



Using Machine Learning Methods to Predict the Effect of Alternative Fuel Mixtures on Exhaust Emissions

Samet Bilban^{1*}, Hasan Aydoğan²

^{1*} Selçuk University, Faculty of Technology, Department of Automotive Engineering, Konya, Turkey (ORCID:0000-0002-9035-0690), sametbilban@hotmail.com

² Selçuk University, Faculty of Technology, Department of Mechanical Engineering, Konya, Turkey (ORCID:0000-0003-1404-6352), haydogan@selcuk.edu.tr

(2nd International Conference on Applied Engineering and Natural Sciences ICAENS 2022, March 10-13, 2022)

(DOI: 10.31590/ejosat.1081539)

ATIF/REFERENCE: Bilban, S. & Aydoğan, H. (2022). Using Machine Learning Methods to Predict the Effect of Alternative Fuel Mixtures on Exhaust Emissions. *Avrupa Bilim ve Teknoloji Dergisi*, (34), 273-279.

Abstract

Machine learning is a science that deals with the design and development processes of algorithms that enable data-based learning. Machine learning methods try to find the most suitable model for new data prediction processes by using the past data. In this study, the data obtained from the engine trials with fuel mixtures of 5%, 10%, 15% by volume using 1-Propanol, 2-Propanol, AVGAS and gasoline fuel were used. Obtained data were compared with 100% gasoline values. In the study, a 4-cylinder engine with direct injection and turbocharging was used. With the obtained measurement results, a database was created to be used in machine learning. With the created database, estimation processes were carried out on ANN, GBA, SVM and AB machine learning models. At the end of the study, it was found that the most suitable model for the estimation of CO, CO₂, HC, O₂ values was ANN with an R² value of 0.9999. For the NO value, it was determined that the AB method was used with an R² value of 0.9996. In the estimation process of the CO value, GBA and AB methods are other machine learning methods that can be used as they have a higher value than 0.99 R². CO₂, HC and O₂, and in the output value estimation process, GBA and AB are other methods that can be used instead of ANN as they have a higher value than 0.99 R². It has been found that there is another machine learning method that can be used for NO value estimation, with an AB 0.99 R² value.

Keywords: Internal combustion engine, Machine learning, Artificial Neural Networks, Engine emissions, Gradient Boosting.

Alternatif Yakıt Karışımlarının Egzoz Emisyonları Üzerindeki Etkisini Öngörmek İçin Makine Öğrenme Yöntemlerinin Kullanılması

Öz

Makine öğrenimi, veriye dayalı öğrenmeyi sağlayan algoritmaların tasarım ve geliştirme süreçleriyle ilgilenen bir bilimdir. Makine öğrenmesi yöntemleri, geçmiş verileri kullanarak yeni veri tahmin süreçleri için en uygun modeli bulmaya çalışır. Bu çalışmada, 1-Propanol, 2-Propanol, AVGAS ve benzin yakıtı kullanılarak hacimce %5, %10, %15 yakıt karışımları ile motor denemelerinden elde edilen veriler kullanılmıştır. Elde edilen veriler %100 benzin değerleri ile karşılaştırılmıştır. Çalışmada direkt enjeksiyonlu ve turboşarjlı 4 silindirli bir motor kullanıldı. Elde edilen ölçüm sonuçları ile makine öğrenmesinde kullanılmak üzere bir veri tabanı oluşturulmuştur. Oluşturulan veri tabanı ile ANN, GBA, SVM ve AB makine öğrenmesi modelleri üzerinde tahmin işlemleri gerçekleştirilmiştir. Çalışma sonunda CO, CO₂, HC, O₂ değerlerinin tahmini için en uygun modelin R² değeri 0,99999 olan YSA olduğu bulunmuştur. NO değeri için R² değeri 0,9996 ile AB yönteminin kullanıldığı belirlendi. CO değerinin tahmin sürecinde GBA ve AB yöntemleri 0,99 R²'den daha yüksek bir değere sahip oldukları için kullanılabilir diğer makine öğrenmesi yöntemleridir. CO₂, HC ve O₂ ve çıkış değeri tahmin sürecinde GBA ve AB 0,99 R²'den daha yüksek bir değere sahip oldukları için YSA yerine kullanılabilir diğer yöntemlerdir. AB 0,99 R² değeri ile NO değeri tahmini için kullanılabilir başka bir makine öğrenmesi yöntemi olduğu tespit edilmiştir.

