



# A Comparative Study of Machine Learning Techniques in Prediction of Exhaust Emissions and Performance of a Diesel Engine Fuelled with Biodiesel Blends

Quang Hung Do\*†, Shih-Kuei Lo\*\* and Jeng-Fung Chen\*\*

\*Faculty of Information Technology, University of Transport Technology, Hanoi 100000, Vietnam

\*\*Department of Industrial Engineering and Systems Management, Feng Chia University, Taichung 40724, Taiwan

†Corresponding author: Quang Hung Do; hungdq@utt.edu.vn

Nat. Env. & Poll. Tech.  
Website: [www.neptjournal.com](http://www.neptjournal.com)

Received: 20-04-2020

Revised: 18-07-2020

Accepted: 20-07-2020

## Key Words:

Exhaust emissions  
Diesel engine performance  
Biodiesel engine  
Machine learning

## ABSTRACT

Biodiesel has been receiving increasing attention because of its fuel properties and compatibility with petroleum-based diesel fuel. Therefore, it is necessary to measure the engine performance and exhaust emissions of engines using petroleum-based diesel fuel and biodiesel blends. The main goal of this study is to investigate the capability of several machine learning (ML) techniques including artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), general regression neural network (GRNN), radial basis function (RBFN), and support vector regression (SVR) for predicting performance and exhaust emissions of the diesel engine fuelled with biodiesel blends. The case application is a Hyundai D4CB 2.5 engine together with B0, B10 and B20 biodiesel blends which are popularly used in Vietnam. The engine process parameters are used as inputs and the outputs include predicted torque and NOx emission. Different predicting models based on ML techniques are developed and validated. The performance of each model is evaluated and compared using root mean squared error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), and correlation coefficient (R). The obtained results indicate that SVR can be used to develop the model for the prediction of performance and exhaust emissions. The study also provides a better understanding of the effects of engine process parameters on performance and exhaust emissions.

## INTRODUCTION

The diesel engines have been proved to be cost-effective and an ideal replacement for steam engines. The increasing use of diesel engines in on-road, off-road vehicles and other fields such as industry and agriculture sector results in the climbing threat of air pollutions. Biodiesel is an alternative, renewable, clean diesel fuel similar to conventional or 'fossil' diesel. The process to produce biodiesel is the conversion from the triglyceride fats to esters which is called transesterification with methanol/ethanol. Biodiesel can be made from natural sources such as vegetable oil, animal oil/fats, tallow and waste cooking oil (Meher et al. 2006, Ramadhas et al. 2005, Van 2005). Biodiesel in blends or neat form is a good successor for petroleum diesel and could be used directly in any diesel engine (Agarwal et al. 2008). However, it is more practical that biodiesel should substitute for a fraction of petroleum diesel in use when the alternative fuel is gradually introduced as blends with diesel (Mamilla & Rao 2016, Shailaja & Raju 2017). Blending biodiesel with diesel in an adequate percentage can achieve a better result than that of diesel or pure diesel. Therefore, it is necessary

to evaluate the blends of biodiesels for their performances and exhaust emissions.

The computational intelligence approach can explore the nonlinear relationship and discover hidden knowledge from the dataset. As a result, the approach has been applied to several practical problems in various scientific disciplines. In recent years, computational intelligence approaches in general, and machine learning (ML) techniques, in particular, have been used to predict engine performance and exhaust emissions (Canakci et al. 2009, Do et al. 2020, Liu et al. 2018). Ghanbari et al. (Ghanbari et al. 2015) used a support vector machine (SVM) to predict the performance parameters and exhaust emissions of a diesel engine operating on nano diesel blended fuels. The results showed that SVM is capable of predicting diesel engine performance and emissions. The use of machine learning methods may be highly recommended to predict engine performance and exhaust emissions instead of having to carry out complex and time-consuming experiments. Other than that, in the field of automotive engineering, ML techniques provided a more accurate and simple method than in the analysis of engine

performance and exhaust emissions. In modelling biodiesel processes, artificial intelligence has been a good replacement for conventional computing based on polynomial and linear regressions in resolving the nonlinear prediction problems (Noor 2014).

In general, the above-mentioned studies have shown that the ML-based models are efficient in predicting water quality. In developing ML-based models, since each dataset has its characteristics and features, no model can be efficiently applied to all types of data. In this study, the main objective is to develop different ML models including ANN, ANFIS, GRNN, RBFN, SVR for predicting performance and exhaust emissions of the diesel engine fuelled with biodiesel blends. In the study, an experimental setup was firstly designed to collect the dataset. The experimenter was conducted in Motors Research Laboratory at the University of Transport Technology. Biodiesel blends including B0, B10 and B20 were used as fuel in a Hyundai engine. The performance and exhaust emissions were then recorded in real-time. This study examined the validity of developed models. The obtained dataset was applied to the developed models to perform the prediction of performance and emissions of the engine. Several performance criteria were utilized to identify which model is the best one that can provide superior predictions of exhaust emissions and performance of a diesel engine fuelled with biodiesel blends.

