International Conference on Engineering, Science and Advanced Technology (ICESAT-2023), Technical Engineering College / Mosul, Iraq

The Application of ANFIS to Prediction of the Amount of Emissions from an Emitter Activity to Reduce Climate Change

Shwan Abdlwahaab Agha
Department of Information Technologies
Altinbas University
İstanbul, Türkiye
shwan@moen.gov.iq

Hasan H. BALIK
Department of Computer Engineering
İstanbul Aydin University
Küçükçekmece, İstanbul, Türkiye
hasanbalik@gmail.com

2023 International Conference on Engineering, Science and Advanced Technology (ICESAT) | 979-8-3503-4440-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICESAT58213.2023.10347325

***Abstract*—Understanding the links between CO2 emissions from fuel consumption and climate change may assist countries in determining the amount of CO2 or CO2 equivalent emissions, and act accordingly by reformulating new energy policies to reduce emissions and achieve sustainable development. This understanding can help guide countries in developing better energy policies that are designed to reduce CO2 and equivalent emissions, while still promoting economic growth and development. To reach the global climate goals set by the United Nations [25], countries must take decisive action to reduce emissions from fuel consumption and energy production while promoting clean energy, which can help mitigate the effects of climate change. Understanding the magnitude of the emissions problem is necessary before coming up with innovative, long-lasting solutions to better comprehend the issue of emissions. An important first step in achieving this understanding is to accurately assess the current levels of emissions from fuel consumption and energy production in a certain country and analyze the amount of these emissions and finally implement strategies to reduce them. Using machine learning technologies such as predictive analytics, and artificial intelligence can help create meaningful insights into the data to understand patterns in emissions that could enable better strategies for reducing emissions. Using a dataset that belongs to the Canadian government, with input data including engine size, number of cylinders, and fuel consumption, an ANFIS model was created to predict emissions. The amount of emissions represents the output. A variety of membership functions were tested in order to gain the best results. Another dataset was generated by applying the Extra-Trees regressor algorithm to the mentioned dataset and retraining the ANFIS model using the generated dataset to achieve better results. The resulting model can then serve as a trustworthy tool to support the application of efficient mitigation strategies and to inform and help policymakers.**

Keywords*—Prediction, CO2 Emissions, Adaptive Neuro-Fuzzy Inference System, Extra Trees Regressor.

I. INTRODUCTION

The greenhouse gas emissions that cause climate change are among the most concerning issues of the twenty-first century because they have a significant impact on nearly every aspect of life on Earth, from the economy to how we will live in the future. This problem also poses a serious threat to both humans and other species. Results from earlier studies show a strong and consistent relationship between the noted points and climate change:

The rising of maximum and minim temperatures.

The rising of sea levels. Thawing permafrost.

We see evidence of the points we mentioned above every day in the form of extreme weather conditions, hurricanes of enormous sizes, floods, and forest fires, as well as other problems like desertification and land degradation. It is becoming increasingly clear that these phenomena are the result of climate change and global warming, both of which have been caused by human activity.

As a result of issues like climate change, desertification, degradation, global warming, and biodiversity loss, machine learning is gaining more academic and commercial interest. Large amounts of data can be analyzed using these algorithms collected over the course of previous years by researchers in machine learning algorithms. Machine learning algorithms have been demonstrated to be effective and sufficient, particularly when dealing with large datasets. The scope of using these algorithms on large datasets may offer a comprehensive solution to a better understanding of some challenging problems, like the problems mentioned before. Researchers should be able to look into ways to determine the most effective methods for calculating the severity of the effects of climate change, confirming these results, and then classifying them.

In this case, it is important to define global warming precisely. Earth's temperature has fluctuated over the years, both upward and downward. And everything is based on the amount of sunlight that the earth receives. An additional element that affected the planet's normal increase and decrease in temperature over the past century has contributed to climate change. This demonstrates the need to be clear about the precise situation within the next century. Many other questions need to be addressed, such as how much more will the climate change? How can we stop climate change?