Anahtar Kelimeler: İçten yanmalı motor, Makine öğrenimi, Yapay Sinir Ağları, Motor emisyonları, Gradyan Artırma.

*Corresponding Author: sametbilban@hotmail.com

1. Introduction

It is seen that the use of internal combustion engines is increasing day by day and the use of individual vehicles increases the use of fossil fuels. It is known that this causes pollution on the atmosphere and global climate changes. For this reason, it becomes very important to reduce the gases released into the atmosphere and the emission values.

In order to reduce the negative effects of internal combustion engines on the world, the search for alternative fuels is being studied. Experiments and tests on alternative fuels and fuel mixtures are of great importance in the world.

While alternative fuels are used in a special way in systems that burn a single fuel, gasoline, which is used as current fuel, can be used as a mixture with other fuels, including LPG or diesel. Oxygen content of gasoline fuel can be increased by using 1-Propanol in order to provide better combustion and lower emission values [1], [2]. 1-Propanol contains three carbons in its structure and is defined by the formula CH_3H_7OH [2], [3]. OH groups in its structure increase the oxygen content of gasoline [4]. At the same time, it increases the knock resistance of gasoline with its high octane number. 1-Propanol, which is widely used as a solvent and anti-icing agent in the industry, contributes to the cleaning of the fuel system [5].

2-Propanol is a colorless flammable and organic compound. 2-Propanol mixes well with many solvents, including water. It is a volatile chemical.

Vapors are heavier than air. It can have a mild irritating effect on the eyes, nose and throat. Its boiling point is 83°C . Its melting point is -89.5°C . It is miscible with alcohol, ether and chloroform in solubility, but it is a chemical that does not dissolve in salt solution. 2-Propanol density is known as 0.875 g/ml at 25°C [6].

AVGAS is an aviation fuel used in spark-ignited internal combustion engines to propel aircraft [7]. It is a fuel obtained by formulating gasoline in aircraft piston engines and called aviation gasoline.

In order to analyze the effects of gasoline additive fuel mixtures on engine performance values, complex mathematical operations and long-term experimental studies are needed. Extra statistical methods are used to reveal the relationship between the features that affect the motor performance values. Measurements can be made in a specific engine speed range and using a specific test setup. Machine learning methods can be used to overcome all these problems.

By using 1-Propanol, 2-Propanol AVGAS gasoline fuel, additive-gasoline mixtures at 5%, 10%, 15% by volume were obtained. These mixtures were tested in the engine and compared with 100% gasoline values. The obtained data set was used in machine learning algorithms. Thus, exhaust emission values could be estimated at different rates and for intermediate turnover values without testing in the laboratory environment.

In this study, machine learning methods such as Artificial Neural Network (ANN), Support Vector Machines (SVM), Gradient Boosting algorithm (GBA) and Adaboost (AB) were used using the estimation of the parameters affecting the emission values (CO , CO_2 , HC , NO , O_2) of alternative fuel mixtures has been carried out. The results of each machine learning methods

are discussed comparatively and their effects on the result are given.

2. Material and Method

In this section, the use and application of parameters affecting the emission values of alternative fuel mixtures on estimation processes are discussed in detail by using ANN, SVM, AB and GBA methods from machine learning methods.

ML consists of methods used to extract information from data. ML methods cover a field of study that combines statistics, artificial intelligence and computer science. Also known as predictive analytics or statistical learning. Many websites and devices are based on ML algorithms [8].

ML methods require a comprehensive data set to predict relationships between data. Machine learning techniques are divided into three subgroups as unsupervised machine learning (training with unclassified or unlabelled data), supervised machine learning (training with classified or labelled data), and reinforced machine learning (includes reward or experience-based methods) as seen in Figure 1 [9].