**PRELIMINARIES**

**Artificial Neural Network (ANN)**

A neural network in which activations spread only in a forward direction from the input layer through one or more hidden layers to the output layer is known as a multilayer feed-forward network. For a given set of data, a multi-layer feed-forward network can give a good non-linear relationship. It is affirmed that a feed-forward network even with

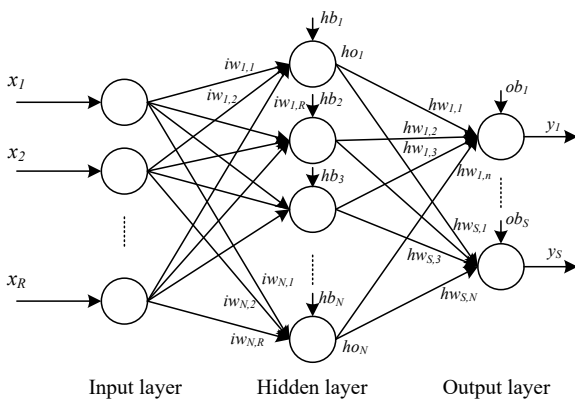


Fig. 1: A feed-forward network with three layers.

only one hidden layer can approximate any continuous function (Hornik et al. 1989, Ken-Ichi 1989). In Fig. 1,  $R$ ,  $N$ , and  $S$  are the numbers of input, hidden neurons, and output, respectively;  $iw$  and  $hw$  are the input and hidden weights matrices, respectively;  $hb$  and  $ob$  are the bias vectors of the hidden and output layers, respectively;  $x$  is the input vector of the network;  $ho$  is the output vector of the hidden layer, and  $y$  is the output vector of the network. A network with a structure that is more complicated than necessary over fits the training data (Caruana et al. 2001).

**Adaptive Neuro-Fuzzy Inference System (ANFIS)**

A fuzzy inference system (FIS) employs fuzzy if-then rules when acquiring knowledge from human experts to deal with imprecise and uncertain problems (Yusof et al. 2012). FISs have been widely used to solve different classification problems (Fakhrahmad et al. 2012). However, fuzzy systems cannot learn from or adjust themselves (Ata & Kocycigit 2010). The system can overcome some limitations of neural networks, as well as the limits of fuzzy systems (Nauck et al. 1997, Singh et al. 2005) when it can represent knowledge in an interpretable manner and the ability to learn. The details of the neuron-fuzzy system were proposed by Takagi and Hayashi (Takagi & Hayashi 1991). Among the neuron-fuzzy systems, ANFIS, introduced by Jang (Jang 1993), has been one of the most common tools.

To present the ANFIS architecture and simplify the explanations, assume that the FIS has two inputs,  $x_1$  and  $x_2$ , two rules, and one output,  $y$  as shown in Fig. 2. A FIS has two inputs and two fuzzy if-then rules that can be expressed as in Equation 1.

Where  $x_1$  and  $x_2$  are the inputs;  $A_1, A_2, B_1, B_2$  are the linguistic labels;  $p_i, q_i$  and  $r_i$  ( $i = 1$  or  $2$ ) are the consequent parameters (Jang 1993) that are identified in the training process; and  $y_1$  and  $y_2$  are the outputs within the fuzzy region. Eq. (1) represents the first type of fuzzy if-then rules, in which the output part is linear. The output part can also be constants (Sugeno 1985) and represented as:

$$\begin{aligned} \text{Rule 1: If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1 \text{ then } y_1 &= C_1 \\ \text{Rule 2: If } x_1 \text{ is } A_2 \text{ and } x_2 \text{ is } B_2 \text{ then } y_2 &= C_2 \quad \dots(1) \end{aligned}$$

Where ( $i = 1$  or  $2$ ) are constant values.

For complicated problems, the first type of if-then rules is widely used to model the relationships of inputs and outputs (Wei et al. 2007). In this research, we also used a linear function for the output.

It can be seen that the ANFIS architecture has two adaptive layers: layer 1 and layer 4. Layer 1 has parameters related to the fuzzy membership functions and layer 4 has parameters  $\{p_i, q_i, r_i\}$  related to the polynomial. Adjusting

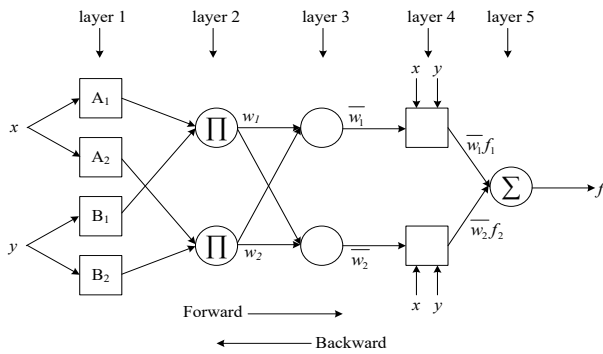


Fig. 2: An ANFIS architecture of two inputs and two rules.

the parameters includes two steps. In the forward pass of the learning algorithm, the premise parameters are fixed, functional signals go forward until layer 4, and the consequent parameters are identified by the least squares method to minimize the measured error. In the backward pass, the consequent parameters are fixed, the error signals go backward, and the premise parameters are updated by the gradient descent method (Jang et al. 1997). This hybrid learning algorithm can decrease the complexity of the algorithm and increase learning efficiency (Singh et al. 2005).

**Radial Basis Function (RBFN)**

The architecture of RBF network includes three layers: the input layer, the hidden layer, and the output layer, as shown in Fig. 3. Although the structure of Radial Basis Function (RBF) neural network is rather simple, the network has a strong generalization ability (Jiang et al. 2016). The RBF neural network has shown a good classification and approximation performance in various applications (Batool et al. 2013, Guan et al. 2016).

As shown in Fig. 3, the estimated output is a weighted summation using the following equation:

$$y_s = \sum_{j=1}^J w_{js} R_j(x), s = 1, 2, \dots, S \quad \dots(2)$$

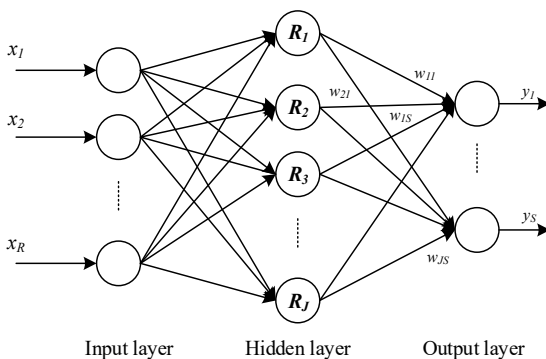


Fig. 3: A RBF network.

