https://communications.science.ankara.edu.tr

Commun.Fac.Sci.Univ.Ank.Series A2-A3
Volume 63, Number 1, Pages 25-31 (2021)
DOI: 10.33769/aupse.525325
ISSN 1303-6009 E-ISSN 2618-6462

Received by the Editors: February 11, 2019; Accepted: February 03, 2021

DETERMINING THE MOST RELEVANT INPUT PARAMETER SET BY USING EXTREME LEARNING MACHINE

SEMRA GÜNDÜÇ[1] and RECEP ERYİĞİT[1]

[1]Ankara University, Computer Engineering Department, Ankara, TURKEY

ABSTRACT. In this work, Extreme Learning Machine (ELM) algorithm is used to estimate the GDP per capita. The amount of electricity production, from four different sources, is chosen as input parameters. To find out the most relevant input data for a reasonable estimation of GDP, different sources introduced separately to ELM. By following the coefficient of determination of estimation, by trial and error, results are obtained. The residuals are also given to show that model perform well. Renewable energy sources produce the best results in the estimation of GDP.

1. INTRODUCTION

Gross domestic product (GDP) which is mostly related to the economic development of a country is an important monetary metric to measure eco- nomic growth. There are many factors affect GDP hence many types of research which aim to understand and show a correlation between GDP and some input parameters, mainly economic parameters, such as production, industry, etc. [1, 2], and non-economic parameters such as education, health, number of scientific paper, etc. [2].

The amount of electricity produced by the country from different sources, particularly for developed countries, is an excellent candidate to use as the input parameter to estimate the GDP [3]. Because there are a considerable number of parameters which can be used as input, it may be difficult to choose the most relevant and proper ones for a good estimation. Moreover, the relation between input and output mostly has a nonlinear structure [4].

In the literature, there is a huge number of researches in different disciplines which purpose is to extract information from a data set to predict the future and find out relations hidden in the data by using machine learning and statistic techniques in the last decades. Before naming this discipline as data mining, by using classical statistical inference, mostly there is a scientific method which drives the process. However, in the real-world examples, the most process includes many different and uncertain number of variables and also there is no specific method to model the process. For this reason, it is the most significant point to decide the number (and the source) of parameters which are relevant (sometimes hidden variables) or exogenous parameters for the case. To collect and save the data and to reduce the computation time, it must be known not only

Keyword and phrases. Gross domestic product estimation, extreme learning machine, electricity production sources

gunduc@eng.ankara.edu.tr-Corresponding author; eryigit@eng.ankara.edu.tr
0000-0002-3811-9547; 0000-0002-4282-6340

which parameters but also how many are leading indicators of the process.

The well-known conventional neural network implementations consume lots of time because of slow gradient-based algorithms and updating all parameters at every iteration. At this point, soft computing introduces some advantages when it is not known internal parameters of the system, and there is no analytical approach through lack of a driving mechanism of the process. For this reason, to apply soft computing methods to introduce a better solution.

This research aims to investigate the relation between the amount of electricity production, from four different sources, and GDP per capita. Extreme Learning Machine (ELM) [5] algorithm is used to estimate GDP. The input data are obtained from the World Bank Database [6] including 38 countries starting from the year 1960 to 2017. The amount of electricity production is used as input data for ELM to estimate GDP.

The graphical structure of the Extreme Learning Machine algorithm, based on Single-Layer Neural Networks, is given in Figure 1. Here ω is the weight vector between the hidden nodes and input nodes, β is the weight vector between output and hidden nodes.

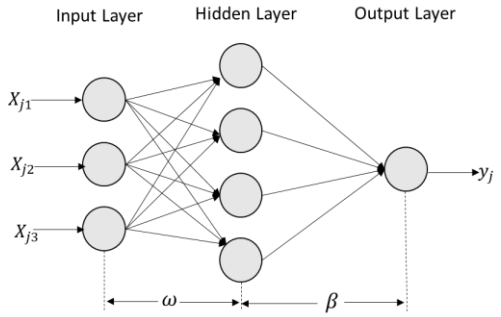


FIGURE 1. A schematics representation of ELM.