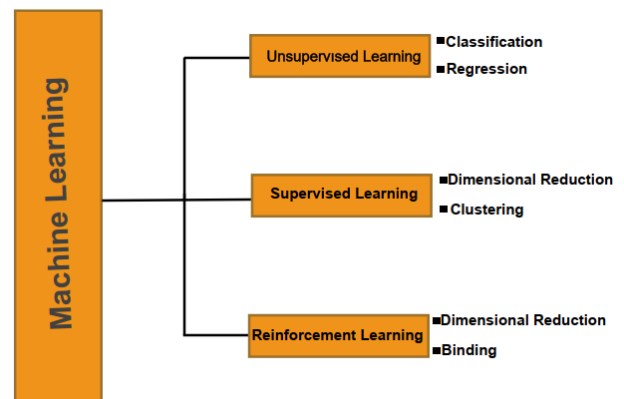


Figure 1. Classification of machine learning approaches

Supervised Learning: In a system where supervised learning method is used, inputs and corresponding outputs are given in ML algorithms and the system tries to establish a relationship between input and output data. The purpose of the established relationship is to estimate the output of unused data. Supervised learning methods can be used to classify data and perform regression operations depending on whether the data is quantitative or qualitative. Machine learning methods that apply qualitative data are used for classification operations. Machine learning methods for quantitative data perform regression operations.

Unsupervised Learning: In unsupervised learning, when a data set is given to the system without input-output information, the system tries to discover whether there is a possible relationship in this data. The main purpose of unsupervised learning is to analyse data by segmenting and clustering the data.

Reinforcement Learning: In reinforcement learning, the system interacts with an environment and learns according to the results of actions, that is, it refers to the learning process based on experience.

2.1. Gradient Boosting Algorithm

The gradient boosting algorithm is known as one of the decision tree ensembles [10]. GBA is one of the machine learning methods used for classification and regression problems. An iterative training process is applied to minimize the value of the lost function. The data set used for training the model is used to estimate the output label value of each data. The dataset is relabelled and errors are eliminated in the estimation process to give more importance to training data with poor results. GBA is used in datasets where prediction results with high accuracy are needed [11]. The use of GBA in multi-parameter problems is increasing day by day due to its high performance values in nonlinear problems [12]. GBA pseudocode is given in Table 1 [12].

Table 1. GBA pseudocode representation

Input: The model training dataset $\{(x_i, y_i)\}_{i=1}^n$, $L(y, F(x))$ is defined as the missing function and the maximum number of iterations M .

- The model is initially initialized with a fixed value:

$$F_0(x) = \underset{\gamma}{\operatorname{arg\,min}} \sum_{i=1}^N L(y_i, \gamma_i)$$
- For $m = 1$ from up to M :

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)} \text{ for } i = 1, 2, 3 \dots n$$

$p_m(x)$ represents the model prediction value.

$$\gamma = \underset{\gamma}{\operatorname{arg\,min}} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \gamma p_m(x_i))$$

$$F_m(x) = F_{m-1}(x) + \gamma p_m(x)$$
- Output value:** $F_m(x)$

2.2. AdaBoost

Adaboost is defined as one of the ensemble regression methods, also known as Adaptive Boost. It is also one of the machine learning methods used in regression and classification processes. Each node on the Adaboost regression tree holds a feature in the dataset, and each branch represents the output from the model testing process.

In the Adaboost regression method, the estimation result is obtained by summing the estimations obtained or by the principle of majority vote. Each boosting operation is performed by applying the weights w_1, w_2, \dots, w_n to the model training data one by one. In the first step, all of the weights are determined as $w_i=1/N$, so the model training process is started on the real data.

As a result of each increment, the weights are updated and the learning algorithm is reapplied. After repeated operations, the

boosted model increased the weights of the training data that could not be predicted correctly and decreased the weight of the data that could not be predicted correctly. The larger the number of increments, the more difficult-to-predict samples will have an increasing effect on the model. The pseudocode containing all these operations is given in Table 2 [12],[13].