Where  $S$  denotes the number of outputs,  $J$  is the number of nodes in the hidden layer and  $w_{js}$  is the connection weight between  $j$ -th node of the hidden layer and  $s$ -th node of the output layer. There are several radial basis functions, the most commonly used one is as follows:

$$R_j(x) = \exp\left(-\frac{\|x-c_j\|^2}{2\sigma_j^2}\right), j=1,2,\dots, J \quad \dots(3)$$

Where  $x$  is the input pattern vector, each input is represented by the  $N$ -dimensional vector;  $c_j$  and  $\sigma_j$  are the centre and the width of RBF, respectively;  $\|x - c_j\|$  is the norm of the vectors  $x$  and  $c_j$ , which can be considered as the distance between the vectors  $x$  and  $c_j$ .

**General Regression Neural Network (GRNN)**

GRNN, a modified version of radial basis function (RBF) networks, is designed for function regression and approximation (Yaseen et al. 2016). GRNN can be successfully applied to deal with various linear and non-linear problems and can accurately perform predictions without requiring large samples (Alilou & Yaghmaee 2015).

GRNN represents a structure that contains along with a hidden radial layer both probabilistic hidden and output layers. These networks are based on the method for approximation of probability density with the help of Gaussian kernel functions.

$$\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \dots(4)$$

GRNN has several outstanding characteristics including the ability to approximate any smooth function and generating consistent forecasts when the training data set size becomes big and the estimation error closes to zero, with only mild restrictions on the function (Cigizoglu & Alp 2006, Kim & Kim 2008). There are four layers in GRNN, including the input, pattern, summation, and output layers. The number of neurons in the input and output layers represent the independent and dependent variables, respectively. The number of neurons in the pattern layer is the number of data patterns in training. The summation layer has the number of neurons that is equivalent to the number of output nodes plus one (Ji et al. 2017).

**Support Vector Regression**

Support Vector Machine (SVM) method is a supervised learning method influenced by advances in statistical learning theory (Abe 2010, Meyer et al. 2003). SVM has been successfully applied in solving classification, prediction and recognition problems (Do & Chen 2013). If a training dataset has feature-label pairs  $(x_i, y_i)$  with  $i = 1, \dots, n$ . The optimum separating hyperplane will be represented as follows:

$$g(x) = \text{sign}\left(\sum_{i=1}^n y_i \alpha_i K(x_i, x_j) + b\right) \quad \dots(5)$$

Where  $K(x_i, x_j)$  denotes kernel function;  $\alpha_i$  is a Lagrange multiplier; and  $b$  is the offset of the hyperplane from the origin. The constraints are  $0 \leq \alpha_i \leq C$  and  $\sum \alpha_i y_i = 0$ , where  $\alpha_i$  is a Lagrange multiplier for each training point and  $C$  is the penalty. Only the training points being close to the support vectors have nonzero  $\alpha_i$ . However, in real-world applications, data are noisy and no linear separation exists in the feature space. Thus, the optimum hyperplane is written as:

$$y_i(w, x_i + b) \geq 1 - \vartheta_i, \vartheta_i \geq 0 \quad \dots(6)$$

Where  $w$  denotes the weight vector that determines the orientation of the hyperplane in the feature space;  $\vartheta_i$  is the  $i$ th positive slack variable that estimates the amount of violation according to the constraints.

## RESEARCH DESIGN

### Experimental Setup

As shown in Fig. 4, the assembled model includes Alpha 160 which is a brake testing device. Moreover, AVL-553S-200 is a temperature controller for cooling water; AVL PLU 160 estimates fuel consumption. BOBCAT is a system for automation of measurement and testbed. I/O Cube is designed to get signals from sensors; FEM is for converting the signal. K57 is a control board; ECU is an engine control unit; throttle pedal is used for the brake and clutch pedal assembly; FTIR (Fourier-transform infrared spectroscopy) is used to obtain an infrared spectrum of absorption or emission; AVL415SE is a smoke meter. Testo 350 is an emission gas analyser and

a picoscope is used to test the control signal.

The study was conducted in Motors Research Laboratory at the University of Transport Technology. The connection of the modules is shown in Fig. 4. The operating parameters and the number of emissions were updated in real-time; a picoscope module connected to a computer was used to record the control signal from the EGR valve.

Motor moment was measured by an eddy current dynamometer (Alpha 160) which has a max power of 160 kW, a max torque of 400 Nm and a max speed of 10000 RPM. With a throttle actuator, testing brake can adjust the driving resistance load and the speed of an automobile.

Fuel flow was measured and analysed by AVL PLU 160 - a high precision fuel consumption measurement system in which a gear meter driven by a servo motor with encoder defines a geometric volume to pulse frequency ratio when gear rotation is adjusted to media flow. The system is a 0-160 litre/hour volumetric flow meter with a maximum error of 2%.

The FT-IR (Fourier transformation infrared spectroscopy) was applied to a continuous measurement of main pollutant emissions from biodiesel engines (i.e., carbon monoxide-CO, hydrocarbons-HC, nitrogen oxides-NOx, carbon dioxide-CO<sub>2</sub>).

Particulate matter-PM was measured and analysed by an AVL Smoke Meter (AVL 415SE) which uses the filter paper method to determine the Filter Smoke Number (FSN defined according to ISO 10054) and the soot concentration in mg/m<sup>3</sup>. The measurement precision is 0.0001 FSN.