Figure 1 shows the increasing average of the temperature according to NASA.

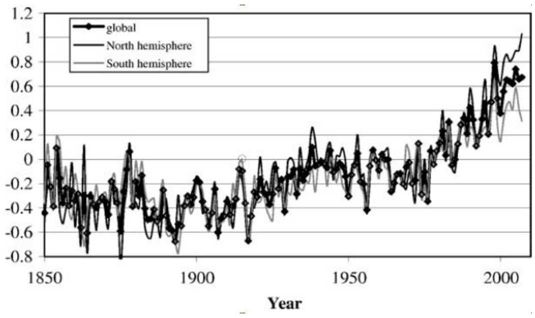

Fig. 1. Rate of land/air temperature increase (°C) [1].

By examining the scope and nature of the issue regarding carbon dioxide emitted from vehicles, this study's main goal is to find solutions to decrease the effects of climate change.

Overall, addressing the problem of carbon dioxide emissions from vehicles will require a combination of technological, behavioral, and policy changes. By exploring the scope and nature of this problem, and identifying and implementing possible effective solutions, it is possible to significantly reduce the impact of transportation in climate change. The goal will therefore be building a model that can predict the amount of emissions, it will be based essentially on machine learning. And it will be targeting precisely the amount of CO2 emissions from vehicles. Basically, the results will represent an estimation of carbon emissions amount and it will not show the exact amount of emissions due to different circumstances.

II. LITERATURE SURVEY

In an effort to improve predictions across a variety of fields, recent developments in artificial neural networks and fuzzy logic research have been increasingly used. Among these technologies, (ANFIS) has emerged as a powerful technique, allowing for the prediction of complex phenomena and complex behavior.

S.Hoda Rahmati et al. in their study compared between ANFIS, Artificial neural networks (ANN), Multivariable Regression (MLR), and Autoregressive Integrated Moving Average (ARIMA) methods to forecast the amount of urban water consumption per capita in Tehran / Iran. They were able to recognize the effects of climate change as data pre-processing methods such as data transformation and input variables selection were applied for improving the accuracy of the models. The results indicated that ANFIS is significantly superior to other methods [2].

In their study on the relationship between CO2 emissions and global warming temperatures, Khan and Khan [3] attempted to accurately model the complex connections between variables by utilizing three advanced statistical tools: Adaptive Neuro-Fuzzy Inference Systems, Artificial Neural Networks, and Fuzzy time series models. Due to the reason that traditional statistical methods can produce results with significant margins of error, they tried to avoid making strict assumptions in their analysis.

H.Qubehar et al. present the design and real-time implementation of an efficient Adaptive Nueral-Fuzzy Inference System for regulating the greenhouse's climate. With the help of their system, it will be easier to monitor the temperature and relative humidity reference levels required for the growth of crops. Techniques from artificial neural networks and fuzzy logic are combined. A training dataset for ANFIS is included, and the robustness of the fuzzy logic controller model produced by MATLAB Simulink has been experimentally tested. The proposed control system is designed to provide good performances in terms of set point tracking, response time, robustness against changes in external parameters, non-linearity, and energy optimization [4].

In a study conducted by Maryam Mokhtarzad et al. [5], a comparison was made between the effectiveness of (ANNs), (ANFIS), and Support Vector Machines (SVMs) as models for predicting drought. Data from the Bojnourd meteorological station from January 1984 to December 2012 were utilized for the study, with input parameters of temperature, humidity, and season precipitation, and the Standardized Precipitation Index (SPI) as the output. Results of the study showed that the SVM model was the most precise and accurate, outperforming both the ANN and ANFIS models in terms of accuracy.

Patel and Parekh in their study regarding rainfall forecasting using ANFIS produced Eight models using various membership functions and climatic parameters as inputs. In this study, the generalized bell-shaped built-in membership function (gbell) has been used as a membership function in both the Hybrid and Back propagation method for ANFIS, the study reveals that the hybrid Model with seven membership functions and using three inputs, temperature, relative humidity, and wind speed gives best result to forecast rainfall for study area [6].