Table 2. Adaboost pseudocode representation

Logic: Training dataset $T = \{(x_1, y_1), \dots, (x_n, y_n)\}$, $x_i \in X$ and starting weights $w_i^1 = \frac{1}{N}, i = 1, \dots, N$ is defined as N represents the number of sample data sets. ca_i value of data i . corresponds to the value of the example in the real category. K represents the total number of samples in the same category. $\Pi(\cdot)$ pointer function and Z_t is called the normalization factor.

for $t=1 \dots T$:

With the SAMME.R algorithm, w^t weights are assigned to the samples and placed in the $h^t(x)$ function. As a result of the regression process of the $h^t(x)$ function, the error value is calculated and the training **error**^t is calculated as error.

$$\mathbf{error}^t = \sum_{i=1}^N w_i^t \Pi(ca_i \neq h^t(x_i))$$

In the regression process, the weights are calculated and kept in the variable a^t .

$$a^t = \ln \frac{1 - \mathbf{error}^t}{\mathbf{error}^t} + \ln(K - 1)$$

Thus, the weights of the samples are recalculated as w_i^t and the update process is repeated.

$$w_i^{t+1} = \frac{w_i^t}{Z_t} \cdot \exp(a^t \Pi(ca_i \neq h^t(x_i)))$$

2.3. Support Vector Machine

A support vector regression method has been developed in order to use the SVM classification method in regression problems. As Vapnik suggested that SVM be used in regression analysis, many researchers started to use it frequently in regression (SVR) processes [14]. While linear regression models try to keep the error in the minimum range, SVR tries to keep the error rate within a certain limit value range. It can be more advantageous than many other machine learning methods because it has good generalization ability on small data sets [15].

While the structure of the SVR and the hyperparameters used determine the support vectors, the kernel functions are used to determine the feature areas. Input parameters, x_i and y_i If we define it as a vector consisting of n , including the total number of data, $\{(x_i, y_i), \dots, (x_n, y_n)\}_{i=1,2,\dots,n}$ is defined as.

SVR generates an output function using the equation given in equation 1 [16].

The $\phi(x)$ given in equation 1 is defined as an area with a large dimension.

$$y = w \cdot \phi(x) + b \tag{1}$$

2.4. Artificial Neural Networks

ANN, one of the application areas of artificial intelligence, is one of the machine learning methods used to model complex systems. ANNs provide convenience in solving many problems with their ability to interpret nonlinear relationships between independent and dependent variables. The ANN structure consists of an input layer, an output layer, and one or more hidden layers.

ANN is the representation of the human nervous system, which consists of neurons that communicate with each other through connections called axons. Neurons consist of one or more connections. Each link is assigned a weight value. The output link is responsible for carrying the incoming signals to other neurons. Signals are aggregated according to the activation threshold level of each neuron. The output signal is determined by passing all collected signals through an activation function. As shown in Equation 2, X is an input vector and y is a perceptron that produces a single output [18]. w is the value of the input weights, the b bias value and Φ a value corresponding to the nonlinear activation function.

$$y = \Phi(\sum w_i x_i + b) \tag{2}$$

The activation function is used to solve nonlinear problems with ANN. Activation function ensures that the output signal is not a linear value [19].

The three most frequently used activation functions in the literature are hyperbolic tangent, ReLU and sigmoid. Let l represent the number of inputs, m the number of hidden units, and n the number of output units. The formula for each hidden layer is defined in equation 3.

$$h_j = f(\sum_{i \rightarrow j} w_{ij}^{(1)} x_i + w_{0j}^{(1)}) \tag{3}$$

When the values obtained in the output layer are compared with the actual values, deviations/errors may occur. Weights, bias value and parameters should be adjusted to reduce the error in the output layer. The error value is calculated as given in equation 4.

$$\Delta_w = e_k * x_{jk} * a \tag{4}$$

e_k represents the error of the output unit, x_{jk} and k values are the input values causing the error and a value is defined as the learning rate, which determines how much the weights need to be changed to correct the error. After the error is calculated, the variation of the weights in the hidden layer is calculated as in equation 5.

$$e_j = w_j * e_k * t(y_k) \tag{5}$$

The weight updates are made by adding the e_j values obtained to the weights of the network determined at the beginning. Thus, this process is repeated by going forward and backward until an optimum and generalized ANN is obtained.