The underlying structure of the relations between electricity production and GDP may provide an opportunity for constructing better energy policies. Besides estimating the GDP, it is also important to forecast the electricity demand for future planning [7].

2. MODEL

In this work, the aim is to use well-known learning algorithm ELM method [5] to understand the relationship between the amount of Electricity production and GDP per capita. Secondly, it is also necessary to specify the number of the most relevant input parameters, source of electricity production, to get the best estimation result for an output parameter, GDP.

In the literature, there is much research which investigates the relation between some parameters and GDP. In [2] the effects of scientific and techno logic factors are studied. Prediction of economic development by considering energy resources is discussed in [8]. Also, the comparison of two methods, ELM and Back Propagation, are investigated in [9]. Even though it is trendy and studied by using many different mathematical models, such as ANN, Backpropagation, linear regression, it is still essential and needs to identify which input parameter is relevant for estimation of GDP, especially when there is no information about how and which parameters will be considered.

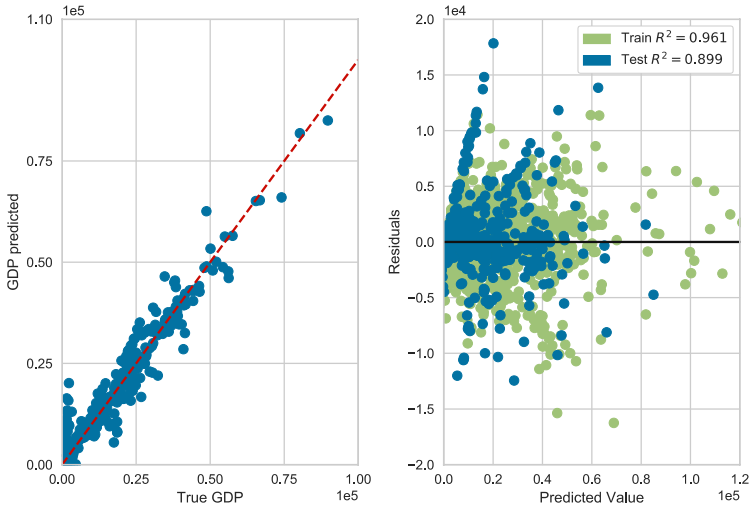
In this work, the amount of electricity production from four different sources, coal, gas, petrol, renewable, are used as input parameters of ELM method to estimate the GDP per capita. The inputs are introduced in sequence, and coefficients of determination, R^2 of estimating are recorded. By using the trial-and-error method the most relevant input parameters are determined. In Table 3 the results are introduced. We interpret the input which result in the best R^2 includes the most valuable data.

3. Result and Discussion

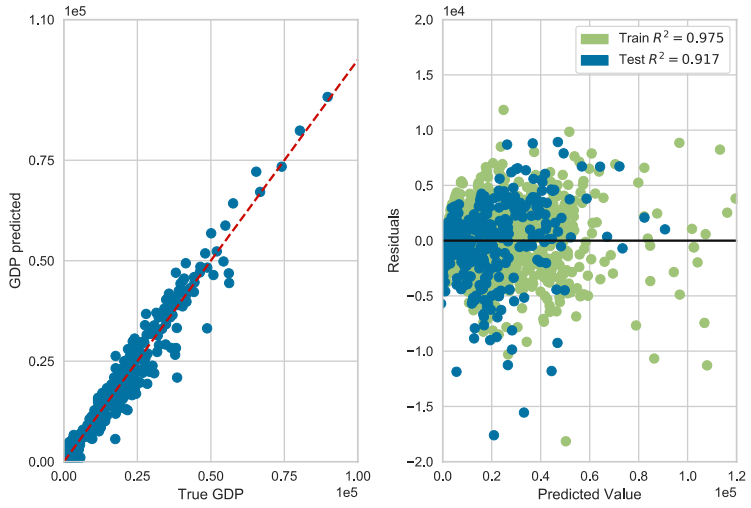
In this work, the ELM method is used to estimate the GDP. The amount of electricity production, from four different sources, are used input values. First to understand the individual contribution, each production type is introduced ELM separately and results, showing the coefficients of determination for ELM are given in Table1. It is seen that the amount of renewable energy production, excluding hydroelectric and including geothermal, solar, tides, wind, biomass and biofuels, is the leading indicator (most relevant input parameter) to estimate GDP. Gas stands for the second useful input variable. As well as to interpret this situation is out of this work, it can be said that some approaches for preserving the environment might be more considered in developed countries.

