between each sector to output the GDP for the next 36 years. The computed GDP demonstrates a satisfactory forecast in this validation series. The deviation near the end of the interval can be attributed to the US economic recession from 2007 to 2009, as this model assumes that no economic crash or random event significantly changes the relationship between the inputs. This demonstrates that the model remains mostly accurate during stable economic periods.

5. CONCLUSION

There by utilizing an artificial neural network with multilayer perceptron architecture, an accurate model of GDP from three economic sectors has been constructed and validated. This model accounts for nonlinearity in the relationship between the three sectors and can be used for forecasting the future GDP by assuming the growth of each sector. For example, if it is predicted that the gross output of the primary sector will increase by 1%, the secondary sector by 2%, and the tertiary sector by 3%, then this model predicts that the GDP would increase by approximately 2% from the latest recorded value (the second quarter of 2017) at $19,437,063.72. If only the tertiary sector is to increase by 5% while the primary and secondary sectors remain the same, this model predicts that the GDP would increase by approximately 3.9% at a value of $19622736.53. If solely the primary sector is to increase by 5%, then this model predicts that the GDP would decrease by less than 0.1%, demonstrating the nonlinear, complex relationship formed within the model.

This model assumes that there is no major event or economic crash in the country that could affect the relationship between the three sectors [10]. The result is found to be satisfactory over stable economic periods. The purpose of the study is to present the application of an artificial neural network as an econometric model. An economist could use such a model to compute GDP following the current trends and predictions of growth in each sector. Economic planners could use this model to decide which industries to allocate resources and keep the economy healthy and growing. This

model can be further developed by providing more possible inputs and can be easily adapted for developing countries by accounting for indicators such as human development.

6. REFERENCES

[1] Coplan, Jill Hamburg. "12 signs America is on the decline. *Fortune*, 20 July 2015, fortune.com/2015/07/20/united-states-decline-statistics-economic/.

[2] "Education Index." *Human Development Reports*, United Nations Development Programme, 15 Nov. 2013, hdr.undp.org/en/content/education-index. Accessed28 Aug. 2018.

[3] Hawksworth, John. "The World in 2050". *PwC*, Feb. 2017, www.pwc.com/gx/en/issues/economy/the-world-in-2050.html

[4] Goodwin, Neva R., et al. *Macroeconomics in Context*. 2nd ed., PDF ed., Abingdon, New York, NY, 2015.

[5] Negnevitsky. "Lecture 7 Artificial Neural Networks." PowerPoint file, 18 Aug. 2014

[6] Attah-Obeng, Prudence, et al. "An Econometric Analysis of the Relationship between Gdp Growth Rate and Exchange Rate in Ghana" *Journal of Economics and Sustainable Development*, PDF ed., vol. 4, no. 9, 2013, pp. 1-8.

[7] Mahmud, Musa Abba. "Econometric Model on Population Growth and Economic Development in India: An Empirical Analysis." Proceedings of the International Symposium on Emerging Trends in Social Science Research. Chennai Symposium, 3 Apr. 2015, Chennai, India

[8] "Falsification in Economics." *Unlearning Economics*, WordPress, 15 Mar. 2013, unlearningeconomics.wordpress.com/tag/the-solow-growth-model/.

[9] Camacho, Máximo, and Jaime Martínez-Martín. "Real-Time Forecasting US GDP from Small-Scale Factor Models." *Banco De España*, PDF ed., 2014.

[10] Sinha, Ashok K. "An Econometric Modeling of Development Process using Artificial Neural Network: A case study of India." *International Journal of Engineering Research and Development*, PDF ed., vol. 13, no. 1, Jan. 2017, pp. 4