2.5. CO, CO₂, HC, NO, O₂ Hyperparameters Used to Estimate Values

Hyper Parameter optimization is the optimum parameter selection process for machine learning algorithms. Hyperparameters are the values that affect the learning process of the model. A machine learning model training process involves choosing the optimal parameters for which the learning algorithm will be used to learn the optimal parameters that accurately map to the input values (independent variable) or target value (dependent variable).

Table 3. Hyperparameters used for all machine learning methods

Model Name	Hyperparameters
AB	<ul style="list-style-type: none"> It is a basic predictive tree. The number of estimators was set to 50. Learning rate fixed at 1.0. The SAMME.R algorithm was used as the boosting method (which achieves a lower test error rate with faster convergence and fewer boost iterations). For the regression process, the missing function was determined as exponential.
SVM	<ul style="list-style-type: none"> The v-SVM parameter is selected as the SVM type. C (regression cost) parameter 4. ν (Determining an upper and lower limit for regression tasks) was determined as 0.5. Kernel function is assigned as Linear. Numerical deviation optimization value (permissible deviation from expected value) value 10^{-4}.
ANN	<ul style="list-style-type: none"> The number of neurons in each hidden layer is set to 10. Activation function Logistic is selected. The weight dissolving function was determined as L-BFGS-B. $\alpha = 10^{-4}$ (L2 punishment arrangement term) The maximum iteration is set at 200.
GBA	<ul style="list-style-type: none"> The number of trees to be included is determined as 100. The learning rate is set at 0.1. The maximum depth of the tree is determined as 3. The smallest number of subsets that can be divided was determined as 2. The number of training samples to be placed on the tree was determined as 1.0.

3. Result and Discussion

Fuel type, engine speed, engine torque, engine power, exhaust gas temperature, noise value and specific fuel consumption were used as input parameters of the machine learning model. Model output parameters, exhaust emissions CO, CO₂, HC, NO and O₂ were determined. The correlation coefficient (R²) value was used from the comparison between the experimental data and the predicted results. For each output value, machine learning models are trained with 80% training and 20% testing of all data. R², MSE and MAE values were calculated over the predicted values of the models.

ANN was determined as the most suitable model for estimating the CO output value. ANN was obtained as a generalized model for CO output value with 0.0001 MSE, 0.0006 MAE and 0.9999 R² value. If the model fitness value is considered as 90% and above, GBA and AB values are other methods with high accuracy in estimating the CO output value. ANN was chosen because it has lower MAE and higher R² value than GBA and AB methods. The ANN-CO predicted-actual value graph is given in Figure 2.

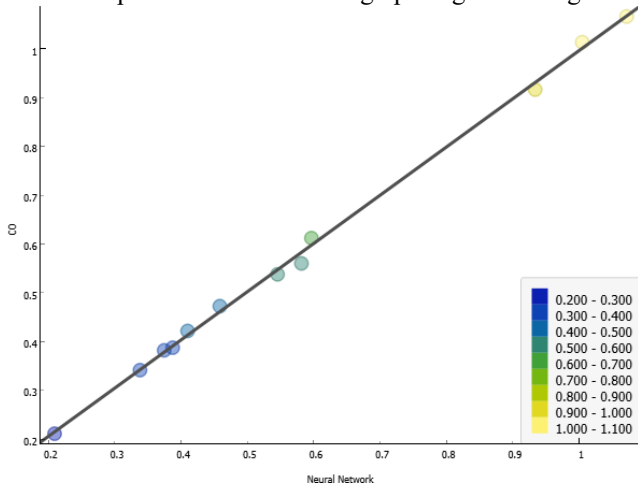


Figure 2. ANN-CO predicted-actual graph

In estimating the CO₂ value, ANN was obtained as a more generalized model compared to other methods, with 0.001 MSE, 0.0024 MAE and 0.9999 R² values. 0.9996 R² GBA and 0.9987 AB values are other machine learning models that follow the ANN model with high accuracy values. ANN-CO₂ predicted-actual value graph is given in Figure 3.