To forecast flooding in regions characterized by flat or concave topography, it is important to account for high levels of precipitation, particularly in tropical areas. In their study, Suparta and Samah pointed out some crucial areas in South Tangerang, one of the cities that are currently developing most quickly, the flooding issue there cannot be disregarded. The study investigated how artificial intelligence methods like ANFIS can be used to predict rainfall. The suggested method combines transparent linguistic representations of fuzzy systems with neural network learning capabilities. The ANFIS model was created, trained, and tested in order to assess a model's potential. It has a diverse input structure and membership functions. It was discovered that 80% of data testing for rainfall predictions based on ANFIS time series are accurate. Decision-makers must have access to accurate multi-step-ahead flood forecasts because they are essential [7].

III. MATERIALS AND METHODS

A. *Artificial Intelligence and Climate Change*

Artificial intelligence can play a major role in contributing to addressing the challenges posed by climate change. For example, machine learning algorithms can be used to analyze large amounts of data and recognize patterns that can inform the development of more effective policies and strategies for mitigating and adapting to the impacts of climate change [8].

B. Fuzzy Inference System (FIS)

Fuzzy Inference Systems (FISs) are computational intelligence tools for modeling and controlling complex systems. They are built upon fuzzy logic, a mathematical framework for representing and processing uncertain, vague, and incomplete information. Fuzzy Inference Systems have been widely utilized across a range of domains, including control engineering, decision-making, pattern recognition, and data mining. The fuzzy rule base consists of a set of fuzzy rules that specify the relationships between the inputs and the output of the FIS. Each fuzzy rule has the form of "IF input1 is A AND input2 is B AND input3 is C THEN output is D", where A, B, C and, D are fuzzy sets. The fuzzy sets represent the membership functions of the inputs and the output, which assign a membership degree to each element of the universe of discourse. The fuzzy inference engine applies the fuzzy rules to the inputs and combines them using a fuzzy operator, such as AND, OR, or MIN. The resulting fuzzy set is the fuzzy output of the FIS [9]. The " (1)" of a "gaussmf" is given by:

$$mu(x) = \exp [-0.5 \square ((x - c) / s) ^2] \quad (1)$$

Where:
mu(x) is the membership degree of x in the fuzzy set described by the Gaussian MF.
x represents the input value.
c represents the center of the Gaussian distribution.
s is the standard deviation of the Gaussian distribution.

C. Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) are computer-based models created on the principles of biological neural systems. ANNs learn through experience by identifying patterns and relationships in data. These networks consist of processing elements or artificial neurons that are interconnected through coefficients known as weights, forming multiple layers of the neural network structure. The computation ability of ANNs arises from these connections. Each processing element has weighted inputs, a transfer function, and a single output. The properties of the ANN are defined by the transfer functions, the learning rule, or the network architecture, with the adjustable parameters of the weights which play a crucial role. The weighted sum of the inputs determines the neuron activation, which is then processed through the transfer function to create the final output. The transfer function makes the network non-linear. The connection weights between neurons are improved during training to reduce prediction errors and achieve the desired targets [10].

D. Adaptive Neuro-Fuzzy Inference System (ANFIS)

One of the machine learning techniques is the Adaptive Neuro-Fuzzy Inference System (ANFIS), which is a type of supervised learning algorithm. According to Jang, Sun, and Mizutani's works from 1997 [11], ANFIS is a hybrid system that uses artificial neural networks and fuzzy logic to learn from training data and make predictions. The algorithm is trained on a dataset in the supervised learning setting where both the inputs and the anticipated outputs are known. By changing its parameters to reduce the difference between the predicted and actual outputs, the model seeks to understand the relationship between inputs and outputs. These results support the notion that ANFIS operates as a supervised learning algorithm [12]. Premise parameters are associated with the fuzzy sets in the input variables and are adjusted during the fuzzification process. These parameters determine the shape and location of the membership functions for the input variables. The premise parameters are adaptive because they are adjusted during the training process based on the error between the actual and predicted output. The adjustment is made using gradient descent optimization methods such as the backpropagation algorithm. Consequence parameters are associated with the fuzzy sets in the output variables and are adjusted during the defuzzification process. These parameters determine the shape and location of the membership functions for the output variables. The consequence parameters are also adaptive and are adjusted during the training process based on the error between the actual and predicted output. The adjustment is made using the least squares method. Fig. 2 shows the structure of ANFIS.