AVL 553S200 is used to bring the engine coolant and excellent temperature stability. This model is designed for the

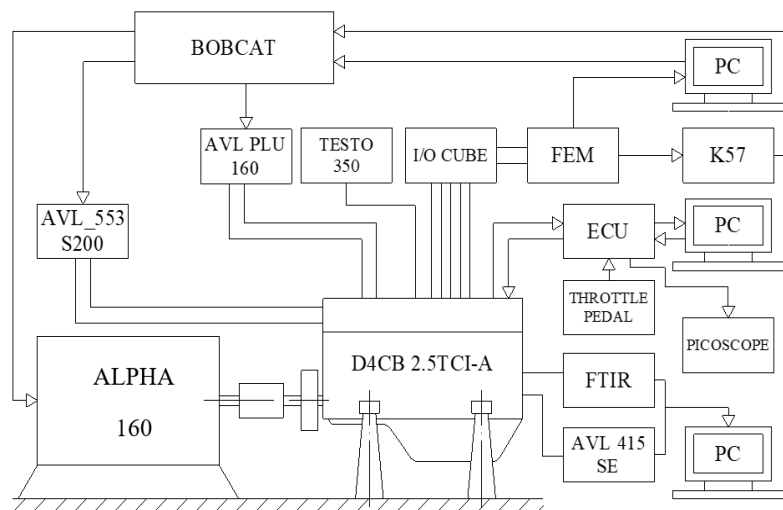


Fig. 4: The experimental setup.

Table 1: Fuel properties of B0, B10 and B20.

Property	B0	B10	B20
Density at 15°C (kg/m <sup>3</sup> )	0.830	0.8355	0.841
Calorific value (MJ/kg)	42.5	41.84	41.18
Kinematic viscosity at 40°C (mm <sup>2</sup> /s)	0.830	0.8355	0.841
Cetane number	48	48.35	48.68

Table 2: Input variables used for predicting exhaust emissions and performance.

No.	Input variables	Abbreviation	Source
1.	Throttle position (%)	Alpha	Sensor
2.	Air intake pressure (bar)	P_INTAKE	Sensor
3.	The speed engine (rpm)	SPEED	Sensor
4.	Cooling water temperature (°C)	T_WATER	Sensor
5.	Intake air temperature (°C)	T_INTAKE	Sensor
6.	Exhausted emissions temperature (°C)	T_EXHAUST	Sensor
7.	EGR temperature (°C)	T_EGR	Sensor
8.	EGR rate (%)	%EGR	Sensor
9.	Lambda	Lambda	Calculated

maximum engine power of 250 kW. Intake air temperature and pressure, exhaust emission, exhaust gas recirculation (EGR) are measured by K-sensor and residual pressure meter.

The control of the test-bed and engine was done by Bobcat 2.0 in which all signals from the test-bed and other modules in the laboratory were recorded. In this study, the engine performance and emissions of biodiesel blends (B10, B20) and diesel fuel (B0) were studied. Fuel properties of B0, B10 and B20 are presented in Table 1.

### Data Description

Descriptive statistics of observed variables including the maximum (Max), minimum (Min), mean and standard deviation (Std) are summarized in Table 3.

The experiment can collect and measure 26 kinds of exhaust emissions including NO, H<sub>2</sub>O, CH<sub>4</sub>, CO, NO<sub>2</sub>, N<sub>2</sub>O, NH<sub>3</sub>, SO<sub>2</sub>, HCHO, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>, NC<sub>8</sub>, C<sub>3</sub>H<sub>6</sub>, C<sub>4</sub>H<sub>6</sub>, NOx, HC, etc. As suggested in the standards developed by national and international bodies, this study focuses on investigating two kinds of emissions including NOx (ppm) and HC (ppm). The performance parameter predicted in this study is Torque (Nm) that is a measure of rotational effort applied on engine crankshaft by the piston. Torque in an engine is the amount of rotational force that the engine produces. Descriptive statistics of dependent variables are represented in Table 4. Fig. 5 includes correlation

Table 3: Descriptive statistics of observed variables.

No.	Input	Min	Max	Mean	Std.
B0					
1.	Alpha	9.94	44.87	26.3	9.246
2.	P_INTAKE	0	1.49	0.331	0.349
3.	P_EGR	0	1.46	0.334	0.34
4.	SPEED	1244.64	2753.25	1684.993	482.888
5.	T_WATER	65.55	96.18	82.438	6.21
6.	T_INTAKE	42.37	170.6	80.428	33.272
7.	T_EXHAUST	148.05	786.74	429.945	164.312
8.	T_EGR	44.9	348.8	139.243	64.642
9.	%EGR	0	56.592	31.349	21.363
10.	Lambda	0.529	5.758	1.951	1.135
B10					
1.	Alpha	14.91	39.82	27.55	6.652
2.	P_INTAKE	0.04	1.09	0.315	0.239
3.	P_EGR	0.05	1.07	0.313	0.234
4.	SPEED	1211.31	2501.56	1823.697	411.672
5.	T_WATER	75.11	111.81	91.879	8.148
6.	T_INTAKE	52.79	150.64	88.741	22.784
7.	T_EXHAUST	280.11	712.43	471.083	116.475
8.	T_EGR	50.9	330.2	143.235	60.123
9.	%EGR	0	44.084	21.784	16.639
10.	Lambda	0.658	3.845	1.605	0.629
B20					
1.	Alpha	9.94	44.86	26.479	9.94
2.	P_INTAKE	0	1.48	0.35	0
3.	P_EGR	0	1.45	0.347	0
4.	SPEED	1209.11	2753.11	1896.502	1209.11
5.	T_WATER	57.79	95.91	82.211	57.79
6.	T_INTAKE	38.75	170.93	84.775	38.75
7.	T_EXHAUST	143.4	775.15	423.069	143.4
8.	T_EGR	42.9	334.3	146.845	42.9
9.	%EGR	0	54.733	27.909	0
10.	Lambda	0.626	6.365	2.056	0.626

Table 4: Descriptive statistics of dependent variables.