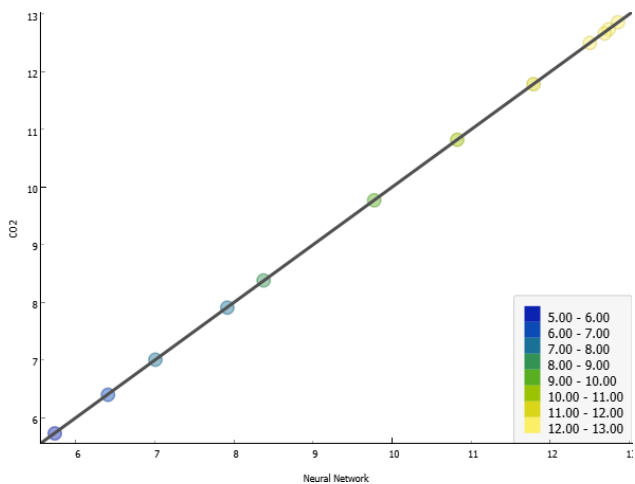


Figure 3. ANN- CO₂ predicted-actual graph

For the HC value estimation process, ANN was found to be the most suitable model with 0.2368 MSE, 0.0001 MAE and 0.9999 R² values. GBA, AB and SVM are other machine learning methods with model fitness values over 90% accuracy. Since ANN has lower model error values and high R² value, it helped to determine it as a more suitable model than other methods. The ANN-HC predicted-actual value graph is given in Figure 4.

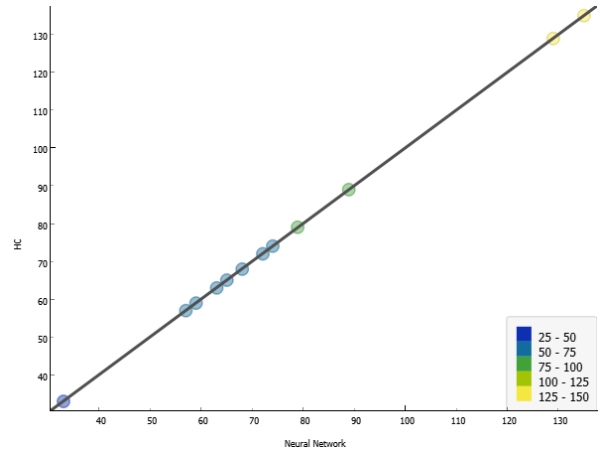


Figure 4. ANN- HC predicted-actual graph

For the estimation of the O₂ model output value, the ANN method is a model with higher accuracy than other methods with 0.008 MSE, 0.0158 MAE and 0.9999 R² values. GBA 0.9997 and AB 0.9989 are other methods that can be used with R² values for the estimation of the O₂ value. ANN-O₂ predicted-actual value graph is given in Figure 5.