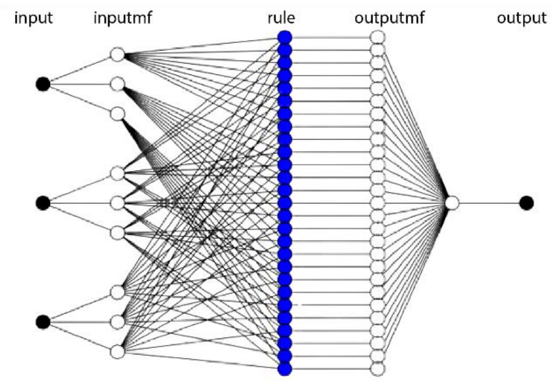

Fig. 2. ANFIS structure.

E. Extra-Trees Regressor

Extra trees regressor is a powerful tool for regression tasks that involve predicting continuous or ordered values [13]. It is an ensemble learning method that combines the predictions of multiple models to obtain a better performance. It is based on decision trees, which are models that make predictions by learning a set of rules from the data. Extra-trees regressor aggregates the predictions of decision trees by averaging or voting, it can handle missing values, multiclass and multilabel tasks, and imbalanced data. It is simple to implement, fast to train, robust to over-fit, accurate, stable, and interpretable. However, it is sensitive to noise and outliers, and may not be as accurate as other methods [14].

IV. PROPOSED METHOD

A. System Outline

A reasonable approach to building the model was the use of multiple membership functions in order to achieve the best results. Membership functions that have been evaluated are ("trapmf", "trimf", "psigmf", "pimf", "gaussmf", "gauss2mf", "dsigmf", "gbellmf"). The training was done with 100 epochs and Fig. 3 shows the system outline.

B. Practical Part

This study recommends using an ANFIS model to forecast how much CO2 is emitted by vehicles. A variety of vehicle characteristics and the corresponding CO2 emissions data will be used to train the model. Engine size, number of cylinders, and fuel consumption serve as the model's inputs. CO2 emissions serve as the model's output. The ANFIS model will be able to adapt to the distinct patterns in the data and make more precise predictions by utilizing a supervised learning approach. The ANFIS model is more robust in real-world scenarios because it can manage input variables uncertainty.

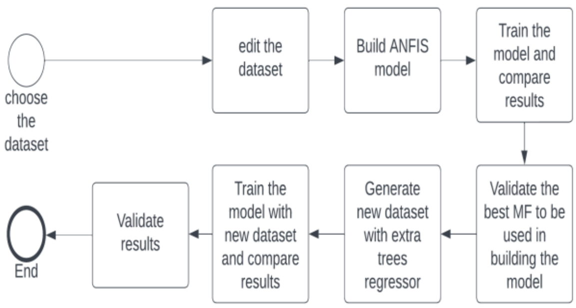

Fig. 3. Proposed system outline.