Input	Min	Max	Mean	Std.
B0				
NOx	22.573	733.963	176.584	160.813
Torque	3.54	211.91	81.826	56.706
B10				
NOx	14.487	704.4	199.178	175.057
Torque	27.87	202.16	94.328	46.407
B20				
NOx	16.362	616.581	154.334	140.244
Torque	4.97	204.67	74.32	55.348

matrixes showing correlation coefficients between sets of variables. Each cell indicates the correlation between two variables.



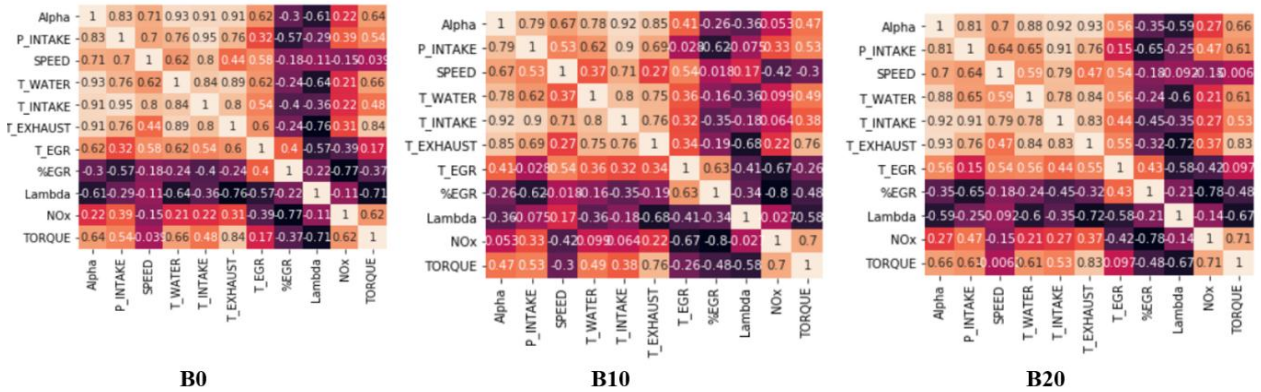


Fig. 5: Correlation matrix.

**Model Development**

The first step of the development of ML-based models is the preparation of the dataset. In this step, the collected dataset is divided into two groups - training and testing datasets. The training and testing dataset are used for the construction and validation of the models, respectively. In this study, 70% and 30% of the dataset is for training and testing, respectively. The next stage is designing the structure of the model and

adjusting the parameters. For example, in developing an ANN-based model, the number of hidden layers, the number of neurons per layer, and the type of transfer function need to be adjusted.

**Performance Evaluation Criteria**

To evaluate the performance of the developed models, several criteria are used. These criteria are applied to the

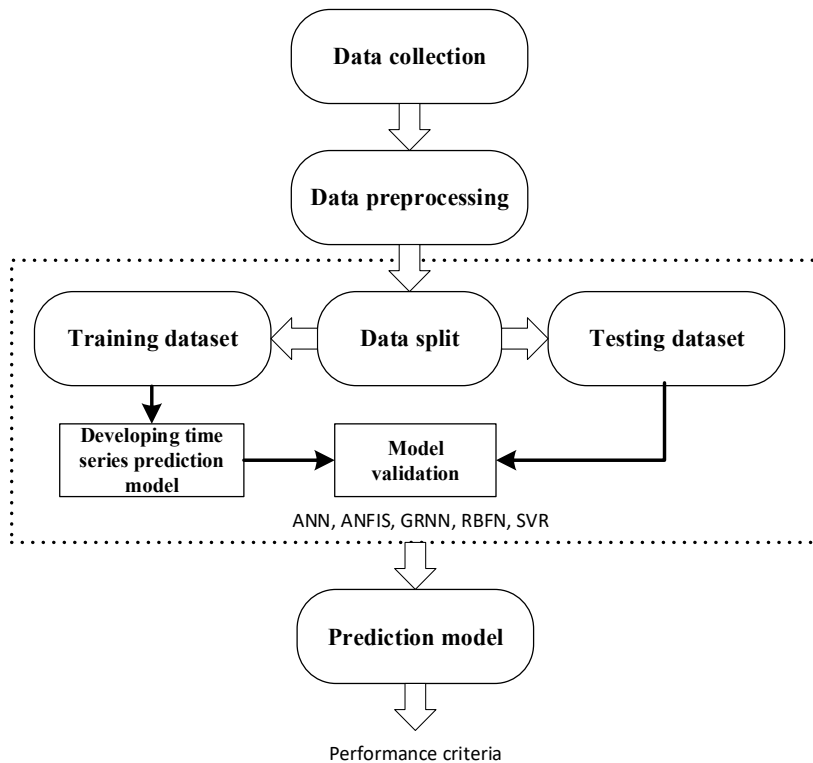


Fig 6: The model development.

model to know how well it worked. The criteria are used to compare predicted values and actual values. They are: Root mean squared error (RMSE), Mean absolute percentage error (MAPE), Mean absolute error (MAE) and Correlation coefficient ( ).

**RESULTS AND DISCUSSION**

All models were coded in Matlab 2015a environment. To avoid the over-fitting problem, the 10-fold cross-validation was utilized. For each technique, various sets of parameters were tried to obtain the best architecture of each classifying model. To evaluate the performance of the classifying model, several performance criteria were used. These criteria were applied to know how well the developed models worked. Several main parameters are set for each model as follows: the architecture of the ANN model is with two hidden layers (the number of hidden nodes are 12 and 6, respectively). The parameters were set as follows: the training algorithm is Levenberg-Marquardt; the number of epochs is set to 100.

For the ANFIS model, the number of fuzzy rules was 2. For the SVR model, the kernel function was Gaussian. For GRNN and RBFN models, the spread constants were 0.12 and 0.2, respectively.