Table 1. Coefficient of determination for ELM.

Source	R^2 Train	R^2 Test
Coal	0.961	0.899
Gas	0.975	0.917
Petrol	0.912	0.871
Renewable	0.898	0.948

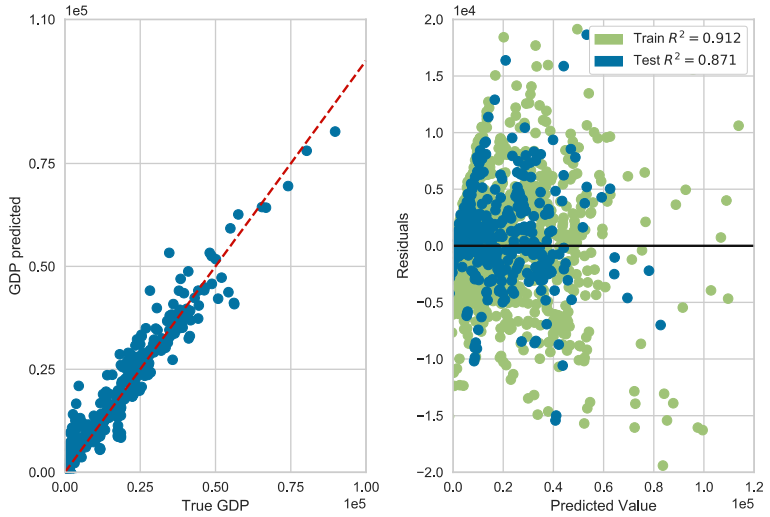


a) Coal sources

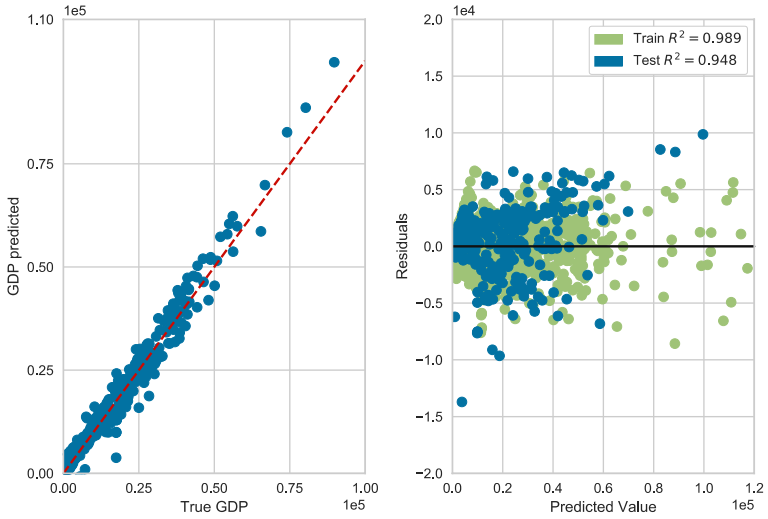


b) Gas sources

FIGURE 2. Prediction of GDP by using ELM from a) coal b) gas sources.



a) Petrol sources



b) Renewable sources

FIGURE 3. Prediction of GDP by using ELM from a) petrol b) renewable sources.

Fig.2 and Fig.3 show the graphical representations of prediction of GDP by considering the amount of electricity production from, coal, gas, renewable and petrol sources, respectively. Fig.4 shows the same situation when all different sources are considered as a single input to estimate GDP.

In each graphic, the left figure shows the prediction and the right one shows the residuals. The residuals, difference between the observed value and the estimated value, are also considered in this work to have a sense of how accurate our model is by relating the amount of electricity production to GDP. Also, residual plots show that our model, with chosen inputs, perform well.