We developed the ANFIS model using Matlab R2022. The (ANFIS) model is trained by first defining the input and output data sets, initializing the ANFIS model with values from the dataset, and then adjusting the ANFIS model's parameter values using a learning algorithm, such as a hybrid learning algorithm, to reduce the difference between the model's output and the desired output. The ANFIS model can also handle relationships between input and output variables that are not linear. Due to the fact that the ANFIS model can handle problems with high dimensions, it is regarded as advantageous for problems with a large number of input variables. We will create the FIS model after loading the data into the workspace of Matlab. An important step in the training process for an (ANFIS) model is selecting the best membership function because it can significantly affect the model's performance. Different types of membership functions have different properties that make them suitable for various types of data and problems. In the proposed study, we began by splitting the dataset into 30% for model testing and 70% for training. In order to be able to test and validate the model, matrices with [2217×4] and [5168×4] were uploaded. To get the best results, we generated the FIS using a variety of membership functions. The evaluated membership functions were ("gbellmf", "gaussmf", "gauss2mf", "trapmf", "pimf", "psigmf", "trimf" and "dsigmf"). The outcome of the training using the aforementioned membership functions, each with 100 epochs, was examined in this study. The "gaussmf" has been noted to have the best error results, so it will be taken into consideration when building the model. With a testing error rate of (11.9468), this model will allow us to predict the emissions from a huge variety of vehicles. The training procedure of the ANFIS model will be optimized by creating a new dataset through by applying an extra-trees regressor to the current dataset in order to produce more accurate results. Using Python and sci-kit learn library, the algorithm is employed as a tool for refining the ANFIS model's parameters. The generated dataset will be up to (56897×4), which is supposed to be more effective for training the model than the previous dataset, which consists of (7384×4). The same training procedure and membership functions were used for validating the membership function with the lowest error rate. The "gaussmf" achieved the best result. With an average testing error for the training of 1.8226 and an average testing error for testing of 1.767, it was clear from the results of this study that the (gaussmf) had once again produced the best results in terms of the accuracy for predicting the amount of the output. After the model is constructed, the membership functions need to be fine-tuned, the tuning of the membership functions resulted in a slight increase in the average error rate for testing to (4.2801). Fig. 4 below shows the model design.

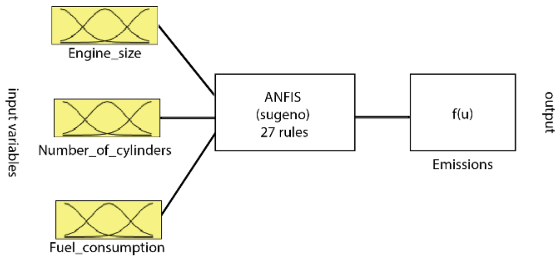

Fig. 4. Proposed ANFIS model.

V. RESULTS AND DISCUSSIONS

Regarding the used dataset, this study utilized a comprehensive dataset from the Canadian government's "Natural Resources of Canada" for its research [24]. A variety of important variables are included in the dataset, which has been carefully selected to provide an answer to the research question. The strength of this dataset lies in its diversity and large sample size with thousands of observations. The variables in the dataset include engine size, number of cylinders, and fuel consumption information for a variety of vehicles both in urban and highway environments. This dataset is considered valuable and reliable, and it is expected to enable the researchers to conduct a thorough analysis and draw meaningful conclusions. 70% of the dataset will be used to train the model and 30% for testing.

The results of training the model indicated that "gaussmf" had achieved the best results with an error rate of 15.3459. Since an error rate of 15% is not considered accurate enough to build a decision on. As a solution, we proposed to modify the dataset by applying the extra-trees regressor on the proposed dataset and generating a new dataset to be used for training the model. The new dataset is around 700% larger in scale than the first dataset which will extend the time of training and testing but will improve the accuracy of the model. With the same division of 70% for training and 30% for testing the model and validating the results. We repeated the procedure of building the model with multiple membership functions, and once again, it was shown that the "gaussmf" had achieved the highest accuracy with an error rate of 1.8226%. In this way, the proposed model was able to decrease the error rate from 15% to 1.8226% as we see in Fig. 5, proving its accuracy and effectiveness compared to the initial unmodified dataset. The main aim of this paper is to develop an efficient and accurate machine learning model that could successfully predict the emissions from associated characteristics of vehicles with a very low error rate. With the modification of the dataset and the use of the extra-trees regressor, we were able to achieve this goal and demonstrate the effectiveness of our proposed model. as show in TABLE I. and TABLE II.

TABLE I. SHOWS THE ERROR RATE FOR THE DIFFERENT MEMBERSHIP FUNCTIONS.