Tables 5-7 show the performance statistics of all models. It is clearly seen that the SVR model has the smallest RMSE, MAPE, and MAE values as well as the biggest R-value. This means that the SVR has a better performance in all criteria.

The comparison between actual values and corresponding forecasting values obtained by the SVR model are shown in Figs. 7-9. The figures present prediction values and actual values. In all scenarios, the forecasting values tend to be close to the actual values. This indicates a good agreement between the prediction values obtained by the developed models and the actual values. This finding also indicates that SVR provides accurate forecasting results.

Based on the results, it can be inferred that the SVR model can be used as a suitable tool in the prediction of exhaust emissions and performance of a diesel engine fuelled

Table 5: Performance criteria on the testing set obtained by different techniques for B0.

Predicted parameter	Model	RMSE	MAPE	MAE	R
NOx	ANFIS	129.8523	0.3673	70.5089	0.7738
	ANN	95.6113	0.3857	51.7597	0.7756
	GRNN	164.2776	0.5149	92.1703	0.5135
	RBFN	212.6085	0.6062	120.7113	-2.5466e-16
	SVR	<b>71.6624</b>	<b>0.4577</b>	<b>47.9384</b>	<b>0.8572</b>
Torque	ANFIS	12.0361	0.1876	7.3354	0.9846
	ANN	10.8662	0.1268	6.0425	0.9840
	GRNN	20.3166	0.1147	11.0365	0.9332
	RBFN	72.2844	1.5029	58.4099	-3.7701e-16
	SVR	<b>8.3087</b>	<b>0.1887</b>	<b>6.5485</b>	<b>0.9875</b>

SVR model for B0

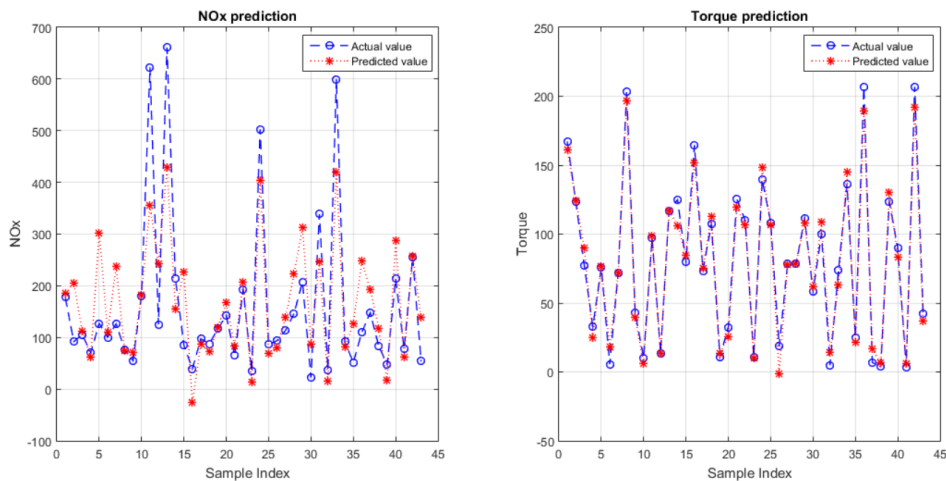


Fig. 7: Performance of SVR model for B0.

Table 6: Performance criteria on the testing set obtained by different techniques for B10.

Predicted parameter	Model	RMSE	MAPE	MAE	R
NOx	ANFIS	95.2373	0.4499	55.6865	0.8160
	ANN	87.6310	0.2139	47.9662	0.9139
	GRNN	153.8600	0.7406	102.2467	0.7405
	RBFN	241.2296	0.6787	162.5765	-1.9499e-16
	SVR	60.5215	0.7055	49.9064	0.9359
Torque	ANFIS	7.1567	0.3762	4.7942	0.9920
	ANN	9.1031	0.0883	5.3696	0.9889
	GRNN	11.6969	0.0881	6.0004	0.9803
	RBFN	74.7863	0.9549	57.7388	1.0034e-16
	SVR	6.5670	0.2766	5.6123	0.9944

SVR model for B10

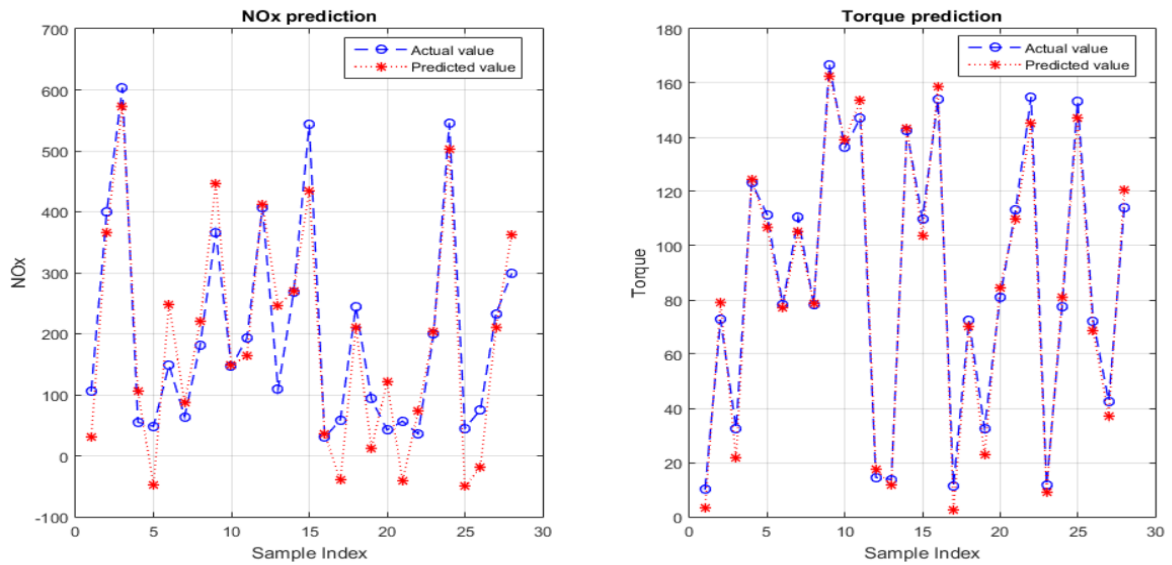


Fig. 8: Performance of SVR model for B10.