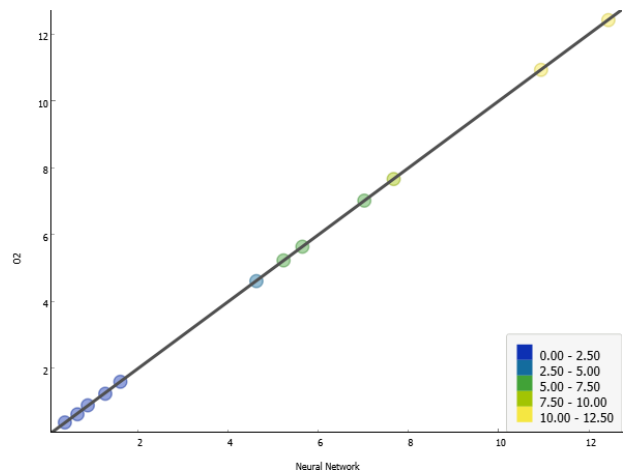


Figure 5. ANN- O₂ predicted-actual graph

The AB method was determined to be the most suitable method for estimating the NO value with 173.5416 MSE, 3.1251 MAE and 0.9996 R² values. It was seen that the most suitable model was the GBA method with 322.3016 MSE, 14.5745 MAE and 0.9994 R² values. It has been determined that the ANN and SVM methods cannot be used in NO estimation since the accuracy value is below 90%. It has been shown that tree-based AB and GBA methods can be used more appropriately than ANN and SVM methods. In the estimation process of the NO value, the MSE and MAE error values were high because the actual values of the NO value vary between approximately 1000-3250 when looking at the database. Therefore, MSE and MAE values were high in NO estimation. The AB-NO predicted-actual value graph is given in Figure 6.

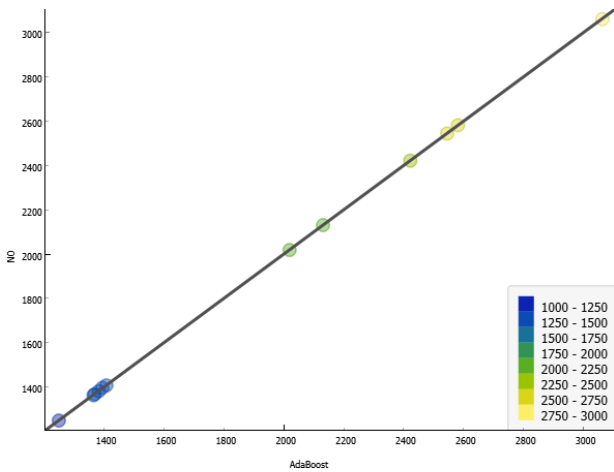


Figure 6. AB-NO predicted-actual graph

Today, it has been determined that there are very few studies conducted in our country and internationally on different fuel mixtures in an engine with TSI technology. In this respect, the experiments and research on the TSI engine, which is a new and up-to-date engine technology, will guide researchers and students who will work in this field.

It has been demonstrated that researchers who will work with machine learning and alternative fuel mixtures can create their own datasets and how to predict and improve prediction processes with high accuracy values without continuous experimental studies in the real test environment.

4. CONCLUSION

In this study, a test process was carried out on a 1.2 TSI engine using 1-Propanol, 2-Propanol and AVGAS-added gasoline blends. In the test processes on the engine, each fuel mixture with additives was tested separately. As a result of each test process, engine speed, engine power, engine torque, exhaust gas temperature, noise, CO, CO₂, HC, NO, O₂ and specific fuel consumption values were taken. After all the testing procedures were completed, the database was created. This database was used to create models in ANN, SVM, GBA and AB machine learning methods. Table 4 shows which generalized model is for each output value.

Table 4. Optimal model representation for all machine learning prediction values

Tahmin	Model	MSE	MAE	R ²
CO	ANN	0.0001	0.0006	0.9999
CO ₂	ANN	0.0001	0.0024	0.9999
HC	ANN	0.2368	0.0001	0.9999
O ₂	ANN	0.0008	0.0158	0.9999
NO	AB	173.5416	3.1251	0.9996

As seen in Table 4, ANN was found to be the most suitable model for the estimation of CO, CO₂, HC, O₂ values. It was found to be AB for the NO prediction value.

It has been seen that using 1-Propanol, 2-Propanol and AVGAS-added gasoline blends, intermediate value estimations can be made without performing real test operations.