MF	**gbell**	**gauss**	**gauss2**	**trap**	**pi**	**psig**	**tri**	**dsig**
error	15.88	15.34	16.67	18.85	21.59	16.83	15.99	16.83

TABLE II. SHOWS THE ERROR RATE FOR THE DIFFERENT MEMBERSHIP FUNCTIONS USING THE ENHANCED DATASET.

MF	**gbell**	**gauss**	**gauss2**	**trap**	**pi**	**psig**	**tri**	**dsig**
error	1.84	1.82	2.53	8.47	9.15	2.14	3.68	2.14

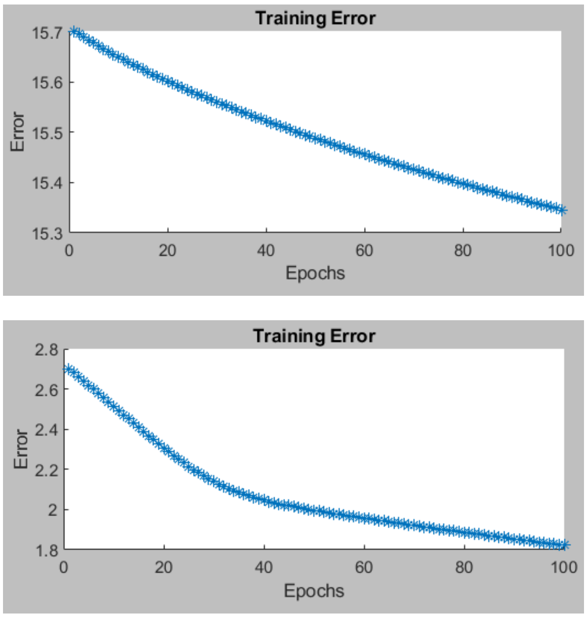

Fig. 5. Proposed ANFIS model accuracy comparison (using the proposed and the generated dataset).

VI. CONCLUSION

The high accuracy of the ANFIS model demonstrates its potential for effectively predicting CO2 emissions from vehicles. The ability of the model to handle non-linear relationships and uncertainty in the data makes it well-suited for this type of problem. Furthermore, the fact that the model considers multiple inputs in its predictions can provide a more complete and accurate representation of the relationship between vehicle characteristics and emissions. In conclusion, the use of ANFIS models for predicting CO2 emissions from vehicles has shown great potential, with high accuracy and the ability to handle complex relationships and uncertainty in the data. These results highlight the importance of utilizing advanced modeling techniques, such as ANFIS, to address the pressing challenges of reducing emissions in the transportation sector and mitigating the impacts of climate change. For future work, the proposed model can be used to make comparisons with different datasets that belong to other countries with different standards.

REFERENCES

[1] G. A. Florides and P. Christodoulides, “Global warming and carbon dioxide through sciences,” Environ. Int., vol. 35, no. 2, pp. 390–401, 2009, doi: 10.1016/j.envint.2008.07.007. (NASA figure adapted from Goddard Institute for Space Studies Surface Temperature Analysis.)

[2] S. H. Rahmati, O. B. Haddad, H. Sedghi, and H. Babazadeh, “a Comparison of Anfis , Ann , Arma & Multivariable Regression Methods for Urban Water-Consumption Forcasting , Considering Impacts of Climate Change : a Case Study on Tehran Mega City,” Indian J. Sci. Res., vol. 7, no. 1, pp. 870–880, 2014.

[3] M. Z. Khan and M. F. Khan, “Application of ANFIS, ANN and fuzzy time series models to CO2 emission from the energy sector and global temperature increase,” Int. J. Clim. Chang. Strateg. Manag., vol. 11, no. 5, pp. 622–642, 2019, doi: 10.1108/IJCCSM-01-2019-0001.

[4] H. Oubehar, A. Selmani, M. H. Archidi, A. Lachhab, and B. Bouchikhi, “Controller for Greenhouse Climate,” 2018 Int. Conf. Electron. Control. Optim. Comput. Sci., pp. 1–4, 2018.