Table 7: Performance criteria on the testing set obtained by different techniques for B20.

Predicted parameter	Model	RMSE	MAPE	MAE	R
NOx	ANFIS	108.4337	0.8895	67.7203	0.7059
	ANN	83.5869	0.5381	49.9563	0.7040
	GRNN	99.4999	0.4343	53.3029	0.7706
	RBFN	150.7342	0.6662	85.4440	0.5930
	SVR	<b>60.8779</b>	<b>0.6113</b>	<b>49.1192</b>	<b>0.8910</b>
Torque	ANFIS	12.2276	0.1232	7.2107	0.9806
	ANN	3.8390	0.0813	2.6034	0.9973
	GRNN	12.5482	0.0682	5.2313	0.9840
	RBFN	79.5727	1.2715	62.7723	-1.7801e-16
	SVR	<b>7.9331</b>	<b>0.2259</b>	<b>6.1559</b>	<b>0.9917</b>



SVR model for B20

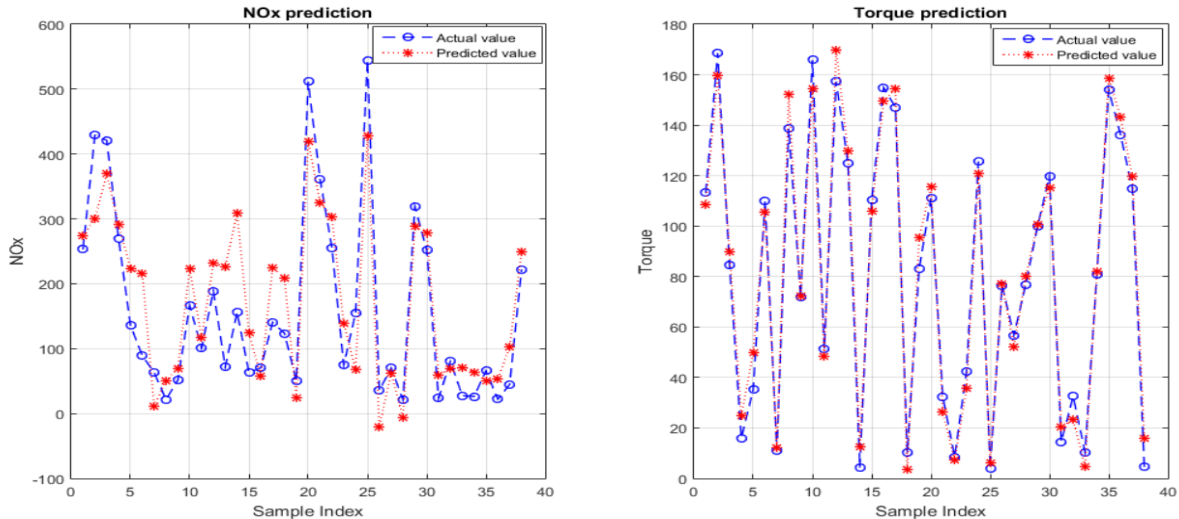


Fig. 9: Performance of SVR model for B20.

with biodiesel blends. Moreover, all the obtained results of developed models are highly correlated and precise. It can be concluded that the SVR-based model can be acceptable in serving as a forecasting method for operational parameters of the diesel engine.

## CONCLUSIONS

In the study, the forecasting models based on ANFIS, ANN, GRNN, RBFN and SVR techniques were proposed for forecasting exhaust emissions and performance of a diesel engine fuelled with biodiesel blends. In the case application, The B0, B10 and B20 biodiesel blends were used in the Hyundai D4CB 2.5 engine. The engine operational parameters were used as inputs and the outputs include forecasting values of the torque and NOx emission. The obtained results showed that the SVR outperformed in all scenarios. These findings also demonstrated the remarkable advantage of ML-based techniques in forecasting for particular applications. The results of the present study also show that a comparative analysis of different techniques is always supportive in enhancing the performance of a forecasting model. This work can thus contribute to the development of selecting the best forecasting model for other purposes. For future research, we will develop and incorporate our forecasting model in a user-friendly software tool to make this forecasting task easier.

## REFERENCES

Abe, S. 2005. Support Vector Machines for Pattern Classification. Springer, London, pp. 44.