References

- [1] Y. Qian, J. Guo, Y. Zhang, W. Tao, and X. Lu, (2018). "Combustion and emission behavior of N-propanol as partially alternative fuel in a direct injection spark ignition engine," *Appl. Therm. Eng.*, vol. 144, pp. 126–136, doi: 10.1016/J.APPLTHERMALENG.2018.08.044.
- [2] M. S. Gökmen, İ. Doğan, and H. Aydoğan, (2021). "Yanıt Yüzey Metodolojisi Kullanılarak 1-Propanol/Benzin Yakıt Karışımlarının Egzoz Emisyonlarına Etkisinin Araştırılması," *Eur. J. Sci. Technol.*, no. 24, pp. 67–74, doi: 10.31590/ejosat.898563.
- [3] G. R. Gawale and G. Naga Srinivasulu, (2020). "Experimental investigation of propanol dual fuel HCCI engine performance: Optimization of propanol mass flow rate, impact of butanol blends (B10/B20/B30) as fuel substitute for diesel," *Fuel*, vol. 279, p. 118535, doi: 10.1016/J.FUEL.2020.118535.
- [4] M. Mourad and K. R. M. Mahmoud, (2018). "Performance investigation of passenger vehicle fueled by propanol/gasoline blend according to a city driving cycle," *Energy*, vol. 149, pp. 741–749, doi: 10.1016/J.ENERGY.2018.02.099.
- [5] X. Liu, H. Wang, Z. Zheng, J. Liu, R. D. Reitz, and M. Yao, (2016). "Development of a combined reduced primary reference fuel-alcohols (methanol/ethanol/propanols/butanols/n-pentanol) mechanism for engine applications," *Energy*, vol. 114, pp. 542–558, doi: 10.1016/J.ENERGY.2016.08.001.
- [6] A. Kimya, "Ataman Kimya," (2019). <https://atamankimya.com>.
- [7] Shell, "Shell," (2010). <https://www.shell.com/business->
- [8] A. C. Müller and S. Guido, (2020). *Introduction to Machine Learning with Python*.
- [9] A. U. Osarogiagbon, F. Khan, R. Venkatesan, and P. Gillard, (2021). "Review and analysis of supervised machine learning algorithms for hazardous events in drilling operations," *Process Saf. Environ. Prot.*, vol. 147, pp. 367–384, doi: 10.1016/J.PSEP.2020.09.038.
- [10] T. Hastie, R. Tibshirani, and J. Friedman, (2008). *The Elements of Statistical Learning*.
- [11] S. Raschka, D. Julian, and J. Hearty, (2016). *Python: deeper insights into machine learning: leverage benefits of machine learning techniques using Python: a course in three modules*.
- [12] U. Ozkaya and L. Seyfi, (2015). "Dimension optimization of microstrip patch antenna in X/Ku band via artificial neural network" *Procedia-Social and Behavioral Sciences*, 195, pp. 2520-2526.
- [13] Q. Liu, X. Wang, X. Huang, and X. Yin, (2020). "Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data," *Tunn. Undergr. Sp. Technol.*, vol. 106, no. August, p. 103595, doi: 10.1016/j.tust.2020.103595.
- [14] V. Vapnik, S. E. Golowich, and A. Smola, (1997). "Support vector method for function approximation, regression estimation, and signal processing," *Adv. Neural Inf. Process. Syst.*, pp. 281–287.
- [15] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, (2019). "Vector field-based support vector regression for building energy

- consumption prediction,” *Appl. Energy*, vol. 242, no. September 2018, pp. 403–414, doi: 10.1016/j.apenergy.2019.03.078.
- [16] B. Dong, C. Cao, and S. E. Lee, (2005). “Applying support vector machines to predict building energy consumption in tropical region,” *Energy Build.*, vol. 37, no. 5, pp. 545–553, doi: 10.1016/J.ENBUILD.2004.09.009.
- [17] J. Lin, C. Cheng, and K.-W. Chau, (2006). “Using support vector machines for long-term discharge prediction,” *Hydrol. Sci. J.*, vol. 51, no. 4, pp. 599–612, doi: 10.1623/hysj.51.4.599.
- [18] P. Golbayani, I. Florescu, and R. Chatterjee, (2020). “A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees,” *North Am. J. Econ. Financ.*, vol. 54, p. 101251, doi: 10.1016/J.NAJEF.2020.101251.
- [19] A. Domínguez-Sáez, G. A. Rattá, and C. C. Barrios, (2018). “Prediction of exhaust emission in transient conditions of a diesel engine fueled with animal fat using Artificial Neural Network and Symbolic Regression,” *Energy*, vol. 149, pp. 675–683, doi: 10.1016/J.ENERGY.2018.02.080.