[5] M. Mokhtarzad, F. Eskandari, N. Jamshidi Vanjani, and A. Arabasadi, “Drought forecasting by ANN, ANFIS, and SVM and comparison of the models,” Environ. Earth Sci., vol. 76, no. 21, pp. 0–9, 2017, doi: 10.1007/s12665-017-7064-0.

[6] Patel, Jignesh, and Falguni Parekh. "Forecasting rainfall using adaptive neuro-fuzzy inference system (ANFIS)." Int. J. Appl. Innov. Eng. Manag , vol. 3.6, pp. 262-269 , 2014.

[7] W. Suparta and A. A. Samah, “Rainfall prediction by using ANFIS times series technique in South Tangerang, Indonesia,” Geod. Geodyn., vol. 11, no. 6, pp. 411–417, 2020, doi: https://doi.org/10.1016/j.geog.2020.08.001.

[8] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, pp. 338-353, 1965.

[9] S. Agatonovic-Kustrin and R. Beresford, “Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research,” J. Pharm. Biomed. Anal., vol. 22, no. 5, pp. 717–727, 2000, doi: 10.1016/S0731-7085(99)00272-1.

[10] J.-S. Roger Jang, “01_NeuroFuzzyApproach.pdf.” p. 614, 2000.

[11] J. S. R. Jang, C. T. Sun, and E. Mizutani, "Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence," Prentice Hall, 1997.

[12] L. E. O. BREIMAN, J. Amer Sfofist. Assoc, vol. 89, pp. 1255–1270.

[13] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely randomized trees,” Mach. Learn., vol. 63, no. 1, pp. 3–42, 2006, doi: 10.1007/s10994-006-6226-1.

[14] A. F. Güneri, T. Ertay, and A. Yücel, “An approach based on ANFIS input selection and modeling for supplier selection problem,” Expert Syst. Appl., vol. 38, no. 12, pp. 14907–14917, 2011, doi: https://doi.org/10.1016/j.eswa.2011.05.056.

[15] D. Holtz-Eakin and T. M. Selden, “Stoking the fires? CO2 emissions and economic growth,” J. Public Econ., vol. 57, no. 1, pp. 85–101, 1995, doi: https://doi.org/10.1016/0047-2727(94)01449-X.

[16] D. Elavarasan and P. M. Durai Raj Vincent, “Fuzzy deep learning-based crop yield prediction model for sustainable agronomical frameworks,” Neural Comput. Appl., vol. 33, no. 20, pp. 13205–13224, 2021, doi: 10.1007/s00521-021-05950-7.

[17] Y. Wang and Y. Chen, “A Comparison of Mamdani and Sugeno Fuzzy Inference Systems for Traffic Flow Prediction,” J. Comput., vol. 9, Jan. 2014, doi: 10.4304/jcp.9.1.12-21.

[18] Requena-Mesa, C., Reichstein, M., Mahecha, M., Kraft, B. & Denzler, J. Predicting landscapes as seen from space from

environmental conditions. In IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 1768–1771, IEEE, 2018.

[19] M. Nuruzzaman, Modeling and Simulation in Simulink for Engineers and Scientists. AuthorHouse, 2004.

[20] Jang, J. S. R. "ANFIS: Adaptive-network-based fuzzy inference system". IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3 , pp. 665-685, 1993.

[21] T. Takagi, M. Sugeno, Fuzzy identification of systems and applications to modeling and control, IEEE Trans. On SMC, vol.15, No. 1, pp. 116-132, 1985.

[22] R. E. Al-Jarrah, and A. H. Al-Salaymeh, “A Novel Hybrid ANFIS Model for Predicting the Maximum Temperature”, Journal of Advanced Research in Dynamical and Control Systems, vol. 11, no. 2, pp. 285-294, 2019.

[23] https://open.canada.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64#comments

[24] What are the Sustainable Development Goals? | Policy and advocacy | Sightsavers