- Agarwal, D., Kumar, L. and Agarwal, A.K. 2008. Performance evaluation of a vegetable oil fuelled compression ignition engine. *Renewable energy*, 33(6): 1147-1156.
- Alilou, V. K. and Yaghmaee, F. 2015. Application of GRNN neural network in non-texture image inpainting and restoration. *Pattern Recognition Letters*, 62: 24-31
- Ata, R. and Koçyigit, Y. 2010. An adaptive neuro-fuzzy inference system approach for prediction of tip speed ratio in wind turbines. *Expert Systems with Applications*, 37(7): 5454-5460.
- Batool, F., Adeel, S., Azeem, M., Khan, A. A., Bhatti, I.A., Ghaffar, A. and Iqbal, N. 2013. Gamma radiations induced improvement in dyeing properties and color fastness of cotton fabrics dyed with chicken gizzard leaves extracts. *Radiation Physics and Chemistry*, 89: 33-37.
- Canakci, M., Ozsezen, A.N., Arcaklioglu, E. and Erdil, A. 2009. Prediction of performance and exhaust emissions of a diesel engine fueled with biodiesel produced from waste frying palm oil. *Expert systems with Applications*, 36(5): 9268-9280.
- Caruana, R., Lawrence, S. and Giles, C.L. 2001. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In: *Advances in Neural Information Processing Systems*, pp. 402-408.
- Cigizoglu, H.K. and Alp, M. 2006. Generalized regression neural network in modelling river sediment yield. *Advances in Engineering Software*, 37(2): 63-68.
- Do, Q.H. and Chen, J.F. 2013. A neuro-fuzzy approach in the classification of students' academic performance. *Computational Intelligence and Neuroscience*, 2013.
- Do, Q.H., Tuan, T.T., Ha, L.T.T., Doan, T.T.H. and Nguyen, T.V.A. 2020. Development of artificial neural networks trained by heuristic algorithms for prediction of exhaust emissions and performance of a diesel engine fuelled with biodiesel blends. In: *Applied Nature-Inspired Computing: Algorithms and Case Studies* (pp. 275-253). Springer, Singapore.
- Fakhrhmad, S.M., Rezapour, A.R., Jahromi, M.Z. and Sadreddini, M.H. 2012. A new fuzzy rule-based classification system for word sense disambiguation. *Intelligent Data Analysis*, 16(4): 633-648.
- Ghanbari, M., Najafi, G., Ghobadian, B., Mamat, R., Noor, M.M. and Moosavian, A. 2015. Support vector machine to predict diesel engine performance and emission parameters fueled with nano-particles additive to diesel fuel. In *IOP Conference Series: Materials Science and Engineering*. IOP Publishing.

- Guan, X., Zhu, Y. and Song, W. 2016. Application of RBF neural network improved by peak density function in intelligent color matching of wood dyeing. *Chaos, Solitons & Fractals*, 89: 485-490.
- Hornik, K., Stinchcombe, M. and White, H. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5): 359-366.
- Jang, J.S.R., Sun, C.T. and Mizutani, E. 1997. Neuro-fuzzy and soft computing - A computational approach to learning and machine intelligence [Book Review]. *IEEE Transactions on automatic control*, 42(10): 1482-1484.
- Jang, J.S. 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3): 665-685.
- Ji, X., Shang, X., Dahlgren, R.A. and Zhang, M. 2017. Prediction of dissolved oxygen concentration in hypoxic river systems using support vector machine: a case study of Wen-Rui Tang River, China. *Environmental Science and Pollution Research*, 24(19): 16062-16076.
- Jiang, J., Cao, D. and Chen, H. 2016. Boundary value problems for fractional differential equation with causal operators. *Applied Mathematics and Nonlinear Sciences*, 1(1): 11-22.
- Funahashi, K.I. 1989. On the approximate realization of continuous mappings by neural networks. *Neural networks*, 2(3): 183-192.
- Kim, S. and Kim, H.S. 2008. Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *Journal of hydrology*, 351(3-4): 299-317.
- Liu, Z., Zuo, Q., Wu, G. and Li, Y. 2018. An artificial neural network developed for predicting of performance and emissions of a spark ignition engine fueled with butanol-gasoline blends. *Advances in Mechanical Engineering*, 10(1).
- Mamilla, V.R. and Rao, G.L.N. 2016. Optimal performance and emission analysis of diesel engine fuelled with palm oil methyl ester with an artificial neural network. *American Journal of Modern Energy*, 2(4): 17-21.
- Meher, L.C., Sagar, D.V. and Naik, S.N. 2006. Technical aspects of biodiesel production by transesterification - A review. *Renewable and Sustainable Energy Reviews*, 10(3): 248-268.
- Meyer, D., Leisch, F. and Hornik, K. 2003. The support vector machine under test. *Neurocomputing*, 55(1-2): 169-186.
- Nauck, D., Klawonn, F. and Kruse, R. 1997. *Foundations of Neuro-Fuzzy Systems*. John Wiley & Sons, Inc.
- Noor, R.M. 2014. Recent developments of neural networks in biodiesel applications. In: *International Conference on Swarm, Evolutionary and Memetic Computing*, pp. 339-350, Springer, Cham.
- Ramadhas, A. S., Jayaraj, S. and Muraleedharan, C. 2005. Biodiesel production from high FFA rubber seed oil. *Fuel*, 84(4): 335-340.
- Shailaja, M. and Raju, A.S.R. 2017. Neural network - Based diesel engine emissions prediction for variable injection timing, injection pressure, compression ratio and load conditions. In: *Emerging Trends in Electrical, Communications and Information Technologies*. Springer, Singapore, pp. 109.
- Singh, T.N., Kanchan, R., Verma, A.K. and Saigal, K. 2005. A comparative study of ANN and neuro-fuzzy for the prediction of dynamic constant of rockmass. *Journal of Earth System Science*, 114(1): 75-86.
- Sugeno, M. 1985. An introductory survey of fuzzy control. *Information sciences*, 36(1-2): 59-83.
- Takagi, H. and Hayashi, I. 1991. NN-driven fuzzy reasoning. *International Journal of Approximate Reasoning*, 5(3): 191-212.
- Van Gerpen, J. 2005. Biodiesel processing and production. *Fuel Processing Technology*, 86(10): 1097-1107.
- Wei, M., Bai, B., Sung, A. H., Liu, Q., Wang, J. and Cather, M.E. 2007. Predicting injection profiles using ANFIS. *Information Sciences*, 177(20): 4445-4461.
- Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J. and El-Shafie, A. 2016. Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. *Journal of Hydrology*, 542: 603-614.
- Yusof, N., Ahmad, N.B., Othman, M.S. and Mohammad, F.A. 2012. A concise fuzzy rule base to reason student performance based on rough-fuzzy approach. In *Fuzzy inference system-theory and applications*. INTECH Open Access Publisher, pp. 63.