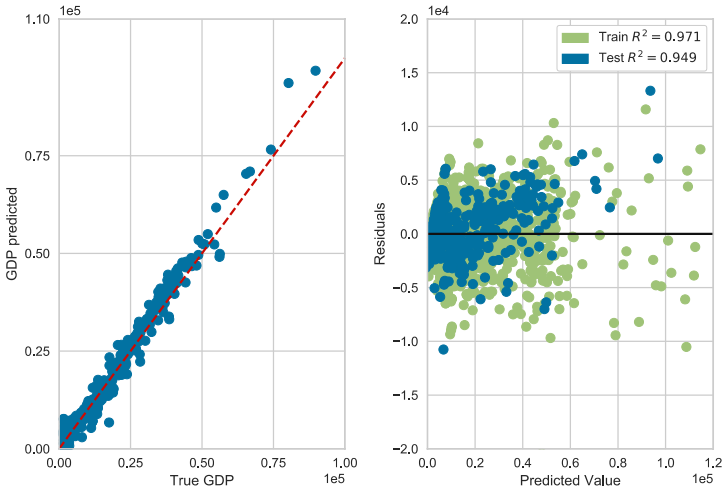


FIGURE 4. Prediction of GDP by using ELM from all (coal + gas + petrol + renewable) sources.

3. CONCLUSIONS

It is a leading problem to choose the most related features of the systems. The success of traditional methods depends on the selection of the best parameters to describe the system. For many systems it is not possible to define the best parameters clearly. For that purpose, we used EML structure to understand the relationship between GDP values of countries and energy consumption. Because EML needs less time for computation and provide accurate results than conventional methods, it is mainly used for classification, regression, forecasting and many other problems in different disciplines.

Authors Contribution Statement The authors have been contributed equally to this paper in writing, designing, analyzing, and collecting data processes.

Declaration of Competing Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

[1] Rakic, G., Milenkovic, D., Vujovic, S., Vujovic, T., Jovic, S., Information system for e-GDP based on computational intelligence approach, *Physica A: Statistical Mechanics and its Applications*, 513 (2019), 418–423. https://doi.org/10.1016/j.physa.2018.09.010

[2] Markovic, D., Petkovic, D., Nikolic, V., Milovancevc, M., Petkovic, B., Soft computing prediction of economic growth based in science and technology factors, *Physica A: Statistical Mechanics and its Applications*, 465 (2017), 217-220. https://doi.org/10.1016/j.physa.2016.08.034

[3] Alshehry, A.S., Belloumi, M., Energy consumption, carbon dioxide emis- sions and economic growth: The case of Saudi Arabia, *Renewable and Sustainable Energy Reviews*, 41 (2015), 237-247. https://doi.org/10.1016/j.rser.2014.08.004

[4] Mbarek, M.B., Abdelkafi, I., Feki, R., Nonlinear causality between renewable energy, economic growth, and unemployment: Evidence from Tunisia, *Journal of Knowledge Economy*, 19 (2018), 694-702. https://doi.org/10.1007/s13132-016-0357-9

[5] Huang, G.B., Zhu, Q.Y., Siew, C.K., Extreme learning machine: Theory and applications, *Neurocomputing*, 70 (2006), 489-501. https://doi.org/10.1016/j.neucom.2005.12.126

[6] World Bank Open Data, https://data.worldbank.org.

[7] Kankal, M., Uzlu, E., Neural network approach with teachinglearning- based optimization for modeling and forecasting long-term electric energy demand in Turkey, *Neural Computing Application*, 28 (2017), 737-747. https://doi.org/10.1007/s00521-016-2409-2

[8] Cogoljevic, D., Alizamir, M., Piljan, I., Piljan, T., Prljicc, K., Zimonjic, S., A machine learning approach for predicting the relationship between energy resources and economic development, *Physica A: Statistical Mechanics and its Applications*, 495 (2018), 211-214. https://doi.org/10.1016/j.physa.2017.12.082

[9] Milacic, L., Jovic, S., Vuovcc, T., Miljkov, J., Application of artificial neu- ral network with extreme learning machine for economic growth estimation, *Physica A: Statistical Mechanics and its Applications*, 465 (2017), 285-288. https://doi.org/10.1016/j.physa.2016.08.040